SOME ELECTRON DIFFRACTION STUDIES

I. ELECTRON DIFFRACTION STUDIES OF SOME HEAVY METAL HEXAFLUORIDES

II. THE MOLECULAR STRUCTURE OF GASEOUS N_2O_4

III. APPLICATIONS OF THE METHOD OF LEAST SQUARES IN ELECTRON DIFFRACTION INVESTIGATIONS

Thesis by

Darwin Waldron Smith

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1959

ACKNOWLEDGEMENTS

It is with great pleasure that I acknowledge the friendly, inspiring guidance and valuable criticism of my research supervisor, Professor Verner Schomaker.

I wish to thank Dr. Kenneth Hedberg for suggesting that I work on his N_2O_4 data which he gathered in Norway, and for several discussions concerning the least squares refinement of sector data. To both cand. real. Lise Hedberg and Dr. Kenneth Hedberg go my thanks for their guidance in the reduction of the N_2O_4 data.

I am very glad to acknowledge collaboration with Dr. Masao Kimura in the refinement of the hexafluoride compounds, and I wish to thank him for our many friendly discussions.

To Professor Jurg Waser and Dr. Richard E. Marsh I express my thanks for reading and criticising several sections of this thesis. My appreciation also goes to Professors Linus Pauling, Harden McConnell, Richard M. Radger and others for helpful discussions on several phases of my work.

I wish to thank Dr. Robert Nathan, Dr. Joel N. Franklin, and Mr. Kindrick J. Hebert for discussions relating to the programming of the Burroughs 205 computer.

To the California Institute, the Allied Chemical and Dye Co., the Eastman Kodak Co., and the E. I. du Pont du Nemours Co. I express my sincere appreciation for the scholarships and fellowships which have made graduate school possible.

My thanks also go to the many staff members and graduate students of the Institute who have made my years here so pleasant

and rewarding.

Finally I wish to dedicate this thesis to my wife Jo Ann, whose love and encouragement mean so much, to our two daughters Shelly and Lorraine, and to our wonderful parents.

ABSTRACT

I. An electron diffraction investigation has been made by the sector-microphotometer method of WF_6, OsF_6, IrF_6, UF_6, NpF_6 and PuF_6. The photographs of all these compounds exhibit a phase shift anomaly which gives rise to apparently assymetric structures, the anomaly setting in at smaller angles θ of scattering for the heavier molecules and at smaller values of $q = (40/\lambda)\sin\theta/2$ the greater the electron wavelength. There is good evidence for the symmetrical octahedral structure of all the compounds. The metal-fluorine distances were found to be (in Angstrom units) 1.833(W-F), 1.831(Os-F), 1.830(Ir-F), 1.996(U-F), 1.981(Np-F), and 1.971(Pu-F).

II. An electron diffraction molecular structure determination of gaseous N_2O_4 has shown that the N-O distance, the N-N distance, and the ONO bond angle are 1.177 A, 1.752 A, and 134.8° respectively. The molecule is planar, which is surprising in view of the long N-N bond.

III. A structure refinement procedure, which was applied to the above compounds, has been programmed for the Burroughs 205 computer. It involves the fitting of a theoretical intensity curve to the experimental one.

TABLE OF CONTENTS

I. ELECTRON DIFFRACTION STUDIES OF SOME HEAVY METAL HEXAFLUORIDES

An electron diffraction investigation has been made of six gaseous heavy metal hexafluorides: those of tungsten, osmium*, iridium, uranium, neptunium, and plutonium. Early electron diffraction work of UF_6 and WF_6 had resulted in anomalous unsymmetrical structures (2), (3), while spectroscopic observations (4) - (7) had favored the symmetrical octahedral structure. However, in 1952 Schomaker and Glauber (8) showed that failure of the first Born approximation, used in computing scattering amplitudes, had led to erroneous electron diffraction results for these and certain other compounds containing both heavy and light atoms. Subsequently UF_6 (9) and WF_6 (10) have been successfully interpreted in terms of a symmetrical model by using angle-dependent phase shifts in the scattering amplitudes.

The present investigation is of interest in confirming the symmetrical octahedral structure of the other four hexafluorides and in continuing the study of the phase shift by providing more extensive data at various electron wavelengths for all six compounds. This investigation was instigated by Dr. Bernard Weinstock of the Argonne National Laboratory who supplied all of the compounds except WF_6. The WF_6 was kindly supplied by Dr. John M. Teem of this Institute. The samples were considered to be sufficiently pure for structure determination.

---

*The volatile yellow solid OsF_6 has been incorrectly called OsF_8 for many years. However Weinstock and Malm (1) have shown by molecular weight determination and chemical analysis that it is the hexafluoride and that the compound previously called OsF_6 was also identified incorrectly.

Table I lists the electron accelerating potentials, wave lengths, and sample temperatures for each compound photographed. The photographs were prepared by the sector method by Professor Schomaker, Dr. Weinstock and Dr. Kimura. Dr. Kimura prepared microphotometer tracings and reduced them to experimental intensity curves. The experimental curves were interpreted by the radial distribution method and by a least squares refinement procedure in which theoretical intensity curves are adjusted to the experimental curves*.

Theory

The scattering of electrons by gas molecules containing both heavy and light atoms has been discussed by Schomaker and Glauber (8), and Hoerni and Ibers (9). For electrons of wave length λ scattered at angle θ, the expression for the total scattered intensity is

$$I_t(s) = B(s) + K\sum_{ij}{}' |f_i(s)||f_j(s)| \cos \Delta\eta_{ij}(s) e^{-\sigma_{ij}^2 s^2/2} \frac{\sin r_{ij}s}{r_{ij}s} \qquad (1)$$

where $s = (4\pi/\lambda) \sin \theta/2$; B is the background or structure independent scattering; K is a constant; $f_j = |f_j| e^{i\eta_j(s)}$ is the complex scattering amplitude (for electrons) of atom j; $\Delta\eta_{ij} = \eta_i - \eta_j$; σ_{ij} is the root mean square variation of r_{ij}, the distance between atoms i and j. The second term (the molecular part) in the right member of (1) may be called I_m so that we have $I_t = B + I_m$.

For precise sector work, especially where a determination of vibrational parameters is desired, theoretical intensity curves should

---

*I gratefully acknowledge collaboration with Dr. Kimura in interpreting the data and refining the results.

TABLE I

ELECTRON ACCELERATION POTENTIALS USED IN MAKING ELECTRON DIFFRACTION PICTURES

Compound	Sample temp.	Potential (kev) and wavelength (A)						
		v = 47.040 λ = 0.0553	v = 39.772 λ = 0.0603	v = 33.600 λ = 0.0658	v = 23.520 λ = 0.0791	v = 11.760 λ = 0.1125	v = 11.468 λ = 0.1139	v = 5.880 λ = 0.1595
WF_6	-55°	x	x	x	x	x		x
OsF_6	-39°		x			x		
IrF_6	-42°		x				x	
UF_6	-25°	x	x				x	
NpF_6	-23°	x	x				x	
PuF_6	-27°	x	x				x	

be calculated from an expression of the form

$$sI(s) \equiv \frac{s(I_t - B)}{K|f_k||f_\ell|} = \sum_{ij}{}' \frac{|f_i||f_j|}{r_{ij}|f_k||f_\ell|} \cos \Delta\eta_{ij} e^{-\sigma_{ij}^2 s^2/2} \sin r_{ij}s. \qquad (2)$$

In this expression $f_k(s)$ and $f_\ell(s)$ may be scattering factors for a particular pair of atoms in the molecule or may be some sort of "average" scattering factors for the molecule (averaged in some way with respect to the atoms). They are introduced in order to make the coefficients in the Fourier sum nearly constant. Such a theoretical intensity curve has the familiar appearance of a damped sine series. In the present case as explained later, it is impossible to determine absolute vibrational parameters; therefore it was considered adequate to calculate theoretical curves from the expression

$$sI(s) = \sum_{ij} \frac{C_{ij}}{r_{ij}} \cos \Delta\eta_{ij}(s) e^{-\sigma_{ij}^2 s^2/2} \sin r_{ij}s, \qquad (3)$$

where C_{ij} is independent of s.

In the case of a symmetrical octahedral metal hexafluoride (MF_6), we have

$$\begin{aligned} sI(s) = {} & \frac{6C_{M\text{-}F}}{r} \cos \Delta\eta \sin rs\, e^{-\sigma_{M\text{-}F}^2 s^2/2} \\ & + \frac{12C_{F\cdot F}}{\sqrt{2}\, r} \sin \sqrt{2}\, rs\, e^{-\sigma_{F\cdot F}^2 s^2/2} \\ & + \frac{3C_{F\cdot F}}{2r} \sin 2rs\, e^{-\sigma_{F\cdot\cdot F}^2 s^2/2} \end{aligned} \qquad (4)$$

where $r = r_{M\text{-}F}$ and $\Delta\eta = \Delta\eta_{M\text{-}F}$. This expression leads to an

apparent split M-F distance, for if we make the approximation $\Delta\eta = \delta s$,

$$\frac{6C}{r} \cos \Delta\eta \sin rs = \frac{3C}{r} \sin (r + \delta)s + \frac{3C}{r} \sin (r - \delta)s \tag{5}$$

$$= \frac{3C}{r+\delta} \sin (r + \delta)s + \frac{3C}{r-\delta} \sin (r - s)s + O(\delta/r^2).$$

The apparent split is $\Delta r = 2\delta$, and early anomalous structures were reported with just such splits. The critical region in the pattern, known as the cut-off point, occurs at that value of s for which $\Delta\eta = \pi/2$. In the split model the two sine terms beat out at $s_c = \pi/2\delta$, and again at $s'_c = 3\pi/2\delta$. The split model can be used for approximate theoretical curves out to about its second cut-off point, but since the data extend beyond $s = 3\pi/2\delta$, (and do not show a second cut-off point in this range), a more accurate treatment than the split model was needed. This will be discussed under "Calculation of Theoretical Curves."

Experimental Intensity Curves

The electron diffraction photographs were made on Kodak process plates, developed in Microdol. The microphotometer used to obtain density data is the Chemistry Division's Sinclair Smith model, made by the Fred C. Henson Company, Pasadena, in which the radiometer detecting unit has been replaced by an RCA 1P39 phototube. An R-C electronic filter circuit serves to damp out part of the noise due to photographic imperfections. The signal is amplified by an Allied Physics Corporation vibrating reed electrometer*, and traced by a

*We wish to thank Professor Thomas Lauritsen for loaning us the electrometer.

Brown potentiometer. The illumination is by parallel light, the lamp being powered by storage batteries.

The photometer has high stability; the drift is negligible under the conditions of illumination and amplification present in making a trace providing the lamp is warmed up for about two hours and the phototube is warmed up by being illuminated with a low light level for 30 minutes to an hour. Exact tests of phototube linearity were not necessary because small deviations are absorbed in the unknown non-linearity of the photographic emulsion, discussed below. That the system is approximately linear is shown by the fact that doubling the slit width showed approximately double the phototube current. Indeed, a plate traced twice at slit widths differing by a factor of two yielded nearly identical density curves.

The slit width was 0. 2 mm as measured on the plate except for low potential (long wavelength) pictures, where, because of the broader rings, 0. 4 mm was used. Such a width and a slit length of 2 mm introduced only a small broadening of features in the pattern. That this broadening is indeed small may be seen by considering the highest frequency to be expected in the pattern: that due to the long F· · · F bond of about 4 A. This bond contributes to the pattern a term of period about 1. 4 mm on the plate when the wavelength is 0. 056 A, the shortest wavelength used. The plate speed was the slowest available, 1. 4 mm/min., and the recorder chart speed was 10 ft/hr.

Pip marks were superimposed over the full length of the trace by impressing the recorder pen with a voltage pulse every 0. 5 mm translation of the plate. This pulse came from a microswitch driven

by an arm on the microphotometer screw. These pip marks enabled distances to be found without the possibility of errors caused by the paper shrinking or stretching.

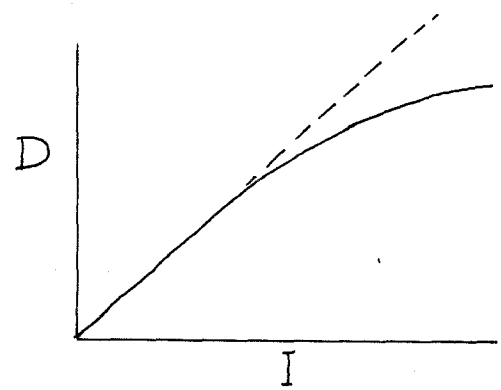
The plates were spun on a rotating stage to average out graininess and defects in the emulsion and to reduce the effect of possible unsymmetrical spurious background caused by scattering from the apparatus. The centering of a plate on the rotating stage was done with the aid of the beam stop shadow. The error in centering the diffraction pattern includes two effects. First, the error in centering the beam stop shadow is about 0.1 mm. The electron beam can be centered on the axis of the sector with an error of about 0.1 mm. The beam stop rotates with the sector, therefore its shadow is concentric with the axis of the sector. Thus the error in centering the pattern on the rotating stage is probably less than 0.2 mm.

Two types of microphotometer traces were obtained from each plate. In one, the range from zero to 100% transmission was recorded. In the second, a greater amplification and a shift of the zero line resulted in an expanded scale, making readings more accurate. The first type of trace was used merely to locate the zero and 100% lines on the expanded record.

The plates were traced along the full diameter of the pattern, giving two identical halves. The center of the trace could be located to well within 0.1 mm either by measuring the positions of maxima and minima on the trace or by making a mark in the clear glass near the center of the beam stop shadow before tracing the plate. A distance scale was marked off on the trace and readings were taken at every

0. 2 mm (as measured on the plate) except for the low potential pictures where 0. 4 mm sufficed. This interval is of the same order of magnitude as the slit width and as the uncertainty in centering. Microphotometer transmission readings were converted to densities.

In sector work it is desirable to convert densities into relative electron intensities as accurately as possible. At low exposures the density D and intensity I are proportional; a plot of D vs. I is a straight line. At higher exposures, the D vs. I curve falls below the straight line:



and a positive correction must be applied to the density. Other corrections which might be applied include a factor of $1/\cos^2\theta$ to correct for the increased distance from the scattering center to the outer parts of the plate, and, assuming a thick emulsion, a factor of $1/\cos\theta$ because of its angle of tilt with respect to a diffracted ray. In order to correct for the non-linearity of the photographic emulsion, one needs two plates at different exposures of the same exact diffraction pattern, developed together (11). Unfortunately in the present work the electron accelerating voltage was changed with every plate in a set because of limited amounts of some of the samples, because of the primary purpose of determining the effect of the voltage on the cut-off point, and because of the necessary overall effective haste in which the work was done.

Therefore the above described correction for the non-linearity of the photographic emulsion could not be made and the amplitudes of the rings in the dense inner part of a picture are too small compared with the amplitudes of the light outer rings (before dividing by the background). Thus the experimental intensity curves (strictly density curves) are in error by being multiplied by an unknown but smooth modification function. This is not really serious however, because although it will render impossible the determination of absolute vibrational parameters, it will not affect bond distances or the position of the cut-off point. Our results should certainly be more accurate than visual results because we still have quantitative measurements of shapes of diffraction rings and of relative amplitudes of a number of neighboring maxima and minima.

However, another method for at least partially correcting for the non-linearity of the emulsion as well as the $1/\cos^3\theta$ effect is to divide by B. The procedure therefore was as follows: densities obtained from microphotometer traces were corrected for the presence of the sector. The total intensity curve I_t thus obtained as a function of r was then plotted and a scale of $q = \frac{40}{\lambda} \sin \frac{\theta}{2} = \frac{10s}{\pi}$ was marked off on it. The intensities were then read off at integral values of q. A smooth experimental background B was drawn in by hand and subtracted from I_t. The result $I_t - B$ was divided by the experimental background as is the general practice in sector work*. When photographic density corrections are possible it seems better to divide by

*See, for example, reference (12).

the theoretical background if it is known, or by the product of scattering powers for a particular pair of atoms as suggested previously under "Theory". An experimental background line is certain to contain errors of judgement in drawing it as well as errors caused by spurious background scattering. Since the theoretical background is not known exactly, especially for our samples, it is difficult to say what the best procedure is. Dividing by the experimental background is perhaps justified in view of the possibility that under ideal experimental conditions, the background obtained may be more nearly correct than the theoretical background, and may approach the shape of the true scattering factors closer than theoretical ones do. In the present experiments, the background was not ideal, but the background used in dividing was monotonic and was adjusted, where necessary, to conform to the general shape of the theoretical background. Let this adjusted divisor be B'. Any detailed deviation of B' from a more proper divisor is of the same general nature as the non-linearity of the emulsion, and dividing by B or B' compensates somewhat for this non-linearity and also for the $1/\cos^3\theta$ factor. The above considerations and the decision to approximate C_{ij} by a constant in the expression (equation 3) for the theoretical curves probably justify the foregoing procedure, but we must keep in mind the possibility that I_m has been multiplied by a non-constant modification function.

The final experimental curves were obtained by multiplying the result of the above procedure by s:

$$sI(s) = \frac{s(I_t - B)}{B'} \qquad (6)$$

to conform with equation 2. Radial distribution (RD) curves were obtained for all of the compounds at several voltages. These RD curves support the octahedral model (see "Evidence for Octahedral Structure") and show two M-F peaks separated by the apparent split distance. From the M-F distance, the apparent split, and the approximate temperature factors as read from the RD curves, preliminary theoretical curves were obtained. They were used as a guide in re-adjusting the background where necessary.

The sector deviates markedly from its intended shape at 1.1 - 1.2 cm and at 1.3 - 1.5 cm. It was designed to have the angular opening α proportional to R for R greater than 1.27 cm. For R less than this value the opening is greater in order to give a flat background with carbon compounds. The turnover point on the sector corresponds to 1.35 cm on the plate, one of the points of difficulty. The sector calibration curve, (a plot of R^3/α), was obtained from readings made with a travelling microscope. For small α large errors are introduced in R^3/α and it is difficult to draw a curve through the widely scattered calibration points.

Visual comparison of non-sector plates with sector plates or with curves obtained from sector plates shows a definite discrepancy near 1.4 cm. Theoretical intensity curves also show the same discrepancy. Therefore, in the process of refining models, the region of the sector calibration curve near 1.4 cm was redrawn a number of times. The attempt was made to obtain a calibration curve that would fit the calibration points fairly well and yet would lead to experimental intensity curves having reasonable agreement with theoretical curves.

Somewhat of a compromise had to be made in the attempt to satisfy these two requirements. The final sector calibration curve that was used for all of the pictures is shown in Fig. 1 along with the measured calibration points and the original curve through the points.

Fig. 1 ⟶

The original (broken line) and final (solid line) sector calibration curves near 1.4 cm. The radius R corresponds to distances on the photographic plate, and α is the angular sector opening in radians.

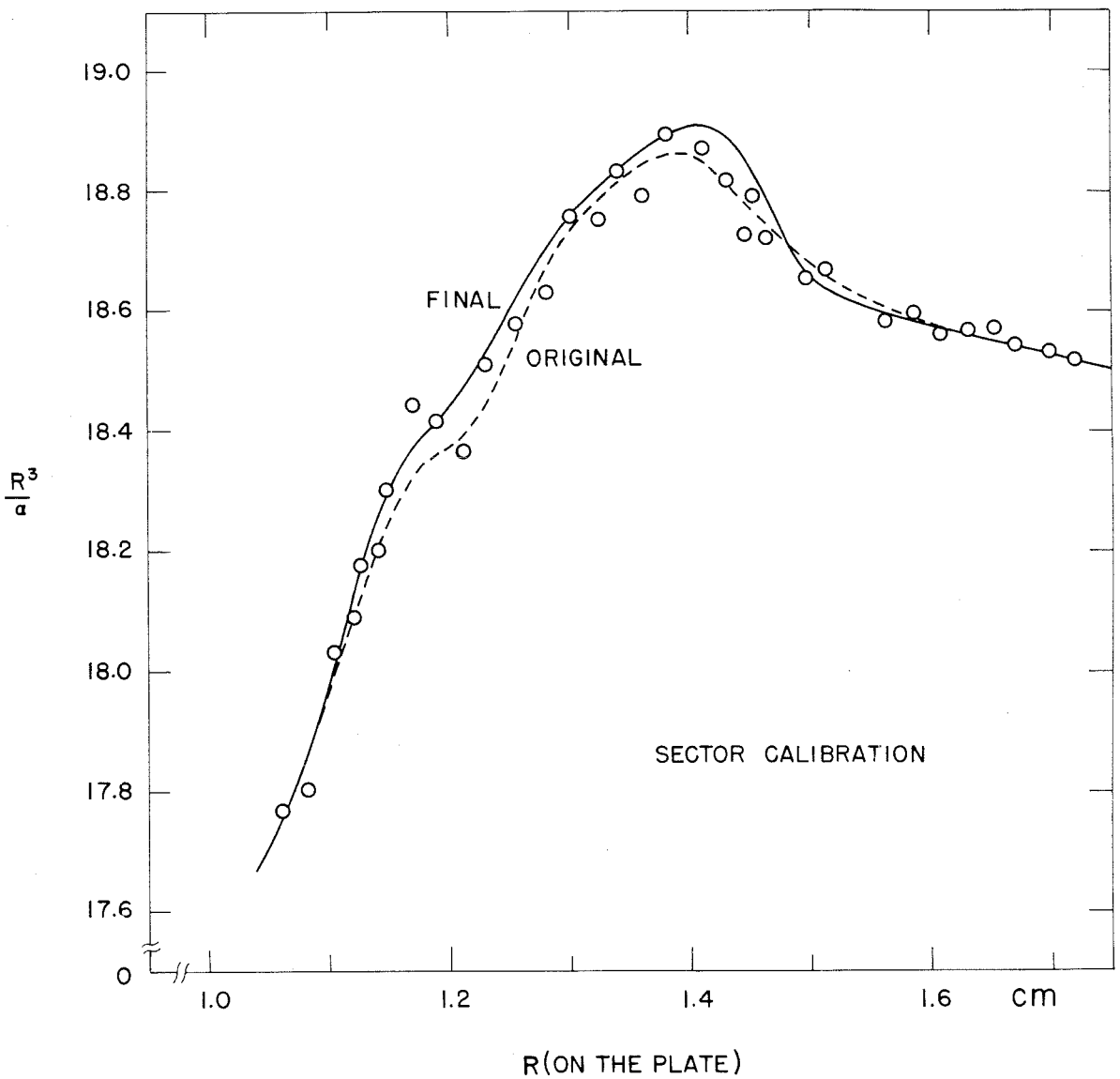

19.0
18.8
18.6
18.4
18.2
18.0
17.8
17.6
0
$\frac{R^3}{a}$
FINAL
ORIGINAL
SECTOR CALIBRATION
1.0
1.2
1.4
1.6
cm
R(ON THE PLATE)

Fig. 1

Calculation of Theoretical Curves

A program for calculating theoretical intensity curves on the Burroughs 205 digital computer is described in the appendix. This is a basic program, written primarily for the calculation of intensity curves I_c of the form

$$sI(s) = I_c = \sum_i A_i e^{-\sigma_i^2 s^2/2} \sin r_i s \qquad (7)$$

where A_i does not depend on s, and the double summation over atoms of equation 2 has been replaced by a single summation over pairs of atoms, i being the i^{th} pair of atoms or r_i the i^{th} interatomic distance. This simplified expression, (equation 7), is satisfactory for most of the compounds studied by electron diffraction, but could not be used without modification in the present investigation. The reason is that the anomalous phase shift requires the use of the factor $\cos \Delta\eta$ in the M-F term as in equation 4. Ibers and Hoerni (13) have calculated for 39.47 kev electrons the atomic scattering amplitudes including the angle-dependent phase shifts $\eta(\theta)$. These calculations were based on the partial-waves scattering theory and were carried out for selected atoms. In Fig. 2 are plotted values of $\Delta\eta = \eta_M - \eta_F$ as a function of $q = \frac{10s}{\pi} = \frac{40}{\lambda} \sin\frac{\theta}{2}$ for the metals with lowest and highest atomic numbers of interest here, tungsten and plutonium. Curves for the other metals fall between these two curves. All these curves have approximately the same shape.

