# I

# THE CRYSTAL STRUCTURES OF MONO- AND THE TWO DIMETHYLUREAS

# II

# THE DETERMINATION AND USE OF CRYSTALLOGRAPHIC PARAMETERS

Thesis by

Albert Hybl

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1961

# ACKNOWLEDGEMENT

It is with pleasure that I acknowledge the direction and guidance of my research advisor Dr. R. E. Marsh. His valuable encouragement and patient criticism have proven invaluable in the development of my scientific understanding, without which this thesis would not have been written. I also wish to thank Dr. R. B. Corey for the opportunity to observe at first hand the progress of his work on the determination of the structure of an important protein.

I wish to thank the Shell Fellowship Committee and the Alfred P. Sloan Foundation for financial assistance during my graduate work.

I wish to thank the authors and the publishers of Acta Crystallographica for figures 15, 18, 21 and 27. Again, I thank Dr. R. B. Corey for permission to use figure 25 and for making available several of his molecular models. I thank Dr. K. Hoogsteen for permission to use figure 22 in advance of publication. Likewise, I thank Dr. R. Gerdil for figure 24.

During the various stages in the crystal structure determinations, I had need to use several computer programs made available to me by several authors. In particular, I wish to acknowledge the following sources: Dr. R. Pasternak for his 205 least-squares and structure factor program, Dr. K. Hoogsteen for his Fourier program, and Mr.

N. Jones for his bond distances and angles program. There are several other programs which deserve acknowledgement; however, I hope the authors will forgive me for having made the omission.

Finally, it is with a deep sense of gratitude that I thank my mother, who has made possible my educational opportunity.

# Abstract

I. The crystal structures of mono-, and the two dimethylureas were determined by x-ray crystallographic methods. The positional and anisotropic temperature parameters of all heavy atoms were refined by the method of least-squares. All of the hydrogen atoms bonded to the amide nitrogen atoms were found in difference map sections computed in the planes of the molecules. Preferred orientations were found for the two methyl groups on N,N-dimethylurea, but the methyl groups in N-methylurea and N,N'-dimethylurea appear to be rotating. The preferred methyl orientations are neither the perfectly staggered nor eclipsed arrangements and appear to depend, in part, on the manner of packing of the molecules to form hydrogen bonded chains in the crystal.

The average bond distances were found to be: C=O, 1.253 Å; $C_1$-N (carbonyl C), 1.336 Å and $C_2$-N (methyl group C), 1.447 Å. The partial double bond character is calculated to be 40% for the C=O bond and 30% for the $C_1$-N bonds. The methyl groups appear to make little change in the contributions of the resonance structures to the hybrid structure. The average out of plane distance for the heavy atoms is 0.02 Å, the maximum being 0.05 Å.

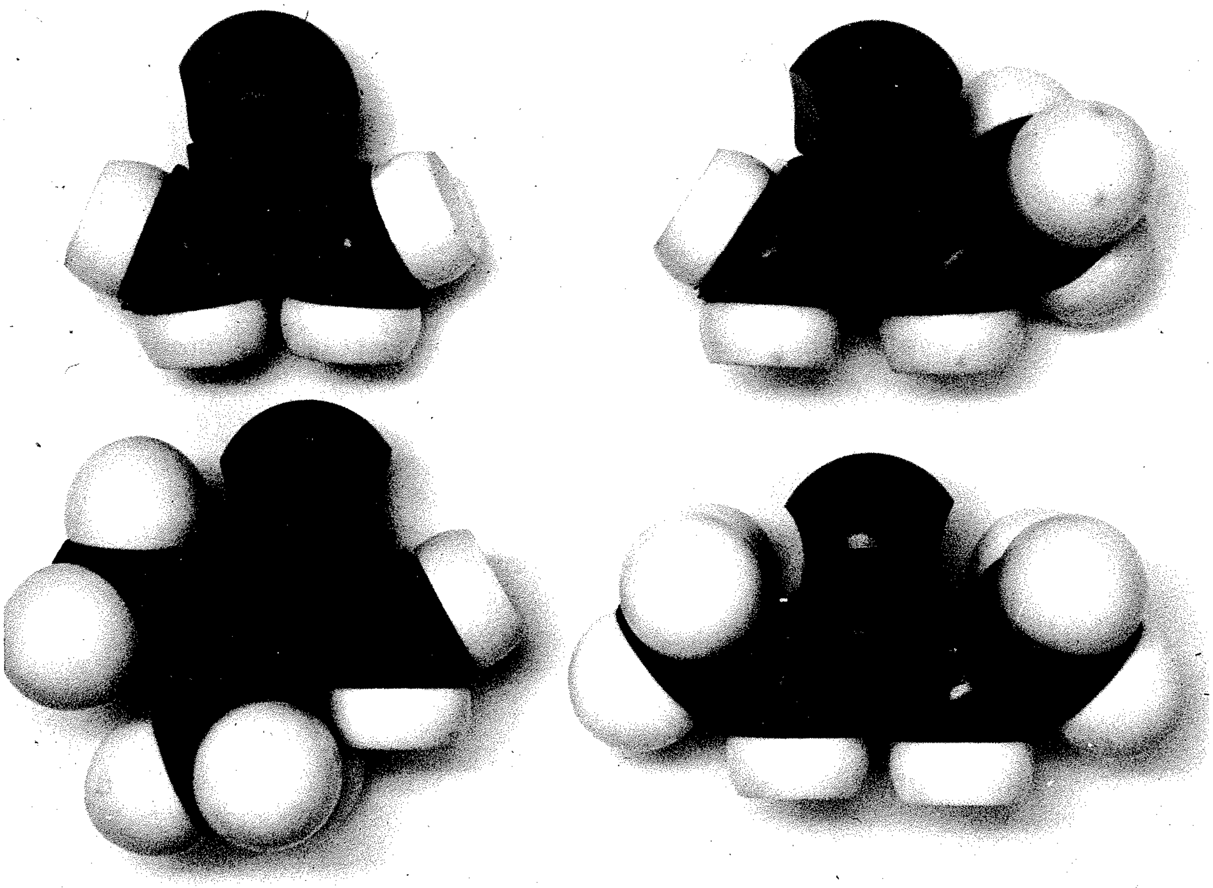
The various types and configurations of cis- and trans-amide hydrogen bonds are reviewed and classified by the type of hydrogen bonded chain structure formed.

II. Expressions employable for the calculation of structure factors and their derivatives for any orthorhombic space group are developed. Complementing this work is a general description of the least-squares method as used to determine a parameter set. Several methods for the conversion of a parameter set into more perceptible quantities, such as interatomic distances, bond angles, planarity of groups of atoms, magnitudes and direction cosines of the principal axes of the vibrational ellipsoids of each atom and the rigid body representation of the thermal displacements and vibrations, are reviewed.

## Table of Contents

# PART I

# THE CRYSTAL STRUCTURES OF MONO- AND THE TWO DIMETHYLUREAS

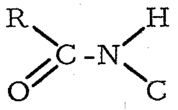


## Introduction

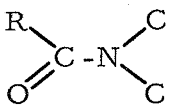
Ten years ago Pauling and Corey (1, 2, 3) proposed several plausible configurations for polypeptide chains in proteins, the most important being the alpha helix structure. One of the ways in which Pauling and Corey attacked the problem of the structure of proteins was to make complete and accurate determinations of the crystal structures of amino acids, small peptides and other simple substances related to proteins, in order that information about interatomic distances, bond angles and other configurational parameters might be obtained that would permit the reliable prediction of reasonable configurations for the polypeptide chain. Their studies started fifteen years prior to their formulation of the alpha helix and are still going on. Recent experimental verification that the alpha helix is an important part of globular proteins has justified this method of attack. The first direct proof that the alpha helix is present in a protein was recently reported by Kendrew (4), who obtained a three-dimensional electron density map of myoglobin at a resolution of 2 Å.

The crystal structure determinations of N-methylurea and of the two dimethylureas were undertaken in part as a continuation of studies on the comparative crystallography of small peptides and related substances initiated by Pauling and Corey. These compounds contain the characteristic configuration of two peptide linkages:

(a)



(b)



, where R in the peptide group is a carbon atom. The peptide linkage (a) is the one found in most peptides, while the linkage (b) is found in peptides containing the amino acid proline or hydroxyproline. Proline is an important constituent of many proteins; for example, approximately 12% of the amino-acid residues in gelatin and collagen are proline and another 10% of the residues are hydroxyproline.

A study of the effect of the methyl groups on the resonance within the molecules, on the manner of hydrogen bonding and on the molecular packing will contribute more knowledge of the complex forces between molecules in crystals. Presumably, the same forces influence the type of folding realized in polypeptide chains in protein molecules and the purine-pyrimidine pairs in nucleic acids, and this information can be used in the deduction of trial model structures of related compounds.

Table 1 lists a comparison of some physical properties of urea and its methyl-substituted derivatives. It has long been recognized that when comparing two compounds within a closely related series, the one with the higher melting point will, in general, have the higher molecular symmetry. As can be seen from the table, N,N-dimethylurea (mp 182°) violates this rule when compared to N,N'-dimethylurea (mp 106°). In addition, the large difference between the densities of

N,N- and N,N'-dimethylurea (1.255 and 1.142 g cm$^{-3}$, respectively) suggests that in the former the molecules have achieved near optimum packing. It was felt that the melting point and density anomalies could be explained on the basis of the crystal structures. A comparative study of these compounds has, indeed, led to a qualitative description of the manner in which these compounds melt.

The contents of the first part of this dissertation have been divided into two major sections. The first section describes the experimental work plus the determination and refinement of the crystal structures of N,N-dimethylurea, N,N'-dimethylurea and N-methylurea, in that order. The second section is devoted to the presentation of the results. General intra-molecular features are discussed, followed by a description of the chain character of the hydrogen bonding. Next, the crystal structures are described in terms of the packing of the chains. Finally the melting point anomaly is examined on a molecular basis.

Table 1. Comparison of some physical properties of urea and its methyl-substituted derivatives.

| Compound | mol. wt. | Density | mp | bp |
|---|---|---|---|---|
| Urea | 60.06 | 1.335 g/cm$^3$ | 132.7°C | d. |
| N-methylurea | 74.08 | 1.204 | 101 | d. |
| N,N-dimethylurea | 88.11 | 1.255 | 182 | |
| N,N'-dimethylurea | 88.11 | 1.142 | 106(102) | 270 |
| N,N,N'-trimethylurea | 102.14 | 1.19 | 75.5 | 232.5 |
| N,N,N',N'-tetramethylurea | 116.16 | (0.972)$^{15}$ | liq. | 177 |
| *cyclo*-Ethyleneurea | 86.10 | 1.45 | 131 | |
| 2-Imidazolone | 84.10 | | *d* 250 | |

## Experimental Work: N,N-dimethylurea

### i) Unit cell and space group.

N,N-dimethylurea crystallizes predominantly in the form of flat plates. A few crystals, however, are in the form of slightly flattened needles. One crystal of the latter type was mounted so that the needle axis [101] was perpendicular to the x-ray beam. Preliminary rotation and equi-inclination Weissenberg photographs showed the crystal to be monoclinic, and revealed the following conditions for non-extinction:

| | |
|---|---|
| $hk\ell$; | no conditions |
| $h0\ell$: | $h + \ell = 2n$ |
| $0k0$: | $k = 2n$ |

On the basis of this system of indexing the space group is $P_{2_1/n}$.

The dimensions of the unit cell were obtained from a rotation photograph about the b-axis and a Straumanis rotation photograph of the [010] zone. They were found to be:

$a = 8.736 \pm 0.003$ Å

$c = 9.294 \pm 0.003$ Å

$b = 6.10 \pm 0.01$ Å

$\beta = 109.58 \pm 0.04$ °

Groth (5) reported the axial ratios 1.2095:1:1.7063 and $\beta = 93.867°$. These figures correspond quite well to an alternate setting in which the a and c directions are taken coincident with the $[10\bar{1}]$ and [101] directions of the $P_{2_1/n}$ space group, respectively. The axial ratios and the monoclinic angle computed from the x-ray data then become 1.2077:1:1.7059 and 93.757°, respectively.

The density reported by Groth is 1.255 g $cm^{-3}$; the density calculated on the basis of four molecules of $C_3H_8N_2O$ in the unit cell is 1.254 g $cm^{-3}$.

ii) Intensity data.

Complete three-dimensional intensity data for Cu $K_\alpha$ radiation were collected on multiple-film equi-inclination Weissenberg photographs taken about the b and the [101] axes. All layer lines having inclination angles less than 40° were recorded (k from 0 to 4 and (h-$\ell$) from 0 to 8). The Ewald sphere contains 1061 possible independent planes; 981 are contained in the effective sphere of reflection with sin $\theta$ = 0.98. Of these, 831 reflections were strong enough to be recorded. For the photographs around the b axis, a lath-shaped crystal was cleaved perpendicular to its needle axis. Another crystal was mounted about its needle axis for the intensity photographs around [101].

The intensity data were estimated by visual comparison with standard scales prepared from the same two crystals. The intensities were corrected for Lorentz and polarization factors using a program written for the Burroughs 205 computer (an analogous program for the 220 is described in appendix II). No corrections for absorption or extinction were made. The corrected intensities from all sets of photographs were correlated to bring them to the same relative scale.

Approximate temperature and scale factors were obtained by Wilson's method (6), and the corrected intensities were put on an absolute scale. The temperature factor B was found to be 1.73 Å $^2$ for the $(h0\ell)$ data and 3.28 Å$^2$ for the complete three-dimensional data. It was decided to use the more reasonable value of 3.28 Å$^2$ for the first structure factor calculations. The atomic scattering curves of McWeeny (7) were used for the early structure factor and least-squares calculations; during the final stages of refinement, the atomic scattering curves of carbon, nitrogen and oxygen were changed to the average of the values given by Berghuis et al. (8) and Hoerni and Ibers (9).

## Determination of the Structure of N,N-dimethylurea.

### i) Derivation of the trial model.

A Patterson vector map projected onto (010) was calculated from the $h0\ell$ intensities. This map indicated that the molecules possess pseudo mm symmetry, the pseudo-mirror planes being coincident with the (101) and $(10\bar{1})$ planes. The Patterson function further suggested that the molecules are oriented in chains along the [101] and $[10\bar{1}]$ directions. Using this information and the Corey and Pauling (10) molecular models an attack was made at the structure.

A satisfactory trial structure was constructed by assuming that the heavy atoms form a planar configuration and that two molecules are hydrogen bonded across a center of symmetry to form a dimer pair.

It was also assumed that these dimer pairs form hydrogen bonded chains extending along the needle axis. The signs of twenty-five $h0\ell$ structure factors were determined by use of structure-factor graphs. A Fourier projection of the electron density onto (010) was computed from these structure factors. The results showed that the assumed structure was basically correct.

The indicated x and z coordinates for the C, N and O atoms were used to compute with greater reliability the signs of the original twenty-five structure factors as well as the signs of a few additional reflections. The scattering curve of nitrogen was used for both the oxygen and $NH_2$ nitrogen atoms since the differentiation between them could not be made at this time. The results of these calculations were used to compute a new electron density map. Some improvement in resolution was apparent and an indication as to which atom might be the $NH_2$ nitrogen and which might be the oxygen was also observable; however, it was decided to continue the investigation for a short while without making the distinction.

This second Fourier map gave good values for the x and z coordinates for all the heavy atoms. In order to obtain approximate y coordinates, a N···O hydrogen-bond distance of 3 Å was assumed between molecules forming the centro-symmetric dimer pairs. This assumption made possible the estimation of the dihedral angle between the molecular plane and xz plane. Estimates of the y coordinates were made graphically by drawing a line in the (101) plane inclined from the $[10\bar{1}]$ axis by the amount of the dihedral angle and passing through the origin. The positions of the atoms in the xz plane were projected onto the $[10\bar{1}]$ axis and the corresponding values of y were taken from the graph. The combined set of coordinates comprised the first set of positional parameters used in the three-dimensional least-squares refinements.

ii) Refinement of parameters.

The first series of three-dimensional refinements of the parameters were made using a Burroughs 205 computer and a structure-factor least-squares program written by Pasternak (11); only the positional parameters were adjusted and a single isotropic temperature factor was used. Eight iterations, including only reflections weighted equally with $\sin \theta$ less than 0.75 (449 in number), reduced the reliability factor, R, from 0.55 to 0.20. Interatomic distances and bond angles computed at this point substantiated the identification of the carbonyl oxygen and the amide nitrogen atoms. All subsequent calculations included the oxygen f-curve assignment. During the next three iterations, all available data were used, and the weighting function was changed to $\sqrt{w} = 0.2 + \sin \theta$. The R factor was reduced from 0.34 to 0.20. In the last four of this series of least-squares iterations, the weighting function was changed to the Hughes scheme (12). The R factor at the end of the isotropic refinement was 0.18.

The structure factor and least-squares program as originally formulated by Pasternak allowed for the inclusion of anisotropic temperature parameters for each atom into the calculation of the structure factor, but did not provide for the least-squares refinement of these parameters. At this time it was decided to modify the program to provide for the calculation of least-squares shifts for the six anisotropic temperature parameters.

After many cycles of anisotropic refinement, the R factor was 0.11. The hydrogen atom coordinates were now added. The coordinates of the two amide hydrogens were obtained from the first five sections (0/25 to 4/25 along b) of an incomplete three-dimensional difference map. Because of machine failure, the remaining sections were not computed. Unfortunately the omitted sections contained the region in which the methyl hydrogens lay. There was, however, some indication on the available sections of the orientations of the methyl groups. The coordinates of the hydrogen atoms of $C_3$ were computed assuming that they are in the staggered position with respect to $H_2$ of the amide group. The coordinates of the hydrogens of the methyl carbon $C_2$ were computed assuming that they are in the eclipsed position relative to the O atom, and thus staggered with respect to the hydrogen atom $H_5$ attached to $C_3$.

iii) Refinement of the hydrogen atom positions.

Several least-squares iterations which included the refinement of the hydrogen parameters were now computed. The final R factor for structure factor calculations on the Burroughs 205 was 0.10.

Since this compound is essentially planar, a three-dimensional Fourier program was written for the Burroughs 220 computer which made possible the calculation of the electron density at the points of a square grid ruled on a general plane. This program is described in appendix III. The electron density section in the least-squares plane of N,N-dimethylurea is shown in figure 1. Contours are at intervals of 1 e. $Å^{-3}$, beginning with 1 e. $Å^{-3}$. The origin of the grid system is at the centroid of the heavy atoms. As in all subsequent general plane Fourier calculations, the points of the square grid are separated by a 0.2 Å interval.

Figure 2 shows a difference Fourier computed in the best plane of the molecule and centered on the mid-point of the two hydrogen atoms of the amide group. Figures 3a and 3b show difference Fourier maps computed in the planes of the hydrogen atoms of the methyl groups. The coefficients of these two planes were determined in the following way. The line representing the bond between the nitrogen and methyl carbon atoms was assumed to be the normal to the desired plane. The C-H distance was assumed to be 1.0 Å and the carbon atom was assumed to be tetrahedrally bonded; accordingly, the center of the plane was placed at 0.33 Å from the carbon atom.

