Acta Cryst. (1964). 17, 538

Least-Squares Refinement of Molecular Structures from Gaseous Electron-Diffraction Sector Microphotometer Data. III. Refinement of Cyclopropane.*

BY O. BASTIANSEN,[†] F. N. FRITSCH AND KENNETH HEDBERG[†]

Department of Chemistry, Oregon State University, Corvallis, Oregon, U.S.A.

(Received 1 May 1963)

The method of least squares presented in the preceding two articles was applied to the cyclopropane structure, using data from the Norwegian diffraction apparatus (Skaneke, 1960). The general refinement (including $H \cdots H$ interactions) led to the following results for the six distances (r and d) and root-mean-square amplitudes of vibration (l) , all in \tilde{A} , and the two interesting angles: $r_{\text{C}(1) \text{C}(1)} = 1.510 \pm 0.002$, $r_{\text{(C)}\text{H}(1)} = 1.089 \pm 0.003$, $r_{\text{C}(1)H(2)} = 2.234 \pm 0.006$, $d_{\text{H}(1)H(1')} = 1.838 \pm 0.008$, $d_{\text{H}(1)H(2)} = 2.521 \pm 0.017, d_{\text{H}(1)H(2')} = 3.120 \pm 0.010, l_{\text{C}(1)C(2)} = 0.048_3 \pm 0.001_7, l_{\text{C}(1)H(1)} = 0.084_3 \pm 0.004_1,$ $l_{\text{C}(1)}_{\text{H}(2)} = 0.114_1 \pm 0.006_8, \quad l_{\text{H}(1)\text{H}(1')} = 0.051_1 \pm 0.023_8, \quad l_{\text{H}(1)\text{H}(2)} = 0.214_8 \pm 0.107_6, \quad l_{\text{H}(1)\text{H}(2')} = 0.087_1 \pm 0.005_4$ $0.043_1, \angle H(1)C(1)H(1') = 115.1 \pm 1.0^{\circ}$ and $\angle C(2)C(1)H(1) = 117.7 \pm 0.4^{\circ}$. The distance values correspond to r_a , which differs slightly from r_e ; the mean amplitudes differ correspondingly from l_e . The standard errors include estimates of all known systematic effects deriving from the experiments and method of data reduction, and estimates of correlation among the intensity data, However, tests of various refinement conditions have revealed certain systematic changes in the values of the mean amplitudes of unknown cause, which suggest caution. No such effects on the distances **were** observed, and we feel they may be accepted with confidence.

Introduction

It has seemed worthwhile to illustrate the working of the least-squares refinement method presented in the preceding two articles (Hedberg & Iwasaki, 1964, and Iwasaki, Fritsch & Hedberg, 1964; hereafter called H & I and IF & H). Cyelopropane, besides being of considerable structural interest, is excellent for this purpose because it permits demonstration of the major features of the method in an especially simple way. Further attractions were the existence of excellent data and a structural analysis giving a very good trial structure (Skancke, 1960). We are grateful to siv. ing. Skancke for use of his intensity data (on which our refinement is based) and for use of his structural results before publication.

The following section summarizes Skancke's results. This is followed by a description of the general refinement (all distances and mean amplitudes included), and then of some other refinements reflecting the effect of starting model, weight matrix, data interval (Δs) , omission of $H \cdots H$ interactions, and others. The symbols used are defined in IF & H and explained there and in H $\&$ I. A brief discussion of the cyclopropane structure will be published in *Acta Chemica Scandinavica.*

Preliminary analysis

The diffraction data were obtained in the Norwegian apparatus and after reduction yielded the experimental intensity data given in Table 1; the corresponding curve is shown in Fig. 2. The data may be regarded as described by equation (14) , H & I, with constant coefficients nZ^2 for $C \cdots C$ interactions and variable coefficients $n(Z-f)_{\text{H}}Z_{\text{C}}^2(Z-f)_{\text{C}}^{-1}$ and

$$
n(Z-f)_H^2 Z_C^2(Z-f)_C^{-2}
$$
 for $C \cdots H$ and $H \cdots H$

interactions respectively.

The experimental radial distribution curves (Fig. 3), one calculated with no damping and one with

Fig. 1. Diagram of cyclopropane molecule showing atom numbering.