Fig. 2

Graph of the dependence of the phase shift on scattering angle:

$\Delta\eta(\theta)$ vs. $q = \frac{40}{\lambda} \sin \frac{\theta}{2}$ for 40 kev electrons.

Pu-F and W-F, (solid lines): from calculations of Ibers and Hoerni (13),

A, (broken line): $\Delta\eta_{W-F}$ approximated by equation 9.

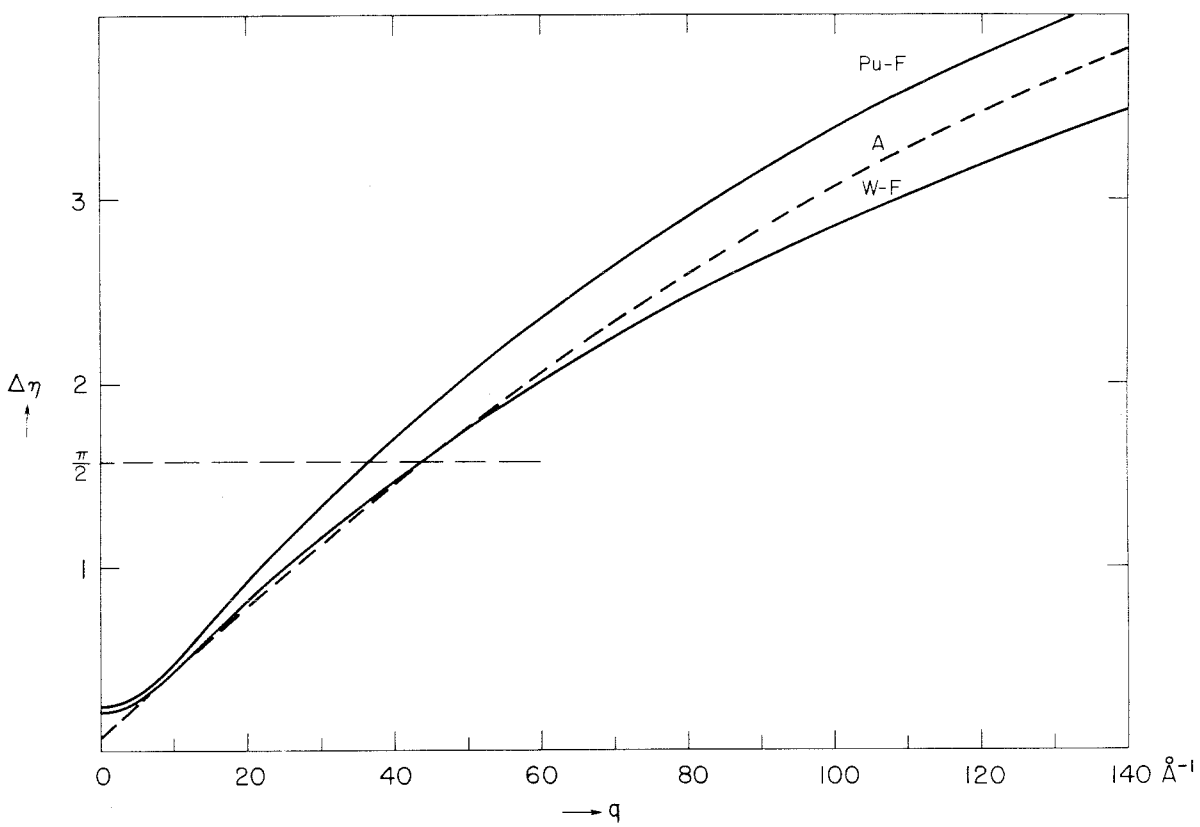


Fig. 2

Their calculations applying directly only to $V = V_o = 39.47$ kev, Ibers and Hoerni proposed the following approximation for V not too different from V_o:

$$\eta(Z, \theta, V) = \eta(Z', \theta', V_o) \quad \text{if}$$

$$Z' = Z(v_o/v); \quad \sin(\theta'/2) = (Z'/Z)^{1/3}(\lambda_o/\lambda) \sin(\theta/2) . \tag{8}$$

Here, v is the velocity of the electrons, $v_o = 1.1148 \times 10^{10}$ cm/sec, and $\lambda_o = 0.06056$ A. It is thus seen that curves of $\Delta\eta$ vs. q for various voltages also have the same general shape.

An analytical expression was used to approximate $\Delta\eta$ in the calculation of intensity curves. The cut-off point $q_c = 10s_c/\pi$ could then be used as an adjustable parameter to be determined from the data and compared with the theoretical predictions. It was convenient to use the expression

$$\Delta\eta(q) = \frac{\pi}{2} + 2\pi[a(q - q_c) + b(q - q_c)^2] \tag{9}$$

with three adjustable parameters, q_c, a, and b. (By $\Delta\eta(q)$ is meant $\Delta\eta(\theta)$ expressed as a function of q.) The data, however, do not justify more than a determination of q_c and a very rough determination of a (which fixes the gradient $\left.\frac{\partial\Delta\eta(q)}{\partial q}\right]_{q=q_c} = 2\pi a$ at cut-off), and it was decided to fix b at $b = -0.0000129$ for every compound at every voltage. This value had been chosen in a preliminary refinement of NpF_6-47 kev to make the second cut-off approximately correct for that particular set of data. It gives roughly the correct curvature to $\Delta\eta$. Any error in b cannot affect r or q_c significantly, and, al-

though it does affect the absolute temperature factors, it should have little effect on their ratios.

In Fig. 2 is included a plot of $\Delta\eta$ as calculated from equation 9, with a and q_c taken from the final least squares refinement of WF_6-40 kev. The deviation from the theoretical $\Delta\eta$ curve for W-F is less than 10% over the whole range of q, with very good agreement near $q = q_c$.

The theoretical intensity curves program was modified to calculate $\cos \Delta\eta$ by equation 9 and to form I_c from the expression

$$I_c = \sum_i A_i e^{-\sigma_i^2 s^2/2} \cos \Delta\eta \sin r_i s \qquad (10)$$

The amplitude coefficient A_i (or $\frac{C_{ij}}{r_{ij}}$ in equation 3) was determined by plotting the ratio $\frac{|f_i||f_j|}{B_{\text{Theoretical}}}$ as a function of q. Although it was not constant, the value of this ratio near $q = 60$ was used to obtain A_i. Any error introduced by assuming that A_i is constant will again affect mainly the absolute values of the temperature factors.

To examine the effect of varying the A_i, WF_6-40 kev was refined by least squares using three choices of the ratio $A_1:A_2:A_3$. The results are listed in the following table, the relative value of A_1 being 1000.

A_2	A_3	σ_1	σ_2	σ_3	q_c	a
300	55	0.0379	0.0944	0.0654	43.9	0.00503
285	50	0.0386	0.0936	0.0650	43.9	0.00496
174	30	0.0395	0.0665	0.0356	43.7	0.00414

The subscripts 1, 2, and 3 refer to W-F, F·F, and F··F respectively; σ_i and a are in Angstrom units and q_c is in reciprocal Angstrom units. In the first row the amplitude ratio corresponds to A_i obtained from the scattering amplitudes of Ibers and Hoerni at q = 50; the second row for q = 65; and in the last row the ratio was estimated from atomic numbers. As expected, the results are insensitive to small changes in the ratio $A_1:A_2:A_3$. There was no change in r which was 1.830 A in every case, and the agreement in q_c is essentially perfect. In the first two rows the differences in the temperature factors are considerably less than their standard deviations. The use of atomic numbers to obtain the scattering amplitudes is seen to cause larger differences, however. The amplitude ratios used in the final least squares refinements were 1000:285:50 for WF_6, OsF_6, and IrF_6; and 1000:278:50 for UF_6, NpF_6, and PuF_6.

Least Squares Refinement

In Part III a basic least squares program is described for the refinement of electron diffraction structures. This program minimizes the weighted sum of squares of the differences $I_e - I_c$ between experimental and calculated intensity curves by adjusting the distances, temperature parameters and amplitude scale factor in I_c. Overlays were written to modify this program for the special case of heavy atom hexafluorides. Since it was desired to adjust q_c and a by least squares, the computer was programmed to calculate the necessary partial derivatives. If I_{M-F} is the M-F contribution to I_c, then

$$I_{M\text{-}F} = \frac{6C_{M\text{-}F}}{r} \cos \Delta\eta \sin rs \, e^{-\sigma^2_{M\text{-}F} s^2/2}$$

$$= A \cos \Delta\eta \sin \frac{rq\pi}{10} \, e^{-\sigma^2 q^2/20} \quad (\text{using } \frac{\pi^2}{100} \approx \frac{1}{10}) \tag{11}$$

$$\frac{\partial I_c}{\partial q_c} = A \sin \Delta\eta \sin \frac{rq\pi}{10} \, e^{-\sigma^2 q^2/20} [\, 2\pi a + 4\pi b(q - q_c)] \tag{12}$$

$$\frac{\partial I_c}{\partial a} = -2\pi(q - q_c) \sin \Delta\eta \sin \frac{rq\pi}{10} \, e^{-\sigma^2 q^2/20} \tag{13}$$

Another modification that was made was provision for semiautomatic recycling of the iterative refinement procedure. At the end of every refinement cycle the computer stops on a breakpoint and the operator inspects the results which have just been printed out. If another cycle is desired, the operator so indicates and the computer automatically applies the least squares adjustment to each independent variable and calculates a new model. Thus the new value of $r_{M\text{-}F}$ is multiplied by $2^{1/2}$ and by 2 to obtain the two new dependent bond distances $r_{F \cdot F}$ and $r_{F \cdot \cdot F}$. Then 0.005 A is added to $r_{M\text{-}F}$ and a second model is calculated as required by the program for approximating the partial derivative of the intensity with respect to $r_{M\text{-}F}$ by taking differences:

$$\frac{\partial I_c}{\partial r_{M\text{-}F}} = \frac{I_c(r_{M\text{-}F} + 0.005) - I_c(r_{M\text{-}F})}{0.005} \tag{14}$$

The scale factor on the calculated intensity I_c is adjusted by multiplying every amplitude A_i by the quantity $(1 + \delta_{scale})$ where

δ_{scale} is the least squares correction to the scale factor. The other independent variables, (σ_i, q_c and a), are adjusted by simple addition of the least squares corrections.

A diagonal weight matrix P is used in the least squares program, with $p_{qq} = (q - 1)e^{-\gamma q^2}$. The same weighting parameter, $\gamma = 0.0012\pi^2 q^2/100$, was used for every compound at every voltage. Our weighting function is very similar to that of Bastiansen, Hedberg and Hedberg (14), who used $P = se^{-0.0012s^2}$.

It had originally been planned to adjust the background by least squares, but this was not attempted. (It was tried unsuccessfully in the refinement of N_2O_4 as described in Part II of this thesis.) Although preliminary theoretical curves were used as a guide in drawing the background, it became evident from $I_e - I_c$ curves obtained during least squares refinement that the background sometimes needed further adjustment. Therefore, where necessary, a smooth background correction line was drawn by eye, using as a guide the $I_e - I_c$ curves. This adjustment to the background was not allowed to follow the fluctuations in $I_e - I_c$, of course, but only its general trend over regions of 30-50 q-units or so. After this final background adjustment, several further least squares refinement cycles were carried out until no significant change in the parameters occurred.

Results of Least Squares Refinement

Figures 3-7 show the final I_e, I_c, and $I_e - I_c$ curves, labeled E (experimental), C (calculated) and E-C. The calculated curves are based on the parameters of Tables II-VI. These tables and figures are the results of the least squares refinement of the structures of WF_6, OsF_6, IrF_6, UF_6, and NpF_6. (The structure of PuF_6 is discussed on page 44.)

Our experimental data do not extend to small scattering angles due to the large beam stops which were used. The curves in Figs. 3-7 are drawn to cover that range in q which was used in the least squares adjustment, except for NpF_6, where the calculated curves are extended down to $q = 0$. The extra range for small q is shown for only one compound as an example of the general effect on the diffraction patterns of all MF_6 compounds of varying the accelerating voltage of the electrons.

In Tables II-VI are listed the best values and the standard deviations of six of the seven parameters adjusted by least squares. The seventh is the scale factor on I_c. The standard deviation is the external estimate, based on the assumption that the observations are independent. However, the observations are not independent; they are the amplitudes at integral q-values of a continuous experimental intensity curve.

Figs. 3-7

Experimental (E), calculated (C), and difference (E-C) curves for WF_6, OsF_6, IrF_6, UF_6, and NpF_6. (The exact values of the accelerating potentials are listed in Table IX.) The arrows indicate the position of the cut-off point on each calculated curve.

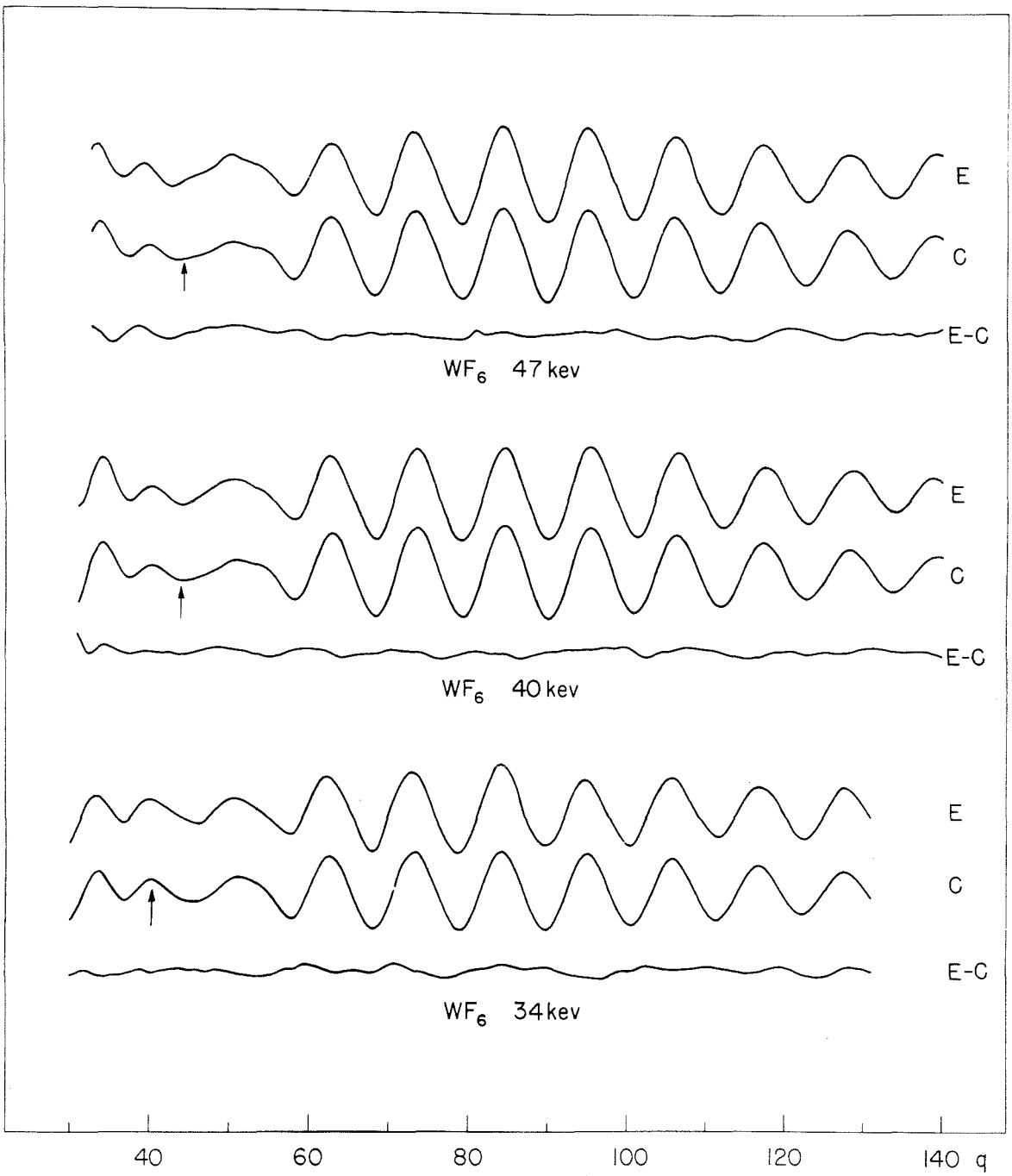

E
C
E-C
WF_6 47 kev
E
C
E-C
WF_6 40 kev
E
C
E-C
WF_6 34 kev
40
60
80
100
120
140 q