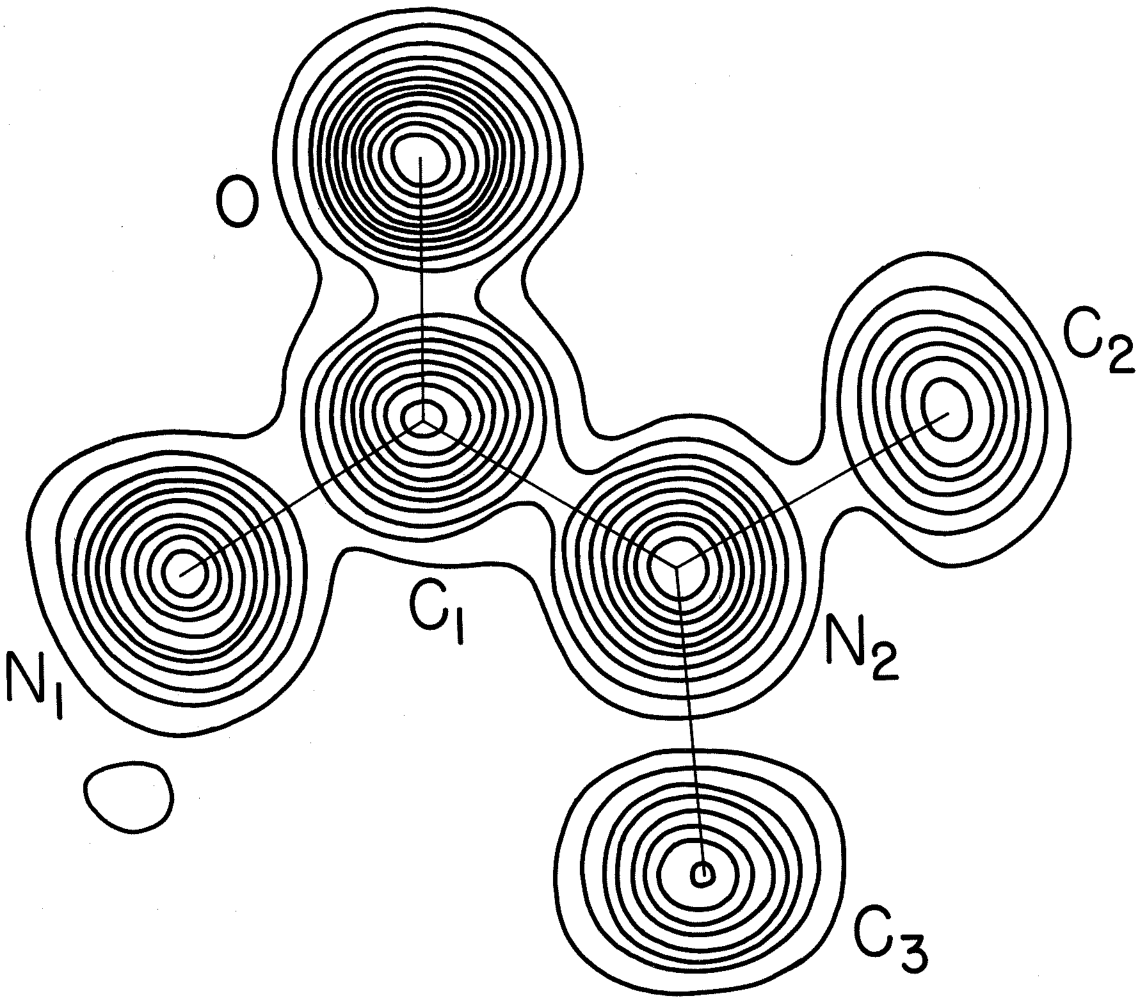


Figure 1. - The electron density section in the least-squares plane of N,N-dimethylurea. Contours at intervals of 1 e. Å$^{-3}$, beginning with 1 e. Å$^{-3}$.

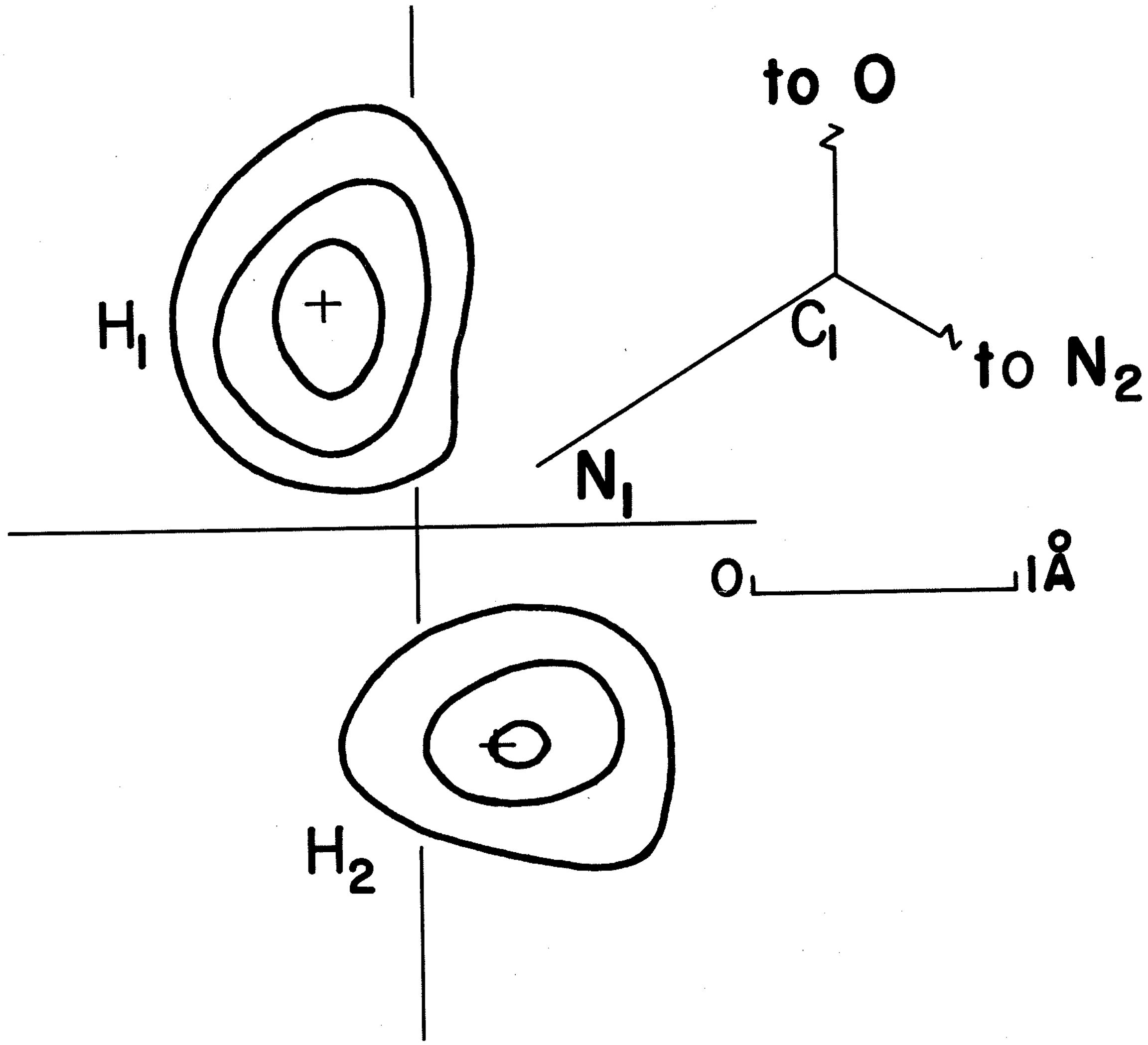
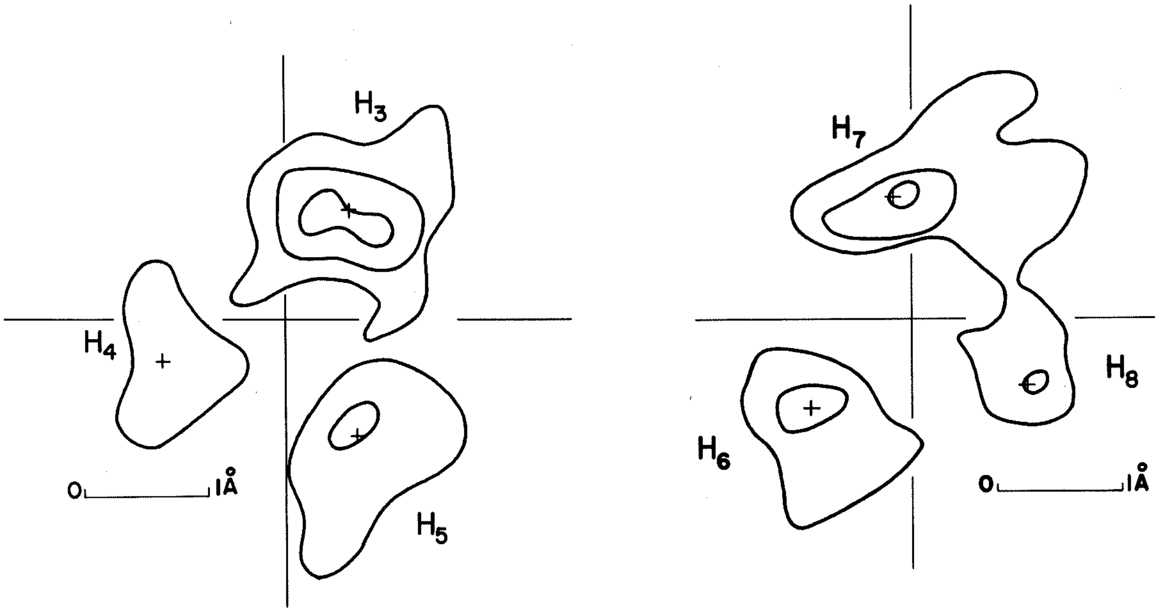
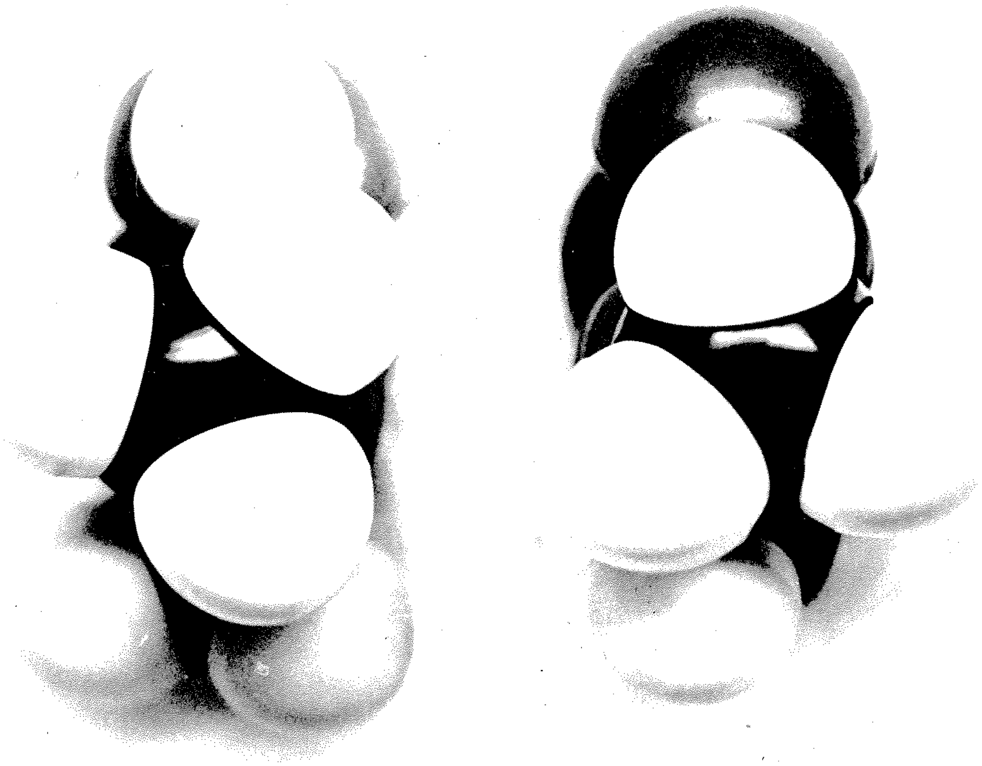


Figure 2. - The difference Fourier in the best plane of N,N-dimethylurea showing only the region containing the amide group. Contours at intervals of 0.2 e. Å$^{-3}$, beginning with 0.2 e. Å$^{-3}$.



(a) (b)

Figure 3. - The difference Fourier maps in the plane of the hydrogen atoms of the methyl groups (a) for methyl carbon $C_3$ (b) for methyl carbon $C_2$. Contours at intervals of 0.2 e. Å$^{-3}$, beginning with 0.2 e. Å$^{-3}$.



(c) (d)

Figure 3. - A photograph of models showing the preferred hydrogen atom orientations of the methyl groups as indicated from the difference maps given in figure 3a and 3b, respectively.

Figure 3a is a view of the plane of the hydrogen atoms of the methyl group $C_3$ (for the numbering of the atoms, see figure 1). The plane of the heavy atoms intersects this plane along the vertical axis. The relative orientation of this plane can be grasped in the following way: Imagine that we are looking through the plane and can see the whole of the molecule. We would see atoms $N_2$ and $C_3$ directly beneath the origin, while the amide nitrogen $N_1$ would be seen near the top of the vertical axis and the other methyl group ($C_2$) toward the bottom. Figure 3c is a photograph of a model illustrating this orientation.

This Fourier shows three peaks which coincide nicely with the final hydrogen coordinates, marked by crosses, obtained from the least-squares process. The peak heights vary from 0.3 to 0.7 e. $Å^{-3}$; the estimated noise level in the difference Fourier is about $\pm$ 0.2 e. $Å^{-3}$. The methyl group is rotated into a position such that $H_3$ is in van der Waals contact (2.00 Å) with $H_2$ of the amide nitrogen.

The third difference Fourier, figure 3b, shows the hydrogen atoms of the other methyl carbon atom $C_2$. The plane of the heavy atoms intersects this plane along the vertical axis with the oxygen atom (not shown) near the top and below the plane. Figure 3d shows a model of this orientation. The methyl group is very nearly eclipsed relative to the oxygen atom, and one of the hydrogen atoms ($H_7$) is bracketed between $H'_2$ and $H'_3$ of an adjacent molecule in the same hydrogen bonded chain. Figures 14 and 16 show this relationship.

The initial examination of these difference Fouriers revealed that one coordinate of $H_2$ was in error; it had been entered into the structure-factor least-squares calculations with the wrong sign. This error was corrected and three more least-squares refinements were made, giving a slight improvement in the R factor. The final R factor is 0.099. At this point the refinement was deemed complete, as no parameter shift was greater than 1/4 of its standard deviation. The final positional parameter and their standard errors are given in table 2. The final structure factor list is given in table 3. The orthogonalized coordinates are given in table 4.

The bond distances and angles in N,N-dimethylurea calculated from the parameters listed in table 2 are given in table 5. A detailed discussion of them will be given in a later section. The best least-squares plane for the heavy atom coordinates as given in table 4 is

$$0.6753\,X - 0.5427\,Y - 0.4994\,Z - 0.1491 = 0.$$

### iv) Temperature parameters discussion.

The magnitudes and direction cosines of the principal axes of the vibration ellipsoids were obtained using the method pointed out by Rollett and Davies (RD) (13). The results are listed in table 7. Table 8 gives the angular deviations of the major axes of the thermal ellipsoids away from the normal of the best molecular plane. It also gives the deviations of the minor axes of the methyl carbon ellipsoids away from their respective N-to-methyl bonds.

Table 2. Final positional parameters and their standard errors for N,N-dimethylurea.

| atom | x | y | z | σx | σy | σz |
|---|---|---|---|---|---|---|
| O | 0.1690 | 0.1974 | 0.0318 | 0.0004 | 0.0005 | 0.0004 |
| $N_1$ | -0.0610 | 0.1792 | -0.1711 | .0003 | .0005 | .0003 |
| $N_2$ | 0.1240 | 0.4571 | -0.1509 | .0004 | .0005 | .0004 |
| $C_1$ | 0.0804 | 0.2765 | -0.0898 | .0004 | .0006 | .0004 |
| $C_2$ | 0.2725 | 0.5738 | -0.0670 | .0006 | .0008 | .0005 |
| $C_3$ | 0.0225 | 0.5659 | -0.2882 | .0005 | .0008 | .0006 |
| $H_1$ | -0.084 | 0.046 | -0.111 | .007 | .010 | .006 |
| $H_2$ | -0.142 | 0.216 | -0.279 | .008 | .011 | .007 |
| $H_3$ | -0.050 | 0.453 | -0.375 | .008 | .011 | .007 |
| $H_4$ | -0.019 | 0.709 | -0.250 | .008 | .011 | .008 |
| $H_5$ | 0.105 | 0.657 | -0.346 | .008 | .010 | .007 |
| $H_6$ | 0.335 | 0.605 | -0.151 | .008 | .010 | .007 |
| $H_7$ | 0.336 | 0.480 | 0.023 | .008 | .012 | .008 |
| $H_8$ | 0.238 | 0.742 | -0.042 | .008 | .012 | .008 |