^{*} This work was supported by the Directorate of Chemical Sciences, Air Force Office of Scientific Research, under contract AF 49 (638)-783.

t On leave from the University of Oslo, Norway.

^{~:} Alfred P. Sloan Foundation Fellow.

					(Skancke, 1960)				
S	$I_{\text{obs}}(s)$	\pmb{s}	$I_{\mathrm{obs}}(s)$	\pmb{s}	$I_{\mathrm{obs}}(s)$	\pmb{s}	$I_{\text{obs}}(s)$	8	$I_{\rm obs}(s)$
$\mathbf 0$	$\bf{0}$	10	-675	21	146	31	65 —	42	64
	819		886		321		145		86
	1429	$\mathbf{11}$	925 -		428	32	223		101
	1779		849		488		257		108
ı	1802		672	22	508		252	43	83
	1531		449		469		213		57
	1061	12	168		382	33	126		37
	514		123		189		27		14
$\boldsymbol{2}$	$\overline{5}$		379	23	4		78	44	$\bf{5}$
	386		575		216		152		33
	637	13	697		325	34	205		50
	734		784		385		234		73
$\bf{3}$	727		831	24	413		221	45	85
	663		794		362		172		90
	591	14	682		231	35	117		53
	541		516		82		53		8
4	502		219	25	90		15	46	12
	450		123		263		83		26
	370	15	423		374	36	133		29 ł.
	217		718		413		160		43
5	$\bf{0}$		918	26	390		163	47	37
	204		999		319		142		34
	436	16	924		220	37	104		28
	578		711 —		102		59		$\boldsymbol{9}$
6	633		374	27	46		21	48	36
	564		24		181		32		46
	$324 -$	17	312		290	38	78		40
	15		569		357		107		43
7	337		729	28	387		130	49	34
	579		767		356		145		32
	686	18	715		276	39	139		19
	642		548		157		108		13
8	404		342	29	23		66	50	$\mathbf{11}$
	51		55		103		15		23
	305	19	203 $\overline{}$		219	40	14		25
	601		380		306		64		$\overline{27}$
9	765		510	30	358		93	51	25
	738		560		362		92		23
	558	20	524		331	41	77		19
	280		421		256		40		11
10	77		230	31	151		1		
	400		34		38		31		

Table 1. *Intensity data for cyclopropane*

damping to reduce series termination errors, show not only the expected three major peaks corresponding to C-H, C-C and $C \cdots H$ distances, but give clear indication of all $H \cdots H$ distances as well. The first column of Table 2 lists the distance and mean amplitude values read from the damped curve, those for C-H, C-C, and $C \cdots H$ by fitting Gaussian curves to three points of each peak. The parenthesized values for $l_{\text{C-H}}$, $l_{\text{C-C}}$, and $l_{\text{C}} \dots$ result from a least-squares fit of these parameters ordy, based on intensity data.

General least-squares refinement

It is obvious from the well-resolved peaks of the radial distribution curves that the structure found by Skancke is a good one, and that refinement can lead only to small changes. There remain some structural reasons for application of least squares, namely the important prospect of refining the $H \cdots H$ mean amplitudes: The $H \cdots H$ interactions themselves were reasonably ignored in Skancke's radial distribution analysis (their total contribution to the intensity is very small as may be seen from Fig. 2), except for a check at the end which revealed good agreement between the observed $H \cdots H$ peak positions and those calculated from the $C(1)\overline{C}(2)$, $C(1)H(1)$ and $C(1)H(2)$ distance values.

The most general refinement of the cyclopropane structure possible with the existing least-squares method, assuming D_{3h} symmetry for the molecule, includes treatment of three of the six different distances as geometrical parameters (these are most conveniently taken as $C(1)C(2)$, $C(1)H(1)$, and $C(1)H(2)$ and all six mean amplitudes as vibration parameters (see H $\&$ I).

The trial structure

The interatomic distance and mean amplitude values for the trial structure are given in Table 2. They differ from the values in the first column of the Table 2. Cyclopropane: distances (r and d) and root-mean-square amplitudes (l) , in A , k in arbitrary intensity units of Table

Fig. 2. Intensity curves for cyclopropane.