Fig. 3

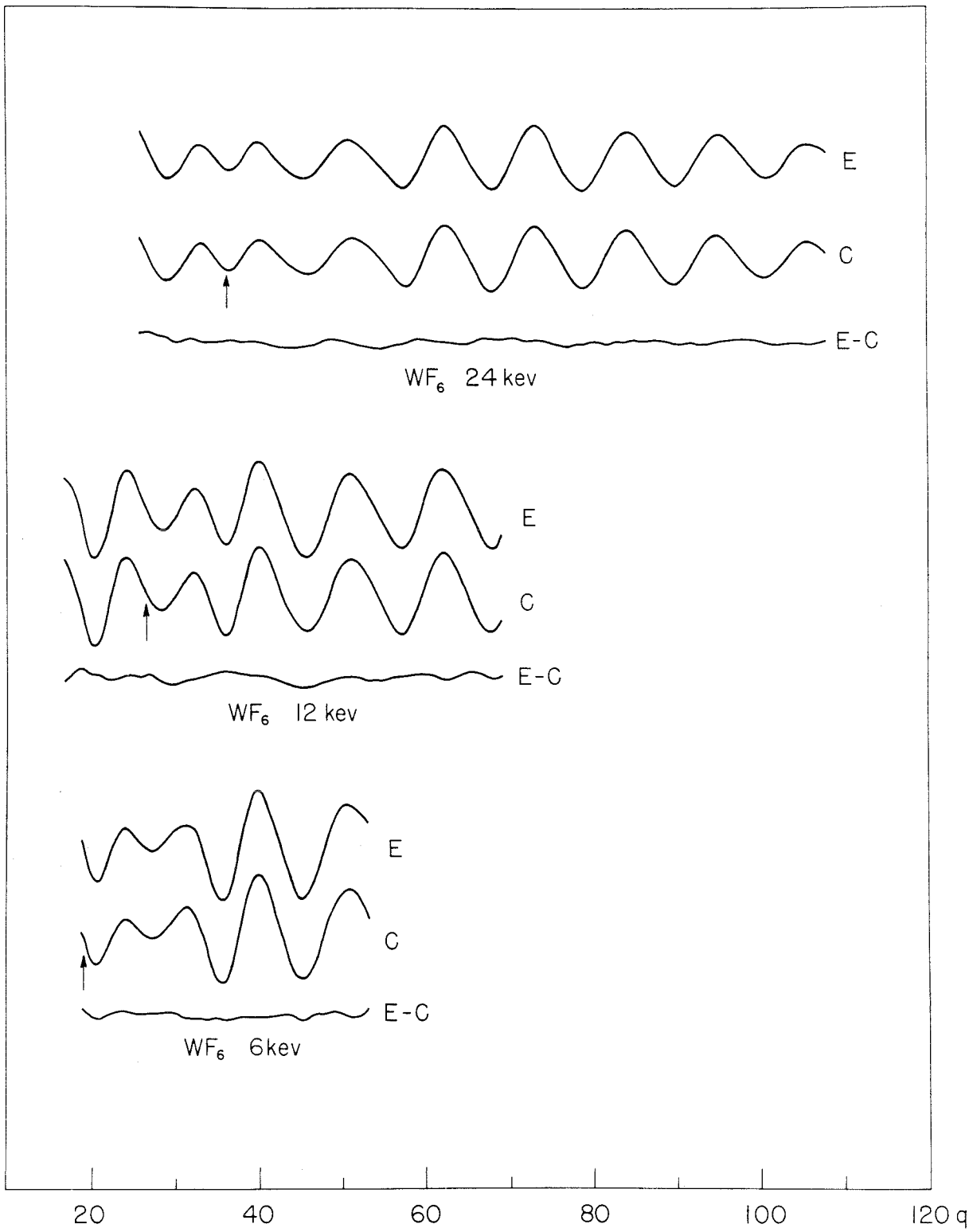

E
C
E-C
WF6 24 kev
E
C
E-C
WF6 12 kev
E
C
E-C
WF6 6kev
20
40
60
80
100
120 q

Fig.4

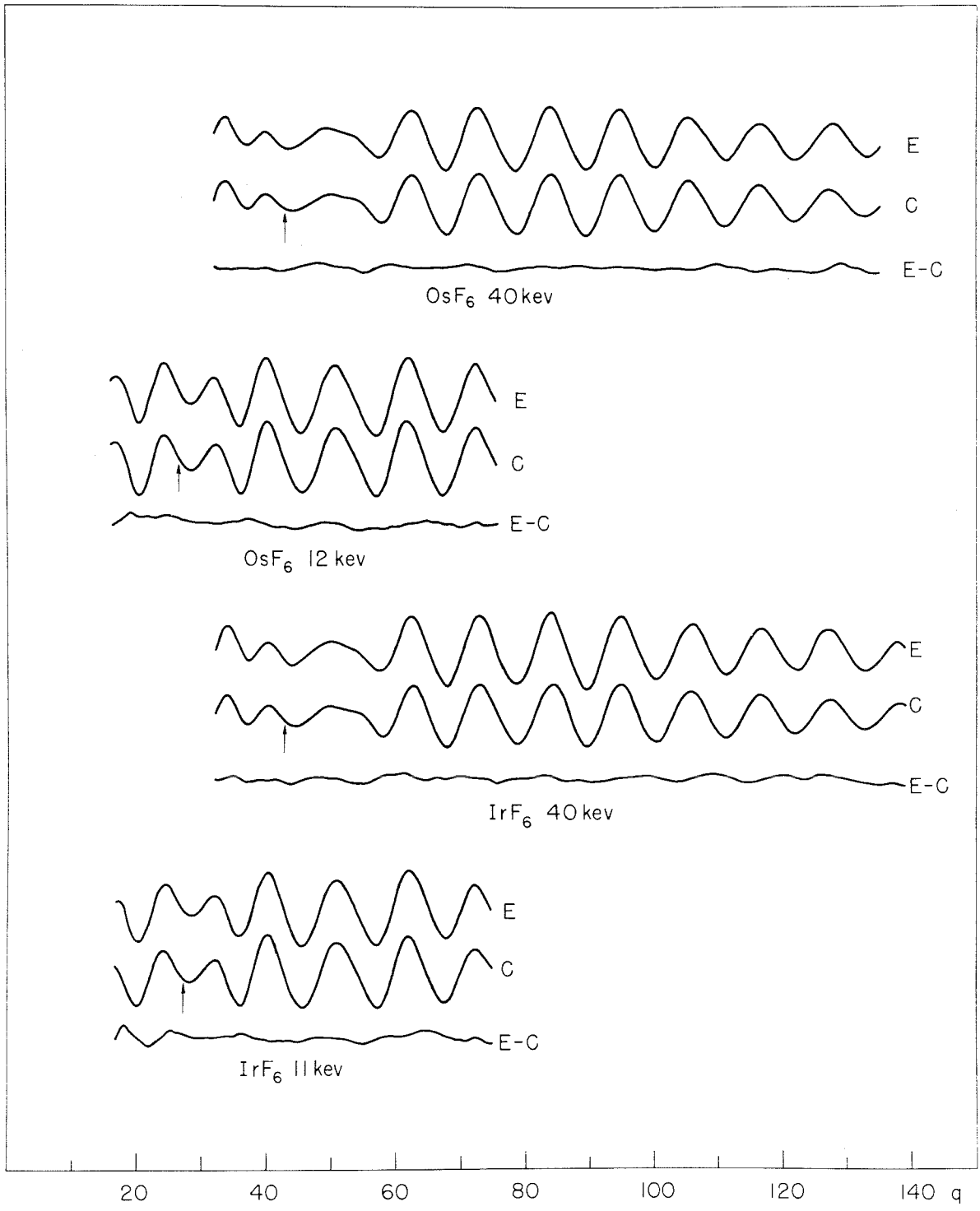

E
C
E-C
OsF_6 40kev
E
C
E-C
OsF_6 12 kev
E
C
E-C
IrF_6 40 kev
E
C
E-C
IrF_6 11 kev
20
40
60
80
100
120
140 q

Fig.5

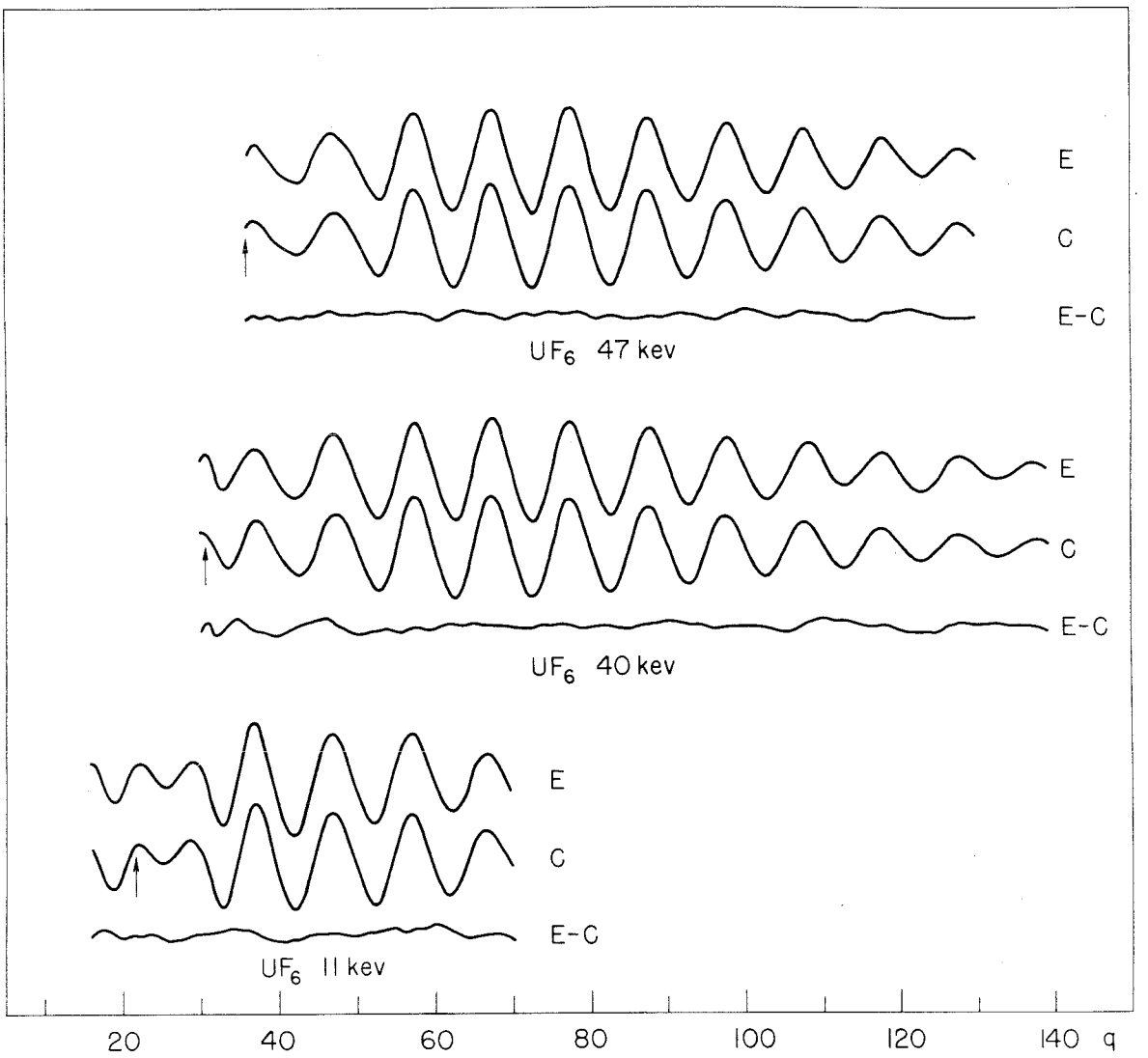

E
C
E-C
UF_6 47 kev
E
C
E-C
UF_6 40 kev
E
C
E-C
UF_6 11 kev
20
40
60
80
100
120
140
q

Fig. 6

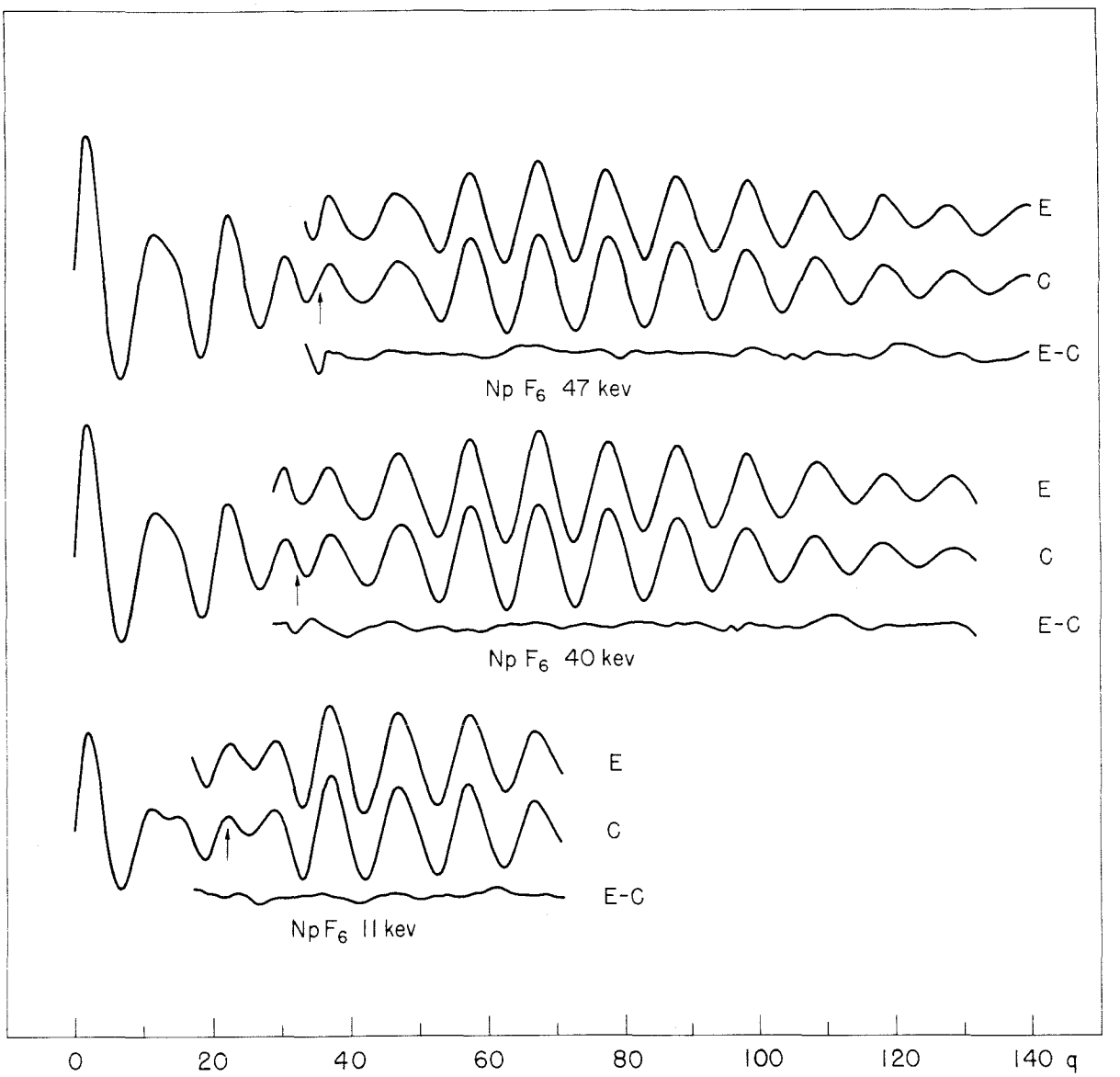

E
C
E-C
Np F6 47 kev
E
C
E-C
Np F6 40 kev
E
C
E-C
NpF6 11 kev
0
20
40
60
80
100
120
140 q

Fig. 7

Tables II - VI

The least squares values of the M-F distance r, the temperature parameters σ_i (i = 1, 2, 3 = M-F, F· F, and F·· F, respectively), the cut-off point q_c, and the parameter a (defined in equation 9). In these tables, r, σ_i, and a are in Angstrom units and q_c is in reciprocal Angstrom units. Note that the individual temperature parameters σ_i are physically meaningless (see page 38). For a summary of these tables, with limits of error, see Table XIII.

TABLE II

LEAST SQUARES REFINEMENT OF WF_6

	WF_6-47 kev		WF_6-40 kev	
	Best value	std. dev.	Best value	std. dev.
r	1.836	0.0007	1.829	0.0006
q_c	44.3	0.6	43.9	0.4
σ_1	0.035	0.004	0.039	0.003
σ_2	0.090	0.008	0.094	0.006
σ_3	0.043	0.012	0.065	0.014
a	0.0050	0.0004	0.0050	0.0004

	WF_6-34 kev		WF_6-24 kev	
r	1.836	0.0007	1.828	0.0009
q_c	40.3	0.6	36.5	0.5
σ_1	0.040	0.004	0.051	0.004
σ_2	0.100	0.009	0.106	0.010
σ_3	0.067	0.018	0.056	0.014
a	0.0046	0.0005	0.0052	0.0006

	WF_6-12 kev		WF_6-6 kev	
r	1.828	0.0016	1.838	0.0019
q_c	27.0	0.6	19.3	0.7
σ_1	0.070	0.004	0.052	0.009
σ_2	0.129	0.007	0.109	0.009
σ_3	0.090	0.016	0.085	0.017
a	0.0031	0.0007	0.0031	0.0006

TABLE III

LEAST SQUARES REFINEMENT OF OsF_6

OsF_6-40 kev

	Best value	std. dev.
r	1.832	0.0005
q_c	43.2	0.4
σ_1	0.044	0.002
σ_2	0.087	0.006
σ_3	0.070	0.013
a	0.0045	0.0004

OsF_6-12 kev

r	1.830	0.0013
q_c	26.8	0.6
σ_1	0.063	0.003
σ_2	0.112	0.008
σ_3	0.073	0.018
a	0.0042	0.0005

TABLE IV

LEAST SQUARES REFINEMENT OF IrF_6

IrF_6-40 kev

	Best value	std. dev.
r	1.832	0.0006
q_c	42.8	0.4
σ_1	0.044	0.002
σ_2	0.093	0.007
σ_3	0.080	0.017
a	0.0043	0.0004

IrF_6-11 kev

r	1.829	0.0014
q_c	26.7	0.5
σ_1	0.064	0.005
σ_2	0.093	0.009
σ_3	0.042	0.024
a	0.0074	0.0007

TABLE V

LEAST SQUARES REFINEMENT OF UF_6

UF_6-47 kev

	Best value	std. dev.
r	1.999	0.0005
q_c	35.9	0.8
σ_1	0.049	0.006
σ_2	0.126	0.011
σ_3	0.073	0.015
a	0.0048	0.0008

UF_6-40 kev

r	1.995	0.0007
q_c	30.6	0.9
σ_1	0.049	0.006
σ_2	0.133	0.013
σ_3	0.073	0.018
a	0.0048	0.0007

UF_6-11 kev

r	1.994	0.0016
q_c	22.0	0.6
σ_1	0.071	0.009
σ_2	0.118	0.013
σ_3	0.074	0.026
a	0.0078	0.0008

TABLE VI

LEAST SQUARES REFINEMENT OF NpF_6

NpF_6-47 kev

	Best value	std. dev.
r	1.983	0.0008
q_c	35.9	1.1
σ_1	0.054	0.005
σ_2	0.117	0.015
σ_3	0.077	0.020
a	0.0042	0.0010

NpF_6-40 kev

r	1.982	0.0006
q_c	32.8	0.6
σ_1	0.053	0.005
σ_2	0.128	0.010
σ_3	0.090	0.020
a	0.0051	0.0006

NpF_6-11 kev

r	1.977	0.0015
q_c	22.2	0.6
σ_1	0.069	0.008
σ_2	0.109	0.012
σ_3	0.077	0.027
a	0.0078	0.0009

In order to correct for this lack of independence in the observations, the estimate is made in Part III that the true standard deviations are larger than the ones obtained by least squares by a factor of roughly $(20/3r)^{1/2}$, where r may be taken as the M-F distance in the present compounds. Since r is about 2 A we would estimate that the correction factor for the standard deviations is roughly 1.8 or 2. Of course systematic errors will also increase our estimates of the standard deviations. The conclusions about the error estimates are discussed later.

Table VII (page 37) gives for each set of data the value of $s = [L'PL/(n-m)]^{1/2}$, the mean error of an observation of weight unity, (not to be confused with $s = (40/\lambda) \sin(\theta/2)$) where L'PL is the sum of weighted squares of residuals, n is the number of observational equations and m is the number of adjustable parameters. Thus s is a measure of the goodness of fit of the observed and calculated curves. However, there is an arbitrary but different amplitude scale for each set of data, and s in the table is based on this arbitrary scale (which corresponds to the scale in Figs. 3-7). Therefore s was put on a standardized scale for all sets of data, and is tabulated s' in Table VII, corresponding to $A_{M-F} = 1000$ in equation 10.

A rough estimate based on s' can be given of the error in I_e expressed as a fraction of the amplitude of a typical ring. In the case of WF_6-40 kev, for example, the maximum of highest amplitude in the photometered pattern is near q = 80. If $A_{M-F} = 1000$, this maximum has an amplitude of about 520. A measure of the expected error in the observations is $s'p^{-1/2}$ where p is the weighting function (page 21).

In the region $q = 40$ to 100, p takes on values between 30 and 40, so the expected error in the observations is about 7% to 9% of the height of the maximum at $q = 80$. Of course such estimates of error are very sensitive to any background corrections, (because of the dependence of $L'PL$ on changes in the background). Although s' has the same order of magnitude for each curve (the range in s' covers barely more than a two-fold spread), the agreement among the curves would probably be improved if further background corrections were attempted.

TABLE VII

MEAN ERROR OF AN OBSERVATION OF WEIGHT UNITY

	s (arbitrary scale)	s' (standardized scale)
WF_6-47 kev	161	320
40	135	240
34	136	280
24	117	279
12	163	147
6	117	143
OsF_6-40	109	192
12	180	218
IrF_6-40	136	241
11	176	321
UF_6-47	122	163
40	155	232
11	185	300
NpF_6-47	39	193
40	167	195
11	217	322

Temperature Parameters

It has been mentioned several times that we cannot determine the absolute temperature factors. This is the case in visual work also, where in drawing a visual curve an effort is made not to show the decreasing amplitudes of the diffraction rings with increasing angle of scattering. The visual method is inherently unable to reveal absolute temperature factors because comparisons of ring amplitudes can be made only among two or three neighboring rings. Although this limitation is no longer inherent in the sector-microphotometer method -- amplitudes can be compared for all the rings in a pattern -- this limitation still exists in the present investigation primarily because the correction for the non-linearity of the photographic emulsion could not be made. Other steps in our procedure which obscure the absolute temperature factors (but could be corrected were it not for the basic factor -- the inability to correct for the photographic emulsion) include division by a somewhat uncertain background and the use of constant coefficients A_i in equation 10.

However, as in the case of visual work, we can still say something about relative temperature factors. Equation 4 can be rewritten as follows:

$$sI(s) = e^{-\sigma^2_{M-F}s^2/2}\left[\frac{6C_{M-F}}{r}\cos\Delta\eta\sin rs + \frac{12\,C_{F\cdot F}}{\sqrt{2}\,r}\sin\sqrt{2}\,rs\,e^{-(\sigma^2_{F\cdot F}-\sigma^2_{M-F})s^2/2} + \frac{3C_{F\cdot F}}{2r}\sin 2rs\,e^{-(\sigma^2_{F\cdot\cdot F}-\sigma^2_{M-F})s^2/2}\right] \qquad (15)$$

and one would anticipate better results for $(\sigma^2_{F\cdot F}-\sigma^2_{M-F})$ and for $(\sigma^2_{F\cdot\cdot F}-\sigma^2_{M-F})$ than for the individual temperature parameters σ_{M-F}, $\sigma_{F\cdot F}$, or $\sigma_{F\cdot\cdot F}$.

Therefore in table VIII are listed both the individual temperature parameters σ_i and the relative temperature parameters (d_i-d_{M-F}), where $d_i = \frac{1}{2}\sigma_i^2$ (often called a_{ij} in the temperature factor expression $e^{-a_{ij}s^2}$).

The agreement among relative temperature parameters of the same compound at different voltages is reasonable when the standard deviations of the individual temperature parameters are considered.

In IrF_6-11 kev, d_3-d_1 is negative (indicated by * in Table VIII) and therefore evidently in error.