| | | |
|---|---|---|
| 0k0 | | |
| 0 | --- | 1922 |
| 2 | 80 | -94 |
| 4 | 110 | 85 |
| 6 | 20 | -19 |
| 0k1 | | |
| 1 | 130 | -138 |
| 2 | 182 | 190 |
| 3 | 154 | -196 |
| 4 | 11 | -16 |
| 5 | 59 | 59 |
| 6 | 20 | -18 |
| 7 | 14 | 15 |
| 0k2 | | |
| 0 | 68 | 66 |
| 1 | 208 | 224 |
| 2 | 161 | -173 |
| 3 | <7 | -9 |
| 4 | 39 | 41 |
| 5 | 47 | 43 |
| 6 | <10 | -3 |
| 7 | 27 | 32 |
| 0k3 | | |
| 1 | 19 | -2 |
| 2 | 106 | -102 |
| 3 | 38 | 36 |
| 4 | 78 | 74 |
| 5 | 36 | 29 |
| 6 | 54 | -50 |
| 7 | <7 | -6 |
| 0k4 | | |
| 0 | 74 | -72 |
| 1 | 212 | -235 |
| 2 | 69 | -60 |
| 3 | 112 | -118 |
| 4 | 56 | -46 |
| 5 | 31 | -28 |
| 6 | <9 | 1 |
| 7 | 46 | -42 |
| 0k5 | | |
| 1 | 154 | 144 |
| 2 | <9 | -5 |
| 3 | 85 | -81 |
| 4 | 164 | 166 |
| 5 | 46 | 43 |
| 6 | 8 | -8 |
| 7 | 27 | -28 |
| 0k6 | | |
| 0 | 62 | 61 |
| 1 | 94 | -90 |
| 2 | 30 | 23 |
| 3 | 51 | 51 |
| 4 | 17 | -14 |
| 5 | 17 | -14 |
| 6 | 51 | 48 |
| 0k7 | | |
| 1 | 38 | -37 |
| 2 | <11 | 4 |
| 3 | 54 | -67 |
| 4 | 16 | -13 |
| 5 | 18 | 19 |
| 0k8 | | |
| 0 | 54 | -59 |
| 1 | 56 | -60 |
| 2 | 14 | 12 |
| 3 | 47 | 40 |
| 4 | <9 | -1 |
| 5 | 33 | -10 |
| 0k9 | | |
| 1 | 27 | 35 |
| 2 | 42 | -40 |
| 3 | 43 | 40 |
| 4 | 18 | -17 |
| 0,k,10 | | |
| 0 | 16 | -9 |
| 1 | 39 | -47 |
| 2 | 12 | -13 |
| 3 | 47 | 39 |
| 1k0 | | |
| 1 | 176 | -196 |
| 2 | 118 | -116 |
| 3 | 195 | -198 |
| 4 | 22 | 19 |
| 5 | 41 | 44 |
| 6 | <10 | 9 |
| 7 | 31 | 28 |
| 1k1 | | |
| 0 | 460 | 552 |
| 1 | 11* | 3 |
| 2 | 44 | -38 |
| 3 | 55 | 63 |
| 4 | 78 | 73 |
| 5 | 14 | -17 |
| 6 | 17 | -16 |
| 7 | 8 | 6 |
| $\bar{1}k1$ | | |
| 0 | 240 | 218 |
| 1 | 361 | 383 |
| 2 | 318 | -375 |
| 3 | 181 | -189 |
| 4 | 38 | 44 |
| 5 | <10 | 0 |
| 6 | 35 | 40 |
| 7 | 47 | 39 |
| 1k2 | | |
| 1 | 111 | -110 |
| 2 | 210 | -221 |
| 3 | 79 | 80 |
| 4 | 38 | -38 |
| 5 | 60 | -55 |
| 6 | 77 | -78 |
| 7 | 13 | 12 |
| $\bar{1}k2$ | | |
| 1 | 450 | 508 |
| 2 | 144 | 121 |
| 3 | 43 | 39 |
| 4 | 110 | -114 |
| 5 | 22 | 13 |
| 6 | 17 | -15 |
| 7 | 54 | -40 |
| 1k3 | | |
| 0 | 21 | 23 |
| 1 | 19 | -6 |
| 2 | 77 | 59 |
| 3 | 146 | 140 |
| 4 | <9 | -4 |
| 5 | 104 | 113 |
| 6 | 44 | -47 |
| 7 | 6 | -9 |
| $\bar{1}k3$ | | |
| 0 | 217 | -248 |
| 1 | 225 | 244 |
| 2 | 79 | -72 |
| 3 | 86 | -73 |
| 4 | 17 | -13 |
| 5 | <10 | 6 |
| 6 | <10 | -1 |
| 7 | <7 | 0 |
| 1k4 | | |
| 1 | 15 | -19 |
| 2 | 96 | -125 |
| 3 | 59 | 52 |
| 4 | 160 | 161 |
| 5 | 33 | 28 |
| 6 | 47 | -48 |
| 7 | 25 | 19 |
| $\bar{1}k4$ | | |
| 1 | 91 | 92 |
| 2 | 62 | 49 |
| 3 | 38 | -36 |
| 4 | 111 | 106 |
| 5 | 52 | 44 |
| 6 | 30 | 28 |
| 7 | 34 | -27 |
| 1k5 | | |
| 0 | 118 | -112 |
| 1 | 138 | -133 |
| 2 | 12 | -15 |
| 3 | 72 | -67 |
| 4 | 79 | -70 |
| 5 | 13 | -9 |
| 6 | 26 | 23 |
| $\bar{1}k5$ | | |
| 0 | 18 | 12 |
| 1 | 102 | -105 |
| 2 | 46 | -41 |
| 3 | 10 | 15 |
| 4 | <11* | -11 |
| 5 | 14 | -9 |
| 6 | 48 | 50 |
| 7 | 17 | -17 |
| 1k6 | | |
| 1 | <10 | -5 |
| 2 | <10 | -4 |
| 3 | 55 | -44 |
| 4 | 51 | 37 |
| 5 | 21 | 20 |
| $\bar{1}k6$ | | |
| 1 | 10 | 7 |
| 2 | 73 | -68 |
| 3 | 161 | -159 |
| 4 | 81 | -67 |
| 5 | 43 | 44 |
| 6 | 27 | -18 |
| 1k7 | | |
| 0 | <12 | 0 |
| 1 | 31 | -18 |
| 2 | 136 | 135 |
| 3 | 19 | -16 |
| 4 | 23 | -21 |
| $\bar{1}k7$ | | |
| 0 | 147 | 146 |
| 1 | 14 | -10 |
| 2 | 15* | -9 |
| 3 | 102 | 92 |
| 4 | 13 | 9 |
| 5 | 11 | -12 |
| 6 | 13 | 12 |
| 1k8 | | |
| 1 | 48 | -44 |
| 2 | 33 | 29 |
| 3 | 27 | 29 |
| 4 | 10 | 11 |
| $\bar{1}k8$ | | |
| 1 | 30 | 31 |
| 2 | 23 | 15 |
| 3 | 15* | 3 |
| 4 | 40 | -40 |
| 5 | 14 | 11 |
| 1k9 | | |
| 0 | 92 | -88 |
| 1 | 43 | -51 |
| 2 | 34 | 29 |
| 3 | 35 | 35 |
| $\bar{1}k9$ | | |
| 0 | 14 | -12 |
| 1 | 27 | -31 |
| 2 | 48 | -47 |
| 3 | 43 | 45 |
| 4 | 14 | 13 |
| $\bar{1}$,k,10 | | |
| 1 | <12 | -12 |
| 2 | 22 | -25 |
| 3 | <16 | 9 |
| $\bar{1}$,k,11 | | |
| 0 | 57 | 54 |
| 1 | <28 | -28 |
| 2 | 14 | -17 |
| 2k0 | | |
| 0 | 138 | 133 |
| 1 | 189 | -219 |
| 2 | 30 | -30 |
| 3 | 23 | 28 |
| 4 | 26 | 15 |
| 5 | 82 | -87 |
| 6 | <10 | 10 |
| 7 | 13 | -13 |
| 2k1 | | |
| 1 | 84 | -91 |
| 2 | 188 | 169 |
| 3 | 96 | 97 |
| 4 | 71 | 58 |
| 5 | 100 | -105 |
| 6 | 39 | 39 |
| 7 | 20 | 21 |
| $\bar{2}k1$ | | |
| 1 | 440 | 446 |
| 2 | 79 | 59 |
| 3 | 32 | 21 |
| 4 | 95 | -93 |
| 5 | <10 | 1 |
| 6 | 10 | -13 |
| 7 | 23 | -23 |
| 2k2 | | |
| 0 | 304 | -338 |
| 1 | 7 | -2 |
| 2 | 93 | 84 |
| 3 | 96 | 94 |
| 4 | 87 | 79 |
| 5 | 36 | 33 |
| 6 | 16 | -17 |
| 7 | <7 | 1 |
| $\bar{2}k2$ | | |
| 0 | 455 | -492 |
| 1 | 544 | 605 |
| 2 | 143 | -120 |
| 3 | 102 | -76 |
| 4 | 46 | -39 |
| 5 | 32 | 22 |
| 6 | 26 | 30 |
| 7 | 27 | 24 |
| 2k3 | | |
| 1 | 154 | -138 |
| 2 | 295 | -293 |
| 3 | 131 | 126 |
| 4 | 35 | -32 |
| 5 | 91 | -91 |
| 6 | 67 | -72 |
| 7 | 32 | 24 |
| $\bar{2}k3$ | | |
| 1 | 93 | -90 |
| 2 | 108 | 105 |
| 3 | 47 | 42 |
| 4 | 105 | -101 |
| 5 | 20 | 22 |
| 6 | 49 | 46 |
| 7 | 14 | -9 |
| 2k4 | | |
| 0 | 91 | -84 |
| 1 | 6 | 2 |
| 2 | 24 | 19 |
| 3 | 18 | 20 |
| 4 | 37 | -34 |
| 5 | 70 | 73 |
| 6 | 13 | -15 |
| $\bar{2}k4$ | | |
| 0 | 59 | -69 |
| 1 | 113 | 71 |
| 2 | 187 | 179 |
| 3 | 93 | -83 |
| 4 | 17 | -13 |
| 5 | 35 | -31 |
| 6 | 22 | -27 |
| 7 | 15 | -16 |
| 2k5 | | |
| 1 | 111 | 105 |
| 2 | 48 | -34 |
| 3 | 49 | 49 |
| 4 | 31 | 28 |
| 5 | <8 | 4 |
| 6 | 13 | -22 |
| $\bar{2}k5$ | | |
| 1 | 60 | -51 |
| 2 | <9 | -9 |
| 3 | 89 | -75 |
| 4 | 45 | -41 |
| 5 | 22 | 17 |
| 6 | 17 | 16 |
| 7 | 20 | -14 |
| 2k6 | | |
| 0 | 101 | -104 |
| 1 | 111 | 110 |
| 2 | 49 | 40 |
| 3 | 20 | -20 |
| 4 | 42 | -38 |
| $\bar{2}k6$ | | |
| 0 | 177 | 179 |
| 1 | 21 | 22 |
| 2 | 17* | 17 |
| 3 | 107 | 110 |
| 4 | 22 | 11 |
| 5 | 21 | 25 |
| 6 | 13 | 15 |
| 2k7 | | |
| 1 | 84 | -89 |
| 2 | 39 | 36 |
| 3 | <20 | 13 |
| 4 | 22 | -20 |
| $\bar{2}k7$ | | |
| 1 | 21* | -15 |
| 2 | 44 | 41 |
| 3 | 60 | -57 |
| 4 | 78 | -66 |
| 5 | 48 | 47 |
| 6 | 21 | 15 |
| 2k8 | | |
| 0 | 51 | -54 |
| 1 | <14 | 0 |
| 2 | 96 | 103 |
| 3 | 34 | -31 |
| 4 | 16 | -16 |
| $\bar{2}k8$ | | |
| 0 | 51 | 45 |
| 1 | 15* | 12 |
| 2 | 91 | -97 |
| 3 | 18 | -17 |
| 4 | 12 | 10 |
| 5 | 49 | -38 |
| 2k9 | | |
| 1 | <12 | -23 |
| 2 | 43 | -6 |
| $\bar{2}k9$ | | |
| 1 | 21 | 21 |
| 2 | 36* | 31 |
| 3 | 29 | 27 |
| 4 | 10 | -12 |
| $\bar{2}$,k,10 | | |
| 0 | 51 | 54 |
| 1 | <10 | -4 |
| 2 | 18 | -18 |
| 3 | 24 | -20 |
| $\bar{2}$,k,11 | | |
| 1 | <12 | -8 |
| 2 | 8 | 3 |
| 3k0 | | |
| 1 | 122 | -111 |
| 2 | 226 | 217 |
| 3 | 66 | 65 |
| 4 | 42 | -33 |
| 5 | 17 | -17 |
| 6 | 56 | 61 |
| 7 | 37 | 30 |
| 3k1 | | |
| 0 | 164 | -169 |
| 1 | 124 | -114 |
| 2 | 170 | 152 |
| 3 | 155 | -150 |
| 4 | 24 | 19 |
| 5 | 119 | -136 |
| 6 | 31 | -34 |
| 7 | <6 | 4 |
| $\bar{3}k1$ | | |
| 0 | 209 | -211 |
| 1 | 120 | 120 |
| 2 | 8 | -6 |
| 3 | 88 | -80 |
| 4 | 91 | -81 |
| 5 | 10 | 6 |
| 6 | 14 | 17 |
| 7 | 16 | -17 |
| 3k2 | | |
| 1 | 150 | -125 |
| 2 | 131 | 123 |
| 3 | 138 | 128 |
| 4 | 33 | 29 |
| 5 | 119 | -128 |
| 6 | 34 | 33 |
| 7 | 20 | 15 |
| $\bar{3}k2$ | | |
| 1 | 39 | 36 |
| 2 | 93 | 95 |
| 3 | 42 | 39 |
| 4 | 53 | -49 |
| 5 | 20 | 21 |
| 6 | 23 | 27 |
| 7 | 7 | -5 |
| 3k3 | | |
| 0 | <14 | -14 |
| 1 | 40 | 36 |
| 2 | 92 | 77 |
| 3 | 11 | 16 |
| 4 | 72 | 69 |
| 5 | 44 | 48 |
| 6 | 19 | -19 |
| $\bar{3}k3$ | | |
| 0 | 64 | -58 |
| 1 | 204 | 190 |
| 2 | 390 | 363 |
| 3 | 98 | -97 |
| 4 | 37 | -35 |
| 5 | 41 | -46 |
| 6 | 44 | -48 |
| 7 | <7 | -5 |
| 3k4 | | |
| 1 | 20 | -21 |
| 2 | 34 | -26 |
| 3 | 19 | 12 |
| 4 | <10 | 2 |
| 5 | 26 | -28 |
| 6 | <7 | -12 |
| $\bar{3}k4$ | | |
| 1 | 231 | -249 |
| 2 | 114 | 108 |
| 3 | 50 | 49 |
| 4 | <10 | 7 |
| 5 | 23 | -19 |
| 6 | 64 | 59 |
| 7 | 12 | 12 |
| 3k5 | | |
| 0 | 77 | -89 |
| 1 | 125 | 129 |
| 2 | 95 | -99 |
| 3 | 65 | -68 |
| 4 | 8 | -9 |
| $\bar{3}k5$ | | |
| 0 | 87 | -73 |
| 1 | 63 | -58 |
| 2 | 114 | 104 |
| 3 | 27 | 20 |
| 4 | <11 | 1 |
| 5 | 14 | 16 |
| 6 | <8 | -4 |
| 7 | 13 | 12 |
| 3k6 | | |
| 1 | 88 | 90 |
| 2 | 24 | 23 |
| 3 | 29 | 31 |
| 4 | 78 | -66 |
| $\bar{3}k6$ | | |
| 1 | 16 | -20 |
| 2 | 15* | -6 |
| 3 | 10 | -7 |
| 4 | 72 | -63 |
| 5 | 13 | -10 |
| 6 | <7 | 2 |
| 3k7 | | |
| 0 | 82 | -65 |
| 1 | 94 | 104 |
| 2 | 40 | 44 |
| 3 | 41 | -40 |
| 4 | 14 | -13 |
| $\bar{3}k7$ | | |
| 0 | 50 | 42 |
| 1 | 20 | 9 |
| 2 | 78 | -77 |
| 3 | 22* | -21 |
| 4 | 15* | -1 |
| 5 | 29 | -30 |
| 6 | 11 | -7 |
| 3k8 | | |
| 1 | 31 | -20 |
| 2 | 26 | 27 |
| 3 | <14 | 13 |
| $\bar{3}k8$ | | |
| 1 | 56 | 57 |
| 2 | 90 | 94 |
| 3 | 51 | 45 |
| 4 | 9 | -10 |
| 5 | <7 | -3 |
| 3k9 | | |
| 0 | 20 | -16 |
| $\bar{3}k9$ | | |
| 0 | 13 | -8 |
| 1 | 51 | 51 |
| 2 | 47 | -47 |
| 3 | 39 | -37 |
| 4 | 39 | 32 |
| 5 | 9 | -9 |
| $\bar{3}$,k,10 | | |
| 1 | 25 | -27 |
| 2 | <16 | 4 |
| 3 | 7 | 7 |
| $\bar{3}$,k,11 | | |
| 0 | 20 | 18 |
| 1 | 25* | 21 |
| 2 | <7 | 1 |
| 3 | 19 | -21 |
| 4k0 | | |
| 0 | 115 | -101 |
| 1 | 200 | 190 |
| 2 | 62 | 60 |
| 3 | 59 | 57 |
| 4 | 89 | -76 |
| 5 | 19 | -19 |
| 6 | 17 | -22 |
| 7 | 42 | 40 |
| 4k1 | | |
| 1 | 94 | -87 |
| 2 | 221 | 211 |
| 3 | 108 | 97 |
| 4 | 61 | -43 |
| 5 | 13 | -11 |
| 6 | 67 | 71 |
| 7 | 21 | 19 |
| $\bar{4}k1$ | | |
| 1 | 141 | 137 |
| 2 | 17 | 4 |
| 3 | <9 | 8 |
| 4 | 189 | 187 |
| 5 | 43 | 39 |
| 6 | <9 | -1 |
| 7 | <6 | -6 |
| 4k2 | | |
| 0 | <14 | -10 |
| 1 | 43 | -35 |
| 2 | 23 | 24 |
| 3 | 71 | -60 |
| 4 | 43 | 35 |
| 5 | 44 | -44 |
| 6 | 26 | -31 |
| $\bar{4}k2$ | | |
| 0 | 107 | -102 |
| 1 | 93 | -83 |
| 2 | 93 | 86 |
| 3 | 45 | -37 |
| 4 | 69 | -60 |
| 5 | 44 | -49 |
| 6 | 22 | 23 |
| 7 | 13 | -14 |
| 4k3 | | |
| 1 | 134 | -130 |
| 2 | <10 | 2 |
| 3 | 35 | -39 |
| 4 | 25 | -18 |
| 5 | <12 | -14 |
| 6 | <6 | 11 |
| $\bar{4}k3$ | | |
| 1 | 189 | -177 |
| 2 | 95 | 87 |
| 3 | 9 | -6 |
| 4 | 10 | -1 |
| 5 | 10 | -2 |
| 6 | 27 | 28 |
| 4k4 | | |
| 0 | 256 | 273 |
| 1 | 44 | 40 |
| 2 | 24 | -26 |
| 3 | 57 | -55 |
| 4 | 35 | 39 |
| $\bar{4}k4$ | | |
| 0 | 195 | -182 |
| 1 | 105 | -95 |
| 2 | 190 | 171 |
| 3 | <10 | -1 |
| 4 | 11* | -3 |
| 5 | 20 | -18 |
| 6 | 15 | -13 |
| 7 | 16 | 15 |
| 4k5 | | |
| 1 | 48 | 50 |