- I_o : Experimental. The dotted innermost portion was obtained from theoretical curves.
- I_1 : Theoretical. r and l values from 3rd cycle of general refinement.
- Theoretical. $H \cdots H$ terms only from 3rd cycle of I_2 : general refinement.
- Theoretical. r and l values from poor trial structure I_{3} : $(Table 2).$

table for no important reason except for the $H \cdots H$ distances which, of course, are calculated from the geometry determined by the distance parameters.

The J matrix

In order to calculate the derivatives with respect to the distance parameters the J matrix must first be calculated. Using the distance values for the trial structure (Table 3) and proceding as indicated in the Appendix of H $\&$ I the following matrices may be formulated.

$$
\mathbf{X} = \{rc-c, rc-H, \varphi\}
$$
\n
$$
\mathbf{Y} = \{rc...H, d_{\mathbf{H}(1)}... \mathbf{H}(1'), d_{\mathbf{H}(1)}... \mathbf{H}(1'), d_{\mathbf{H}(1)}... \mathbf{H}(2')\}
$$
\n
$$
\boldsymbol{\xi} = \{rc-c, rc-H, rc...H\}
$$
\n
$$
\boldsymbol{\eta} = \{d_{\mathbf{H}(1)}... \mathbf{H}(1'), d_{\mathbf{H}(1)}... \mathbf{H}(2), d_{\mathbf{H}(1)}... \mathbf{H}(2'), \varphi\}
$$
\n
$$
\mathbf{D}_{x}^{\xi} = \begin{pmatrix}\n1 & 0 & 0 \\
0 & 1 & 0 \\
0.9030 & 0.8031 & -0.5349\n\end{pmatrix},
$$
\n
$$
\mathbf{D}_{\xi}^{x} = \begin{pmatrix}\n1 & 0 & 0 \\
0 & 1 & 0 \\
\frac{0}{1.688} & 1.501 & -1.870\n\end{pmatrix},
$$

previous cycle.

Fig. 3. Radial distribution curves for cyclopropane. Experimental curves were calculated from $P(r)/r = \Sigma I(s) \exp(-bs^2) \sin sr$,

8

theoretical curves from

$$
P(r)/r = \sum_{n} \frac{1}{2} A_n (nk_n)^{-\frac{1}{2}} \exp \left[-(r_n - r)^2 / (4k_n) \right] \text{ where } k_n = \frac{1}{2} l^2_n + b.
$$

1. Experimental, $b=0$.
2. Experimental, $b=0$.

Experimental, $b=0.0012$.

3. Theoretical, r and l values from 3rd cycle of general refinement, $b=0.0012$.

4. Theoretical, r and l values from poor trial structure (Table 2), $b=0.0012$.

$$
\mathbf{D}_x^\eta=\begin{pmatrix} 0 & 1\cdot 681 & 1\cdot 180 \\ 1 & 0\cdot 938 & 1\cdot 586 \\ 0\cdot 810 & 1\cdot 745 & -0\cdot 594 \\ \hline 0 & 0 & 1 \end{pmatrix}, \\ \mathbf{D}_\xi^\eta=\begin{pmatrix} 1\cdot 992 & 3\cdot 452 & -2\cdot 207 \\ -1\cdot 677 & -1\cdot 443 & 2\cdot 966 \\ -0\cdot 193 & 0\cdot 853 & 1\cdot 111 \\ \hline 1\cdot 688 & 1\cdot 501 & -1\cdot 870 \end{pmatrix},
$$

The partitioning corresponds to that given in $H \& I$; thus, the J matrix is the first three rows of \mathbf{D}_{ξ}^{n} , the elements of the first row, for example, being

 $\frac{\partial r_{\text{H}(1)}\dots_{\text{H}(1')}}{\partial r_{\text{C-C}}}, \frac{\partial r_{\text{H}(1)}\dots_{\text{H}(1')}}{\partial r_{\text{C-H}}}$ and $\partial r_{\text{H}(1)}\dots_{\text{H}(1')}/\partial r_{\text{C}}\dots$ $_{\text{H}}$.