TABLE VIII

TEMPERATURE PARAMETERS

The subscripts 1, 2, 3 refer to M-F, F·F, F··F, respectively

		47	40	34	24	12, 11	6 kev
WF_6	σ_1	0.035	0.039	0.040	0.051	0.070	0.052
	σ_2	0.090	0.094	0.100	0.106	0.129	0.109
	σ_3	0.043	0.065	0.067	0.056	0.090	0.085
	d_2-d_1	0.0034	0.0036	0.0042	0.0043	0.0034	0.0046
	d_3-d_1	0.0003	0.0013	0.0014	0.0029	0.0016	0.0022
OsF_6	σ_1		0.044			0.063	
	σ_2		0.087			0.112	
	σ_3		0.070			0.073	
	d_2-d_1		0.0028			0.0043	
	d_3-d_1		0.0015			0.0007	
IrF_6	σ_1		0.044			0.064	
	σ_2		0.093			0.093	
	σ_3		0.080			0.042	
	d_2-d_1		0.0034			0.0023	
	d_3-d_1		0.0022			*	
UF_6	σ_1	0.049	0.049			0.071	
	σ_2	0.126	0.133			0.118	
	σ_3	0.073	0.073			0.074	
	d_2-d_1	0.0067	0.0067			0.0045	
	d_3-d_1	0.0015	0.0015			0.0002	
NpF_6	σ_1	0.054	0.053			0.069	
	σ_2	0.117	0.128			0.109	
	σ_3	0.077	0.090			0.077	
	d_2-d_1	0.0055	0.0067			0.0035	
	d_3-d_1	0.0016	0.0026			0.0006	

Systematic Errors

The standard deviations reported in tables II-VI make no allowance for systematic errors. By systematic errors is meant those which have a non-random effect on the observation equations. Probably the most important systematic error is that in the wavelength determination, because for the present least squares procedure a simultaneous measurement must be made of the q-value and amplitude of every point in the diffraction pattern. It is assumed that the q-values are known exactly and that errors exist only in the amplitudes. This is a fairly good assumption provided a separate allowance is made for the uncertainty in the wavelength (and the resulting error in the q-scale).

The coordinate q depends on λ through $q = \frac{40}{\lambda} \sin \frac{\theta}{2}$, and λ is determined from the accelerating voltage of the electron beam. The voltage is monitored by a potentiometer connected across a voltage divider. The potentiometer was calibrated with ZnO pictures made at 47, 40 and 34 kev. Zinc oxide pictures made at 24 kev were not used for calibration because their rings were broadened to the extent that measurements were very uncertain. At 12 and 6 kev, the rings were so broad that measurements could not be made at all.

The potentiometer reading f and the quantity $\frac{1}{L\lambda}$ determined from the zinc oxide pictures at camera distance L are given in table IXa.

TABLE IX

WAVELENGTH CALIBRATION

a. Determination of $K = V/f$ from ZnO data

f(volts)	$1/L\lambda$	V_{obs}(kev)	K	λ_{obs}(A)
1.4000	1.8832	47.250	33.750	0.05516
1.1837	1.7210	39.740	33.573	0.06036
1.0000	1.5742	33.488	33.488	0.06599

$\overline{K}$ (best value) = 33.600

b. Determination of λ from $\overline{K}f = \overline{V}$

f(volts)	$\overline{V}$ (kev)	V^{*}_{approx}	λ_{calc}(A)	discrepancy in λ
1.4000	47.040	47	0.05529	0.23%
1.1837	39.772	40	0.06033	0.05%
1.0000	33.600	34	0.06583	0.24%
0.7000	23.520	24	0.07906	
0.3500	11.760	12	0.11252	
0.3413	11.468	11	0.11389	
0.1750	5.880	6	0.15947	

*V_{approx} is the value of the voltage rounded off for quotation in the text and figures.

The "observed" potential V_{obs} was found from $\frac{1}{L\lambda}$ using $L = 9.627$ cm as measured with a cathetometer, and $\lambda = h/\sqrt{2m_o eV(1+eV/2m_o c^2)}$, the relativistic expression for the electron wave length. Thus the proportionality constant $K = \frac{V}{f}$ could be determined from ZnO data. Its average value, $\overline{K} = 33,600$, was used to determine both the best values of the three highest potentials and also the values of the lower potentials for which ZnO data were not obtained. In table IXb, λ was determined from $\overline{V} = \overline{K}f$ and this value of λ was used to fix the absolute scale in all of the molecular structure determinations.

The discrepancy between the observed and calculated values of λ for the three highest potentials is less than 0.24% which is slightly higher than one might expect, since measurements on two different ZnO plates at the same voltage agree to within 0.1%. However, our value of λ differs from that of several previous workers in this laboratory, who found $\lambda = 0.6056$ A for the potentiometer reading $f = 1.1837$ v. This discrepancy of 0.36% in two independent determinations of λ may be largely due to uncertainty in the camera distance. This writer estimates that 0.4% is a reasonable value for the systematic error in the scale of the molecules due to uncertainties in λ. It is interesting to note that the M-F distance as determined at several different wave lengths shows good internal consistency with an average deviation of 0.2% for WF_6 and 0.1% or less for the other compounds.

Certain types of background errors might be called systematic and others random. However, it is very difficult to analyze the effect of background errors. Several generalizations are discussed in Part III.

Other errors which may be called systematic are those due to uncertainty in the sector calibration, lack of correction for the non-linearity of the photographic emulsion, and the various approximations used in the calculation -- the approximation that the coefficients A_i of equation 10 are constant and the use of equation 9 to calculate $\Delta\eta$. However these errors will have little effect on the scale of the molecules. The main source of systematic error in the M-F distance is probably the uncertainty in λ. Therefore 0.4% will be used in estimating limits of error (see "Conclusions").

Plutonium Hexafluoride

Because of the poor quality of the PuF_6 pictures, least squares refinement was impossible on the 47 and 40 kev data, and only partially successful on the 11 kev data.

The 47 kev experimental curve is shown in Fig. 8. From the results on UF_6 and NpF_6 at the same voltage, one would expect the cut-off point to be in the neighborhood of 30-36 q-units. Thus we may refer to calculated curves in which q_c = 30.6 (UF_6-40 kev), and q_c = 36.1 (UF_6-47 kev) in Fig. 6. In the experimental PuF_6 curve, the nearly symmetrical shape of minimum 6 and the relatively large excursion from maximum 6 to minimum 7 (which is possibly in error) favors the higher value of q_c. In order to find q_c by the correlation method, comparisons should be made among the amplitude of the 4th, 5th, and 6th rings (see Fig. 7 where the theoretical curves extend to smaller scattering angles). Thus the experimental PuF_6-47 kev curve is compatible with q_c = 30 to 36, but the lack of data for the 4th ring prevents a correlation treatment. The excessive amplitudes of maximum 6 and

minimum 7 and the peculiar appearance of the 14th ring seem to indicate large errors in the experimental curve in these regions. The absolute amplitudes of "maxima" and "minima" on the photometer trace are small compared to those of the other compounds, and it is impossible to get a good photometer trace.

The experimental curve for PuF_6-40 kev (Fig. 8) is also poor. Because there are no data below $q = 40$, about all that can be said is that q_c is less than 40. The unexpected shelf at $q = 52$ is not due to the troublesome 1.4 cm region of the sector because this shelf corresponds to $R = 1.54$ cm. Maximum 6 and minimum 7 have excessive amplitudes in this case also.

The Pu-F bond distance was found by comparing the positions of maxima and minima of the 47 and 40 kev experimental curves with those of the corresponding calculated curves for NpF_6. These results are given in Tables X and XI.

The results of the least squares refinement of PuF_6-11 kev are shown in Fig. 8 and Table XII. The best theoretical curve, C, is in satisfactory agreement with the experimental curve, E, from the standpoint of the visual method: the height of maximum n can in general only be compared with the average of the heights of maxima (n+1) and (n-1). However, the difference curve, E-C, leaves much to be desired. The resulting parameters are much more uncertain than those of the other compounds. The temperature parameters are especially unreliable -- the least squares adjustment to $\sigma_{F\cdot\cdot F}$ makes the temperature factor for the long F··F distance an increasing rather than a decreasing function of q. However, in the calculated curve, $\sigma_{F\cdot\cdot F}$ = zero was used.

Fig. 8

Experimental curves for PuF_6 at 47 and 40 kev; and experimental (E), calculated (C) and difference (E-C) curves at 11 kev.

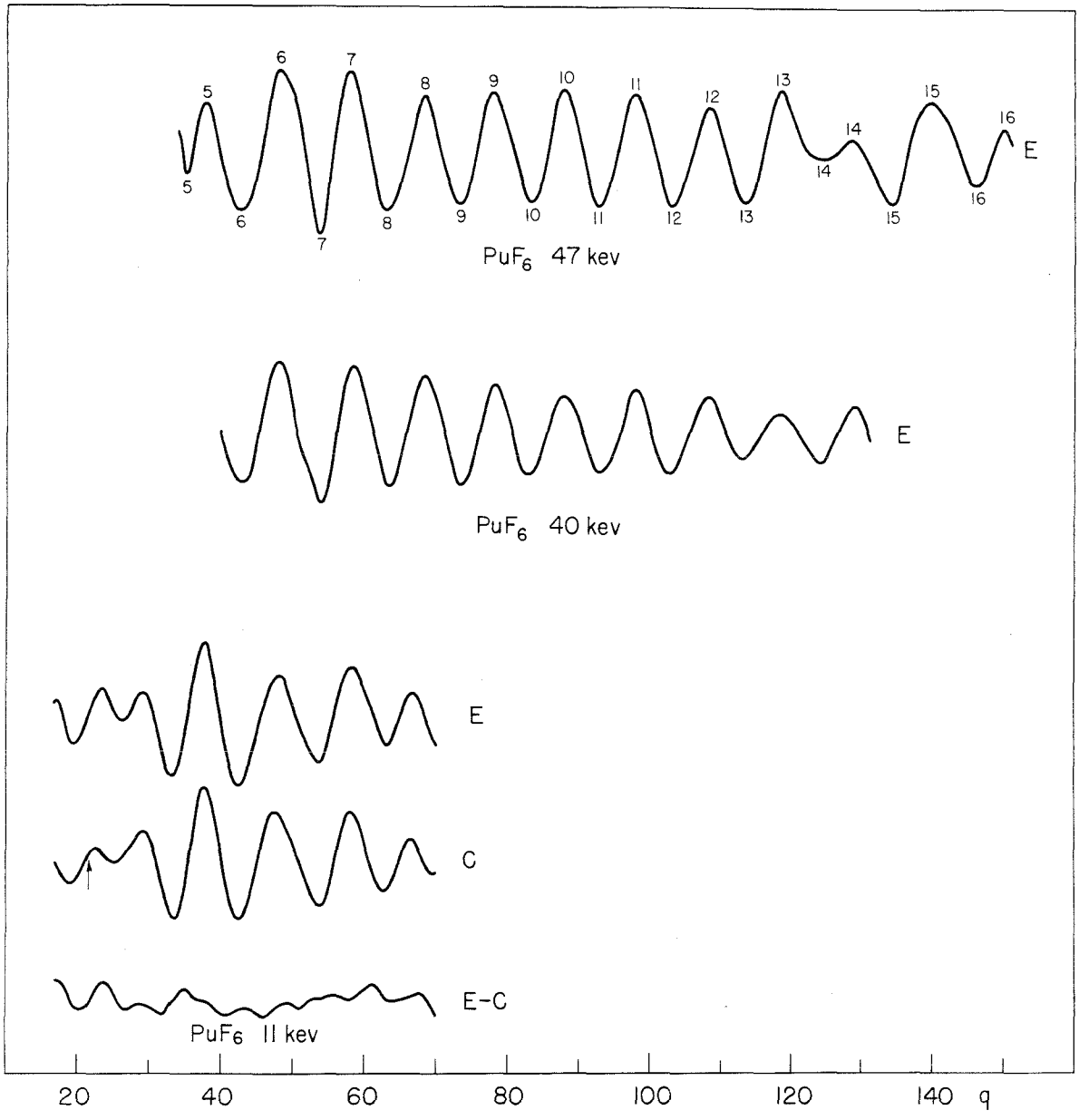

5
6
7
8
9
10
11
12
13
14
15
16
E
PuF6 47 kev
E
PuF6 40 kev
E
C
E-C
PuF6 11 kev
20
40
60
80
100
120
140
q

Fig. 8

TABLE X

ELECTRON DIFFRACTION DATA FOR PuF_6-47 kev

	Min.		Max.	
No.	q_{obs}	q/q_{obs}	q_{obs}	q/q_{obs}
5	35.18	(0.971)	37.83	0.994
6	42.76	0.995	48.20	0.986
7	53.84	0.998	58.18	0.999
8	63.44	0.998	68.64	0.992
9	73.77	0.996	78.21	1.001
10	83.72	0.997	88.15	1.002
11	93.13	1.004	98.05	1.005
12	103.62	1.000	108.56	1.001
13	113.75	1.000	118.44	1.004
14	124.02	0.999	128.64	1.000
15	134.54	0.995	139.38	0.997
16	145.83	0.987	150.14	0.992

Average 23 features 0.997_4

Average deviation 0.004

$r_{Pu-F} = 1.983$ A x 0.997 = 1.977 A

TABLE XI

ELECTRON DIFFRACTION DATA FOR PuF_6-40 kev

	Min.		Max.	
No.	q_{obs}	q/q_{obs}	q_{obs}	q/q_{obs}
5			39.06	(0.960)
6	42.54	1.007	47.69	1.003
7	54.06	0.986	58.57	0.990
8	63.63	0.992	68.68	0.989
9	73.67	0.992	78.37	0.996
10	83.01	1.003	88.04	1.002
11	93.17	1.000	98.01	1.002
12	102.61	1.007	108.47	0.999
13	113.44	0.999	118.29	1.009
14	124.38	0.993	128.96	0.996

Average, 18 features 0.998_0

Average deviation 0.006

$r_{Pu-F} = 1.982\ A \times 0.998 = 1.978A$

TABLE XII

LEAST SQUARES REFINEMENT OF PuF_6

PuF_6-11 kev

	Best value	std. dev.
r	1.958 A	0.0026
q_c	21.6 A^{-1}	1.0
σ_1	0.031 A	0.020
σ_2	0.043 A	0.015
σ_3	--	--
a	0.0111 A	0.0004

Evidence for Octahedral Structure

Radial distribution (RD) curves were prepared from diffraction data at 40 and 11 or 12 kev for all compounds. Preliminary theoretical curves, and later final ones, were used to supplement the experimental data for the inner part of the pattern.

These RD curves appear to furnish good evidence for the octahedral symmetry because the doubled M-F peak is explained by the phase shift, whereas the two F-F peaks have the expected appearance and position for a regular octahedron. This evidence is somewhat weakened however by the fact that the inner parts of the experimental intensity curves do not extend to as low a value of q as would be desired, and therefore the RD curves are heavily influenced by the inner rings of the intensity curve calculated assuming an octahedral model. Thus although it is common practice to supplement the first one or two rings with a calculated curve, it proved dangerous in the present case. It was necessary to supply from calculated curves for the 40 kev data three rings and sometimes part of a fourth. Perhaps more convincing results could be obtained by using experimental intensities alone, but multiplied by a modification function to smooth out the ripples in the RD caused by omitting the inner rings.

Both RD and experimental intensity curves clearly show that the cut-off occurs at a lower value of q; the lower the accelerating voltage, that is, the longer the electron wavelength. Such a dependence of the cut-off on the wavelength (which was observed for UF_6 by Hoerni and Ibers (9)) eliminates any possibility that the anomalous diffraction pattern is due to an asymmetric structure. The fact that we now have

intensity data extending to $3q_c$ without showing a second cut-off is further evidence against the structural asymmetry. It is also evidence that $\Delta\eta = \delta s$ can hold only as a rough approximation, and that the conclusions of Ibers and Hoerni are correct as to the general shape of the $\Delta\eta$ curves (see Fig. 2).

The experimental curves seem to be completely compatible with a symmetrical octahedral model for all the compounds (with the reservations noted above for PuF_6). There remains little doubt that all six M-F bonds are of equal length, but it is hard to determine the F-F distances because of the relatively low scattering power of F. However, to the degree that the F· F and F·· F distances do not fit the data these terms will be effectively eliminated through being assigned large temperature parameters in the least squares refinement. That the temperature parameters are reasonable and do not approach infinity is evidence for at least an approximately octahedral structure.

Conclusions

The results of this investigation are summarized in Table XIII. For each compound, the average of the values of the M-F distance r obtained by least squares from data at different voltages is given as the best value. Presumably a weighted average would be better, but the large systematic error mentioned below would make the weights nearly equal.

To find limits of error one must consider both the least squares standard deviations and the systematic errors. The least squares standard deviation of r, (multiplied by a factor of 2 to allow for lack

of independence in the observations), ranges from 0.0010 to 0.0052 A for different compounds at different voltages. On the other hand, the systematic errors are essentially 0.4% due to uncertainty in λ. The total standard deviation s is given by

$$s^2_{\text{total}} = s^2_{\text{least squares}} + s^2_{\text{systematic}}$$

assuming that the covariance of the two errors is zero. Thus s_{total} is less than 0.0090 A for all compounds. It has been estimated (16) that a reasonable limit of error is twice the standard deviation. This criterion would give 0.018 A, as the limit of error for r, and this has been rounded to 0.02 A in Table XIIIa.

In Table XIIIb the experimental and theoretical cut-off points q_c are compared. The systematic errors in q_c (estimated to be 0.4% here also) are very small compared to the standard deviations obtained by least squares. Therefore the least squares standard deviation was multiplied by a factor of 2 to correct for lack of independence in the observations, and by another factor of 2 to convert standard deviations to limits of error. Thus $4(s_{\text{least squares}})$ is reported in Table XIIIb as the limit of error in q_c.

The theoretical q_c were obtained by graphical interpolation from the table of $\Delta\eta$ given by Ibers and Hoerni (13) and the approximation formula (equation 8). Their table does not extend high enough in atomic number to calculate q_c for the lower voltages by the approximation formula. For UF_6-11 kev the theoretical q_c was obtained from an earlier paper by Hoerni and Ibers (9).

TABLE XIII

THE STRUCTURES OF SIX HEAVY METAL HEXAFLUORIDES

a. M-F bond distances (r), and temperature parameter ($d_i = \frac{1}{2}\sigma_i^2$)

Compound	r	limit of error	d_2-d_1	ave. dev.	d_3-d_1	ave. dev.
WF_6	1.833 A	0.02 A	0.0039	0.0005	0.0012	0.0006
OsF_6	1.831	0.02	0.0035	0.0008	0.0011	0.0004
IrF_6	1.830	0.02	0.0029	0.0006	0.0011	0.0011
UF_6	1.996	0.02	0.0063	0.0012	0.0011	0.0006
NpF_6	1.981	0.02	0.0052	0.0012	0.0016	0.0007
PuF_6	1.971	0.02	--	--	--	--

b. Experimental and theoretical cut-off points

		Exp.	limit of error	Theor.
WF_6	47 kev	44.3 A^{-1}	2.4 A^{-1}	49.5 A^{-1}
	40	43.9	1.6	43.7
	34	40.3	2.4	40.9
	24	36.5	2.0	37.0
	12	27.0	2.4	--
	6	19.3	2.8	--
OsF_6	40	43.2	1.6	42.7
	12	26.8	2.4	--
IrF_6	40	42.8	1.6	42.2
	11	27.6	2.0	--
UF_6	47	35.9	1.6	41.9
	40	30.6	3.6	37.0
	11	22.0	2.4	23.0
NpF_6	47	35.9	4.4	41.5
	40	32.8	2.4	36.8
	11	22.2	2.4	--
PuF_6	47	--	--	41.1
	40	--	--	36.5
	11	21.6	4.0	--

The present investigation of UF_6 completes the preliminary work of Hoerni and Ibers (9) on this compound. These authors found fairly satisfactory agreement with visual curves obtained from 40 kev and 11 kev photographs by using a symmetrical octahedral model with following parameters: $r_{U-F} = 2.00$ A, $d_{F\cdot F}-d_{U-F} = 2.2\times10^{-3}$ A^2, $d_{F\cdot\cdot F}-d_{U-F} = 0.75\times10^{-3}$ A^2. They also used their complex scattering amplitudes corresponding to the theoretical cut-off points in Table XIII. Our results seem to agree well with these earlier ones at 11 kev, but our calculated curve for 40 kev is in better agreement with both our sector data and the visual curve of Dr. Otto Bastiansen which Hoerni and Ibers tried to fit. This indicates that the theoretical cut-off point is too high in the case of UF_6-40 kev.

Our results on WF_6-40 kev are in satisfactory agreement with those of Nazarian (10), who found $r_{W-F} = 1.834$ A, $d_{F\cdot F}-d_{W-F} = 2.2\times10^{-3}$ A^2, $d_{F\cdot\cdot F}-d_{W-F} = 0.8\times10^{-3}$ A^2 and $q_c = 42.9$ A^{-1} by the correlation method with visual data.

The limits of error on r and q_c in Table XIII are believed by this writer to be reasonable estimates, and the limits of error on the temperature parameters are probably somewhat higher than the average deviations in Table XIII.

II. THE MOLECULAR STRUCTURE OF GASEOUS N_2O_4

Introduction and Theory

The expression that describes the scattering of fast electrons by gas molecules is*

$$I'(s) = \frac{K}{s^4}\left[B + \sum_{ij}{}' (Z-f)_i(Z-f)_j e^{-a_{ij}s^2} \frac{\sin r_{ij}s}{r_{ij}s} \right], \qquad (1)$$

where

$I'(s)$ = intensity of electrons scattered at angle θ

K = constant

$B = \sum \left[(Z-f)_i^2 + S_i \right]$

S = incoherent scattering function (15)

Z = atomic number

f = atomic scattering factor for X-rays

$s = (4\pi \sin \frac{\theta}{2})/\lambda$, λ is the deBroglie wave length

a_{ij} = one-half the mean square variation of r_{ij}

r_{ij} = distance between atoms i and j

The term B represents the purely atomic contribution to the diffraction pattern, while the summation (in which the prime indicates that terms with $i = j$ are not to be included) represents the molecular contribution.