| 2 | 76 | 73 |
| 3 | 37* | -34 |
| 4 | 26 | -24 |
| $\bar{4}k5$ | | |
| 1 | 37 | 34 |
| 2 | 41 | -36 |
| 3 | 112 | 103 |
| 4 | 21 | -25 |
| 5 | 83 | -78 |
| 6 | 8 | -12 |
| 4k6 | | |
| 0 | 25 | -32 |
| 1 | 115 | 122 |
| 2 | 88 | -70 |
| 3 | 42 | -38 |
| 4 | 17 | 15 |
| $\bar{4}k6$ | | |
| 0 | 103 | -104 |
| 1 | 71 | -68 |
| 2 | 44 | -27 |
| 3 | 71 | 68 |
| 4 | <11 | 2 |
| 5 | 39 | 36 |
| 6 | <7 | 6 |
| 4k7 | | |
| 1 | 18 | 16 |
| 2 | 24 | 20 |
| 3 | <14 | 11 |
| $\bar{4}k7$ | | |
| 1 | 42 | 41 |
| 2 | 19 | 17 |
| 3 | 44 | 43 |
| 4 | 21 | 14 |
| 5 | 16 | -17 |
| 6 | 21 | 11 |
| 4k8 | | |
| 0 | 54 | -31 |
| 1 | <10 | -4 |
| 2 | <9 | -2 |
| 3 | 67 | -46 |
| $\bar{4}k8$ | | |
| 0 | 27 | -24 |
| 1 | 35 | 33 |
| 2 | 55 | -57 |
| 3 | 46 | -40 |
| 4 | 15 | 14 |
| 5 | 15 | -16 |
| $\bar{4}k9$ | | |
| 1 | 22 | -23 |
| 2 | <11 | 1 |
| 3 | 20 | 18 |
| 4 | 32 | -24 |
| 5 | 30 | -25 |
| $\bar{4}$,k,10 | | |
| 0 | 22 | 11 |
| 1 | 47 | 54 |
| 2 | 41 | 40 |
| 3 | 18* | -17 |
| 4 | 20 | 16 |
| $\bar{4}$,k,11 | | |
| 1 | <12 | -16 |
| 2 | 19 | -12 |
| 5k0 | | |
| 1 | 97 | 87 |
| 2 | 28 | 23 |
| 3 | 9 | -2 |
| 4 | 165 | -169 |
| 5 | 41 | 35 |
| 6 | 15 | 14 |
| 5k1 | | |
| 0 | 134 | -125 |
| 1 | 29 | 24 |
| 2 | 61 | -53 |
| 3 | 22 | 19 |
| 4 | 25 | -26 |
| 5 | 29 | -36 |
| 6 | 7 | -12 |
| $\bar{5}k1$ | | |
| 0 | 91 | -71 |
| 1 | 225 | -226 |
| 2 | 44 | -52 |
| 3 | 37 | -31 |
| 4 | 56 | -45 |
| 5 | 37 | -34 |
| 6 | 46 | 54 |
| 7 | 24 | -31 |
| 5k2 | | |
| 1 | 75 | 70 |
| 2 | 7 | 6 |
| 3 | 22 | 26 |
| 4 | <10 | -5 |
| 5 | <8 | 12 |
| 6 | 13 | 31 |
| $\bar{5}k2$ | | |
| 1 | 66 | 66 |
| 2 | 37 | -34 |
| 3 | 113 | -116 |
| 4 | 14 | 11 |
| 5 | 26 | 27 |
| 6 | 16 | -20 |
| 7 | 19 | -24 |
| 5k3 | | |
| 0 | 186 | 199 |
| 1 | 47 | -45 |
| 2 | 66 | -67 |
| 3 | 52 | 56 |
| 4 | 31 | 31 |
| $\bar{5}k3$ | | |
| 0 | 82 | 71 |
| 1 | 77 | -66 |
| 2 | 86 | 86 |
| 3 | 102 | 102 |
| 4 | 15 | 16 |
| 5 | 10 | -7 |
| 6 | 16 | 19 |
| 5k4 | | |
| 1 | 74 | -73 |
| 2 | <14 | 1 |
| 3 | 99 | -96 |
| 4 | 16 | -15 |
| $\bar{5}k4$ | | |
| 1 | <10 | 2 |
| 2 | 32 | -28 |
| 3 | 41 | 40 |
| 4 | 36 | -42 |
| 5 | 35 | -40 |
| 6 | 11 | -11 |
| 5k5 | | |
| 0 | 174 | 170 |
| 1 | 39 | 41 |
| 2 | 61 | -59 |
| 3 | <17 | -15 |
| 4 | 41 | 22 |
| $\bar{5}k5$ | | |
| 0 | 173 | -164 |
| 1 | 109 | -105 |
| 2 | 41 | -29 |
| 3 | 105 | 104 |
| 4 | <11 | 14 |
| 5 | 43 | 40 |
| 6 | <7 | 4 |
| 5k6 | | |
| 1 | 21 | 9 |
| 2 | <12 | -6 |
| 3 | <15 | -2 |
| 4 | 42 | -35 |
| $\bar{5}k6$ | | |
| 1 | 27 | 18 |
| 2 | 93 | -90 |
| 3 | 45 | 45 |
| 4 | 19 | 18 |
| 5 | 25 | -33 |
| 6 | 15 | -14 |
| 5k7 | | |
| 0 | 14 | 13 |
| 1 | 23 | 18 |
| 2 | <10 | -5 |
| 3 | <13 | -4 |
| $\bar{5}k7$ | | |
| 0 | 57 | 44 |
| 1 | 39 | -42 |
| 2 | 23 | -23 |
| 3 | 16* | -5 |
| 4 | 9 | -11 |
| 5 | 11 | -11 |
| 6 | 13 | -12 |
| 5k8 | | |
| 1 | 15 | 11 |
| $\bar{5}k8$ | | |
| 1 | 19 | -16 |
| 2 | <12 | 0 |
| 3 | <14 | 4 |
| 4 | 34 | 31 |
| 5 | <6 | -6 |
| $\bar{5}k9$ | | |
| 0 | 23 | 21 |
| 1 | 46 | 46 |
| 2 | 36 | 34 |
| 3 | <11 | -8 |
| 4 | 14 | 13 |
| $\bar{5}$,k,10 | | |
| 1 | 43 | -43 |
| 2 | 23* | -17 |
| 3 | 7 | 5 |
| $\bar{5}$,k,11 | | |
| 1 | <10 | 8 |
| 2 | 29 | 28 |
| 6k0 | | |
| 0 | 37 | -27 |
| 1 | 52 | 46 |
| 2 | 28 | 23 |
| 3 | 37 | 36 |
| 4 | 15 | -10 |
| 5 | 8 | 7 |
| 6 | 10 | 11 |
| 6k1 | | |
| 1 | 69 | 66 |
| 2 | <11 | 3 |
| 3 | <11 | 9 |
| 4 | 23* | 23 |
| 5 | <8 | 18 |
| 6 | <5 | 4 |
| $\bar{6}k1$ | | |
| 1 | 41 | 30 |
| 2 | 47 | -37 |
| 3 | 110 | -108 |
| 4 | 32 | 28 |
| 5 | 31 | 26 |
| 6 | <7 | 2 |
| 6k2 | | |
| 0 | 98 | -100 |
| 1 | 132 | -135 |
| 2 | 108 | -103 |
| 3 | 30 | 33 |
| 4 | <12 | 1 |
| $\bar{6}k2$ | | |
| 0 | 180 | 177 |
| 1 | 52 | -50 |
| 2 | <11 | 2 |
| 3 | 127 | 134 |
| 4 | 16 | 16 |
| 5 | 13 | 14 |
| 6 | 43 | 38 |
| 6k3 | | |
| 1 | 77 | 76 |
| 2 | 46 | -49 |
| 3 | 45 | -40 |
| 4 | 25 | 21 |
| $\bar{6}k3$ | | |
| 1 | 46 | -45 |
| 2 | 30 | -27 |
| 3 | 94 | -103 |
| 4 | 87 | -81 |
| 5 | 36 | 38 |
| 6 | <7 | -8 |
| 6k4 | | |
| 0 | 158 | 148 |
| 1 | 49 | -48 |
| 2 | 43 | -46 |
| 3 | 30 | 28 |
| 4 | 31 | 17 |
| $\bar{6}k4$ | | |
| 0 | 60 | 57 |
| 1 | 26 | -21 |
| 2 | 59 | -54 |
| 3 | 55 | 56 |
| 4 | 17 | 18 |
| 5 | 15 | -17 |
| 6 | 11 | -9 |
| 6k5 | | |
| 1 | 16 | -11 |
| 2 | <12 | 8 |
| 3 | 18 | -13 |
| 4 | <14 | -6 |
| $\bar{6}k5$ | | |
| 1 | 38 | 33 |
| 2 | 28 | -28 |
| 3 | 49 | 51 |
| 4 | <11 | 4 |
| 5 | 12 | -12 |
| 6k6 | | |
| 0 | <18 | 10 |
| 1 | <11 | 14 |
| 2 | 17 | -15 |
| 3 | 63 | 41 |
| $\bar{6}k6$ | | |
| 0 | 75 | 68 |
| 1 | 63 | -63 |
| 2 | 24 | -24 |
| 3 | 14 | 7 |
| 4 | <11 | -3 |
| 5 | <8 | -7 |
| 6 | 21 | -22 |
| 6k7 | | |
| 1 | 31 | -22 |
| $\bar{6}k7$ | | |
| 1 | 35 | -28 |
| 2 | 10 | -8 |
| 3 | 24 | -11 |
| 4 | 86 | 83 |
| $\bar{6}k8$ | | |
| 0 | 59 | 54 |
| 1 | <12 | -7 |
| 2 | <12 | -7 |
| 3 | 24 | -24 |
| 4 | <9 | -2 |
| 5 | <5 | -4 |
| $\bar{6}k9$ | | |
| 1 | 13 | -14 |
| 2 | 15* | -15 |
| 3 | 22 | -21 |
| $\bar{6}$,k,10 | | |
| 0 | 23 | -23 |
| 1 | 11 | 12 |
| 2 | 27 | 25 |
| 3 | <8 | -6 |
| $\bar{6}$,k,11 | | |
| 1 | 21 | 15 |
| 2 | <7 | -2 |
| 7k0 | | |
| 1 | 76 | -82 |
| 2 | 41 | -31 |
| 3 | 23 | -26 |
| 4 | <10 | -2 |
| 5 | 10 | 15 |
| 7k1 | | |
| 0 | 47 | -47 |
| 1 | 103 | -96 |
| 2 | 64 | 67 |
| 3 | 20 | 23 |
| 4 | 11 | -12 |
| $\bar{7}k1$ | | |
| 0 | 68 | 70 |
| 1 | 33 | -29 |
| 2 | 47 | 46 |
| 3 | 24 | 23 |
| 4 | <10 | -2 |
| 5 | 7 | -6 |
| 6 | <9 | 4 |
| 7k2 | | |
| 1 | 57 | 59 |
| 2 | 27 | -30 |
| 3 | <20 | 5 |
| 4 | 86 | 61 |
| $\bar{7}k2$ | | |
| 1 | 26 | -21 |
| 2 | 21 | -19 |
| 3 | 75 | -75 |
| 4 | 69 | -59 |
| 5 | 39 | 34 |
| 7k3 | | |
| 0 | 37 | -27 |
| 1 | 96 | -97 |
| 2 | 40 | -37 |
| 3 | 34 | 29 |
| 4 | <15 | -7 |
| $\bar{7}k3$ | | |
| 0 | 153 | 149 |
| 1 | <12 | 13 |
| 2 | 67 | -65 |
| 3 | 36 | 28 |
| 4 | 9 | 4 |
| 5 | 29 | -28 |
| 7k4 | | |
| 1 | <12 | 2 |
| 2 | 45 | 29 |
| 3 | <14 | -18 |
| 4 | 27 | 23 |
| $\bar{7}k4$ | | |
| 1 | 48 | 45 |
| 2 | 79 | 80 |
| 3 | 26 | 28 |
| 4 | <12 | -12 |
| 5 | 21 | 23 |
| 6 | 22 | 22 |
| 7k5 | | |
| 0 | <13 | -4 |
| 1 | 36 | -32 |
| 2 | <6 | 5 |
| 3 | 16 | -12 |
| $\bar{7}k5$ | | |
| 0 | 10 | 5 |
| 1 | 10 | 6 |
| 2 | 78 | -86 |
| 3 | 37 | -41 |
| 4 | 16 | 16 |
| 5 | 22 | -23 |
| 7k6 | | |
| 1 | 25 | -16 |
| $\bar{7}k6$ | | |
| 1 | 18 | -11 |
| 2 | <12 | 6 |
| 3 | <11 | 10 |
| 4 | 40 | 44 |
| $\bar{7}k7$ | | |
| 0 | 70 | 65 |
| 1 | 18 | -13 |
| 2 | 14 | -15 |
| 3 | 31 | -33 |
| 4 | <9 | 4 |
| 5 | <5 | -2 |
| $\bar{7}k8$ | | |
| 1 | 26* | 19 |
| 2 | 9 | -11 |
| 3 | 26 | -37 |
| 4 | 19 | 16 |
| $\bar{7}k9$ | | |
| 0 | <12 | 1 |
| 1 | 19 | 23 |
| 2 | <10 | -3 |
| 3 | 22 | 16 |
| 4 | <6 | 2 |
| $\bar{7}$,k,10 | | |
| 1 | 21 | 14 |
| 2 | 8 | -3 |
| $\bar{7}$,k,11 | | |
| 0 | 38 | -36 |
| 1 | 15 | -13 |
| 8k0 | | |
| 0 | 20 | -28 |
| 1 | <14 | 13 |
| 2 | 88 | 94 |
| 3 | 28 | 33 |
| 4 | <10 | -9 |
| 8k1 | | |
| 1 | 62 | -65 |
| 2 | 55 | -53 |
| 3 | <14 | 13 |
| 4 | 22 | 18 |
| $\bar{8}k1$ | | |
| 1 | 17 | -17 |
| 2 | 16 | 14 |
| 3 | 18 | 22 |
| 4 | 10 | -13 |
| 8k2 | | |
| 0 | 57 | -54 |
| 1 | 45 | -50 |
| 2 | 36 | 29 |
| 3 | 39 | 31 |
| 4 | 30 | -30 |
| $\bar{8}k2$ | | |
| 0 | 9 | 5 |
| 1 | 13 | -17 |
| 2 | 36 | -37 |
| 3 | 29 | 31 |
| 4 | <9 | -3 |
| 5 | 10 | 12 |
| 8k3 | | |
| 1 | 20 | 24 |
| 2 | 35 | -12 |
| 3 | <12 | -1 |
| $\bar{8}k3$ | | |
| 1 | 40 | 42 |
| 2 | 51 | 51 |
| 3 | 16 | 14 |
| 4 | 21 | -21 |
| 5 | 6 | 6 |
| 8k4 | | |
| 0 | <12 | -11 |
| 1 | 9 | -14 |
| 2 | 28 | 26 |
| $\bar{8}k4$ | | |
| 0 | 14 | -10 |
| 1 | 38 | 38 |
| 2 | 87 | -94 |
| 3 | 48 | -51 |
| 4 | 17 | 18 |
| 5 | 32 | -28 |
| $\bar{8}k5$ | | |
| 1 | 20 | 22 |
| 2 | 35 | 34 |
| 3 | 30 | 29 |
| 4 | 14 | -15 |
| 5 | 16 | -17 |
| $\bar{8}k6$ | | |
| 0 | 24 | 24 |
| 1 | 41 | 45 |
| 2 | <12 | 9 |
| 3 | 22 | -24 |
| 4 | 29 | 26 |
| 5 | 13 | 14 |
| $\bar{8}k7$ | | |
| 1 | 14 | -13 |
| 2 | 18* | -18 |
| 3 | 20 | -22 |
| 4 | <6 | 4 |
| $\bar{8}k8$ | | |
| 0 | 9 | 4 |
| 1 | 24 | 26 |
| 2 | <10 | -1 |
| 4 | <5 | 3 |
| $\bar{8}k9$ | | |
| 1 | 19* | 15 |
| 2 | 10 | -9 |
| 3 | 29 | -25 |
| $\bar{8}$,k,10 | | |
| 0 | 9 | -10 |
| 1 | 6 | 6 |
| 2 | <5 | 0 |
| 9k0 | | |
| 1 | 32 | -30 |
| 2 | 14 | -16 |
| 3 | 38 | 31 |
| 9k1 | | |
| 0 | 25 | -27 |
| 1 | <11 | 9 |
| 2 | 72 | 55 |
| 3 | <14 | -3 |
| $\bar{9}k1$ | | |
| 0 | 83 | -81 |
| 1 | 31 | -37 |
| 2 | <10 | 4 |
| 3 | 36 | 34 |
| 9k2 | | |
| 1 | <12 | 12 |
| $\bar{9}k2$ | | |
| 1 | 25 | 29 |
| 2 | 29 | -31 |
| 3 | 31 | 31 |
| 4 | 8 | -18 |
| 9k3 | | |
| 0 | 56 | -38 |
| 1 | 48 | 36 |
| $\bar{9}k3$ | | |
| 0 | 9 | 1 |
| 1 | <11 | 3 |
| 2 | 35 | -39 |
| 3 | <9 | -1 |
| 4 | <7 | 6 |
| $\bar{9}k4$ | | |
| 1 | 9 | 8 |
| 2 | 23 | 23 |
| 3 | 22* | 14 |
| $\bar{9}k5$ | | |
| 0 | <19 | 6 |
| 1 | 54 | 65 |
| 2 | 12 | 14 |
| 3 | 17 | -20 |
| 4 | 26 | 24 |
| $\bar{9}k6$ | | |
| 1 | 44 | -50 |
| 2 | 19 | -17 |
| 3 | 16* | 1 |
| $\bar{9}k7$ | | |
| 0 | <15 | -5 |
| 1 | 25 | 28 |
| 2 | 36 | 37 |
| 3 | 9 | -9 |
| $\bar{9}k8$ | | |
| 1 | <11 | 9 |
| 2 | <9 | -9 |
| $\bar{9}k9$ | | |
| 0 | 14 | -12 |
| 1 | <9 | 7 |
| 2 | <7 | 2 |
| 10,k,0 | | |
| 0 | 19 | -18 |
| 1 | 19 | 19 |
| 2 | 9 | 14 |
| 10,k,1 | | |
| 1 | 19 | -20 |
| $\overline{10}$,k,1 | | |
| 1 | <10 | 8 |
| 2 | 36 | -34 |
| 3 | 30 | 34 |
| $\overline{10}$,k,2 | | |
| 0 | <14 | -4 |
| 1 | 17 | -27 |
| 2 | 12* | -12 |
| 3 | 19 | 23 |
| $\overline{10}$,k,3 | | |
| 1 | 12 | -12 |
| 2 | <10 | -11 |
| 3 | 7 | 5 |
| $\overline{10}$,k,4 | | |
| 0 | 32 | 30 |
| 1 | 15* | 15 |
| 2 | <8 | 4 |
| 3 | 23 | -22 |
| $\overline{10}$,k,5 | | |
| 1 | 30 | -32 |
| 3 | <9 | -1 |
| $\overline{10}$,k,6 | | |
| 0 | <15 | -11 |
| 1 | 20 | 26 |
| 2 | 43 | 41 |
| 3 | 10 | -14 |
| $\overline{10}$,k,7 | | |
| 1 | <9 | -2 |
| 2 | <8 | -5 |
| $\overline{10}$,k,8 | | |
| 0 | 42 | -42 |
| 1 | <7 | -9 |
| 2 | <4 | 7 |
| $\overline{11}$,k,2 | | |
| 1 | 10 | -3 |
| $\overline{11}$,k,3 | | |
| 0 | 35 | 35 |
| 1 | 14 | -19 |
| $\overline{11}$,k,4 | | |
| 1 | <4 | -3 |
| $\overline{11}$,k,5 | | |
| 0 | 8 | -3 |
| 1 | 13 | 14 |