Refinement cycling

The conditions selected for the refinement were $P=E$, $s_{\text{min}}=1.75$, $s_{\text{max}}=51.75$, and $\Delta s=0.25$. The input data included this information, F_{obs} (Table 1), $n=6$, $i=3$, $j=6$, the scale factor, distance parameters, dependent distances, and mean amplitude parameters of the trial structure (Table 2), and the third row of

 \mathbf{D}_{x}^{ξ} and the first three rows of \mathbf{D}_{x}^{η} shown above. As may be seen from Table 2, the results of the first cycle include an impossible value for $l_{H(1)}\ldots_{H(1')}$. The large change responsible for this, although unreasonable, is not altogether surprising in view of the small weight of the $H \cdots H$ terms. Accordingly, the second cycle was carried out using values for the $l_{\text{H}}\ldots$ H' 's which included only small fractions of the changes derived from the first, that is, reduced shifts. The values used are those parenthesized. The results of the second refinement justified this procedure, which was repeated for the last cycle. Only very small changes in the distance parameters were obtained, and therefore the same J matrix was used for all cycles. Convergence of the process is evident from the distance and amplitude values in Table 2, and from the values of V'PV, which steadily approach the values of N'PN in the preceding column on which the cycle is based.

The error matrix M and final results

The error matrix, diagonal elements of which are squares of the standard errors and off-diagonal elements the average products of the standard errors of the corresponding two parameters $(H \& I(20))$, is

Table 3. *Error matrix M* $(\times 10^6)$ *from*

To include an estimated effect of correlation among the intensity data the value of each

given in Table 3. It must be emphasized that these values are derived solely from the data of Table 1 assuming no correlation, and therefore they are certainiy too small. It must be further emphasized that in order to arrive at realistic estimates of the standard errors the values obtained from the diagonal elements of M into which the effect of correlation has been introduced must be increased by amounts reflecting the effect of sources of systematic error.

These considerations may be applied to the cyclepropane results as follows. We assume arbitrarily that only about one-half of the original data are independent and thereby double the magnitude of each element of M. No strong argument in support of this arbitrary factor can be advanced, but it is not contradicted by the results of test refinements made with data at different intervals *As* which could be expected to reveal evidence of correlation. The effect of sources of systematic error may then be estimated by considering separately those which affect primarily the mean amplitude parameters (photographic blackness correction, scattering factors, and sector calibration) and those which affect the scale of the molecule (electron wave length and camera distances). The first of these may be taken as about 0.02 l_{ij} , a quantity which was obtained by careful study in another investigation (Hedberg & Iwasaki, 1962) and if anything should be generous in the case of cyclepropane. The second may be fairly estimated as about 0.0005r, from records of calibrations made about the time of the cyclopropane experiment. The final results derived from these considerations are presented in Table 4. The error matrix itself would seem to be largely unaffected by uncertainty in the wave length and camera distances and, except possibly for the $H \cdots H$ mean amplitude parameters, little affected by the uncertainty in the scattering factors and in the blackness and sector corrections. Except for these $H \cdots H$ mean amplitudes then, the error matrix, doubled to take account of correlation in the original data, is felt to present a fair estimate of the error correlation and we offer it together with Table 4 as the fullest statement of our results.

Table 4. *Final results for cyclopropane*

Values are those obtained directly from treatment described in text and are not better than three places of decimals for distances and amplitudes and one place of decimals in degrees

* The distance results are r_a values (derived from intensity curve refinement). Both they and the mean amplitude values will differ slightly from *re* and *4.*

Other refinements

A number of refinements were carried out in addition to that described above in order to gain some experience with the working of the method. Of considerable interest is the speed of refinement and quality of result from a poor trial structure. To test this a starting model with values of the major parameters quite different from the best values and excluding the $H \cdots H$ interactions was refined. Other conditions were as in the general refinement. The starting values and results of the third cycle are shown in Table 2; intensity and radial distribution curves for the trial structure are shown in Figs. 2 and 3. The agreement of these results with those from the general refinement is excellent and provides demonstration of the ease with which simple structures may be refined. Indeed, after only one cycle all distance and mean amplitude values were within $0.01~\text{\AA}$ of the best values from the general refinement.