In the intensity expression for the compound N_2O_4, $(Z-f)_N$ is very nearly $(Z-f)_O Z_N/Z_O$. Therefore, we may write

---

*The factor $\cos \Delta\eta$ discussed in Part I need not be included here -- nitrogen and oxygen are both light atoms.

$$I'(s) = \frac{K}{s^4}\left[B + \left(\frac{Z-f}{Z}\right)_O^2 I(s)\right], \tag{2}$$

with

$$I(s) = \frac{1}{s}\sum_{ij}{}' \frac{Z_i Z_j}{r_{ij}} e^{-a_{ij}s^2} \sin r_{ij}s. \tag{3}$$

The structure sensitive expression $sI(s)$ is used in the approximation to the radial distribution integral (RDI):

$$rD(r) = \sum_{s=0}^{s_{max}} sI(s) \sin rs. \tag{4}$$

Solving (2) for $sI(s)$ we obtain

$$sI(s) = (s^4 I'(s) - B)s/(Z-f)_O^2 \tag{5}$$

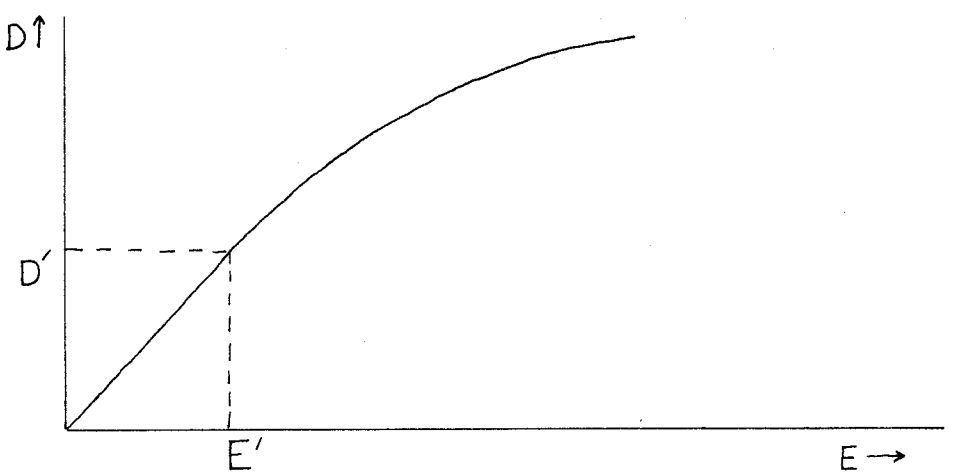
in which we have neglected any constant which affects only the scale intensity.

Reduction of Microphotometer Traces

The data consisted of microphotometer traces of electron diffraction photographs obtained by the sector method. There were three sets of photographic plates corresponding to three camera distances of approximately 49, 19 and 12 cm. The plates were oscillated about the center of diffraction pattern during the process of being traced by the microphotometer in order to smooth out granularity and imperfections of the plate. In spite of the oscillation, the traces still showed high frequency fluctuations, so a smooth curve was drawn by hand through each trace. It was desired to obtain intensity data at integral q-values ($q = 10s/\pi$), so a q-scale was marked directly on each photometer trace.

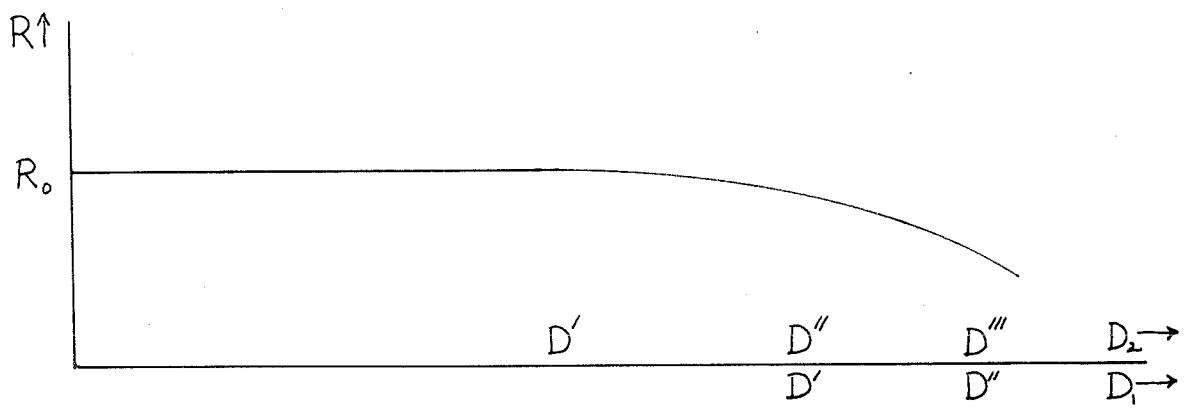
Densities were obtained from the microphotometer traces by the relation $-\log T = D$, which defines the photographic density, D, in terms of the transmission, T, which is proportional to the deflection of the recording needle.

It was desired to convert photographic densities into relative electron intensities. At low exposures, say up to E', the relation between the exposure, E, and the density, D, is linear for practical purposes. At exposures above E' a plot of D vs. E deviates from linearity as is shown schematically in the diagram. The reciprocity law may be assumed valid: The exposure is the product of intensity and time. Therefore two plates differing only in time of exposure will have an approximately constant ratio of densities measured at the same q's, provided both density ranges fall on the linear part of the D-E curve. The ratio will begin to deviate, however, as the exposures on the heavy plate exceed E'.



Typical plot of density vs. exposure

Karle and Karle (11) have discussed the calibration of photographic plates in electron diffraction and the present method is based on theirs. For each set of plates (a set being those made at one camera distance) a table was made of the ratio $R = D_2/D_1$ of the density D_2 of the heaviest plate to that (D_1) of the lightest plate for each q. Then R was plotted against D_1 and D_2 as shown in the diagram:



Plot of ratios of densities vs. densities

The point $D_2 = D'$ where R begins to deviate from a constant value R_o represents the upper limit of the linear portion of the D-E curve. Since we are interested only in relative, not absolute intensities, we may let $D = I_p$, the electron intensity at the plate, for values of D up to D'. In the range $D_2 = D'$ to $D_2 = D''$, the values of R may be used to correct D_2 to an electron intensity scale because D_1 is still on the linear part of the D-E curve. In the above range, letting $I_{p,1}$

be the value of I_p on plate 1, we have the relation

$$\frac{I_{p,2}}{I_{p,1}} = \frac{I_{p,2}}{D_1} = R_o; \qquad I_{p,2} = R_o I_{p,1} = \frac{R_o}{R} D_2 \tag{6}$$

Thus R/R_o was plotted against D_2, giving a calibration curve good up to $D = D''$. Then D_1 was corrected over the range $D_1 = D'$ to D''. The whole process was repeated as many times as was necessary to convert D into I_p for both plates; once or twice was sufficient. The calibration curve was then used to obtain I_p for the other plates in the set. I_p was then averaged point by point for all plates of a set to obtain the final I_p curve for each camera distance.

The above procedure certainly does not take into account all of the factors involved in the conversion of density into intensity. Among other things a complete treatment would consider the following variations from plate to plate: the time of development, the amount of gas in the nozzle region of the camera, the strength of the electron beam, instrumental background, and chemical fog.

Inside the electron diffraction camera a rotating sector modified the exposure in such a way as to partially compensate for the steeply falling background. Accordingly, I_p must be divided by a, the angular sector opening.

Since the plate is flat, not spherical, a correction must be made for the increased distance of the outer parts of the plate as well as its angle of tilt with respect to a diffracted ray. Thus, assuming a thick emulsion, I_p must be divided by a factor of $\cos\theta$ because of the angle of the plate, and a factor of $\cos^2\theta$ because of the increased

distance. The above corrections lead to the expression

$$I'(s) = \frac{I_p(s)}{a(s)\cos^3\theta} = \frac{I_p(s)}{s^3}\left(\frac{r^3}{a(s)}\right)\frac{s^3}{r^3\cos^3\theta} \qquad (7)$$

where the factor $r^3/a(s)$ appears because, due to the shape of the sector, $a(r)/r^3$ or its reciprocal was a convenient representation of the experimental calibration of the sector. Here, r is a radius of the sector. Ignoring a constant factor, one obtains the relation

$$s^4 I'(s) = sI_p(s)\left(\frac{r^3}{a(s)}\right)\frac{1}{\cos^3\theta/2} \; . \qquad (8)$$

Equation 5 may be written

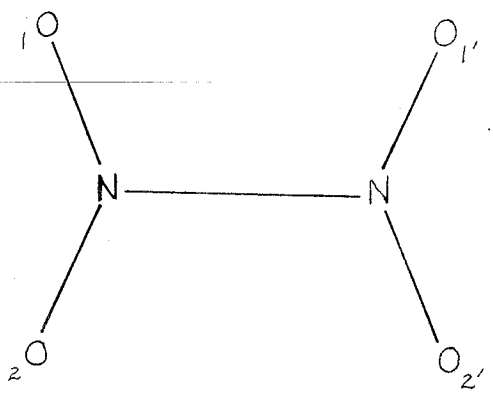
$$sI(s) = \left[\frac{sI_p(s)}{\cos^3\theta/2}\left(\frac{r^3}{a(s)}\right) - B\right]\frac{s}{(Z-f)_o^2} \; . \qquad (9)$$

The background term B was calculated by the theoretical expression, and, when multiplied by an empirical scale factor, could be used in subtracting the background for the 49 cm. plates. The backgrounds for the other two camera distances, however, could not be fitted by a calculated B function, so a smooth background was drawn through a plot of $s^4 I'(s)$ in such a way as to divide the curve into approximately equal positive and negative areas over a range of s corresponding, say, to several major features. Then the background was subtracted out and the result was multiplied by $s/(Z-f)_o^2$. Presumably it would have been better to divide by $(Z-f)_o^2$ before making the above subtraction.

Having obtained sI(s) with a different arbitrary scale for each camera distance, it was desired to join the curves together. There was considerable overlap between the three curves: The 49, 19 and 12 cm. plates covered q ranges of 4-60, 25-150, and 73-195 respectively. Two of the curves were multiplied by an appropriate factor to put all three on the same amplitude scale by taking ratios of amplitudes near maxima and minima in the overlap region. The curves were then averaged together in their overlap ranges to obtain the final experimental sI(s) curve. This curve is analogous to a temperature factored visual curve. The first peak, from $q = 0$ to $q = 5$ was drawn in by hand.

Analysis of Data

The RDI was obtained from sI(s) in the usual way. It has six maxima which can be interpreted in terms of a planar D_{2h} structure:



	RDI maxima	Possible planar model
N-O	1.179 ± 0.002 A	1.177 A
N-N	1.755 ± 0.005	1.750
$O_1 \cdots O_2$	2.175 ± 0.005	2.175
$N \cdots O$	2.455 ± 0.005	2.454
$O_1 \cdots O_{1'}$	2.655 ± 0.005	2.650
$O_1 \cdots O_{2'}$	3.430 ± 0.010	3.428

The above set of parameters which were read from the RDI are consistent with a planar model to within 0.005 A in all six bond lengths.

A confirmation of the above assignment of peaks can be given in terms of peak areas in the RDI provided an allowance is made for a certain concentration of NO_2. Comparison of calculated $(Z_i Z_j / r_{ij})$

and measured peak areas lead to the value 22 mole % NO_2 by taking the average of the relative excess areas of the measured N-O and O_1-O_2 peaks. The following table shows the agreement between calculated and measured areas, where 22% NO_2 has been included.

Peak	Calculated Area	Measured Area (relative scale)	% Discrepancy
N-O	217	227	+4%
N-N	28	27	-4
$O_1 \cdots O_2$	67	64	-5
$N \cdots O$	91	94	+3
$O_1 \cdots O_{1'}$	48	49	+2
$O_1 \cdots O_{2'}$	37	36	-3

The discrepancy can be accounted for as being due to the uncertainty in the base line of the peaks.

Assuming gaussian peaks, the peak half width, $w_{1/2}$, is related to the vibration parameter a by the expression

$$w_{1/2}^2 = 4(\ln 2)a .$$

The half widths were measured and the following values of a were obtained: 0.0008 (N-O), 0.0033 (N-N), 0.0011 ($O_1 \cdots O_2$), 0.0025 ($N \cdots O$), 0.0052 ($O_1 \cdots O_{1'}$), 0.0025 ($O_1 \cdots O_{2'}$). The value of a_{N-N} was believed to be too large, so a smaller value was used in calculating preliminary theoretical curves.

The intensity curve for the best preliminary (planar) model, the experimental curve, and the RDI curve are shown in the Letter to the Editor of the Journal of Chemical Physics on the next page. This model was refined by least squares as described in the next section. The refinement was restricted to conserve the symmetry D_{2h}-mmm.

Least Squares Refinement

The least squares program for the Burroughs 205 computer described in Part III was modified so that the concentration of NO_2 in the sample of N_2O_4 could be adjusted as a parameter. Thus the approximate curve contained, in addition to the terms representing the structure of N_2O_4, the following two terms*

$$w\left[\frac{2x7x8}{1.187}\exp(-1.0x10^{-3}s^2)\sin 1.187s + \frac{8x8}{2.182}\exp(-1.3x10^{-3}s^2)\sin 2.182s\right]$$

with a composition parameter w corresponding to 22 mole percent NO_2. During the refinement w was adjusted by least squares but the other NO_2 parameters were not varied.

Another modification in the program was the provision for automatically repeating the least squares cycles in a manner analogous to that described on page 20 for MF_6 compounds. A sub-routine was included in which the three long distances of N_2O_4 were calculated from the three short distances (N-O, N-N, and $O_1 \cdots O_2$).

The original experimental intensity curve was used in the early cycles of least squares refinement in which the following parameters

*These NO_2 parameters were supplied by Dr. Hedberg (17).

Reprinted from THE JOURNAL OF CHEMICAL PHYSICS, Vol. 25, No. 6, 1282–1283, December, 1956
Printed in U. S. A.

Molecular Structure of Gaseous Dinitrogen Tetroxide

DARWIN W. SMITH, *Gates and Crellin Laboratories on Chemistry,* Pasadena, California*

AND

KENNETH HEDBERG, *Department of Chemistry, Oregon State College, Corvallis, Oregon*

(Received October 5, 1956)

THE N_2O_4 molecule in the crystal has been shown[1] to be coplanar and to consist of $-NO_2$ groups joined by a very long (1.64 A) N−N bond. The preliminary results of our electron-diffraction investigation of N_2O_4 reveal that the molecule is also coplanar in the gas phase but with an even longer (1.75 A) N−N bond.

Excellent diffraction photographs were made in the new Norwegian apparatus[2] using a rotating sector, 34-kv electron accelerating potential, and a special nozzle in which the gas sample was cooled to -20 ± 5°C. The emergent gas at the scattering point apparently contained less than about 25% NO_2. Plates from three camera distances (12, 19, and 48 cm) were carefully selected and microphotometered while being oscillated to reduce the effect of graininess; the resulting traces led by the usual methods to an experimental intensity function (Fig. 1) for the large range $q=5-195$ ($q=10\pi^{-1}s=40\lambda^{-1}\sin\nu/2$, where λ is the electron wavelength and ν the scattering angle).

The radial distribution curve (Fig. 1), obtained by Fourier inversion of the experimental intensity function without use of a modification function, has six well-resolved peaks which correspond to within a few thousandths of an angstrom to a coplanar model having the distance values shown in the diagram. The theoretical intensity curve (Fig. 1), calculated for pure N_2O_4 using these distance values and the vibration parameters $10^4a=10^4\langle\delta l_{ij}^2\rangle_{Av}/2=9$ (N−O), 22 (N−N), 11 ($O_1\cdots O_2$), 25 (N$\cdots$O), 36 ($O_1\cdots O_1'$), and 25 ($O_1\cdots O_2'$), is in very good agreement with the experimental curve over the entire range. If allowance is made in the calculations for presence of NO_2, the agreement is even better.

The N−O distance and O−N−O bond angle in gaseous N_2O_4 are about 0.01 A longer and 8° larger than in the crystal; indeed, the dimensions of the $-NO_2$ grouping are very close to those of NO_2 itself: N−O=1.188±0.004 A, ∠O−N−O=134°4′±15′.[3] The weakness of the N−N bond expected from the chemistry of N_2O_4 is in accord with its extraordinary length (0.27 A longer than the Schomaker-Stevenson[4] covalent single-bond radius sum) and its large vibration parameter, which is in fact comparable to those found for nonbonded interactions through one bond angle. The coplanarity of the N_2O_4 molecule is remarkable in view of the long, weak N−N bond. An attractive possible explanation is that

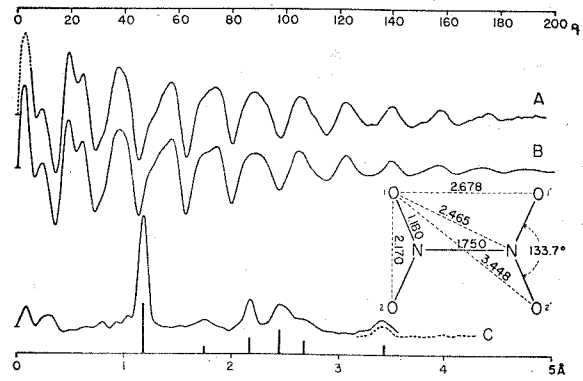


FIG. 1. Electron-diffraction curves for N_2O_4. Curve *A*, experimental intensity; curve *B*, theoretical intensity; curve *C*, radial distribution. The vertical bars beneath curve *C* correspond to the interatomic distances shown in the diagram and have heights proportional to the weights of the terms. The dotted part of curve *C* is part of a second calculation made necessary by our punch-card system in order to obtain points beyond about 3.5 A.

the bond is largely of π- rather than σ type; the length and weakness of the bond are reflections of the poor overlap properties of the orbitals in the bond direction. We are currently refining our structural results; however, we do not expect the distances to change by more than a few thousandths of an angstrom.

We are greatly indebted to cand. real. E. Risberg for design and construction of the low-temperature nozzle, to cand. real. A. Almenningen for preparation of the photographs, and to L. Hedberg for preparation of the figure. One of us (K.H.) wishes to express his gratitude to the John Simon Guggenheim Foundation and to the U. S. Educational Foundation in Norway (Fulbright) for fellowships during which part of this work was carried out. He is especially grateful to Professor O. Hassel and Professor O. Bastiansen for many kindnesses during his stay in Norway.

* Contribution number 2137.

[1] J. S. Broadley and J. M. Robertson, Nature **164**, 914 (1949).
[2] Bastiansen, Hassel, and Risberg, Acta Chem. Scand. **9**, 232 (1955).
[3] G. E. Moore, J. Chem. Phys. **43**, 1045 (1950).
[4] V. Schomaker and D. P. Stevenson, J. Am. Chem. Soc. **63**, 37 (1941).

were adjusted: the three independent distances, the six temperature parameters, the overall scale factor, and the composition parameter. Throughout the refinement, the weight function $p = (q-1)\exp(-0.0012\pi^2q^2/100)$ was used. The resulting difference curve E-C (representing the difference between the experimental and adjusted theoretical curves) indicated that the original background was too high in the regions q = 10 to 40, and 65 to 100, and too low in the region q = 120-180.

An attempt was therefore made to find a suitable background correction by least squares. It is to be expected that the background curve which was drawn visually during the data reduction may contain errors. However the background errors should be "smooth" since the background curve was drawn with much less curvature than the fluctuations due to molecular scattering. One hopes that there will not be any Fourier components of significant weight in the background curve which correspond to molecular distances. Therefore a plausible type of background correction ΔB might be a cosine series having only very low frequency terms. The form chosen was

$$\Delta B = \sum_{n=1}^{9} b_n \cos(n\pi q/200),$$

the coefficients b_n to be adjusted by least squares to fit $(E - \Delta B)$ and C. Only nine parameters could be adjusted, since the program provides for a maximum of 20 adjustable parameters and 11 molecular parameters were already being used. Such a form for ΔB has the essential property of not changing those details of the experimental curve which are important in structure determinations. Thus the lowest frequency component in the experimental curve expected from molecular scattering corresponds to the 1.18 A N-O bond. This component has a period of about 17 q-units. The highest frequency in the

above expression for ΔB has a period equal to 45 q-units and therefore such a background correction will have a negligible effect both on the shapes of maxima and minima, and on the amplitudes of rings as measured from a maximum to neighboring minima.

Although there does seem to be hope for finding a suitable method of adjusting the background by least squares, the ΔB which was obtained by the method described above was felt to be unsatisfactory. It did fit the E-C curve fairly well in the region where the latter was especially poor, but ripples were introduced making the background worse in other regions.

The writer felt that a better background correction curve could be drawn by hand. Therefore, a smooth ΔB curve was drawn in the manner described on page 21 for the MF_6 compounds. This ΔB curve was subtracted from the original experimental curve and the result is shown in Fig. 9, curve E. The experimental intensities of the innermost part of the pattern, q = 5 and 6, are not shown in Fig. 9 and were not used in the succeeding least squares refinement because of their large residuals and the high uncertainty of measurements at such very small scattering angles. The observations covered the wide range q = 7 to 195.

Curve C_1 in Fig. 9 represents the final calculated curve, adjusted by least squares. It corresponds to 7 mole percent NO_2, whereas curve C_2 corresponds to pure N_2O_4 with the same parameters as those for C_1. As one would expect, the 7% NO_2 has little effect, and both C_1 and C_2 fit the experimental curve very nicely. The parameters of N_2O_4 are listed in Table XIV. Also included are the least

squares estimates of the standard deviations, which are probably too small (because of lack of independence in the observations; see Part III) by a factor of about $(20/3 \times 1.2)^{1/2} = 2.4$. In addition the wavelength uncertainty for the Norwegian apparatus was estimated to be 0.2% (14) introducing a possible error of this magnitude in the scale of the molecule. Combining these two effects (see page 53) and multiplying by a factor of two gave the limits of error in Table XIV.