Table 3. Observed and calculated structure factors for N,N-dimethylurea. The three columns in each group contain the values, reading from left to right, of k, $10F_o$ and $10F_c$. Reflections indicated by an asterisk (*) were given zero weight in the final least-squares calculation.

Table 4. Orthogonalized coordinates.

| atom | X | Y | Z |
|---|---|---|---|
| O | 1.3772 | 1.2041 | 0.2789 |
| $N_1$ | -0.0006 | 1.0934 | -1.4981 |
| $N_2$ | 1.5528 | 2.7884 | -1.3212 |
| $C_1$ | 0.9821 | 1.6867 | -0.7863 |
| $C_2$ | 2.5891 | 3.5000 | -0.5865 |
| $C_3$ | 1.0938 | 3.4521 | -2.5239 |
| $H_1$ | -0.3880 | 0.2782 | -0.9731 |
| $H_2$ | -0.3738 | 1.3150 | -2.4392 |
| $H_3$ | 0.7265 | 2.7614 | -3.2801 |
| $H_4$ | 0.6110 | 4.3264 | -2.1918 |
| $H_5$ | 1.9929 | 4.0100 | -3.0305 |
| $H_6$ | 3.3984 | 3.6917 | -1.3234 |
| $H_7$ | 2.8679 | 2.9260 | 0.1975 |
| $H_8$ | 2.2087 | 4.5250 | -0.3674 |

Orthogonalization equations

$X = 8.7358\,x - 3.1138\,z$

$Y = 6.10\,y$

$Z = 8.7565\,z$

Inverse equations

$x = 0.1145\,X + 0.0407\,Z$

$y = 0.1639\,Y$

$z = 0.1142\,Z$

Table 5. Interatomic distances and angles in N,N-dimethylurea.

| Distances | | Angles | |
|---|---|---|---|
| $C_1$-O | 1.234 Å | O-$C_1$-$N_1$ | 121.1° |
| | | O-$C_1$-$N_2$ | 121.7 |
| $C_1$-$N_1$ | 1.351 | $N_1$-$C_1$-$N_2$ | 117.2 |
| $C_1$-$N_2$ | 1.351 | $C_1$-$N_2$-$C_2$ | 120.0 |
| $C_2$-$N_2$ | 1.456 | $C_1$-$N_2$-$C_3$ | 124.6 |
| $C_3$-$N_2$ | 1.448 | $C_2$-$N_2$-$C_3$ | 114.9 |
| | | $C_1$-$N_1$-$H_1$ | 110.4 |
| $N_1$-$H_1$ | 1.04 | $C_1$-$N_1$-$H_2$ | 130.3 |
| $N_1$-$H_2$ | 1.04 | $H_1$-$N_1$-$H_2$ | 119.3 |
| | | $N_2$-$C_3$-$H_3$ | 113.2 |
| $C_3$-$H_3$ | 1.09 | $N_2$-$C_3$-$H_4$ | 105.3 |
| $C_3$-$H_4$ | 1.05 | $N_2$-$C_3$-$H_5$ | 109.5 |
| $C_3$-$H_5$ | 1.17 | $N_2$-$C_2$-$H_6$ | 105.5 |
| $C_2$-$H_6$ | 1.11 | $N_2$-$C_2$-$H_7$ | 108.1 |
| $C_2$-$H_7$ | 1.01 | $N_2$-$C_2$-$H_8$ | 107.8 |
| $C_2$-$H_8$ | 1.12 | | |

Hydrogen bonding distances and angles:

| | | | | |
|---|---|---|---|---|
| O···H-N' | thru center | 2.943 | $C_1$-$N_1$-O' | 117.7 |
| O···HN'' | via glide | 3.063 | O'-$N_1$-O'' | 86.6 |
| | | | O''-$N_1$-$C_1$ | 132.8 |
| | | | $C_1$-O-$N_1$' | 120.9 |
| | | | $C_1$-O-$N_1$'''' | 141.9 |
| | | | $N_1$'-O-$N_1$'''' | 93.4 |

Packing distances:

| | | |
|---|---|---|
| $C_2$ to $C_3$ | via screw axis, down | 3.955 |
| $C_2$ to $C_3$ | via screw axis, up | 4.029 |
| $C_2$ to $C_3$ | via n-glide + <u>b</u> translation | 3.538 |

Table 6. Final anisotropic temperature factor parameters.

The temperature factors are in the form of

$$T_i = \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}k\ell).$$

| atom | $B_{11}$ | $B_{22}$ | $B_{33}$ | $B_{12}$ | $B_{13}$ | $B_{23}$ |
|---|---|---|---|---|---|---|
| O | 0.0121 | 0.0278 | 0.0111 | -0.0037 | -0.0017 | 0.0089 |
| $N_1$ | 0.0109 | 0.0238 | 0.0116 | -0.0092 | -0.0026 | 0.0058 |
| $N_2$ | 0.0106 | 0.0219 | 0.0113 | -0.0048 | 0.0023 | 0.0060 |
| $C_1$ | 0.0095 | 0.0212 | 0.0083 | 0.0016 | 0.0020 | 0.0003 |
| $C_2$ | 0.0136 | 0.0288 | 0.0160 | -0.0182 | 0.0034 | 0.0005 |
| $C_3$ | 0.0174 | 0.0273 | 0.0119 | 0.0004 | -0.0009 | 0.0131 |

Table 7. Magnitudes and direction cosines of the principal axes of the vibration ellipsoids. The direction cosines are given relative to a Cartesian coordinate system obtained by the following orthogonalization relations:

$$X = h\,a^* \sin\beta^*$$
$$Y = k\,b^*$$
$$Z = h\,a^* \cos\beta^* + \ell\,c^* .$$

The X and Y axes are coincident to a and b, respectively, and Z is coincident with c*.

| Atom | Axis i | $B_i$ | $g_i^{(1)}$ | $g_i^{(2)}$ | $g_i^{(3)}$ |
|---|---|---|---|---|---|
|  | 1 | 6.079 | -0.661 | 0.516 | 0.544 |
| O | 2 | 3.504 | 0.545 | 0.829 | -0.124 |
|  | 3 | 2.244 | 0.515 | -0.215 | 0.830 |
|  | 1 | 6.136 | -0.697 | 0.455 | 0.554 |
| $N_1$ | 2 | 2.971 | 0.128 | 0.839 | -0.529 |
|  | 3 | 2.022 | 0.705 | 0.298 | 0.643 |
|  | 1 | 4.914 | -0.614 | 0.507 | 0.605 |
| $N_2$ | 2 | 2.708 | -0.028 | -0.780 | 0.625 |
|  | 3 | 2.530 | 0.789 | 0.367 | 0.493 |
|  | 1 | 3.458 | -0.786 | -0.356 | 0.505 |
| $C_1$ | 2 | 3.127 | -0.234 | 0.928 | 0.290 |
|  | 3 | 2.130 | 0.572 | -0.110 | 0.813 |
|  | 1 | 6.797 | -0.690 | 0.549 | 0.471 |
| $C_2$ | 2 | 4.684 | 0.134 | -0.542 | 0.829 |
|  | 3 | 2.121 | 0.711 | 0.636 | 0.301 |
|  | 1 | 6.962 | 0.792 | -0.362 | -0.493 |
| $C_3$ | 2 | 4.365 | 0.560 | 0.752 | 0.347 |
|  | 3 | 2.234 | 0.245 | -0.551 | 0.798 |

The direction cosines are given relative to a Cartesian coordinate system in which the X and Y axes are coincident to a and b, respectively, and the Z axis is coincident with c*; this axial system coincides with that used for the orthogonalization of the positional parameters. It differs from the one used by RD by a rotation of $\beta$-90 around the b axis.

To obtain the constants of the ellipsoids from the parameters given in table 5, the following transformations are made:

$$X = h\, a^* \sin \beta^*$$

$$Y = k\, b^*$$

$$Z = h\, a^* \cos \beta^* + \ell\, c^* .$$

Writing $-\ln T = (B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}k\ell) = (B'_{11}X^2 + B'_{22}Y^2 + B'_{33}Z^2 + B'_{12}XY + B'_{13}XZ + B'_{23}YZ)$ ,

we have

$$B'_{11} = B_{11}a^2 + B_{13}ac \cos \beta + B_{33}c^2 \cos^2 \beta$$

$$B'_{22} = B_{22}b^2$$

$$B'_{33} = B_{33}c^2 \sin^2 \beta$$

$$B'_{12} = B_{12}ab + B_{23}bc \cos \beta$$

$$B'_{13} = 2B_{33}c^2 \cos \beta \sin \beta + B_{13}ac \sin \beta$$

$$B'_{23} = B_{23}bc \sin \beta$$

The $B'_{ij}$ terms are then used to set up an eigenvalue equation as explained by R&D (see also section vii in part II of this thesis).

The libration of the molecule would be expected to be predominantly an out-of-plane motion with $C_1$ as a pivotal position. This implies that the angle between the major axis of each thermal ellipsoid and the normal to the plane should be small except perhaps for atom $C_1$. These angles are indeed small (see table 8).

The libration of the methyl group would be expected to be a minimum along the N to methyl bonds, and the angle between the minor axes of the thermal ellipsoids and the bond direction should therefore be small; this has been found to be the case.

Table 8. Orientation of thermal librations.

| Atom | Acute angle between the major axes of the thermal ellipsoids and the plane normal |
|---|---|
| O | 3.1° |
| $N_1$ | 6.0 |
| $N_2$ | 7.3 |
| $C_1$ | 53.8 (36.3 for semi-major axis) |
| $C_2$ | 1.9 |
| $C_3$ | 12.4 |

| Direction cosines of the $N_2$ to $C_2$ methyl bond direction | | Acute angle between the bond direction and the minor ellipsoid axis of the $C_2$ methyl group |
|---|---|---|
| g(1) | -0.7117 | |
| g(2) | -0.4887 | 14.4° |
| g(3) | -0.5046 | |

| Direction cosines of the $N_2$ to $C_3$ methyl bond direction | | Acute angle between the bond direction and the minor ellipsoid axis of the $C_3$ methyl group |
|---|---|---|
| g(1) | 0.3169 | |
| g(2) | -0.4583 | 7.0° |
| g(3) | 0.8304 | |

## Experimental Work: N,N'-dimethylurea.

### i) Unit cell and space group.

Crystals of N,N'-dimethylurea are hydroscopic and extremely soft. Initial attempts to obtain good single crystals by evaporation of several solvents were unsuccessful partly due to the hydroscopic nature of the material. Eventually very good crystals were obtained by vacuum sublimation at 111°C. Two different types of crystal habits were observed. First, rectangular prisms condensed on the sides of the collection tube. Attempts to remove these crystals generally failed because the slightest pressure would smear the compound over the glass, like smearing soft butter on toast. Second, several crystals were found hanging from the top of the collection chamber like small icicles. The smallest of these, having a maximum diameter of 0.5 mm and a length of about 2 cm, was mounted in a Lindemann glass capillary by carefully dropping it vertically into the glass cylinder. The end of the capillary was sealed off and the crystal was mounted for x-ray examination. This crystal was used for all the intensity photographs taken during the course of this investigation.

Several methods were used to obtain the unit cell dimensions. The results of each method are given in table 9. The rules for non-extinction of reflections were observed to be: $(hk\ell)$, $h + k = 2n$, $k + \ell = 2n$, $(h + \ell = 2n)$; $(h0\ell)$, $h + \ell = 4n$; $(0k\ell)$, $k + \ell = 4n$. These conditions establish the space group as Fdd2. The calculated density,

Table 9. Unit cell dimensions of N,N'-dimethylurea.

| Method | Radiation | a | b | c |
|---|---|---|---|---|
| Weissenberg 0-layer | Cu $K_\alpha$ | 20.52 Å | 11.55 Å | --- |
| Rotation | Cu $K_\alpha$ | --- | --- | 4.596 Å |
| Precession (h0ℓ) | Mo | 19.99 | --- | 4.521 (4.565)* |
| Precession (0kℓ) | Mo | --- | 11.29 | 4.518 (4.561)* |
| Straumanis rotation | Cr | 20.185 | 11.401 | --- |
| Adopted: | | 20.185 | 11.401 | 4.563 |
| Estimate of error | | ±0.005 | ±0.005 | ±0.01 |

*Re-computed using the Straumanis values for a and b to calibrate the camera constant against systematic errors such as film shrinkage.

assuming 8 molecules per unit cell, is 1.114 g cm$^{-3}$. Groth (5) reported a density of 1.142 g cm$^{-3}$.

These findings are not in agreement with the data reported by H. Mark (14). He found the space group to be $Pmn2_1$ with $\underline{a}$ = 4.53 Å, $\underline{b}$ = 10.9 Å and $\underline{c}$ = 5.14 Å; Z = 2. Groth reported the axial ratios 0.4394:1:1.0600 and a melting point of 102°. These axial ratios agree with the values of H. Mark when one computes the ratios as 2a/c:c/c:2b/c = 0.440:1:1.0603. Using the axial lengths found in this research and computing the ratios in the manner 2c/a:a/a:2b/a one obtains the values 0.452:1:1.130. The melting point determination of the sublimed crystals of N,N'-dimethylurea used in this research observed under a hot stage microscope were found to be 107-108°C. This figure agrees with one of the two values (106 and 102°C) reported in the Rubber Handbook of Chemistry and Physics.

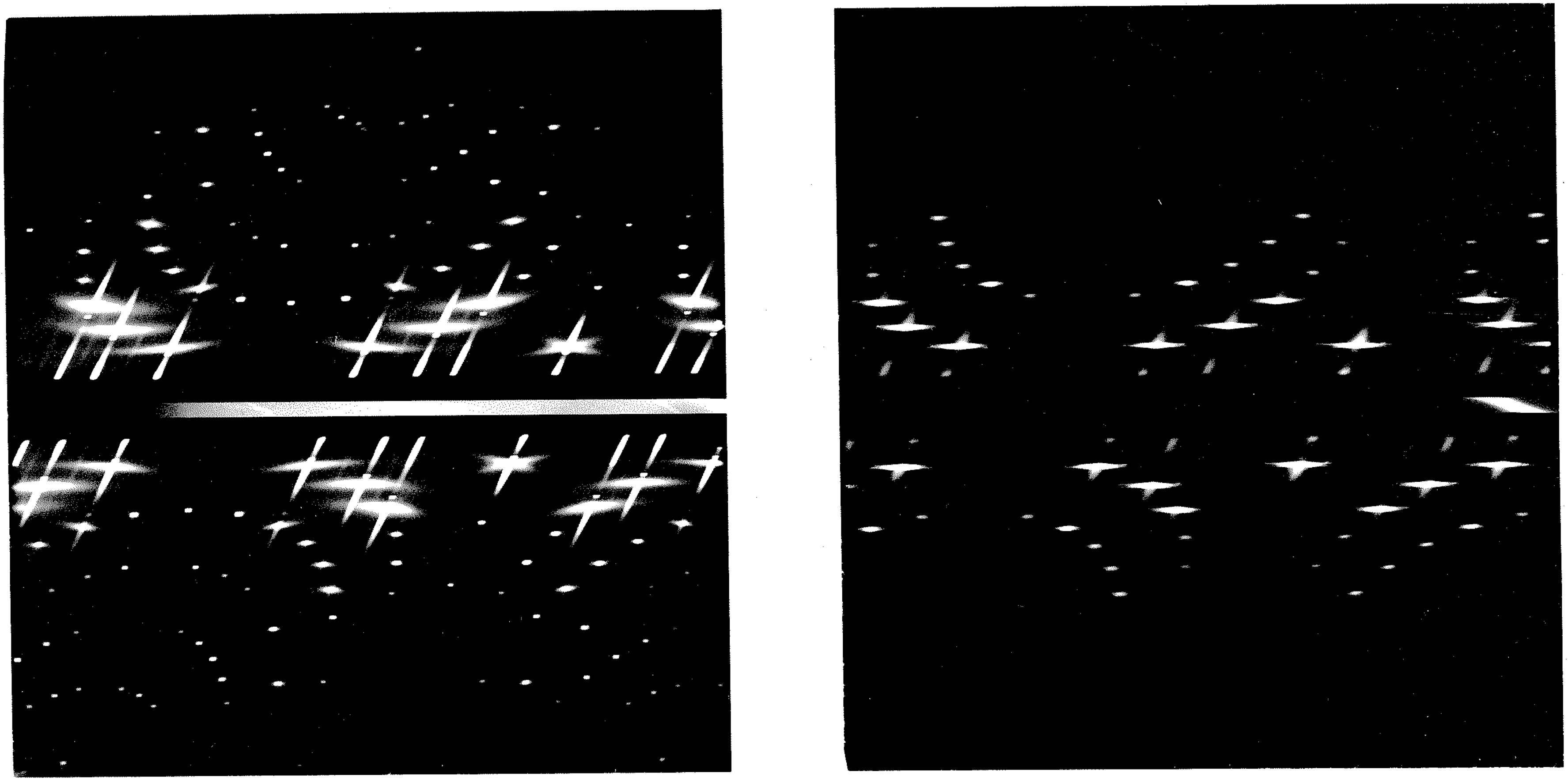
In view of the differences observed in the density, the melting point, the space group and the axial ratios, it seems likely that the crystals used in this research are different from those of H. Mark and Groth. H. Mark does not describe how he prepared his crystals. Groth, on the other hand, reports that crystals of this compound can be obtained by either evaporation of an ether solution or from a closed vessel containing a chloroform solution carefully covered by a layer of ether. It may be that crystals obtained from these solutions are crystallographically different from crystals obtained by sublimation.

ii) Intensity data.

Three-dimensional intensity data for Cu $K_{\alpha}$ radiation were collected on multiple-film equi-inclination Weissenberg photographs taken about the c-axis. All layer lines having inclination angles less than 45° were recorded ($\ell$ from 0 to 4 inclusive). The intensity data were estimated by visual comparison with a standard scale prepared from the same crystal. The intensities were corrected for Lorentz and polarization factors using a program written for the Burroughs 220 computer (see appendix II for a description of the program). No corrections for absorption or extinction were made.

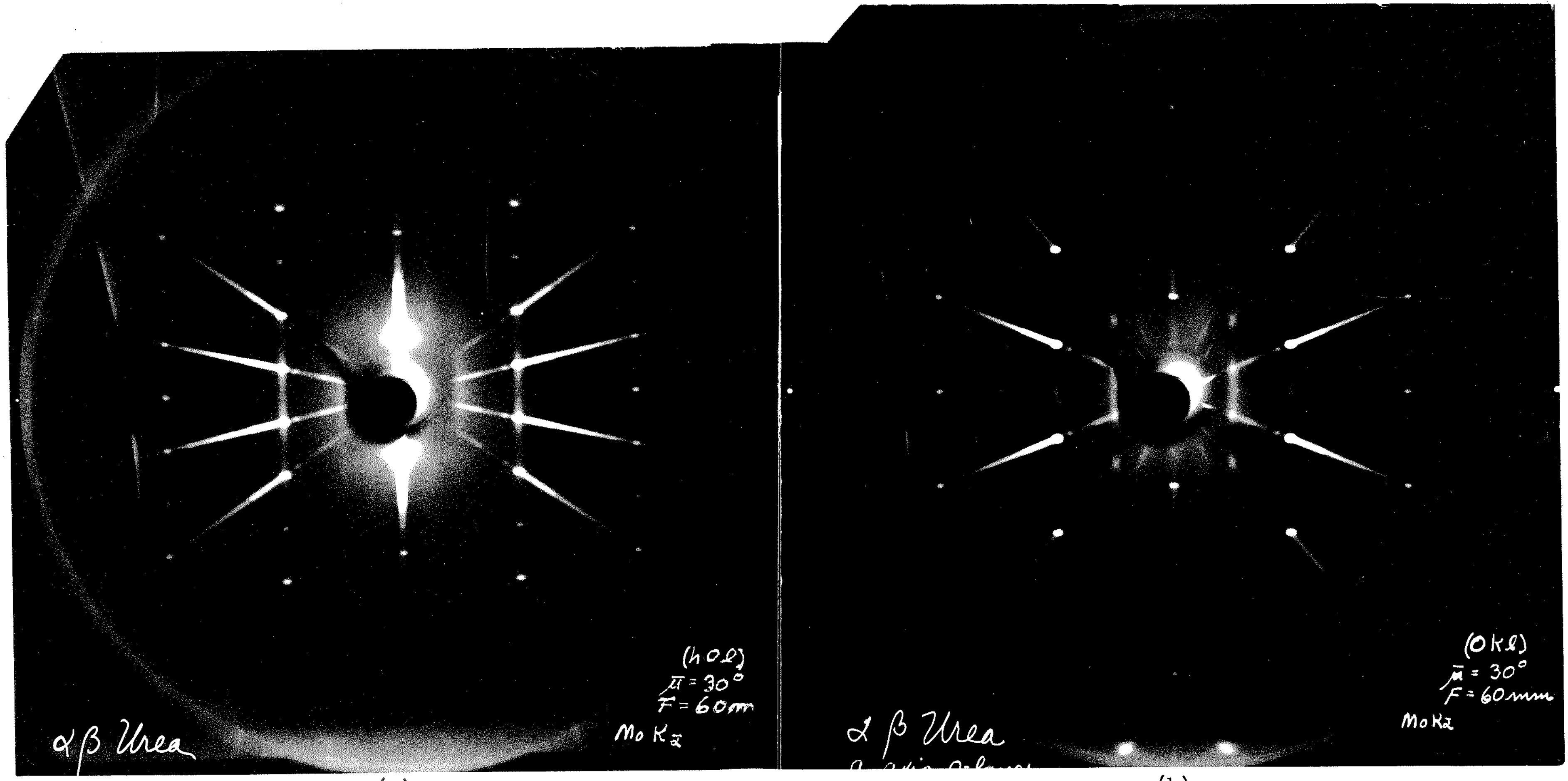
The intensity photographs revealed an interesting disorder effect in the crystal. Figure 4 shows zero-layer c-axis Weissenberg photographs of the same crystal taken at two different times. The one on the left was taken shortly after mounting in January 1958; the one on the right was taken two years later in January 1960. Figure 5 shows precession photographs taken in January 1958 of the (0k$\ell$) and the (h0$\ell$) zones of the same crystal. The precession photographs show diffuse streaking along the h and k row lines but not elsewhere; this indicates that there is no significant disorder along the c-axis and that the cause of the streaking is restricted to the ab-plane.