The effect of weighting of the observations was tested by refinements omitting $H \cdots H$ interactions in which the weight matrix was calculated from $P=\mathscr{N}s$ exp($-bs^2$) instead of being taken as unity. Other conditions were as before. In general, it was found that a choice of b which emphasized the outer data at the expense of the inner led to an increase

"th cycle of general refinement

.... ent should be doubled. r , d , and l values in \dot{A} , angles in degrees, k dimensionless

in the values of the mean amplitudes and the CH distances while leaving the $C(1)C(2)$ distance practically unchanged. With larger values of b , emphasizing intermediate and inner data (but not innermost), the value of the $C(1)C(2)$ mean amplitude was markedly reduced (e.g., $b = 0.004$ gave $l_{\text{C}(1)C(2)} = 0.0429 \pm 0.0027$) while affecting other amplitudes and distances only slightly. An example in which the outer data have been emphasized is shown in Table 2.

The effect of taking data at larger intervals *As* is of interest. Refinements again omitting $H \cdots H$ interactions, but otherwise with conditions as before, were attempted with $\Delta s = 0.5$, 1.0, 2.0, and 4.0, but the last of these could not be made to converge for any of the parameters except, apparently, $r_{\text{C}(1)(C(2))}$ and $l_{\text{C}(1)\text{C}(2)}$. This result is not surprising since the C(1)H(1) and $C(1)H(2)$ terms are damped to 10% of their initial values at s values of about 25 and 18 respectively, by which time they are represented in only five or six data. The parameter values from the refinements with $\Delta s = 0.5$ and 1.0 are essentially the same as is obtained with $\Delta s = 0.25$, while that with $\Delta s = 2.0$ gives slightly larger values, but except for $r_{C(1)H(2)}$ with associated standard errors overlapping those obtained with $\Delta s = 0.25$. The standard errors themselves, of course, are larger. The results for the refinement with $\Delta s = 2.0$ are given in Table 2.

Discussion

The dependence of the refinement results for cyclepropane on the different conditions of refinement described above are typical of what we have observed in other, quite different, cases. Our experience can be summarized in respect to its bearing on the values of parameters determined by this method in general; the remarks apply as well to the cyclopropane results.

The distance parameter values and errors may be accepted with confidence, we feel, keeping in mind that the quantities measured *(ra)* actually differ from those with which they tend to be compared, *e.g., re* values from spectroscopic investigations and r_g (center of gravity) values from electron-diffraction radialdistribution curve analyses. This feeling derives from

the lack of any trend in the distance values with systematic changes in the refinement conditions. The mean amplitude parameters on the other hand, which also differ from *le* values, together with their associated errors do not merit the same confidence. The decreasing trend in these values, as data at small scattering angles are increasingly emphasized, invites caution. For, despite appropriate allowance for the factors known to affect these parameters particularly, it must be admitted that the cause of the trend is unknown and the error allowance may be insufficient. This effect was first noted in an investigation of phosphorus trichloride (Hedberg & Iwasaki, 1962); if it proves to be common to experiments in other laboratories as well, the mean amplitude determination by electron diffraction will need careful investigation.

Each cycle of the general refinement of cyclopropane required about an hour on ALWAC III-E, or about three hours in all. The speed advantage of the method over conventional methods is obvious, even with a slow computer. Moreover, for simple structures like cyclopropane the method is exceptionally easy to apply. Thus, the results of the general refinement could surely have been obtained without use of the radial distribution curve by making rough but reasonable guesses for the values of the $C(1)C(2)$, $C(1)C(1)$ and C(1)H(2) distances and their associated mean amplitudes, and refining the resulting poor trial structure; using these values the general refinement could then be carried out introducing rough trial values for the $H \cdots H$ mean amplitudes.

K. H. wishes to thank the Royal Norwegian Council for Scientific and Industrial Research for a fellowship during which this series of articles was prepared.

References

- HEDBERG, K. & IWASAKI, M. (1964). *Acta Cryst.* **17**, 529. HEDBERO, K. & IWASAKI, M. (1962). *J. Chem. Phys.* **36, 589.**
- IWASAKI, M., FRITSCH, F. N. & HEDBERG, K. (1964). *Acta Cryst.* **17, 533.**
- SKANCKE, P. N. (1960). Thesis, Norwegian Technical University, Trondheim, Norway.