Fig. 9 N_2O_4

E	Experimental
C_1	Calculated, including 7% NO_2
C_2	Calculated, pure N_2O_4
$E-C_1$	Difference
R_{20}	Internal rotation, r.m.s. amplitude 20°
R_{40}	Internal rotation, r.m.s. amplitude 40°
R_F	Internal rotation, free

Critical marks

Curve	Mark	Meaning
R_{20}	Dot	Creation operator; 6 max. too low; 10 min. too low, etc.
R_{20}	Dot circle	Destruction operator; 11 max. too high on right.
R_{40}	Cross	Completely unacceptable features.

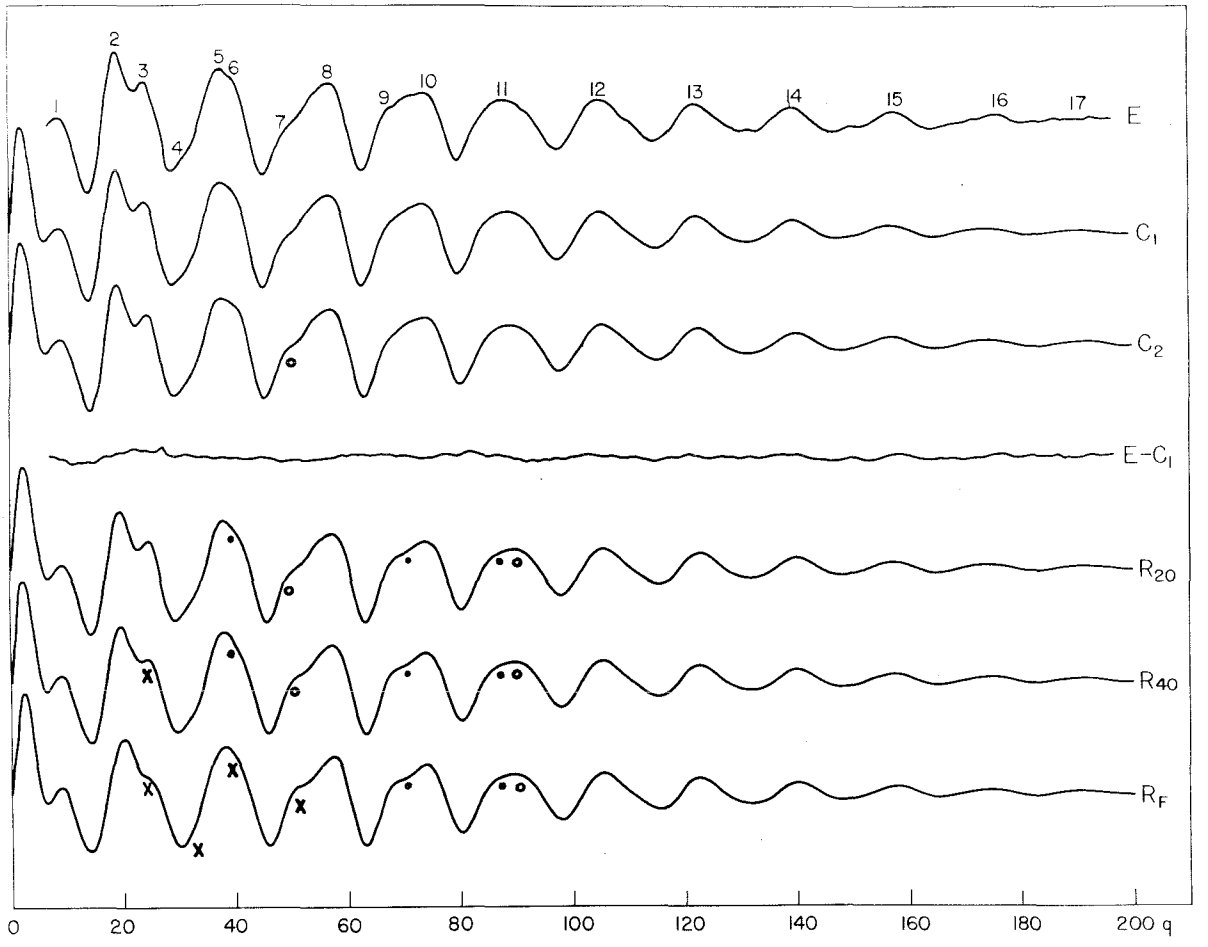

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
E
C1
C2
E-C1
R20
R40
RF
0
20
40
60
80
100
120
140
160
180
200 q

Fig. 9

TABLE XIV

LEAST SQUARES REFINEMENT OF N_2O_4

Parameters for Curves C_1 and C_2, Fig. 9

	r	std. dev.	lim of error	σ	std. dev.
N-O	1.177	0.001	0.003	0.0418	0.0006
N-N	1.752	0.005	.013	.0750	.0064
$O_1 \ldots O_2$	2.173	0.002	.007	.0493	.0018
N...O	2.458			.0778	.0029
$O_1 \ldots O_1'$	2.657			.0958	.0063
$O_1 \ldots O_2'$	3.443			.0737	.0047

$< ONO = 134.8^{\circ}$

Internal Rotation in N_2O_4

Although the theoretical curve based on a planar N_2O_4 model is in excellent agreement with the experimental curve, it was desirable to determine whether introduction of internal rotation about the N-N bond would destroy this excellent agreement. Curves were therefore calculated with varying amplitudes of internal rotation and three such models are listed in Table XV. Assuming a quadratic potential function, the dihedral angle ϕ between the $-NO_2$ groups has a gaussian distribution in the ground state of this hindered rotation. Models R_{20} and R_{40} correspond to root mean square angular displacements σ_ϕ of 20° and 40°, while free rotation was assumed in Model R_F. In these three models, the N-O, N-N, $O_1 \cdots O_2$ and $N \cdots O$ distances and all six temperature parameters are those of the planar model C_2. Since there is negligible difference between C_1 and C_2, no NO_2 was included in the R models.

All of the internal rotation models are distinctly less satisfactory than the planar model, the free rotation model being completely unsatisfactory. In curves R_{20} and R_{40}, relative amplitudes of the main features are satisfactory, but the shapes of a number of these main features are in disagreement as indicated by the critical marks (18) in Fig. 9. A model with $\sigma_\phi = 10^\circ$ is expected to fit the experimental curve in an acceptable manner because during such a small oscillation the most probable values of the long $O \cdots O$ distances do not change by more than 0.003 A from their planar values. At $\sigma_\phi = 20^\circ$ however, the shift is about 0.007 A and the corresponding theoretical curve begins to disagree with the experimental one. However it is possible that a

readjustment of temperature parameters and distances might bring curves with $\sigma_\phi = 20^\circ$ or 40° into agreement. From the above considerations this writer estimates that σ_ϕ is less than 40°.

TABLE XV

INTERNAL ROTATION IN N_2O_4

$$sI(s) = \sum \frac{C_i}{r_i} e^{-\sigma_i^2 s^2/2} \sin r_i s$$

Models R_{20} and R_{40}				Model R_F		
R_{20}	R_{40}					
r	r	C	σ	r	C	σ
1.177 A	1.177 A	5326	0.0418 A	1.177 A	5326	0.0418 A
1.752	1.752	1164	0.0750	1.1752	1164	0.0750
2.173	2.173	3041	0.0493	2.173	3041	0.0493
2.458	2.458	5326	0.0778	2.458	5326	0.0778
3.433	3.433	622	0.0737	3.433	304	0.0737
2.657	2.657	622	0.0958	2.657	"	0.0958
3.427	3.412	1094	0.0737	3.427	"	0.0737
2.663	2.684	1094	0.0958	2.664	"	0.0958
3.412	3.351	752	0.0737	3.412	"	0.0737
2.684	2.759	752	0.0958	2.684	"	0.0958
3.386	3.257	404	0.0737	3.386	"	0.0737
2.716	2.871	404	0.0958	2.716	"	0.0958
3.351	3.135	168	0.0737	3.351	"	0.0737
2.759	3.002	168	0.0958	2.759	"	0.0958
				3.307	"	0.0737
				2.811	"	0.0958
				3.257	"	0.0737
				2.871	"	0.0958
				3.198	"	0.0737
				2.935	"	0.0958
				3.135	"	0.0737
				3.002	"	0.0958
				3.069	"	0.0737
				3.069	"	0.0958

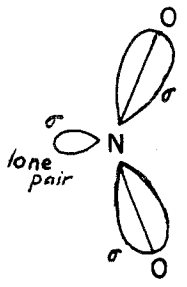
The N-N Bonding in N_2O_4

The least squares refinement confirms the conclusions made earlier* about the structure of N_2O_4. Of special interest is the co-planarity of the molecule in spite of the long (1.75 A) N-N bond. If the length can be explained in terms of weak bonding due to adjacent positive charges on the nitrogen atoms, (19) then the planarity of the molecule needs explanation: One would expect either free rotation or perhaps hindered internal rotation about an equilibrium dihedral angle of 90° because of repulsions of oxygen atoms or of π-bond pairs. On the other hand if the planarity can be explained in terms of π bonding, then one would expect a shorter N-N bond. Smith and Hedberg* suggested that the N-N bond might be largely of π- rather than σ-type: the π-character would keep the molecule planar whereas the lack of σ-character would mean a long bond.

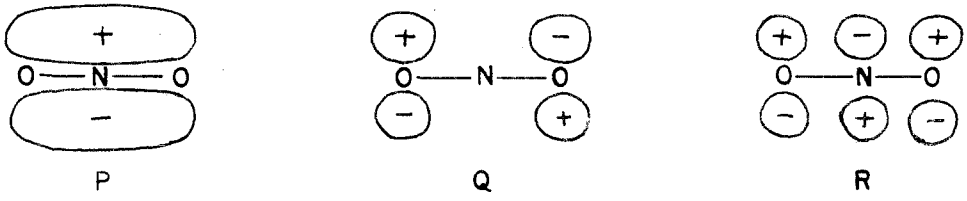
Coulson and Duchesne (20) have even proposed a new type of chemical bond: a "π-only" bond. The N-N bond in N_2O_4 is said to have partial π-character but not any σ-character at all. Let us first consider the molecular orbital (MO) description of NO_2 given by Coulson and Duchesne: There are 13 valence electrons, omitting the K shells and the oxygen 2s electrons. Ten of these occupy five σ orbitals: the two non-bonding oxygen pσ orbitals, the N-O σ-bond orbitals, and the lone-pair σ orbital of nitrogen. This last orbital is isolated on the nitrogen atom and directed away from the oxygen atoms:

---

*Page 65.

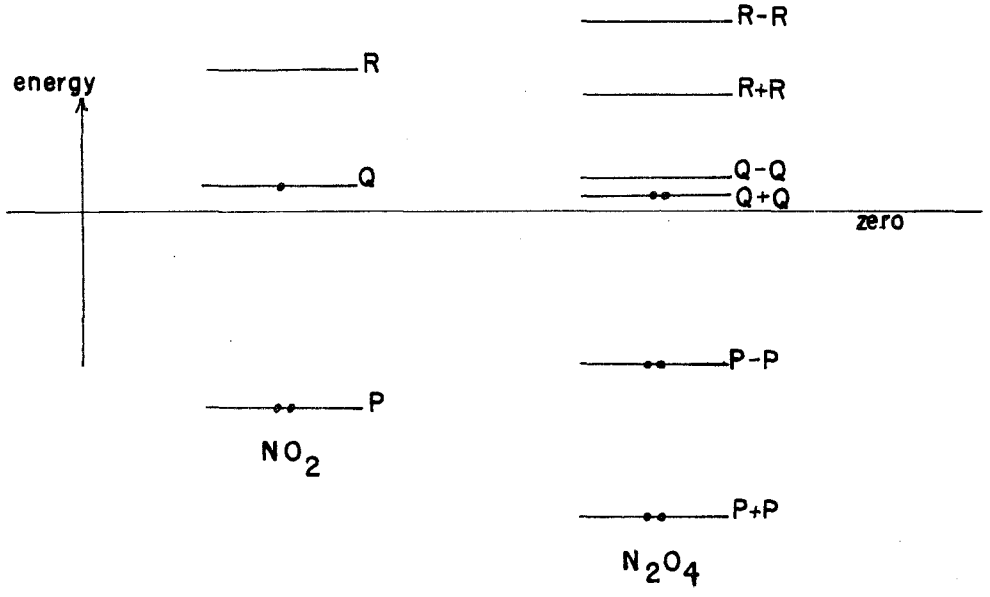


The three remaining electrons are placed two in the bonding molecular π orbital (P in the diagram below) and one in the non-bonding molecular π orbital Q:



Here, the C_2 axis is perpendicular to the paper.

In the above described "lone-pair" model of NO_2, the bonding in dimerization is due to interactions of the π electrons alone because there are no σ orbitals available for bond formation. The expected MO energy levels in the dimer are shown in the diagram below:



The interactions between R+R and P+P and between R-R and P-P due to the identical symmetry of each pair serve to stabilize P+P and P-P and to raise the levels R+R and R-R. The binding energy of N_2O_4 is the difference between the energy of the molecular configuration $(P+P^2(P-P)^2(Q+Q)^2$ and twice the energy $(P)^2(Q)$. The N-N bond is a fractional π bond, but without any σ-character.

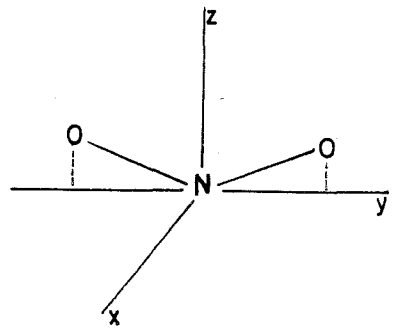
There would be no advantage in promoting one electron from each σ lone-pair orbital of nitrogen in order to get a σ-type N-N bond because this would place two electrons in Q-Q which is antibonding. Coulson and Duchesne point out that such a change in the electronic structure of each $-NO_2$ group would lead to distances and force constants appreciably different from what they are in the monomer, whereas they are nearly identical. Also, a σ-type N-N bond would have a force constant considerably higher than the value 1.29×10^5 dynes/cm calculated by Coulson and Duchesne.

The π-only bond is an attractive explanation of the planarity of N_2O_4 in spite of its long weak N-N bond. It is based on a "lone-pair" model for NO_2. However there is an alternative assignment of electrons in NO_2 which would lead to σ-character in the N-N bond of N_2O_4.

The electronic structure of polyatomic molecules has been discussed by Mulliken (21) in terms of delocalized molecular orbitals. The nuclei of NO_2 are in a bent symmetrical configuration, point group C_{2v}. The approximate MO's are therefore either unchanged or change sign only under the symmetry operations of C_{2v}. If the molecule is oriented in the yz plane with the C_2 axis coinciding with the z-axis,

the MO's will have the following symmetry properties:

	$C_2(z)$	$\sigma_v(xy)$	example: AO on N
a_1	+	+	s, p_z
a_2	+	-	...
b_1	-	+	p_x
b_2	-	-	p_y



Thus, for example, the MO b_1 is antisymmetric with respect to rotation about the two-fold axis and symmetric with respect to reflection in the xz plane as is the p_x atomic orbital on nitrogen.

Mulliken gave the following MO assignment to NO_2 in 1942 (21a):

$$(1a_1)^2(1b_2)^2(2a_1)^2(2b_2)^2(1b_1)^2(3a_1)^2(1a_2)^2(3b_2)^2(4a_1)$$

Walsh (22) also discussed the electronic structure of NO_2 and his assignment (in 1953) was very similar to Mulliken's:

$$(1a_1)^2(1b_2)^2(2a_1)^2(1b_1)^2(3a_1)^2(2b_2)^2(1a_2)^2(3b_2)^2(4a_1)$$

Although they order the orbitals differently, Mulliken and Walsh agree, largely from qualitative arguments, that the odd electron is in orbital $4a_1$, which corresponds to the lone-pair σ orbital on nitrogen. The following methods are useful in determining the unpaired electron distribution in a molecule: (1) Quantitative numerical calculations of the energy levels, such as a self-consistent field, linear combination of atomic orbitals (SFC), (LCAO) MO treatment; (2) Analysis of the electronic spectrum. In addition, the symmetry of the total electronic

wave function (and thus the symmetry of the half-filled orbital) may be revealed by rotational spectroscopy; (3) Determination of the unpaired electron spin density at certain nuclei by electron paramagnetic resonance spectroscopy. These three considerations will be discussed in that order.

To this writer's knowledge, no SFC, LCAO treatments have been made on NO_2. However, such calculations have been made for CO_2 by Mulligan (23) and for O_3 by Fischer-Hjalmers (24). The MO energy levels of CO_2 and O_3 are shown in Fig. 10*, and the interpolation has been made for NO_2. The energy levels have been drawn with zero slope at 180 because, as Walsh (22) has pointed out, there must be either a maximum or a minimum here by symmetry. Although in the case of O_3, the lone-pair σ orbital $4a_1$ falls below $1a_2$ (Ω in the nomenclature of Coulson and Duchesne), it seems unlikely that it would do so in NO_2 according to Fig. 10.

The ionization energies of the various MO's which are indicated in Fig. 10 are the calculated values (23), (24), except for $2\pi_u$, $1\pi_g$ and $1\pi_u$, where the experimental values were used in drawing the levels. (The calculated value is indicated in parentheses in each of these three cases.) Although the interpolation for NO_2 was made on the basis of the bond angles, the bond lengths are also included for comparison.

---

*This figure was suggested by similar ones given by Mulliken (21a), (25) and Walsh (22).

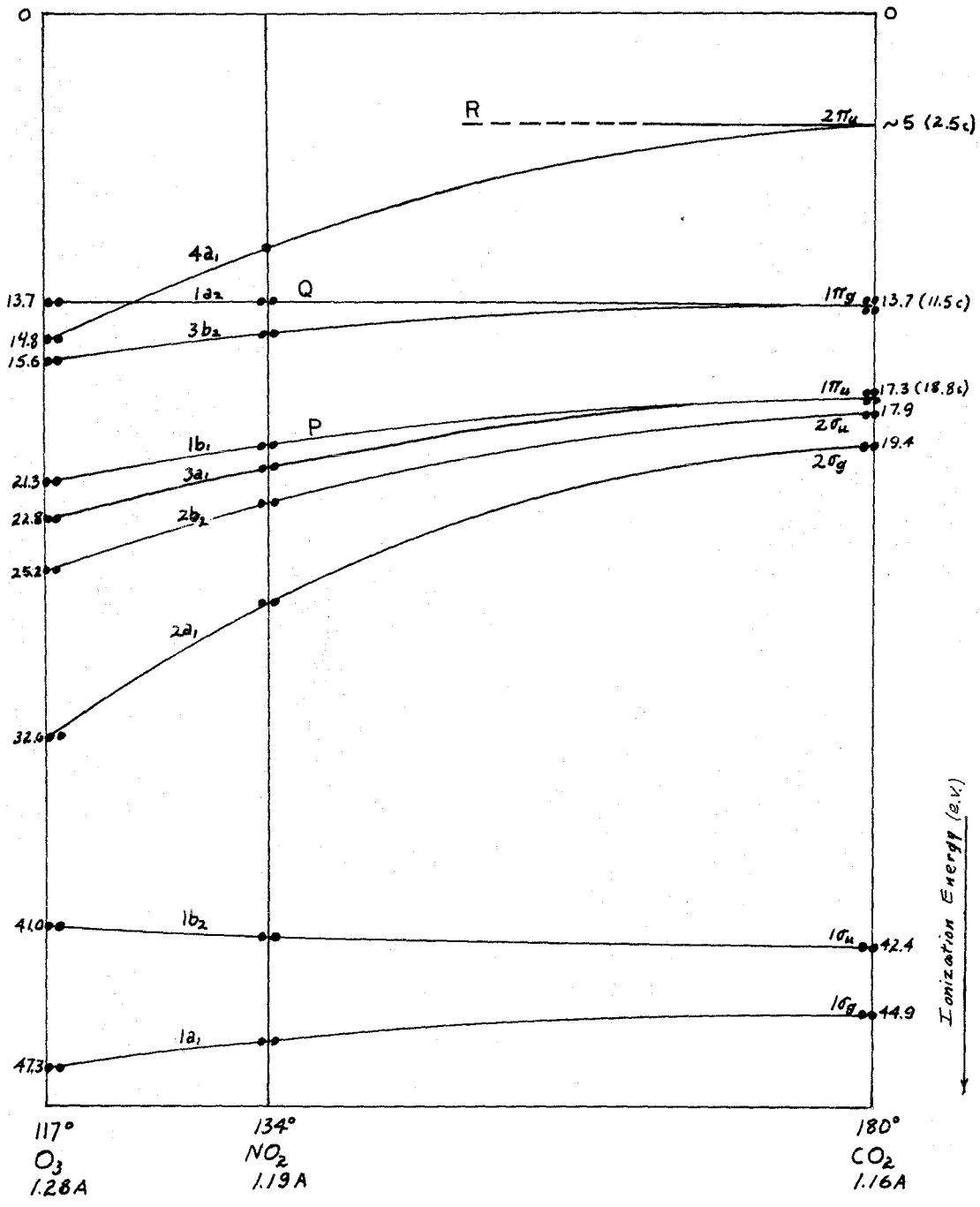


Fig. 10. Molecular Orbital Energy Levels

Discussions of the electronic spectrum of NO_2 have been given by Mulliken (25); Harris, King, Benedict, and Pearse (26); Price and Simpson (27); Walsh (22); and Mori (28). Only Price and Simpson (in 1941) have placed the odd electron in a π orbital; and perhaps current opinion is best expressed by Mulliken (in 1958) who gives an interpretation of the electronic spectrum in terms of an assignment of the odd electron to $4a_1$.