The beautiful diffuse arcing (asterism), as shown by some of the reflections in figure 4a is caused by external forces acting upon the peripheral sections of the crystal so as to bend the corresponding reticular planes. The normals to the respective planes span continuously



(a) (b)

Figure 4. - Zero-layer Weissenberg photographs of a crystal of N,N'-dimethylurea prepared (a) in January 1958 and (b) in January 1960 using the same crystal. The rotation axis is c.



(a) (b)

Figure 5. - Precession photographs taken in January 1958. (a) The (0kℓ) zone. (b) The (h0ℓ) zone. The crystal is the same as that used to prepare the photographs in figure 4.

over a range of up to about ± 10° as shown by the horizontal components of the asterism. These forces undoubtedly arise from the fact that the soft crystal was dropped into the glass capillary; and the effects probably extend only a short way into the crystal, since the high order reflections are present and exhibit only slight diffuseness.

By superpositioning a reciprocal lattice row line chart over the Weissenberg photograph in figure 4a, one finds that the asterism of all effected reflections lie along constant row lines of a pair of orthogonal axes, a' and b' (hereafter referred to as the deformation axes), related to a and b of the space group by a rotation of approximately 45° about the c-axis. (The a' and b' axes do not lie along principal directions in the unit cell.) The streaking of reflections occurring close to one of the deformation axes in reciprocal space runs perpendicular to that axis in reciprocal space. The (400) reflection which is located equidistant from both of the deformation axes shows asterism in both of the two orthogonal directions. The crystal deformation is such that the deformation axes appear to remain perpendicular to the curvature of the bend.

Figure 29 shows a model of the structure of N,N'-dimethylurea viewed along the c-axis. As can be seen, the elongated direction of the two molecules related by the diamond glide plane are nearly coincident with the deformation axes (deviating by only 8°). It is reasonable to believe that little or no deformation of the crystal will occur from compression along these directions; instead, one may expect a deformation in which these directions are made to radially converge or diverge by small amounts, resulting in the symmetric bending of the crystal with respect to the deformation axes.

The nature of the change in the asterism of the reflections appearing on the two Weissenberg photographs in figure 4 show that after two years the curvature of the reticular planes is replaced by multiple discontinuous changes of orientation brought about by polygonization of the deformed crystal. The horizontal asterism of the reflections in figure 4b occurs because the striae of the blocks resulting from the polygonization are so numerous and so close together that they cannot be separated. The c-axis orientations of the polygonized subcrystals are all parallel while the other two axes in the subcrystal lattices have slightly disordered orientations relative to each other. The angular disorientation of the reflections shown in figure 4b can be represented by an oscillation of up to 10° of the undistorted reciprocal lattice about its origin.

Careful examination of the precession photographs indicates that associated with the disorder, the density of the strained regions may also have varied from a higher density at the point of contact with the glass capillary to the equilibrium density at the interior of the crystal. The evidence for this is that the disorder maxima on the precession photographs do not occur exactly at the point of intersection of the lattice rows, particularly in the higher orders.

Several attempts were made to determine the relative scale factors for the different layer lines. The nature of the crystal made it impossible to cut or cleave it perpendicular to the c-axis, thus preventing any attempts to get Weissenberg photographs around the other axes. Precession photographs taken of the (h0ℓ) and (0kℓ) zones could not be used to give reliable scale factors because there were too few spots on the films and the range of intensity was too large to permit correlation of the recorded spots. Accordingly, inter-layer scaling was made on the basis of subsequent structure-factor calculation.

The atomic scattering curves used in this investigation for carbon, nitrogen and oxygen were taken to be the average of the values given by Berghuis _et al._ (8) and Hoerni and Ibers (9). The hydrogen scattering curve was that of McWeeny (7).

## Determination of the Structure of N,N'-dimethylurea.

### i) Derivation of the trial model.

The space group Fdd2 requires 16 asymmetric units in general positions and the only special positions are on the two-fold axes of the space group. Since only 8 molecules of N,N'-dimethylurea are present in each unit cell, each molecule must have a two-fold symmetry axis that is utilized by the space group. This restriction immediately indicated that the C=O bond is coincident with the two-fold axis parallel to c, and hence that the two atoms are at the origin of the (001) projection. In addition, the c-axis repeat distance of 4.563 Å by analogy to urea (c = 4.712 Å) indiated that the N,N'-dimethylurea molecules also form hydrogen bonded chains along the c-axis. The methyl groups are, therefore, in the cis positions with respect to the oxygen atom. The a and b cell dimensions indicated that the molecule is probably elongated in the [120] direction.

The presence of the $C_1$ and O atoms at the origin of the (001) projection is like having a heavy atom there; that is, they contribute heavily and with positive sign to the structure factors for which $h + k = 4n$ and nothing to those for which $h + k = 4n + 2$. Thus three of the strongest hk0 reflections in the former group were given positive signs; the reflections are (400), (220) and (620). These three reflections and the (420), the strongest hk0 reflection in the second group, were used to compute a rough electron density Fourier. (The sign of the (420) was arbitrarily taken as negative for this calculation since the alternative choice of sign represents only an interchange of the y and -y directions.) The result, a large peak elongated in the [120] direction, substantiated the proposed trial model. Using known interatomic distances and bond angles, the most probable projected distances for the $C_1$-N and N-$C_2$ bonds were marked off on the line drawn in the direction of the Fourier peak elongation. Initial values for the x and y cordinates for atoms N and $C_2$ were taken from this drawing.

These x and y cordinates were refined by the method of least-squares on the Burroughs 205 computer using all (hk0) data. The R factor converged from 0.38 to 0.16 in about ten refinement cycles. Preliminary z coordinates were now computed on the basis of predicted interatomic distances. (Since the c-axis is a polar axis, the initial z coordinate of $C_1$ was arbitrarily chosen as zero.) The three-dimensional refinement was carried out on the Burroughs 220 computer using the orthorhombic program described in appendix I.

ii) Refinement of parameters.

The first three-dimensional structure factor calculation was used to adjust the layer scale factors. Initial least-squares refinement failed to behave, in that the z coordinates failed to converge. It was found empirically that a factor of 0.4 had to be applied to the indicated shifts of the z parameters least-squares results before they could be used, and even with this shift factor the convergence was sporadic. During these early refinements, the individual layer scale factors were allowed to change after each cycle. It soon became apparent that the temperature factors, the <u>z</u> parameters and the individual layer scale factors were all overcompensating for each other. It was decided to keep the individual layer scale factors constant relative to each other over several refinement cycles, and from this point on the refinement was smoother.

The least-squares refinement was started giving constant weights to all reflections. This was changed to $\sqrt{w} = \sqrt{w'}/(0.44 + F_o)$ when the anisotropic temperature factors were introduced into the least-squares. The quantity being minimized was $\sum_q w_q (F_o^2 - F_c^2)_q^2$ and this last weighting scheme presumably reflects the pattern of reliability of the observed F's.

When the R factor was reduced to 0.11, the electron density section in the least-squares plane of N,N'-dimethylurea was computed and is shown in figure 6. Figure 7 shows the difference Fourier in the same plane and figure 8 shows the difference Fourier calculated in the plane of the hydrogen atoms of the methyl group. The constants defining the plane for the second difference Fourier were computed in the same way as for the corresponding hydrogen plane in N,N-dimethylurea. The view is through the plane with the plane of the heavy-atoms intersecting vertically and the oxygen atom (not shown) at the top of the axis.

It is evident that the methyl group is not preferentially oriented. Either a free rotation about the N-Me bond or a statistical distribution

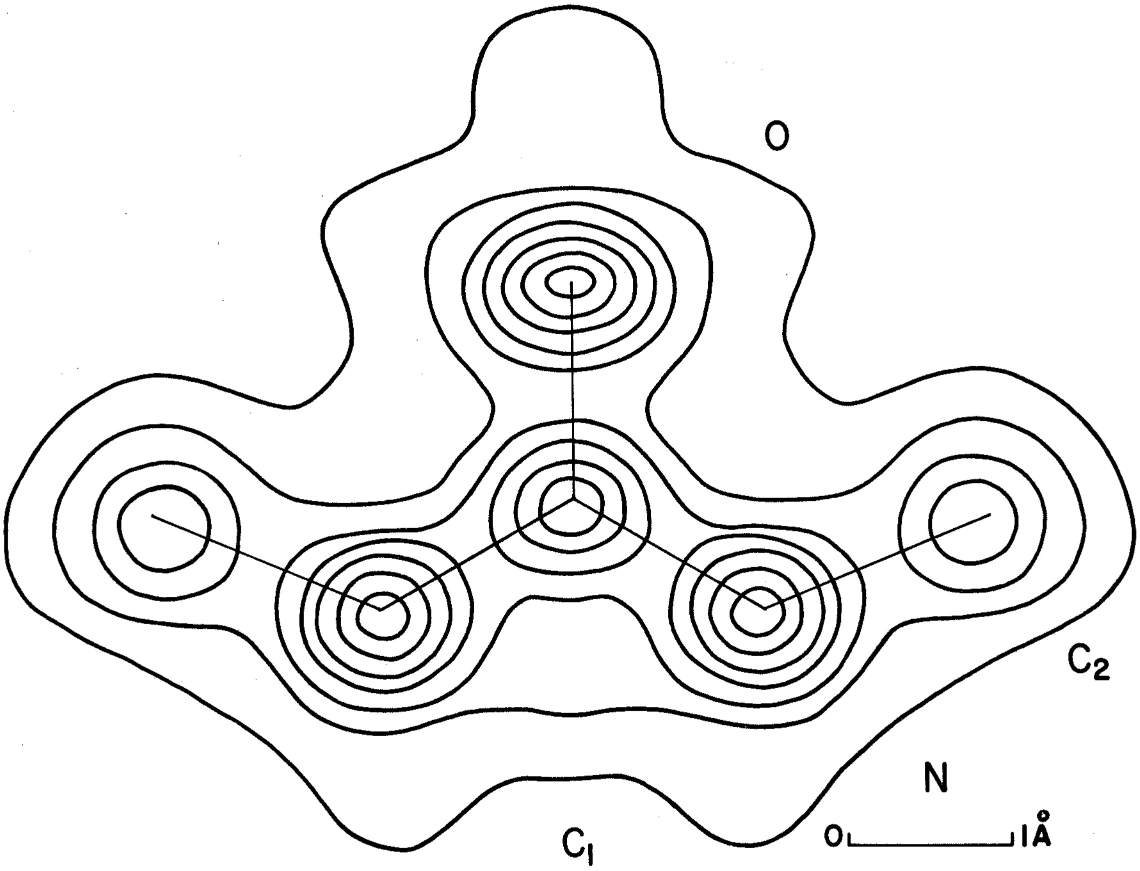


Figure 6. - The electron density section in the least-squares plane of N,N'-dimethylurea. Contours at intervals of 1 e. Å$^{-3}$, beginning with 1 e. Å$^{-3}$.

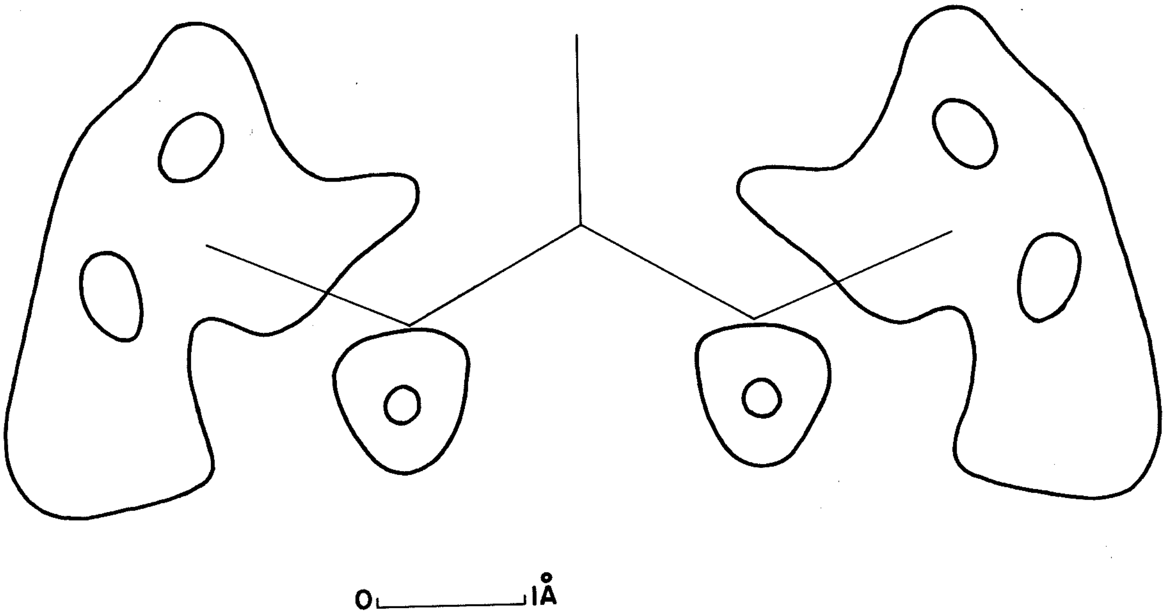


Figure 7. - The difference Fourier in the best plane of N,N'-dimethylurea. Contours at intervals of 0.2 e. Å$^{-3}$, beginning with 0.2 e. Å$^{-3}$.

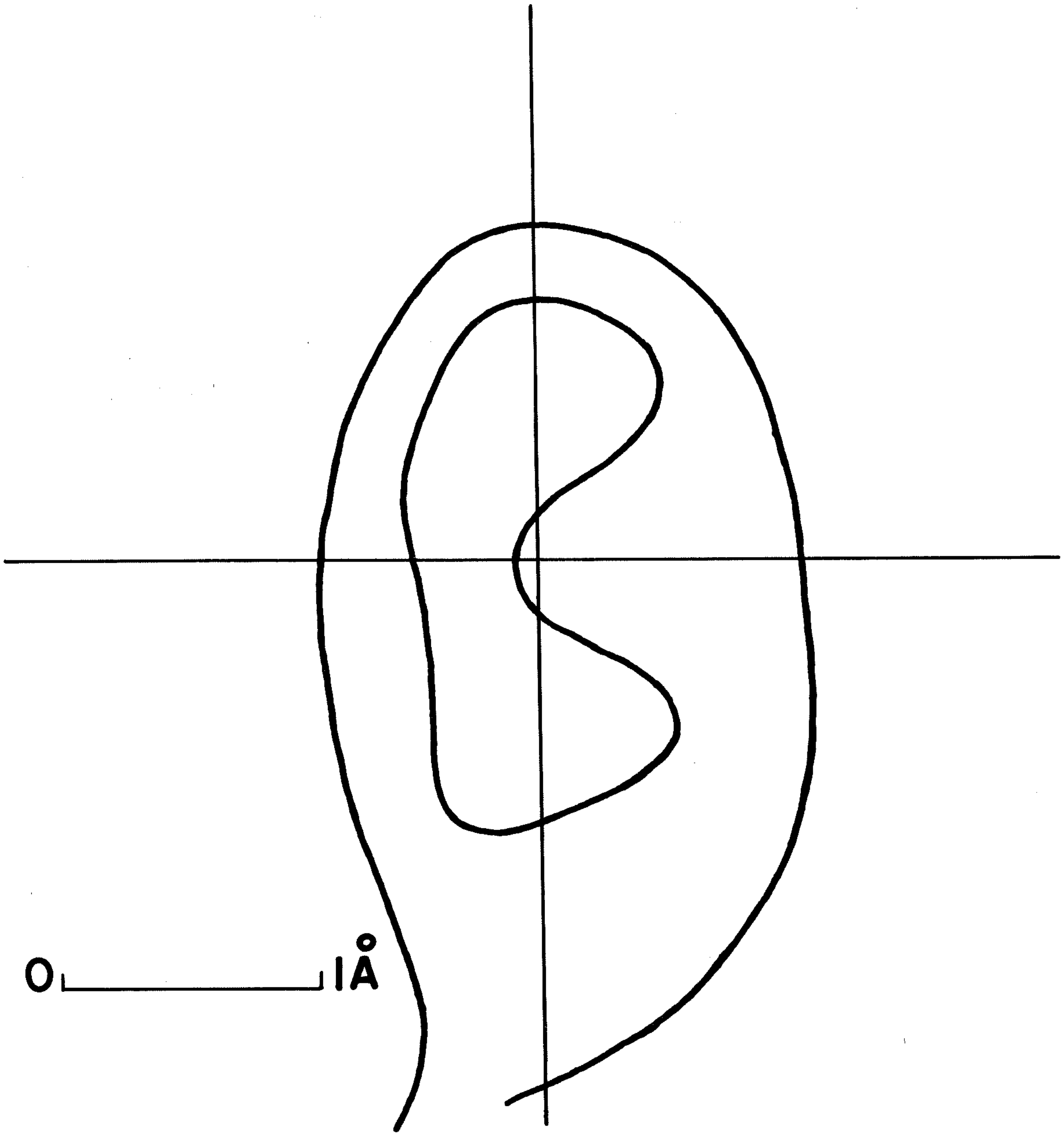


Figure 8. - The difference Fourier map in the plane of the hydrogen atoms of the methyl group. Contours at 0.2 and 0.3 e. Å$^{-3}$.

among two or more orientations can give rise to the observed electron density distribution. On the basis of this observation, it was decided to approximate the torus of electron density by twelve quarter hydrogen atoms equally spaced on the circumference of a circle of radius 0.94 Å. The plane of the circle was placed at 0.335 Å from the carbon atom and its origin at the point of intersection with the extension of the N-Me bond.

Four least-squares refinement cycles were now made with the hydrogen atom contributions included in the structure factors. The hydrogen positions were not refined. At this time the reflections (220), (400) and (420) were given zero weight because the previous structure factor calculations indicated that they may be affected by extinction (or perhaps by the disorder described earlier). Since the quantity being minimized was $\sum_q w_q(F_o^2 - F_c^2)_q^2$, it was felt that an appropriate weighting scheme must reflect the pattern of reliability of the observed $F^2$'s and not the F's as in the previous scheme. Thus the weighting function was changed to $\sqrt{w} = \sqrt{w'}/(0.5 + 0.4586 F_o + 0.04136 F_o^2)$. This scheme corresponds to taking the percentage probable error in the limit of large intensities as constant and the absolute probable error constant in the limit of small intensities with the region between connected by a continuous function. The coefficients in this expression are empirical. The fourth refinement cycle showed negligible shifts and had an R factor of 0.085.

At this point the refinement was deemed complete. The final heavy atom positional parameters and their standard errors are given in table 10. The validity of the standard errors with respect to the polar axis (z parameters) is questionable, as has been pointed out by Templeton (15). The author believes that the C-N bond distances are probably accurate to within 0.01 Å. The accuracy of the C=O bond is more difficult to establish. The intensity data were collected in layers perpendicular to the c-axis, and hence perpendicular to the C=O bond. The accuracy of the C=O bond depends on the resolution of the data along the c-axis and with only 5 layers of data ($\ell$ = 0 to 4) one can expect some series termination errors. The "hat" that the oxygen atom is wearing in the electron density map shown in figure 6 may be caused by series termination, since it does not appear in the corresponding difference Fourier in figure 7. Because of these facts, the author believes that the C=O bond distance is probably accurate to within 0.04 Å.

Table 12 gives the final structure factor list. The coordinates in Angstrom units are listed in table 13, and the bond distances and angles are given in table 14. The least squares plane of the whole molecule (two asymmetric units) is $-0.7990X + 0.6013Y = 0$. Further details of the structure are reserved for a general discussion of all three compounds in this series.

Table 10. Final position parameters and their standard error.

| Atom | x | y | z | σ(x) | σ(y) | σ(z)? |
|---|---|---|---|---|---|---|
| O | 0 | 0 | 0.2782 | 0 | 0 | 0.0016 |
| N | 0.0322 | 0.0815 | -0.1455 | 0.0002 | 0.0003 | 0.0013 |
| $C_1$ | 0 | 0 | 0.0016 | 0 | 0 | 0.0019 |
| $C_2$ | 0.0737 | 0.1708 | -0.0134 | 0.0004 | 0.0005 | 0.0023 |

Table 11. Coordinates* assigned to the hydrogen atoms.

| Atom | x | y | z |
|---|---|---|---|
| $H_1$ | 0.033 | 0.082 | -0.356 |
| $H_2$ | 0.078 | 0.237 | -0.161 |
| $H_3$ | 0.060 | 0.249 | -0.097 |
| $H_4$ | 0.048 | 0.245 | -0.004 |
| $H_5$ | 0.045 | 0.228 | 0.092 |
| $H_6$ | 0.053 | 0.200 | 0.169 |
| $H_7$ | 0.068 | 0.171 | 0.202 |
| $H_8$ | 0.088 | 0.146 | 0.184 |
| $H_9$ | 0.106 | 0.135 | 0.120 |
| $H_{10}$ | 0.118 | 0.138 | 0.026 |
| $H_{11}$ | 0.121 | 0.156 | -0.071 |
| $H_{12}$ | 0.113 | 0.184 | -0.145 |
| $H_{13}$ | 0.097 | 0.213 | -0.178 |

*The methyl group has been shown to be disordered and the torus of electron density is represented by the 12 quarter hydrogen atoms $H_2$ to $H_{13}$ inclusive.

Table 12. - Observed and calculated structure factors for N,N'-dimethylurea. The five columns in each group contain the values, reading from left to right, of k, $10|F_o|$, $10|F_c|$, $10A_c$ and $10B_c$. Reflections indicated by an asterisk (*) were given zero weight in the final least-squares calculation.

| | | | | |
|---|---|---|---|---|
| **0k0** | | | | |
| 0 | ---* | 3845 | 3845 | 0 |
| 4 | 90 | 88 | 88 | 0 |
| 8 | 37 | 31 | -31 | 0 |
| 12 | 80 | 68 | 68 | 0 |
| **2k0** | | | | |
| 2 | 882* | 979 | 979 | 0 |
| 4 | 180 | 202 | 202 | 0 |
| 6 | 82 | 80 | 80 | 0 |
| 8 | 17 | 12 | -12 | 0 |
| 10 | 79 | 70 | 70 | 0 |
| 12 | <12 | 3 | 3 | 0 |
| 14 | 15 | 9 | 9 | 0 |
| **4k0** | | | | |
| 0 | 990* | 1236 | 1236 | 0 |
| 2 | 1023* | 1112 | -1112 | 0 |
| 4 | 377 | 384 | 384 | 0 |
| 6 | 61 | 59 | -59 | 0 |
| 8 | 65 | 64 | 64 | 0 |
| 10 | 100 | 82 | 82 | 0 |
| 12 | 38 | 30 | 30 | 0 |
| 14 | 12 | 8 | -8 | 0 |
| **6k0** | | | | |
| 2 | 984 | 973 | 973 | 0 |
| 4 | 244 | 253 | -253 | 0 |
| 6 | 59 | 44 | -44 | 0 |
| 8 | 107 | 101 | 101 | 0 |
| 10 | 78 | 66 | 66 | 0 |
| 12 | ---* | 0 | 0 | 0 |
| 14 | 13 | 9 | 9 | 0 |
| **8k0** | | | | |
| 0 | 89 | 73 | 73 | 0 |
| 2 | 219 | 208 | -208 | 0 |
| 4 | 280 | 255 | 255 | 0 |
| 6 | 18 | 6 | -6 | 0 |
| 8 | 141 | 121 | 121 | 0 |
| 10 | 44 | 33 | 33 | 0 |
| **10,k,0** | | | | |
| 2 | 215 | 199 | 199 | 0 |
| 4 | 323 | 291 | -291 | 0 |
| 6 | 177 | 169 | 169 | 0 |
| 8 | 129 | 95 | 95 | 0 |
| 10 | 23 | 17 | 17 | 0 |
| **12,k,0** | | | | |
| 0 | 79 | 83 | 83 | 0 |
| 2 | 68 | 60 | -60 | 0 |
| 4 | 123 | 115 | 115 | 0 |
| 6 | 44 | 45 | 45 | 0 |
| 8 | 72 | 57 | 57 | 0 |
| **14,k,0** | | | | |
| 2 | 25 | 22 | -22 | 0 |
| 4 | 17 | 11 | -11 | 0 |
| 6 | 173 | 138 | 138 | 0 |
| 8 | <15 | 9 | 9 | 0 |
| **16,k,0** | | | | |
| 0 | 19 | 8 | -8 | 0 |
| 2 | 28 | 25 | -25 | 0 |
| 4 | 75 | 75 | 75 | 0 |
| 6 | 25 | 16 | 16 | 0 |
| 8 | 24 | 20 | 20 | 0 |
| **18,k,0** | | | | |
| 2 | <15 | 10 | 10 | 0 |
| 4 | 47 | 40 | 40 | 0 |
| 6 | 34 | 33 | 33 | 0 |
| 8 | <13 | 7 | -7 | 0 |
| **20,k,0** | | | | |
| 0 | 23 | 18 | -18 | 0 |
| 2 | 25 | 18 | 18 | 0 |
| 4 | 30 | 32 | 32 | 0 |
| 6 | <13 | 0 | 0 | 0 |
| 8 | 12 | 9 | 9 | 0 |
| **22,k,0** | | | | |
| 2 | <16 | 13 | 13 | 0 |
| 4 | 15 | 9 | 9 | 0 |
| 6 | <12 | 5 | 5 | 0 |
| **24,k,0** | | | | |
| 0 | <12 | 9 | 9 | 0 |
| 2 | <13 | 12 | 12 | 0 |
| 4 | <10 | 4 | 4 | 0 |
| **1k1** | | | | |
| 1 | 734* | 759 | 177 | -738 |
| 3 | 317 | 330 | -314 | 102 |
| 5 | 432 | 392 | 300 | 251 |
| 7 | 185 | 216 | -154 | 152 |
| 9 | 34 | 39 | 24 | 31 |
| 11 | 34 | 40 | 39 | 7 |
| 13 | 11 | 15 | 2 | -15 |
| **3k1** | | | | |
| 1 | 866 | 792 | 31 | 791 |
| 3 | 146 | 167 | 19 | 166 |
| 5 | 218 | 236 | -67 | 226 |
| 7 | 69 | 73 | 44 | 58 |
| 9 | <14 | 10 | -8 | -7 |
| 11 | 31 | 35 | 31 | -17 |
| 13 | 13 | 12 | 6 | 10 |
| **5k1** | | | | |
| 1 | 341 | 307 | -260 | -164 |
| 3 | 653 | 654 | 212 | 619 |
| 5 | 261 | 255 | 24 | 254 |
| 7 | 113 | 117 | -105 | 52 |
| 9 | 84 | 93 | 90 | -23 |
| 11 | 14 | 16 | 14 | -8 |
| **7k1** | | | | |
| 1 | 287 | 260 | -39 | 257 |
| 3 | 37 | 21 | 7 | 19 |
| 5 | 213 | 215 | -99 | 191 |
| 7 | 63 | 71 | 69 | 14 |
| 9 | 28 | 31 | 31 | -1 |
| **9k1** | | | | |
| 1 | 259 | 239 | 100 | 217 |
| 3 | 432 | 381 | 65 | 376 |
| 5 | 80 | 91 | -89 | -21 |
| 7 | 39 | 52 | 52 | -4 |
| 9 | 57 | 53 | 48 | -22 |
| **11,k,1** | | | | |
| 1 | 139 | 138 | -10 | 137 |
| 3 | 112 | 103 | -87 | 55 |
| 5 | 117 | 108 | 41 | 100 |
| 7 | 69 | 70 | 65 | -27 |
| **13,k,1** | | | | |
| 1 | 106 | 118 | 74 | 92 |
| 3 | 79 | 83 | -36 | 75 |
| 5 | 56 | 59 | 17 | -56 |
| 7 | 48 | 47 | 47 | 7 |
| **15,k,1** | | | | |
| 1 | 81 | 93 | -54 | 75 |
| 3 | 16 | 13 | -3 | 13 |
| 5 | 67 | 66 | 64 | 16 |
| 7 | 30 | 34 | 6 | -33 |
| **17,k,1** | | | | |
| 1 | 19 | 23 | 14 | 18 |
| 3 | <11 | 9 | 5 | 7 |
| 5 | 31 | 34 | 26 | -23 |
| **19,k,1** | | | | |
| 1 | 20 | 26 | -23 | 11 |
| 3 | 30 | 34 | 30 | -15 |
| 5 | 17 | 16 | 15 | -5 |
| **0k2** | | | | |
| 2 | 518 | 465 | -362 | -292 |
| 6 | 214 | 207 | 150 | 143 |
| 10 | 25 | 30 | -11 | -28 |
| **2k2** | | | | |
| 0 | 626 | 602 | -57 | -599 |
| 2 | 240 | 239 | -173 | 164 |
| 4 | 80 | 81 | -54 | 61 |
| 6 | 26 | 28 | -11 | 25 |
| 8 | 36 | 41 | -7 | 41 |
| 10 | 20 | 28 | 13 | -25 |
| 12 | 17 | 23 | 6 | -22 |
| **4k2** | | | | |
| 2 | 285 | 261 | -95 | -243 |
| 4 | 244 | 235 | 182 | 149 |
| 6 | 95 | 91 | -9 | 91 |
| 8 | 76 | 80 | -61 | -52 |
| 10 | 17 | 21 | 10 | -18 |
| **6k2** | | | | |
| 0 | 354 | 363 | -326 | -160 |
| 2 | 306 | 249 | 11 | 249 |
| 4 | 78 | 61 | 58 | 21 |
| 6 | 29 | 38 | -38 | 2 |
| 8 | 32 | 38 | 36 | -10 |
| 10 | 31 | 35 | -8 | -35 |
| **8k2** | | | | |
| 2 | 75 | 73 | 61 | -40 |
| 4 | 179 | 168 | -24 | 166 |
| 6 | 43 | 43 | -38 | -19 |
| 8 | 48 | 53 | -7 | -53 |
| 10 | <16 | 10 | 7 | -8 |
| **10,k,2** | | | | |
| 0 | 63 | 56 | -5 | 56 |
| 2 | 186 | 170 | 113 | 127 |
| 4 | 28 | 40 | -11 | -38 |
| 6 | 18 | 16 | -10 | -13 |
| 8 | 24 | 26 | 13 | -23 |
| 10 | 27 | 30 | -24 | -18 |
| **12,k,2** | | | | |
| 2 | 57 | 49 | 5 | 49 |
| 4 | 57 | 59 | -35 | 48 |
| 6 | 59 | 60 | 30 | -53 |
| 8 | 22 | 29 | -13 | -26 |
| **14,k,2** | | | | |
| 0 | 113 | 114 | 89 | 71 |
| 2 | 18 | 23 | -6 | 22 |
| 4 | 26 | 33 | -4 | -33 |
| 6 | <17 | 12 | -6 | -11 |
| **16,k,2** | | | | |
| 2 | 29 | 30 | 12 | 27 |
| 4 | <17 | 17 | 10 | -13 |
| 6 | 31 | 32 | 7 | -31 |
| **18,k,2** | | | | |
| 0 | 32 | 36 | 14 | 34 |
| 2 | 17 | 22 | -19 | -11 |
| 4 | 16 | 20 | 13 | -15 |
| **22,k,2** | | | | |
| 0 | <13 | 5 | 3 | 5 |
| **1k3** | | | | |
| 1 | 208 | 205 | 180 | -98 |
| 3 | 66 | 63 | -4 | -63 |
| 5 | 208 | 208 | 188 | 89 |
| 7 | 110 | 116 | 65 | -96 |
| 9 | <13 | 5 | 4 | 3 |
| 11 | 10 | 13 | -6 | -12 |
| **3k3** | | | | |
| 1 | 215 | 200 | -158 | -124 |
| 3 | 232 | 211 | 208 | -32 |
| 5 | 168 | 164 | 148 | -70 |
| 7 | 94 | 91 | 39 | 83 |
| 9 | 47 | 43 | 7 | -42 |
| 11 | <8 | 9 | 3 | -9 |
| **5k3** | | | | |
| 1 | 138 | 148 | 141 | -44 |
| 3 | 126 | 110 | 98 | -51 |
| 5 | 151 | 133 | 133 | 8 |
| 7 | 77 | 75 | 11 | -75 |
| 9 | 17 | 12 | 3 | 12 |
| 11 | <7 | 10 | -3 | -10 |
| **7k3** | | | | |
| 1 | 33 | 26 | 18 | -19 |
| 3 | 274 | 250 | 250 | -4 |
| 5 | 47 | 41 | 38 | -13 |
| 7 | 13 | 11 | 9 | 6 |
| 9 | 50 | 49 | 5 | -48 |
| 11 | 8 | 8 | -7 | 2 |
| **9k3** | | | | |
| 1 | 142 | 145 | 144 | -18 |
| 3 | 74 | 73 | 72 | -8 |
| 5 | 42 | 44 | 44 | -4 |
| 7 | 38 | 38 | 11 | -36 |
| 9 | 18 | 11 | -7 | 9 |
| 11 | 9 | 11 | -1 | -11 |
| **11,k,3** | | | | |
| 1 | 114 | 113 | 107 | -34 |
| 3 | 76 | 86 | 86 | -8 |
| 5 | 20 | 25 | -8 | -24 |
| 7 | 13 | 18 | 7 | -16 |
| 9 | 18 | 20 | -7 | -19 |
| **13,k,3** | | | | |
| 1 | 78 | 90 | 78 | 45 |
| 3 | 20 | 16 | 15 | 5 |
| 5 | 23 | 27 | 22 | -16 |
| 7 | 12 | 19 | -6 | -18 |
| **15,k,3** | | | | |
| 1 | 59 | 60 | 41 | -43 |
| 3 | <13 | 8 | 8 | 1 |
| **17,k,3** | | | | |
| 1 | 37 | 38 | 13 | 35 |
| 3 | 20 | 21 | 8 | -19 |
| **19,k,3** | | | | |
| 1 | 15 | 18 | 8 | -16 |
| 3 | 9 | 6 | 1 | 6 |
| **21,k,3** | | | | |
| 1 | 6 | 6 | 1 | 6 |
| 3 | 8 | 14 | 1 | -14 |
| **0k4** | | | | |
| 4 | 142 | 142 | 138 | 36 |
| 8 | 54 | 50 | 49 | 9 |
| **2k4** | | | | |
| 2 | 122 | 132 | 59 | 118 |
| 4 | 74 | 77 | 71 | -30 |
| 6 | 165 | 146 | 146 | -12 |
| 8 | 21 | 25 | -22 | 11 |
| 10 | 16 | 12 | 7 | 9 |
| **4k4** | | | | |
| 0 | 129 | 134 | 33 | 130 |
| 2 | 41 | 43 | 38 | -20 |
| 4 | 148 | 134 | 132 | 25 |
| 6 | 17 | 9 | 7 | 6 |
| 8 | 47 | 45 | 45 | 4 |
| 10 | <9 | 8 | -5 | 6 |
| **6k4** | | | | |
| 2 | 117 | 120 | 108 | 52 |
| 4 | 95 | 96 | 88 | -37 |
| 6 | 58 | 48 | 45 | 16 |
| 8 | 24 | 29 | -24 | 16 |
| 10 | 19 | 13 | 10 | 7 |
| **8k4** | | | | |
| 0 | 89 | 87 | 64 | 60 |
| 2 | 95 | 105 | 96 | -42 |
| 4 | 89 | 75 | 72 | 22 |
| 6 | 10 | 7 | 6 | 4 |
| 8 | 23 | 24 | 23 | 6 |
| **10,k,4** | | | | |
| 2 | 75 | 82 | 81 | 13 |
| 4 | 30 | 28 | 26 | -10 |
| 6 | 28 | 21 | 12 | 17 |
| 8 | 14 | 14 | -8 | 12 |
| **12,k,4** | | | | |
| 0 | 99 | 107 | 107 | -8 |
| 2 | 40 | 46 | 42 | -20 |
| 4 | 23 | 22 | 13 | 18 |
| 6 | <11 | 3 | 2 | 3 |
| 8 | 13 | 10 | 1 | 10 |
| **14,k,4** | | | | |
| 2 | 35 | 39 | 39 | 3 |
| 4 | 9 | 8 | 8 | -2 |
| 6 | 12 | 12 | 2 | 12 |
| **16,k,4** | | | | |
| 0 | 49 | 53 | 51 | -13 |
| 2 | <12 | 9 | -9 | 3 |
| 4 | 9 | 9 | 4 | 8 |
| **18,k,4** | | | | |
| 2 | 15 | 18 | 18 | -2 |

Table 13. - The normalized coordinates,in Angstrom units, for the heavy atoms.

| Atom | X | Y | Z |
|---|---|---|---|
| O | 0.0000 | 0.0000 | 1.2693 |
| N | 0.6506 | 0.9297 | -0.6639 |
| $C_1$ | 0.0000 | 0.0000 | 0.0073 |
| $C_2$ | 1.4884 | 1.9475 | -0.0613 |

| Normalization equations | Inverse equations |
|---|---|
| $X = 20.185x$ | $x = 0.04954X$ |
| $Y = 11.401y$ | $y = 0.08771Y$ |
| $Z = 4.563z$ | $z = 0.21915Z$ |

Table 14. - Interatomic distances and angles.

| | | | |
|---|---|---|---|
| $C_1$-O | 1.262 Å | O-$C_1$-N | 120.6° |
| $C_1$-N | 1.318 | N-$C_1$-N' | 118.8 |
| $C_2$-N | 1.449 | $C_1$-N-$C_2$ | 124.7 |

Hydrogen bond distance N···O' 2.864 Å

iii) Temperature parameter discussion.