Infrared and microwave studies of the rotational levels of NO_2 can give information about the symmetry of the electronic wave function. However on this basis it is impossible to say whether $4a_1$ or $1a_2$ is the half-filled orbital because they are both symmetric with respect to interchanging the oxygen nuclei.*

Bird, Baird and Williams (31) have studied the electron paramagnetic resonance (EPR) spectra of solutions of NO_2 in CCl_4 and CS_2. They observed a triplet with a splitting of 107 ± 3 gauss (300 ± 9 mc) between successive components for NO_2 at 0.3% by volume in both solvents. Because of rapid molecular tumbling in solution, the spin-spin interaction Hamiltonian reduces to the spherically symmetric Fermi interaction (32). Thus the interaction Hamiltonian is

---

*The infrared spectrum of NO_2 has been studied by Moore (29), who found from the alternation of intensities of rotational lines that $K = 0$ is permitted only for even J in the ground vibrational state; therefore the rotational levels (and hence the electronic state) are symmetric with respect to interchanging the O^{16} nuclei. This is in agreement with either $4a_1$ or $1a_2$ for the odd electron. The assignments of microwave rotational lines by Bird (30) confirm the symmetry of the rotational levels.

$$H' = -\frac{16\pi}{3} g_I \mu_o \mu_N \psi^2(o) I \cdot S$$

where $\psi^2(o)$ is the odd-electron wave function at the nitrogen nucleus. The value of $\psi^2(o)$ in NO_2 was thus reported to be 6.3×10^{24} cm^{-3}, or 18% of the value (34×10^{24} cm^{-3}) estimated by Dousmanis (33) for a nitrogen atom 2s state.*

The large value of $\psi^2(o)$ in NO_2 indicates that the lone-pair σ orbital $4a_1$ rather than the π orbital $1a_2$ contains the unshared electron.

We have thus attempted to show from an MO correlation diagram, from electronic spectroscopy, and from EPR measurements that the odd electron is a σ rather than a π electron, and therefore that the N-N bond in N_2O_4 may be expected to have σ-character. That it is a weak bond may be due to the delocalization of $4a_1$ in NO_2. (In O_3, Fischer-Haljmars (24) found that there are large contributions to $4a_1$ from the in-plane p orbitals on the end atoms.) Perhaps the planarity of N_2O_4 can be explained in terms of the π-electron interactions discussed by Coulson and Duchesne (20), and the length of the N-N bond in terms of the adjacent charge rule, and the de-localization of the $4a_1$ MO in NO_2.

---

*I find $\psi^2(o) = 31 \times 10^{24}$ cm^{-3} by extrapolating to r = 0 the nitrogen 2s atomic wave function given by Hartree (34).

III. APPLICATIONS OF THE METHOD OF LEAST SQUARES IN ELECTRON DIFFRACTION INVESTIGATIONS

Introduction

In recent years, the method of least squares has been applied to several electron diffraction structure investigations. The least squares method in the analysis of visual data was discussed in detail by Hamilton and Schomaker (16), and was applied by Jones, Hedberg and Schomaker (35) in the structure investigations of formyl fluoride. An application of the method of least squares to sector-microphotometer data was described by Bastiansen, Hedberg and Hedberg (14) and was used by them in the refinement of 1, 3, 5, 7-cyclooctatetraene. This writer has written a least squares program for the Burroughs 205 digital computer and has used it in the refinement of the structures of several heavy metal hexafluorides and of N_2O_4.

As applied to sector data, the least squares method involves the direct fitting of a theoretical intensity curve to the experimental one. This method has many advantages over both the correlation method commonly used in electron diffraction studies, and the method of least squares on visual data. In the correlation method, a number of calculated curves are compared with the experimental one. It becomes prohibitively laborious when limits of error are to be found for a large number n of parameters, since about 3^n curves must be calculated. In the method of least squares as applied to visual data in these laboratories, the observations are the measured positions of maxima and minima and

qualitative estimates of the relative heights of two or three neighboring features.

In the present application of least squares, the data representing a continuous intensity distribution are used, not just certain special features which happen to be easily compared by eye. That is, every point on an electron diffraction ring carries equal weight--the whole ring and its shape are used, not just the position of special features and the relative heights of neighboring maxima and minima.

In the correlation method, great weight is placed on those portions of the pattern which are especially sensitive to parameter changes. However, such emphasis on a small part of the pattern can be very misleading, especially if that same sensitive region is subject to large experimental error. The latter is just what happened in many of the heavy metal hexafluoride patterns discussed in Part I of this thesis: The part of the pattern which was most sensitive to the cut-off point unfortunately coincided in several instances with the bad place in the sector. If the correlation method had been used with these sector data, it is very likely that errors in q_c would have been considerably larger. But in the least squares procedure, such regions of high uncertainty can be given small weight if desired. *

---

* In the MF_6 determinations this was not done because data near the cut-off point were so valuable, and in many cases we did not have enough data below q_c to be able to afford to lose the data near q_c. If there had been more data below q_c then it would probably have been better to use a special weighting function. However, a standardized weighting function was used for all sets of data. The important thing is that excessive over-emphasis was not placed on this region of poor data.

The least squares procedure to be described here is essentially a refinement procedure: it is assumed that one has a preliminary theoretical intensity curve which is in approximate agreement with the experimental curve. Furthermore it is assumed that the experimental curve I_e has been reduced to a form which can be represented by the following expression for the calculated intensity curve I_c:

$$I_e \approx I_c(q) = sI(s) = K\sum_i A_i e^{-\sigma_i^2 s^2/2} \sin r_i s, \quad q = 10s/\pi. \quad (1)$$

The notation is that of equation 7, page 14, except that a scale factor K has been introduced here. It is desired to minimize the sum of weighted squares of residuals $\sum p(q)(v(q))^2$ where the residuals $v(q) = I_c(q)-I_e(q)$ are the differences, at integral q values, between the adjusted theoretical intensity I_c and the experimental intensity I_e. The parameters to be adjusted are the interatomic distances r and temperature parameters σ of the molecule under investigation, and the scale factor K.

In order to obtain equations of condition which are linear with respect to these parameters, a set of approximate parameters $\bar{x}$ is needed. (See Appendix III.) Let $I_c(q, \bar{x})$ be a preliminary intensity curve which is in reasonable agreement with $I_e(q)$. The equations of condition in linear form are (equation 3, Appendix III)

$$\sum_j \frac{\partial I_c(q, \bar{x})}{\partial \bar{x}_j} \Delta x_j = I_e(q) - I_c(q, \bar{x}) + v(q), \quad q = q_{min}, q_{min}+1, \cdots q_{max} \quad (2)$$

In matrix language, equation 2 becomes (equation 4, Appendix III)

$$AX = L + V,$$

where

$$A = (\partial I_c(q, \bar{x})/\partial \bar{x}_j),$$
$$X = (\Delta x_j),$$
$$L = (I_e(q) - Ic(q, x)).$$

Part 1 of this program will generate the matrices A and L and a diagonal weight matrix P. The solution of the least squares problem $AX = L + V$ with weight matrix P is described in Appendix III, which is incorporated as part 2 of the present program.

Description of the Program

The input data are I_e, the parameters x_j needed for calculating I_c, and a weighting parameter γ from which P is calculated. The computer calculates the partial derivitives of I_c with respect to r_j, σ_j and K:

$$I_c(q) = K \sum_j A_i \exp(-\sigma_i^2 q^2 \pi^2/200) \sin(r_i q\pi/10) \quad (4)$$

$$\partial I_c/\partial K = \sum_j A_i \exp(-\sigma_i^2 q^2 \pi^2/200) \sin(r_i q\pi/10) \quad (5)$$

$$\partial I_c/\partial r_j = [I_c(r_i, r_j + \Delta r_j) - I_c(r_i, r_j)]/\Delta r_j \quad (6)$$

$$\partial I_c/\partial \sigma_j = -(2\sigma_j \pi^2 q^2/200) K A_j \exp(-\sigma_j^2 q^2 \pi^2/200) \sin(r_j q\pi/10) \quad (7)$$

These derivatives are calculated for the m independent variables K, r_j and σ_j, at each of n values of q; all of these partial derivatives forming the nxm matrix A. This matrix is stored on magnetic tape along with the n-vector $L = (I_e - I_c)$ and the diagonal nxn weight matrix P. The latter is computed from the function $P_{qq} = p(q) = (q-1)e^{-\gamma q^2}$ which, with $\gamma \approx 1.2 \times 10^{-4}$, has been proposed (14) to represent the quality of the data.

For molecules containing some symmetry, as is usually the case in electron diffraction investigations, some of the r's will be restricted by assumed geometrical relations. It is thus convenient to choose certain of these r's as independent variables to be adjusted; the remaining r's to be dependent variables--determined by the independent variables and the assumed symmetry. In equation 6, $\partial I / \partial r_j$ is approximated by differencing, and the program requires that for each independent $r = r_j$, a complete set of distances be entered. In each such set, one independent distance is increased by a small amount, nominally 0.005 Å, the remaining independent distances are unchanged, and the dependent distances are of course based on the independent distances of that set.

Matrices A, P, and L are calculated and stored on magnetic tape 20 rows at a time. When all n rows have been calculated, part 2 of the program is automatically read onto the drum from magnetic tape. Part 2 is identical with the least squares program described in Appendix III except that it has been modified to accept magnetic

tape input rather than paper tape input. Thus part 2 will call in from magnetic tape observational equations in sets of 20, form the normal equations and solve them. The least squares answers are printed out as described in Appendix III.

Because of the non-linearity, the least squares refinement is an iterative procedure, the set of adjusted parameters $(\overline{x}+\Delta x)$ obtained from one refinement cycle forms the basis for the input to the next cycle. Thus, for each refinement cycle one must calculate the new dependent distances based on the least squares corrections to the independent distances, as well as the dependent distances needed for differencing in equation 6. The iterative process is usually a convergent one. In practice, about three or four stages of iteration are sufficient if the starting model was reasonably good.

Interpretation of the Standard Deviations

Caution must be exercised in interpreting the standard deviations which are calculated in the program because they are based on the assumption that the observations are independent, whereas actually a large correlation is to be expected in the observations because they are the ordinates at a fairly close spacing ($\Delta q = 1$) of a continuous experimental intensity curve.

In order to assess the effect of this correlation, let us assume that the true intensity curve can be represented as a sum of sinusoidal terms. The intensity pattern is dominated by one or

several terms with large coefficients, and in order to characterize the pattern intensity readings must be made at a spacing Δq smaller than one-half of this dominant wavelength. This writer estimates that a rough rule would be that a spacing $\Delta q = L/3$ would give independent observations, where L is a dominant wavelength. It is not proposed that observations be made at this spacing; to do so would probably be to lose information about the longer distances. Rather it is suggested that observations be made at $\Delta q = 1$ for convenience in molecules containing distances less than 10 A, and that the computed standard deviations be corrected on the basis that $\Delta q = L/3$ would give independent observations.

The computed standard deviation of x_j is

$$s(x_j) = (V'PV(B^{-1})_{jj})^{1/2}(n-m)^{-1/2} \qquad (8)$$

and if the intensity curve is completely defined by n equally spaced observations, taking $2n$ observations at half the previous spacing will give a new estimate

$$s'(x_j) = (2V'PV(B^{-1})_{jj}/2)^{1/2}(2n-m)^{-1/2} \approx s(x_j)2^{-1/2}. \qquad (9)$$

Thus if the number of observations is increased by a factor of g, the new computed standard deviation is too small by a factor of $g^{1/2}$. From the foregoing considerations it is estimated that the standard deviations in r_j resulting from observations at $\Delta q = 1$ are too small by a factor of $(20/3r_j)^{1/2}$ where r_j corresponds to a dominant term.

Background errors complicate the discussion of standard deviations in the parameters. By background errors are meant errors in drawing the background in the process of reducing the microphotometer traces. Such errors are expected to be of constant sign and slowing varying amplitude over ranges in q which are large compared with the periods of terms arising from molecular scattering. Consider the effect of adding a constant background error Δ to I_e which previously gave a good fit with I_c. The matrix L becomes $L + \Delta$ and the least squares solution X becomes $X = B^{-1}A'PL + B^{-1}A'P\Delta$. The last term will be very small because $(A'P)\Delta$ is the dot product of a vector Δ of constant components with one $(A'P)$ whose components oscillate in a sinusoidal manner. In other words no change in parameters can simulate a constant background increment to the pattern and therefore the parameters will not change much. However, although X is not appreciably affected by the background error, $L'PL$ will be increased by an amount roughly $\Delta'P\Delta$. Thus it seems that the main advantage to be gained from adjusting the background is to reduce the error estimates on the parameters rather than to reduce the errors themselves.

APPENDIX I.

THEORETICAL INTENSITY CURVES PROGRAM FOR THE BURROUGHS 205 COMPUTER*

Summary

This program will calculate the theoretical intensity curves which are used in the electron diffraction structure determinations of gaseous molecules.

Introduction

The theoretical intensity function I_c for a molecule containing R distinct distances may be written

$$I_c(q) = \sum_{j=1}^{R} \frac{C_i}{r_i} e^{-\sigma_i^2 \pi^2 q^2/200} \sin(r_i q\pi/10) = \sum_{j=1}^{R} J_i(q) \qquad (1)$$

where C_i is a constant depending on r_i but not on q, r_i is the i^{th} distinct interatomic distance $q = (40/\lambda) \sin \theta/2$, and the temperature parameter σ_i is the root mean square variation of r_i. It is desired to find I_c for integral values of q.

Scheme of Calculation

A sine table and an exponential table form part of the program tape and are stored in memory as follows:

$\sin(2\pi x/1000)$ is stored in cell $(0000 + x \bmod 1000)$, x = integer;

$\exp(-y^2)$ is stored in cell $(1000 + 1000y/3)$, $0 \leq y \leq 2.997$.

---

* From a report submitted to the Electro Data Division of the Burroughs Corporation.

Thus the cells 0000 to 0999 contain sines of angles from zero to 2π, and cells 1000 to 1999 contain exponential functions from e^{-0} to about e^{-9}.

The computer first calculates $J_1(q)$ (which is the contribution to I_c from the first term of equation 1) for all values of q. Then $J_2(q)$ is found and added to $J_1(q)$, the process continuing until all n terms have been computed and accumulated.

Preparation of Data Tape

1. Store a calculation number of up to 4 digits in the least significant portion of cell 3019 to identify the output data.
2. Store the molecular parameters in cells 3200 to (as a maximum) 3399, using the following form:

   (3200) = $C_1 \times 10^{-10}$; for example, 00000 01000 if

   $C_1 = 1000.$

   (3201) = σ_1, r_1; for example, 04230 01177 if

   $\sigma_1 = 0.0423$ and $r_1 = 1.177$ A.

   (3202) = C_2

   (3203), σ_2, r_2

   etc.
3. Store q_{max} in cell 3018; for example 00000 00140 if $q_{max} = 140$. q_{max} must be a multiple of 20. If nothing is stored in 3018 then q_{max} is automatically 200.

4. The above constitutes one set of input data for one intensity curve. At the end of each set of input data the command

6 CUB 3150

should appear, followed by blank tape. After the last set of input data there must be, in addition to the 6 CUB 3150, blank tape and finally

6 STOP .

Variations

If it is desired to enter the amplitudes as $A_i = \frac{C_i}{r_i}$ rather than C_i, store the A_i in 3200, 3202, etc., and replace 6 CUB 3150 by 6 CUB 2340.

Operation of the Program

Control Switches:

Breakpoint: OFF

Output: TAPE

Skip: (not used)

Format:
Grouping and Counters: OFF
Zero suppress: ON

1. Read in the program tape. Following a sum check and some other short operations, the computer will stop with STOP 1234 in C. (If the sum check is unsuccessful, STOP 5555 will stand in C.)

2. Read in the data tape, which will stop after the input data for one curve has entered the computer. The first intensity curve will be calculated and punched and then the next set of data will be called in automatically. This process continues until the end of the data tape. Be sure there is a 6 STOP there. (If the breakpoint switch is set to 4, the computer will halt on a 7 breakpoint at the end of each calculation.)

Print-out

The print-out includes :

1. Calculation number
2. Input parameters
3. The number of terms n
4. I_c (answers) at integral values of q.

APPENDIX II.

RADIAL DISTRIBUTION FUNCTION PROGRAM FOR THE BURROUGHS 205 COMPUTER*

Summary

This program will calculate the radial distribution function which is used in the electron diffraction structure determinations of gaseous molecules. This function is an approximation to the Fourier transform of the intensity pattern.

Introduction

The RD function to be calculated is

$$rD(r) = \sum_{q} I(q) e^{-aq^2} \sin(rq\pi/10)$$

where $I(q)$ is an experimental intensity function which is known for integral values of $q = \frac{40}{\lambda} \sin\frac{\theta}{2}$; e^{-aq^2} is a convergence factor, often called the artificial temperature factor; and peaks in $rD(r)$ located at r' represent interatomic distances of length r'.

The input data consist of the following: a table of $I(q)$, the artificial temperature parameter a, and certain information about the number of points in $rD(r)$ and the desired spacing of these points.

Scheme of Calculation

The $I(q)$ table is first converted into a table of $A(q)$ defined by $A(q) = I(q)\, e^{-aq^2}$. This step is based on the relation

---

* From a report submitted to the ElectroData Division of the Burroughs Corporation

$$e^{-aq^2} = e^{-a} \approx 1-a \text{ when } q = 1,$$

$$e^{-a(q+1)^2} = e^{-aq^2} e^{-2aq} e^{-a}$$

The calculation of $A(q) \sin(rq/10)$ makes use of a sine table similar to that used in Appendix I. The value of $rD(r)$ is calculated and punched for the first value or $r_1 = \Delta r$ where Δr is the desired spacing. Then $rD(r)$ is found for $r_2 = 2\Delta r$, and so on until r_{max} is reached. Provision is made for variable Δr spacing in different regions of r, enabling one to have a fine spacing in the regions where resolution is important, but saving calculation time by having a coarse Δr spacing in regions where no peaks are expected.

Preparation of Data Tape

1. Store the calculation number (up to four digits) in the least significant portion of cell 3500 to identify the output data.
2. Store the intensity data $I(q)$ in cells 2000 to (as a maximum) 2399. The data must be in fixed point; a convenient range being 00000 00000 to 00000 01000. (Overflow will occur if the numbers are needlessly large.) Store $I(1)$ in cell 2000, $I(2)$ in 2001, etc., $I(q)$ in $2000 + (q-1)$.

   Values of $I(q)$ for q less than q_{min} are usually guessed or supplied from theoretical curves. If it is desired to use $I(q) = 0$ for q less than q_{min} it is not necessary to store zeros in the corresponding cells because 2000-2399 are automatically

cleared before the data tape is read in.

3. Store the artifical temperature parameter a x 10^0 in cell 3139.

4. Store in cells 2980 to (as a maximum) 2999 the code words described below which indicate the desired number of points in rD(r) and their spacing. The form of a code word is

ttt0 00 00uu

The first three digits specify the number of points, 250 if ttt = 250. The last two digits specify the Δr spacing, 0.02 A if uu = 02. Δr must be a multiple of .02: .02, .04, etc. but not .01, .03, ... If variable r spacing is not desired, then, in the example given, 250 points will be computed with 0.02 A spacing between points.

If variable r spacing is desired, a separate code word is needed for each change in Δ r. Thus if one wants 250 points with Δr = 0.02 followed by 100 points with Δr = 0.04, etc., then

(2980) = 2500 00 0002,

(2981) = 1000 00 0004, etc.

5. The last word on the data tape is

6 CU 7000

Operation of the Program

Control switches:

Breakpoint:	OFF
Output:	TAPE
Skip:	(not used)
Format:	Grouping and counters: OFF Zero suppress: ON

1. Read in the program tape. Following a sum check and some other short operations, the computer will stop with STOP 1234 in C. (If the sum check is unsuccessful, STOP 0019 will stand in C.)
2. Read in the data tape twice. A sum check on the data tape will be made each time, and if they agree, the calculation will proceed. If not, then control is transferred to clear the data storage region and the computer stops with STOP 1234 in C. Try again.
3. After the print-out, the complete program and all storage domains will be returned to their original condition and a sum check will be made on the program, and the computer will again stop with STOP 1234 in C. This signifies that the computer is ready to compute another RD.

Print-out

The print-out includes:

1) Calculation Number
2) Input data
3) Spacing code word ttt0 00 00uu
4) rD(r) (answers) for ttt points with $\Delta r = 0.uu$
5) Second spacing code word (if used) vvvo oo oovw
6) rD(r) for second region: vvv points with $\Delta r = 0.ww$ etc., until r_{max} is reached.

APPENDIX III.

A LEAST SQUARES PROGRAM FOR THE BURROUGHS 205 COMPUTER*

Summary:

The purpose of this program is to fit an analytical function to a set of observations. Specified coefficients in the analytical function are adjusted by the method of least squares: the weighted sum of squares of deviations of the observations from their predicted values is minimized. The computer will form the least squares normal equations, and solve them by a complete matrix inversion. The printout includes the least squares solutions, their standard deviations, and the estimated mean error of an observation of weight unity.

System Make-up:

Burroughs 205 with A. F. P. and paper tape input and output.

I. Introduction:

It is desired to solve n equations of condition,

$$\sum_{j=1}^{m} a_{ij}x_j = \ell_i + v_i, \qquad i = 1, 2, 3, \ldots n, \qquad (1)$$

in order to express some observed quantities ℓ_i as a linear combination of the known quantities a_{ij}. The $m < n$ unknown coefficients x_j and the n residuals v_i are to be determined by the method of least squares in that the sum of weighted squares of residuals $\sum_{i=1}^{m} p_i v_i^2$

*From a report submitted to the ElectroData Division of the Burroughs Corporation.

shall be a minimum. The weight p_i is defined to be proportional to the reciprocal of variance of error of the observation ℓ_i. If a linear relation such as equation 1 does not exist, it is usually possible to linearize the problem. Thus if the equations of condition are

$$f_i(x_1 x_2 \ldots x_m) = \ell_i + v_i, \qquad i = 1, 2, 3, \ldots n, \tag{2}$$

the f_i may be expanded in a Taylor's series about an approximate set $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots \bar{x}_m)$ of the unknowns. Then one may write

$$f_i(x) = f_i(\bar{x}) + \sum_{j=1}^{m} \frac{\partial f_i(\bar{x})}{\partial \bar{x}_j} \Delta x_j = \ell_i + v_i, \text{ or}$$

$$\sum_{j=1}^{m} \frac{\partial f_i(\bar{x})}{\partial \bar{x}_j} \Delta x_j = \Delta \ell_i + v_i \tag{3}$$

where higher order terms have been neglected and $\Delta \ell_i = \ell_i - f_i(x)$. Equation 3 now has the same form as equation 1, and the Δx_j, the unknown coefficients, are the corrections to the approximate $\bar{x}_j$.

This program is most easily described in matrix language since it consists largely of matrix subroutines. The reader who is not familiar with matrix language will still find this program easy to use. Let matrix A be the array of the a_{ij}'s of equation 1. It is a rectangular array of n rows and m columns and is loaded into the computer by columns as described later. Let L be the column of ℓ_i's and V the column of v_i's. A description of least squares in matrix language is given by Arley and Buch (36).

In matrix language, the equations of condition become

$$AX = L + V \tag{4}$$

where A is the n x m matrix (a_{ij}), X is the m x 1 matrix or vector (x_j), and so on, L and V each having n components. The minimized sum of weighted squares is V'PV where V' is the transpose of V and P is a diagonal weight matrix with $p_{ii} = p_i$. The least squares normal equations are expressed by the matrix equation

$$A'PAX = A'PL, \quad \text{or}$$

$$BX = A'PL, \quad \text{with solution} \tag{5}$$

$$X = B^{-1}A'PL \tag{6}$$

where B is written for A'PA. The sum of weighted squares is

$$V'PV = L'PL - X'A'PL \tag{7}$$

The external estimate of the mean error of an observation of weight unity is

$$s = (V'PV/n-m)^{1/2}. \tag{8}$$

The estimate of the standard deviation of x_j is

$$s(x_j) = ((B^{-1})_{jj})^{1/2} s. \tag{9}$$

II. Limitations:

The number n of equations of condition is unlimited but the number m of coefficients x_j must be ≤ 20. All data must be in FLOATING POINT.

III. Preparation of Data Tape:

A. Initial (instruction) section:

1. $(m-1) \times 10^{-10}$ is stored in cell 1521; for example, 00000 00017 if m = 18.

2. n - m is stored in cell 1524 in floating point notation; for example, 53148 00000, if n = 166 and m = 18.

3. 6 CU 1530 is then punched, followed by blank tape.

4. Optional features:

   a. If P = unity (all equations have equal weight), replace 6 CU 1530 by 6 CU 1533.

   b. If it is desired to suppress the sum check on the input data, replace 6 CU 1530 by 6 CU 1534. (P may or may not be unity.)

B. Data sections:

If n is greater than 20 there will be several sections of data, each section save the last containing data for 20 equations. (Only for the special case of n = 0 (mod 20) will the last section also contain 20 equations.) The first 20 rows of A are loaded, by columns, starting at 3600. Thus

(3600) = a_{11} IN FLOATING POINT

(3601) = a_{21}

(3602) = a_{31}

etc.

(3620) = a_{12}

(3621) = a_{22}

etc., cells 3600 to (as a maximum) 3999 being used.

The first 20 components of L are stored in consecutive cells starting at 3000. Thus

(3000) = ℓ_1

(3001) = ℓ_2

etc., cells 3000-3019 being used.

The first 20 diagonal elements of P are stored from 1560-

1579.

(1560) = P_1

(1561) = P_2

etc.

For the special but common case of all observations having equal weight, it is not necessary to store anything in 1560-1579, as these cells will automatically contain 51100 00000.

Unless the sum check on the input data is to be suppressed, the sum (mod 1) of all the elements in the first 20 rows of each matrix or vector must be obtained. For the purpose of the sum check, the floating point numbers are considered to have their decimal point immediately to the right of the sign position.* These sums are stored as follows:

(1527) = sum (mod 1) of all elements in 20 rows of L,

(1528) = sum (mod 1) of all elements in 20 rows of A,

(1529) = sum (mod 1) of all elements in 20 rows of P.

If all observations are of equal weight nothing is stored in cell 1529. At the end of this section of data for 20 equations, the command

6 CU 1531

appears, followed by blank tape.

Exactly the same form is followed for each succeeding

---

*The following examples will illustrate:

.50 1234 5678	.50 1234 5678
.59 1111 1111	-.59 1111 1111
.09 2345 6789	-.08 9876 5433

We have dropped any integers and retained only the decimal fraction.

set of 20 equations, 20 rows of A into 3600, 20 elements of L into 3000, and 20 elements of P into 1560, the sum check sums into their appropriate cells and the command 6 CU 1531 followed by blank tape. In the last section, which will be the only section if $n \leq 20$, a special precaution is needed unless $n = 0 \pmod{20}$. The number of rows less one for the last section must be stored in cell 1520. Thus if $n = 80$, nothing is stored in 1520. If $n = 83$, 00000 00002 is stored in 1520 in the last section.

Following the last data section is again the command

6 CU 1531

and the blank tape, and then a final command

6 CU 1532.

IV. Operation of the Program:

The control switches are set as follows:

Breakpoint Switch:	OFF
Output:	TAPE (or PAGE for a short problem)
Skip Switch: (Optional)	ON for short print-out OFF for long print-out
Flexowriter:	Grouping and Counters OFF

The program tape is read in, and, following a sum check and some other short operations, the computer stops with STOP 1234 in C. The data tape, prepared as previously described, is now read in and a sum check is made on the first 20 equations. If successful, the computer will proceed with the calculations and call in the next set of 20 equations. On a sum check failure, the computer will stop with STOP 5555 in C. The

print-out of part I occurs as described below. At the end of part I the computer stops with STOP 2124 in C. For part II (which is optional) the rewound data tape is placed in the optical reader and the command

6 CUB 0174

is entered in the keyboard. The computer will stop with PTR 3600 in C. The Clear Button is pressed and the data read in a second time. At the end of part II the computer stops with STOP 0000 in C. A second problem may be run after either the short or long print-out of part I or at the end of part II provided the same options are desired regarding the weighting and the sum check. To run another problem, the command

6 CUB 0000

is entered on the keyboard, and, after the computer stops with STOP 1234 in C, the new data tape is read in.

V. Print-out, Part I:

A. The short print-out (Skip Switch ON):

X, designated M. 04 (Matrix 4), the unknown coefficients.

L'PL, designated l'pl, the sum of weighted squares of the observations.*

V'PV, designated v'pv, the sum of weighted squares of residuals.

s, the standard deviation of an observation of unit weight.

$s(x_j)$, designated sdx, the standard deviations of the x_j's.

---

*The quantity L'PL is of interest when a non-linear problem is treated by least squares. Our x's become corrections to approximate values of the unknowns and the least squares solution is obtained by iteration. One may compare L'PL with the V'PV of the previous cycle of iteration to see if much trouble is caused by nonlinearity. Also, a comparison of L'PL with V'PV for a particular cycle will indicate how far one is from convergence.

B. The long print-out (Skip Switch OFF), in addition to all the above items:

B, designated M. 03, the coefficients of the left-hand side of the normal equations.

A'PL, designated M. 05, the right-hand side.

BX-A'PL, designated M. 00, a measure of the errors in solving the normal equations, obtained by inserting the computed X into the normal equations, and finding the difference between the right and left members.

B^{-1}, designated M. 02, the inverse of B.

VI. Print-out, Part II (Optional):

v_i, the residuals of each observation equation, printed 20 at a time, without designation but with an extra carriage return between each group of 20.

V'PV, designated v'pv dir, the sum of weighted squares of residuals calculated directly from the v's obtained in Part II.

diff, the difference between the directly calculated V'PV and that calculated in Part I from equation 5. This difference should be negligible; it is a check on the formation and solution of the normal equations.

VII. Description of the Code:

In order to handle an unlimited number m of observations, provision has been made for entering the data into the computer for 20 condition equations at a time. After the first set of 20 equations has been read in, its contribution to the normal equations and to L'PL is found. Only 20 equations are read in at a time in order to avoid relying on magnetic tape storage. At this point, the first set is destroyed and the second set of 20 equations is read in. Its contribution to the normal equations and to L'PL is found and added to that of the first set, so that one now has the normal equations corresponding to the first 40 condition

equations. This process is continued until all of the data are read in. A set may actually contain any number h of equations (provided $h \leq 20$ and $(h-1) \times 10^{-10}$ is stored in cell 1520) but it is most convenient to take 20 equations for each set save the last. The partitioning described above is possible since equation 5 (the normal equations) may be written

$$\sum_k A_k' P_k A_k X = \sum_k A_k' P_k L_k \qquad (10)$$

where A_k, P_k, and L_k are sub-matrices or sub-vectors obtained by partitioning by rows. The kth sub-matrix of A must of course have the same number of rows as the kth sub-matrix of P and as the kth sub-vector of L.

The final command 6 CU 1532 on the end of the data tape signifies to the computer that the read-in is complete. The matrix B is now inverted and its inverse refined. The inversion is obtained by a SAC* matrix sub-routine using stright elimination with positioning in

*SAC, semi-automatic coding, is a system of coding which automatically compiles relatively coded sections into a program with absolute addresses. The sections may all be written by the coder or some of them may be sub-routines taken from the SAC sub-routine library. The present program makes extensive use of SAC sub-routines.

Perhaps the most useful feature of SAC is the simplicity which it lends to the writing of programs that have matrix operations, since a number of matrix sub-routines are available. For example, in the multiplication of two 20 by 20 matrices the 400 elements to be found are each the sum of 20 products. These 8000 multiplications and 8000 additions are programmed by the coder who uses SAC by writing a simulated matrix command containing only three words (after some definitions have been given to the computer):

0000 31 8501	Matrix
0000 00 8512	Multiply
0102 03 0000	Matrix 1 times Matrix 2. Store the product in the domain of Matrix 3.

For a description of SAC, see reference (37).

rows. To obtain a better approximation to B^{-1} than the computed inverse R, consider the error matrix E:

$$E = RB - I \text{ (definition of } E) \tag{11}$$

where I is the identity matrix. We may then write

$$(I + E)B^{-1} = R ; \tag{12}$$

$$B^{-1} = (I + E)^{-1}R \approx (1-E)R, \tag{13}$$

where the approximation may be verified by substituting $R - ER$ for B^{-1} in equation 12, and is valid provided E is small.

Let $R_1 = R$ be the first approximation and then R_2 may be defined as $R_1 - E_1R_1$. Let $E_2 = R_2B - I$ and so on. Then the $i + 1^{st}$ approximation is

$$R_{i+1} = R_i - E_iR_i. \tag{14}$$

The number of refinements is thus $i - 1$ and is specified by storing $(i - 2)x10^{-10}$ in cell 1523. If no specification is made, cell 1523 will contain zero and one refinement is automatic. The refinement of B^{-1} is often necessary because of round-off errors or ill-conditioning. An idea may be had of the success of the inversion and refinement by comparing BX-A'PL with A'PL. If B was not accurately inverted, X will be inaccurate and BX-A'PL will not be small compared with A'PL. If B is singular or nearly so an overflow will occur within the SAC inversion sub-routine.

For matrices encountered by the writer in refining electron diffraction structures of molecules by least squares, one refinement gave more than sufficient accuracy in the results. However in certain cases,

for example in the fitting of power series polynomials to experimental data, a program with double-precision matrix inversion, (that is, operations with 18 rather than eight decimal digits) may be needed. Perhaps orthogonal polynomials could be used, however.

After B^{-1} is refined, X is computed and the print-out begins. Since the input data are destroyed during part I, the data tape must be read in a second time if the residuals for each condition equation are desired.

VIII. Program Execution Times:

As an example of the time required for the complete program, 134 equations in 19 unknowns required about 35 minutes for reading in, calculation, and punching out answers on the high speed punch.

IX. Variations:

The program may be modified to print A or L or both in part II. To print A,

0000 31 1020 is stored in cell 0187;

to print L,

0000 03 0510 is stored in cell 0206 and

0000 03 0810 is stored in cell 0209.

For each set of 20 condition equations, 20 rows of A will be printed as matrix 7. Then, if L is called for, 20 lines will be printed, each line having an l_i followed by its corresponding v_i.

To obtain an extra refinement of B^{-1}, 00000 00001 is stored in 1523.

REFERENCES

1. B. Weinstock and J. C. Malm, J. Am. Chem. Soc. 80, 4466-4468 (1958).

2. S. H. Bauer, J. Chem. Phys. 18, 27-41 (1950).

3. O. Bastiansen, private communication to Professor V. Schomaker, 1950.

4. J. Bigeleisen, M. G. Mayer, P. C. Stevenson, and J. Turkevich, J. Chem. Phys. 16, 442-445 (1948).

5. K. N. Tanner and A. B. F. Duncan, J. Am. Chem. Soc. 73, 1164-1167 (1951).

6. T. G. Burke, D. F. Smith, and A. H. Nielsen, J. Chem. Phys. 20, 447-454 (1952).

7. J. Gaunt, Trans. Faraday Soc. 49, 1122-1131 (1953).

8. V. Schomaker and R. Glauber, Nature 120, 290-291 (1952); R. Glauber and V. Schomaker, Phys. Rev. 89, 667-671 (1952).

9. J. A. Hoerni and J. A. Ibers, Phys. Rev. 91, 1182-1185 (1953); J. A. Ibers, Ph.D. Thesis, Calif. Inst. Tech., 1954.

10. G. Nazarian, Ph. D. Thesis, Calif. Inst. Tech., 1957.

11. J. Karle and I. L. Karle, J. Chem. Phys. 18, 957-962 (1950).

12. F. A. Keidel and S. A. Bauer, J. Chem. Phys. 25, 1218-1227 (1956).

13. J. A. Ibers and J. A. Hoerni, Acta Cryst. 7, 405-408 (1954).

14. O. Bastiansen, L. Hedberg, and K. Hedberg, J. Chem. Phys. 27, 1311-1317 (1957).

15. B. v. Borries, Z. Naturforsch. 4A, 51-70 (1949); H. Koppe, Z. Physik 124, 658-664 (1948); F. Lenz, Z. Naturforsch. 9A, 186-204 (1954).

16. W. C. Hamilton, Ph.D. Thesis, Calif. Inst. Tech., 1954; W. C. Hamilton and V. Schomaker, unpublished work.

17. K. Hedberg, private communication.

18. W. F. Sheehan, Jr., and V. Schomaker, J. Am. Chem. Soc. 74, 4468-4469 (1952).

19. L. Pauling, The Nature of the Chemical Bond, Cornell University Press, Ithaca, 1948, p. 271.

20. C. A. Coulson and J. Duchesne, Bull. classe sci., Acad. roy. Belg. 43, 522-532 (1957).

21. R. S. Mulliken, Phys. Rev. 43, 279-302 (1933).

21a. R. S. Mulliken, Revs. Mod. Phys. 14, 204 (1942).

22. A. D. Walsh, J. Chem. Soc., 1953, 2260-2331.

23. J. F. Mulligan, J. Chem. Phys. 19, 347-362, 1428-1429 (1951).

24. I. Fischer-Hjalmars, Arkiv for Fysik 11, 529-565 (1958).

25. R. S. Mulliken, Can. J. Chem. 36, 10-23 (1958).

26. L. Harris, G. W. King, W. S. Benedict, and R. W. B. Pearse, J. Chem. Phys. 8, 765-774 (1940); L. Harris and G. W. King, J. Chem. Phys. 8, 775-784 (1940).

27. W. C. Price and D. M. Simpson, Trans. Far. Soc. 37, 106-113 (1941).

28. K. Mori, Sci. of Light 3, 62-69 (1954); 4, 130-147 (1955).

29. G. E. Moore, J. Opt. Soc. Am. 43, 1045-1050 (1953).

30. G. R. Bird, J. Chem. Phys. 25, 1040-1043 (1956).

31. G. R. Bird, J. C. Baird, and R. B. Williams, J. Chem. Phys. 28, 738-739 (1958).

32. S. I. Weissman, J. Chem. Phys. 22, 1378-1379 (1954).

33. G. C. Dousmanis, Phys. Rev. 97, 967-970 (1955).

34. D. R. Hartree, The Calculation of Atomic Structures, John Wiley and Sons, Inc., New York, 1957.

35. M. E. Jones, K. Hedberg, and V. Schomaker, J. Am. Chem. Soc. 77, 5278-80 (1955).

36. N. Arley and K. R. Buch, Introduction to the Theory of Probability and Statistics, John Wiley and Sons, Inc., New York, 1950, Chap. 12.

37. Warga, Semi-automatic Coding for the Datatron, Report No. 1.20.020, ElectroData, Division of Burroughs; K. Hebert and H. Fox, Semi-automatic Coding for the Datatron with Simulated Matrix Commands, Report No. 2.40.010-2.47.190, ElectroData, Division of Burroughs; K. Hebert, SAC for Magnetic Tape, Caltech Computing Center.

PROPOSITIONS

1. In the analysis of electron diffraction sector data, a background line must be drawn in a somewhat arbitrary manner, leading to correlated observational equations in the least squares refinement. Two methods are proposed for obtaining observational equations which are independent of the background.

   a. The first and second derivatives of the experimental curves may be taken as the observations.

   b. The radial distribution function of the observed and calculated intensity curves may be fitted by least squares, with large weight near the peaks and small weight in the other regions.

2. Dipole moment measurements of the N_2O_4-NO_2 system at various temperatures seem to indicate that one of two alternatives exists (1). Either (a) the dipole moment of NO_2 has a large temperature dependence and the dipole moment of N_2O_4 is zero; or (b) the dipole moment of NO_2 does not have a measurable dependence on temperature, but the moment of N_2O_4 is non-zero. It is proposed that neither of these alternatives is true. A study of the dipole moment of NO_2 by means of molecular beam spectroscopy should show that (a) cannot hold, while the fact that N_2O_4 is symmetric rules out (b).

3. Collin and Lossing (2) observed an ionization potential of NO_2 by electron impact. They pointed out that their value, 13.98 ev,

is not the lowest ionization potential. It is proposed that 13.98 corresponds to the removal of an electron from the $3b_2$ molecular orbit.

4. The N-N bond in N_2O_4 has been discussed in this thesis. Pauling (3) proposed that an oxygen bridge structure should be more stable than the symmetrical structure on the basis of the adjacent charge rule.

(a) The stabilities of these two structures are discussed.

(b) It is proposed that a detailed molecular orbital treatment of NO_2 would help us to understand the structures of both NO_2 and N_2O_4.

5. A modified free electron model for the 2s and 2p electrons of triatomic molecules of first row elements is proposed. It enables qualitative predictions to be made about the number of de-localized MO's of each symmetry species, their general shapes, and relative energies.

6. In many problems of mathematical physics, such as those involving partial differential equations, the solutions cannot be expressed in terms of familiar or named functions. In this age of computers, however, the possibility of obtaining such solutions to any degree of accuracy becomes a reality.

Several suggestions are offered for making the computational results and programs of one computing laboratory more available to other scientists. For example, computer programs could be written in a universal language and special translating programs could

translate from this universal language to the machine language of any particular computer.

7. Simple valence-bond theory predicts a resonance energy for cyclobutadiene comparable with that of benzene. Craig (4) has compared the valence-bond and free-electron treatments for benzene and cyclobutadiene and argues that the value of the exchange integral α appropriate to the former is not valid for the latter. He claims that a much smaller magnitude should be used. However, it is proposed that the same value of α should be used for both compounds, and that the failure of valence-bond theory to explain the lack of stability of cyclobutadiene is due to another cause.

8. A problem in topology is proposed - that of finding the total number of possible Kekulé-like structures for a condensed polynuclear hydrocarbon. A preliminary study of this problem has revealed the following partial result:

$$K = 1 + C + (1/2)\sum_{i=1}^{C-2} C_i(C - i - 1) + \ldots$$

where K is the number of Kekulé structures, and C and C_i are determined from a particular Kekulé structure as follows: C is the total number of closed circuits of alternating single and double bonds; and C_i is the number of such circuits which have bonds in common with i other circuits. No higher terms in the expression are needed if the starting structure has at most two independent circuits. It is usually easy to find a starting structure satisfying this restriction for molecules of up to about six rings.

9. Cubic ice is stable at very low temperatures and hexagonal ice at higher temperatures. Frank (5) attempts to explain these facts, and also the variation with temperature of the axial ratio c/a in hexagonal ice, in terms of ionic and covalent character of the hydrogen bond. His arguments are based on the work of Keffer and Portis (6) who showed that when nearest neighbors in a wurzite-type crystal of a binary compound are oppositely charged, the distortion of the lattice which lowers c/a also lowers the electrostatic energy (stabilizes the crystal). It is proposed that Frank's explanation is unsound because the results of Keffer and Portis are not applicable to ice.

10. Arguments have been presented that the origin of life did not require an extremely improbable or cataclysmic event; see, for example, "Chemical Evolution and the Origin of Life," an address by Melvin Calvin (7). It is proposed that in spite of such arguments, God is necessary in the explanation of the origin of life.

REFERENCES FOR PROPOSITIONS

1. C. C. Addison and J. Lewis, Journal of the Chemical Society, (1953), pp. 1869-1874.

2. J. Collin and F. P. Lossing, Journal of Chemical Physics, 28, (1958), pp. 900-901.

3. L. Pauling, The Nature of the Chemical Bond, Cornell University Press, Ithaca, (1948), p. 271.

4. D. P. Craig, Journal of the Chemical Society, (1951), p. 3175-3182.

5. H. S. Frank, Proceedings of the Royal Society (London), A247, pp. 481-492, (1958).

6. F. Keffer and A. M. Portis, Journal of Chemical Physics, 27, (1957), pp. 675-682.

7. M. Calvin, American Scientist, 44, (1956), pp. 248-263.