The final anisotropic temperature factor parameters for the heavy atoms are given in table 15. These parameters were used to obtain the magnitudes and direction cosines of the principal axes of the vibration ellipsoids listed in table 16. The direction cosines are given relative to a Cartesian coordinate system in which X, Y and Z axes are coincident with a, b and c, respectively.

Tables 16 and 17 show that the anisotropic thermal motion can be described by two librations and one displacement motion. The more significant libration is about an axis through the two methyl groups and the second libration is about the two-fold molecular axis. There also appears to be some parallel displacement of the molecular plane. All of these motions are predominantly out-of-plane and are recognized in part by the small angle that the major thermal axis of each atom makes with the normal to the plane. The libration of the methyl group would be expected to be a minimum along the N-to-methyl bond, and hence the angle between the bond direction and the minor thermal axis of $C_2$ should be small. This angle is found to be 10°. (Probably none of these deviations are greater than the experimental error.)

These thermal effects are compatible with the nature of the disorder found in the crystal. The external forces acting on a molecular level would produce displacements of and librations about the three principal axes of inertia. The axis with the smallest moment of inertia

Table 15. - Final anisotropic temperature factor parameters. The temperature factors are in the form of:

$$T_i = \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}k\ell) .$$

| Atom | $B_{11}$ | $B_{22}$ | $B_{33}$ | $B_{12}$ | $B_{13}$ | $B_{23}$ |
|---|---|---|---|---|---|---|
| O | 0.0095 | 0.0228 | 0.0327 | -0.0072 | 0.0000 | 0.0000 |
| N | 0.0057 | 0.0162 | 0.0473 | -0.0029 | 0.0013 | 0.0001 |
| $C_1$ | 0.0056 | 0.0163 | 0.0538 | 0.0014 | 0.0000 | 0.0000 |
| $C_2$ | 0.0069 | 0.0186 | 0.0907 | -0.0040 | 0.0015 | -0.0063 |

Table 16. - Magnitudes and direction cosines of the principal axes of the vibration ellipsoids.

| Atom | Axis i | $B_i$ | $g_i^{(1)}$ | $g_i^{(2)}$ | $g_i^{(3)}$ |
|---|---|---|---|---|---|
| O | 1 | 21.130 | 0.796 | -0.605 | 0 |
|  | 2 | 6.984 | 0.605 | 0.796 | 0 |
|  | 3 | 2.761 | 0 | 0 | 1 |
| N | 1 | 11.953 | 0.762 | -0.646 | 0.044 |
|  | 2 | 6.337 | 0.636 | 0.759 | 0.138 |
|  | 3 | 3.937 | -0.122 | -0.077 | 0.989 |
| $C_1$ | 1 | 10.413 | 0.790 | 0.612 | 0 |
|  | 2 | 7.648 | 0.612 | -0.790 | 0 |
|  | 3 | 4.545 | 0 | 0 | 1 |
| $C_2$ | 1 | 14.885 | 0.759 | -0.627 | 0.174 |
|  | 2 | 7.891 | 0.464 | 0.335 | -0.820 |
|  | 3 | 6.407 | 0.456 | 0.703 | 0.546 |

Table 17. - Orientation of the thermal librations.

| Atom | Acute angle between the major axes of the thermal ellipsoids and the plane normal |
|---|---|
| O | 1.12° |
| N | 4.21 |
| $C_1$ | 74.70 (semi-major $\theta = 15.30$) |
| $C_2$ | 10.21 |

| Direction cosine of the N to $C_2$ methyl bond direction | Acute angle between the bond directions and the minor ellipsoid axis of the $C_2$ methyl group |
|---|---|
| $g(1) = 0.5782$ | |
| $g(2) = 0.7024$ | 10.10° |
| $g(3) = 0.4159$ | |

about it should have the largest libration; the smallest axis of inertia runs very nearly through the two methyl groups and indeed the largest libration is about this axis. The semi-major axis of inertia is coincident with the two-fold molecular axis; the second largest libration is found to be about this axis. The parallel displacement of the whole molecule may be in part the result of the bending of the crystal.

## Experimental Work: N-methylurea.

### i) Introduction.

The earliest crystallographic observations on N-methylurea were made by Mez (16). He reported the compound to be "rhombic" with axial ratios a:b:c = 0.9904:1:1.2127, the density to be 1.204 g cm$^{-3}$ and the melting point to be 102°C. The earliest x-ray investigation was made by H. Mark (14), who reported the crystals to be rhombic disphenoidal belonging to the space group $V^4(P_{2_1 2_1 2_1}-D_2^4)$ and gave the unit cell dimensions a = 5.63 Å, b = 5.64, c = 4.70. His unit cell dimensions are apparently wrong but the space group is correct. The next investigators to look at N-methylurea were Corey and Wyckoff (17), (hereafter CW); they used diffraction data obtained from oscillation and rotation photographs to show that the crystals have hemihedral orthorhombic symmetry. They confirmed the space group $P_{2_1 2_1 2_1}-D_2^4$ and reported the dimensions a = 6.89 Å, b = 6.96, and c = 8.45 with four molecules per unit cell.

CW proposed a structure in which the plane of the urea group was nearly perpendicular to the ab plane, making an angle of about 34° with the b-axis. The methyl group was thought to lay considerably outside the plane of the urea nucleus. The agreement of observed and calculated $(0k\ell)$ and $(hk0)$ structure factors of higher order was not satisfactory, however, and they were prompted to suggest that the proposed structure was not the correct one. (In the notation of CW, the Miller indices h and k and the atomic parameters x and y refer to the b and a-axes respectively; this notation has been inverted here in order to conform to the present standard notation.)

Gordon (18,19), while attempting to correlate the diamagnetic anisotropy of crystals of urea and N-methylurea, considered an alternative structure. He presumed that the molecules are linked by N-H$\cdots$O hydrogen bonds to form chains with the individual C=O bonds nearly coincident with the two-fold axes parallel to c. To account for the relatively large published value of F(020), he presumed that the molecules lay in planes nearly perpendicular to b and separated by b/2. In an attempt to verify this structure he obtained x-ray diffraction data for the [010] projection in which the atoms were expected to be well resolved; in addition, he made new observation of the (hk0) and (ok$\ell$) reflections. These observations revealed a serious error in the value of |F(012)| and a less serious error in |F(024)| as reported by CW. These changes placed the results of CW in doubt.

Gordon terminated his work by computing the Fourier syntheses of the [001] and the [100] projections using in each case only the dominant structure factors ( |F(020)| , |F(012)| , |F(110)| , |F(011)| and |F(120)| ); for each set of structure factors he selected that combination of signs which led to the most reasonable electron density projection. In the [100] projection a more or less circular peak was found at Y = 0, Z = 5/8; in the [001] projection the maximum was at X = 3/4, Y = 0 and was elongated along [120].

ii) Unit cell and space group.

In the present investigation, crystals of N-methylurea were obtained by vacuum sublimation at 110°C. The majority of the crystals were elongated along the c-axis and showed pseudo-tetragonal symmetry. One of these crystals was mounted with its c-axis perpendicular to the x-ray beam. Equi-inclination Weissenberg photographs confirmed the orientation and the space group $P_{2_1 2_1 2_1}$-$D_2^4$ . Preliminary measurements of the unit-cell parameters from the O-layer Weissenberg photograph and from a rotation photograph agreed within experimental error with the dimensions reported by CW; the values of CW were adopted since they were computed from powder diagrams and the axial ratios, $a/b{:}b/b{:}c/b = 0.9914{:}1{:}1.2141$, agree very well with the optical measurements of Mez (16).

iii) Intensity data.

Three-dimensional intensity data for Cu $K_\alpha$ radiation were collected on multiple-film equi-inclination Weissenberg photographs taken about the c-axis. All layer lines having inclination angles less than 22° were recorded ($\ell$ from 0 to 4 inclusive). The intensity data were estimated by visual comparison with a standard intensity scale prepared from the same crystal. The intensities were corrected for Lorentz and polarization factors using a program written for the Burroughs 220 computer (see appendix II for a description of the program). No corrections for absorption or extinction were made.

Approximate layer scale factors were obtained by comparison with the data of CW and of Gordon. These values were improved by comparison with the calculated structure factors during the least-squares refinements of the parameters.

The atomic scattering curves used in this investigation for carbon, nitrogen and oxygen were taken as the average of the values given by Berghuis *et al.* (8) and Hoerni and Ibers (9). The hydrogen scattering curve was that of McWeeny (7).

## Determination of the Structure of N-methylurea.

### i) Derivation of the trial model.

About a dozen cork-ball molecules of N-methylurea were constructed using straight pins for the covalent bonds. In this model each molecule was assumed to be planar with the methyl group cis to the oxygen atom. This latter assumption was based on the consideration that the hydrogen atoms on N and N' would form hydrogen-bonded chains as in urea and N,N'-dimethylurea.

The next step in the formation of the trial model was to construct the molecular chains. The diamagnetic anisotropy measurements of N-methylurea made by Gordon (18,19) implied that the C=O bonds are nearly parallel to c. In order to satisfy the symmetry as well as the unit cell identity distance along c we need two molecules of N-methylurea per repeat unit along this axis. However, unlike N,N'-dimethylurea, these molecules could not lie in the same plane, since a repeat distance of about 9.2 Å rather than the observed value of 8.45 Å would be needed. Thus, it was necessary to make saw-tooth creases in the molecular chains, with an angle between the planes of adjacent folds in the chain of about 60°, each fold being inclined by 30° away from the c-axis. (The final value for the angle between two adjacent planes was found to be 56°.)

The crease in the plane of the molecular chain was assumed to be such that an oxygen atom of one molecule would be very nearly in the plane of the adjacent molecule, so as to allow two N-H···O hydrogen bonds to be formed. The cork ball molecules were now pierced by long piano wires, which represented the screw axes, so as to form five chains of molecules with two or three molecules of N-methylurea per chain. The last step in construction of the trial model was to find the method of packing of the chains.

Four of the chain models were placed so as to outline the ab face of the unit cell. The fifth chain, with the C=O bonds pointed in the opposite direction from the other four, was placed at the center of the ab face. The chains were now rotated so that the planes of the molecules were nearly perpendicular to the b-axis. At this point the relative height of the fifth chain was adjusted so that its methyl group fitted into the hollows formed by the folds of the adjacent chains. This required twisting the molecular planes somewhat away from perpendicular to b. The amide groups were now found to be oriented favorably for formation of cross-chain hydrogen bonds, using the second hydrogen atom of the $NH_2$ group. In this case the additional N-H··· bond is pointed toward the apex of a fold of a neighboring chain, where an oxygen atom is assumed to be located. This completed the construction of the trial model; it remained only to choose the correct origin and to test the model.

### ii) Two-dimensional refinement.

The structure was refined in the [010] zone by alternate calculation of structure factors and Fourier syntheses. The observed ($h0\ell$) structure factors of Gordon (18) were used throughout the two-dimensional investigation. Although the final structure did not differ qualitatively from the first trial model, it turned out to be very difficult to verify the model because several structures close to the correct one converged to false minima during the application of least squares procedures.

The initial set of $x$ and $z$ coordinates were taken from the cork ball trial model, the origin of the unit cell being taken at the center of symmetry of the projection. These coordinates were used to compute the $h0\ell$ structure factors, the calculated signs were then assigned to the observed structure factors and the electron density Fourier synthesis was computed. Only a small stretch of the imagination was necessary to find a molecule of N-methylurea in the map. New coordinates were read off the Fourier map and were used to compute new structure factors and a few cycles of least-squares were run. Convergence stopped with $R$ at about 0.35, and a new Fourier was calculated. Although the new map had no spurious peaks, the overall situation was not good since no further refinement could be obtained. This was the first of many similar repetitions of an encouraging beginning without a satisfactory ending.

The next idea was to put the methyl group on the other nitrogen atom. After that the molecule was moved a little one way and then a little the other, a little up and a little down; it was rotated a little clockwise and a little counterclockwise. The results were more or less the same: discouraging. The methyl group was moved to the trans positions and then to some of the intermediate non-coplanar positions. A few entirely different arrangements were tried, giving, in general, much worse results. At this low point in the investigation one could hardly disagree with CW when they said: "...when evaluating structures of organic crystals based on limited data it is important to remember that it has not been hard to find an arrangement for methyl urea which though incorrect can explain quantitatively its simple reflections."

Since all of the two-dimensional refinement work was being done on the Burroughs 205 computer, the first draft of a program for this computer was written which would select a random orientation for the planar molecule lying within the limits giving an arrangement compatible with the qualitative features of the initial trial model. It was intended to work in the following way: A set of coordinates for the predicted planar configuration of N-methylurea was computed using a two-dimensional Cartesian coordinate system in which the C=O bond was taken coincident with the Z-axis; the atom $C_1$ was placed at the origin and the horizontal axis was taken coincident with the Y-axis. It was felt that the molecule could be translated by $-0.347 \leqq W \leqq 0.617$ Å along the Z axis. It was also felt that the normal to the plane of the molecule must be at an angle $20 \leqq \phi \leqq 30°$ away from the X-axis and at an angle $110 \leqq \theta \leqq 120°$ away from the Z-axis, where $\phi$ and $\theta$ are the spherical coordinates of the normal to the plane. These conditions defined the acceptable orientations of the trial model. The program was designed to calculate random values for W, $\phi$ and $\theta$ within the ranges specified above and then translate, rotate and de-orthogonalize the planar coordinates so as to give a set of fractional coordinates which could be used as a trial model. One additional check of these coordinates was planned; that is, the cross chain hydrogen bond distance was to be computed and only those configurations with this distance lying between 2.6 and 3.3 Å were to be used. However, before this avenue of approach was tried, a systematic pseudo-equality between the intensities of reflections of the type $h0\ell$ and $\ell 0h$ was noted, and it was felt that this feature ought to be investigated first.

Table 18 lists the $(h0\ell)$ data of Gordon (18) with $|F(h0\ell)|$ and $|F(\ell 0h)|$ pairs in parallel columns in order to show the pseudo-equality between them. In table 19, three special arrangements of atoms were considered in order to find their effect on the structure factors. Although none of the cases I, II or III could be used individually to explain the pseudo-equality, it seemed reasonable that some combination could. Because the pseudo-equality was as good for the $h + \ell$ odd as for the $h + \ell$ even structure factors, case II had to be excluded and a combination of I and III was investigated. A Fourier synthesis based on such an arrangement had previously been calculated. The trial structure giving

Table 18. - Observed structure factors for N-methylurea as reported by Gordon. The structure factors are listed to show the pseudo-equality between $|F_{h0\ell}|$ and $|F_{\ell 0h}|$ pairs.

| hkℓ | $\|F_o\|$ | hkℓ | $\|F_o\|$ | hkℓ | $\|F_o\|$ | hkℓ | $\|F_o\|$ |
|---|---|---|---|---|---|---|---|
| 102 | 19.5 | 201 | 7.7 | 305 | 8.8 | 503 | 5.9 |
| 103 | 2.2 | 301 | 3.1 | 306 | 2.2 | 603 | < 2.0 |
| 104 | 27.9 | 401 | 31.8 | 307 | 7.2 | 703 | - |
| 105 | 5.6 | 501 | 4.0 | 400 | 29.3 | 004 | 32.6 |
| 106 | 15.3 | 601 | 11.6 | 405 | 10.6 | 504 | 7.8 |
| 107 | 3.2 | 701 | - | 406 | 11.1 | 604 | 10.3 |
| 200 | 17.5 | 002 | 25.8 | 407 | 8.0 | 704 | - |
| 203 | 16.1 | 302 | 16.1 | 506 | 2.3 | 605 | 4.1 |
| 204 | 9.8 | 402 | 10.1 | 507 | 3.3 | 705 | - |
| 205 | 7.4 | 502 | 7.7 | 600 | 11.1 | 006 | ≤ 4.2 |
| 206 | 12.6 | 602 | 5.4 | 607 | 2.3 | 706 | - |
| 207 | 7.2 | 702 | - | 800 | - | 008 | 6.4 |
| 304 | 23.4 | 403 | 11.7 | | | | |

Table 19. - Some special relations in plane group Pgg and their effect on the observed structure factors.

Case I. An atom at $(x, x)$ and equivalent positions.

Case II. An atom at $(x, \frac{1}{2} - x)$ and equivalent positions.

Case III. One atom at $(x, \frac{1}{2} - x)$ and another at $(\frac{1}{2} - x, x)$ and the corresponding equivalent positions.

Equivalent positions for plane group Pgg: $\pm \left| x, z; \frac{1}{2} + x, \frac{1}{2} - z \right|$

| Case | $h + \ell = 2n$ | $h + \ell = 2n + 1$ |
|---|---|---|
| I | $A = Tc$ | $A = Ts$ |
| II | $A = Tc \cos \ell\pi$ | $A = -Ts \cos \ell\pi$ |
| III | $A = Tc(\cos \ell\pi + \cos h\pi)$ | $A = 0$ |

where $Tc = 4 \cos 2\pi hx \cos 2\pi \ell x$ and $Ts = -4 \sin 2\pi hx \sin 2\pi \ell x$

| Case | $h + \ell = 2n$ | $h + \ell = 2n + 1$ |
|---|---|---|
| I | $A(h0\ell) = A(\ell 0h)$ | $A(h0\ell) = A(\ell 0h)$ |
| II | $A(h0\ell) = A(\ell 0h)$ | $A(h0\ell) = -A(\ell 0h)$ |
| III | $A(h0\ell) = A(\ell 0h)$ | 0 |
| I + II | $F(h0\ell) = F(\ell 0h)$ | $F(h0\ell) \neq F(\ell 0h)$ |
| I + III | $F(h0\ell) = F(\ell 0h)$ | $F(h0\ell) = F(\ell 0h)$ |

this early Fourier had no atoms over 0.6 Å away from the correct structure and yet it was locked in a false minimum by the least-squares and the Fourier gave no indication as to which direction the atoms should be moved. Thus one can understand the frustration experienced during this part of the investigation.

The Fourier indicated that the oxygen atom and the $NH_2$ nitrogen should be moved toward the line $z = x$, thus satisfying the conditions of Case I. (The molecule considered is the one in the lower left corner of figure 9 with the origin moved to -1/4, 1/4.) The NH nitrogen was moved toward the line $z = 1/2 - x$ and the carbonyl carbon was placed slightly above the centroid of the triangle formed by the two nitrogens and the oxygen atoms. The $N\text{-}CH_3$ bond was drawn parallel to the $C\text{-}NH_2$ bond and extended until it intersected the nearest diagonal. As can be seen in figure 9, this placed a methyl group, from the symmetry related molecule in the lower right, on the line $z = 1/2 - x$ at a point corresponding roughly to one of the positions of case III; the other position being occupied by the NH nitrogen.

The coordinates obtained from these considerations proved to be a good enough starting point for the least-squares process, and five cycles brought the R factor from 0.48 to 0.16. Further refinement made use of the three dimensional data.

During the latter stages of the two-dimensional investigation, an infra-red spectral study of the mono- and dimethylureas was

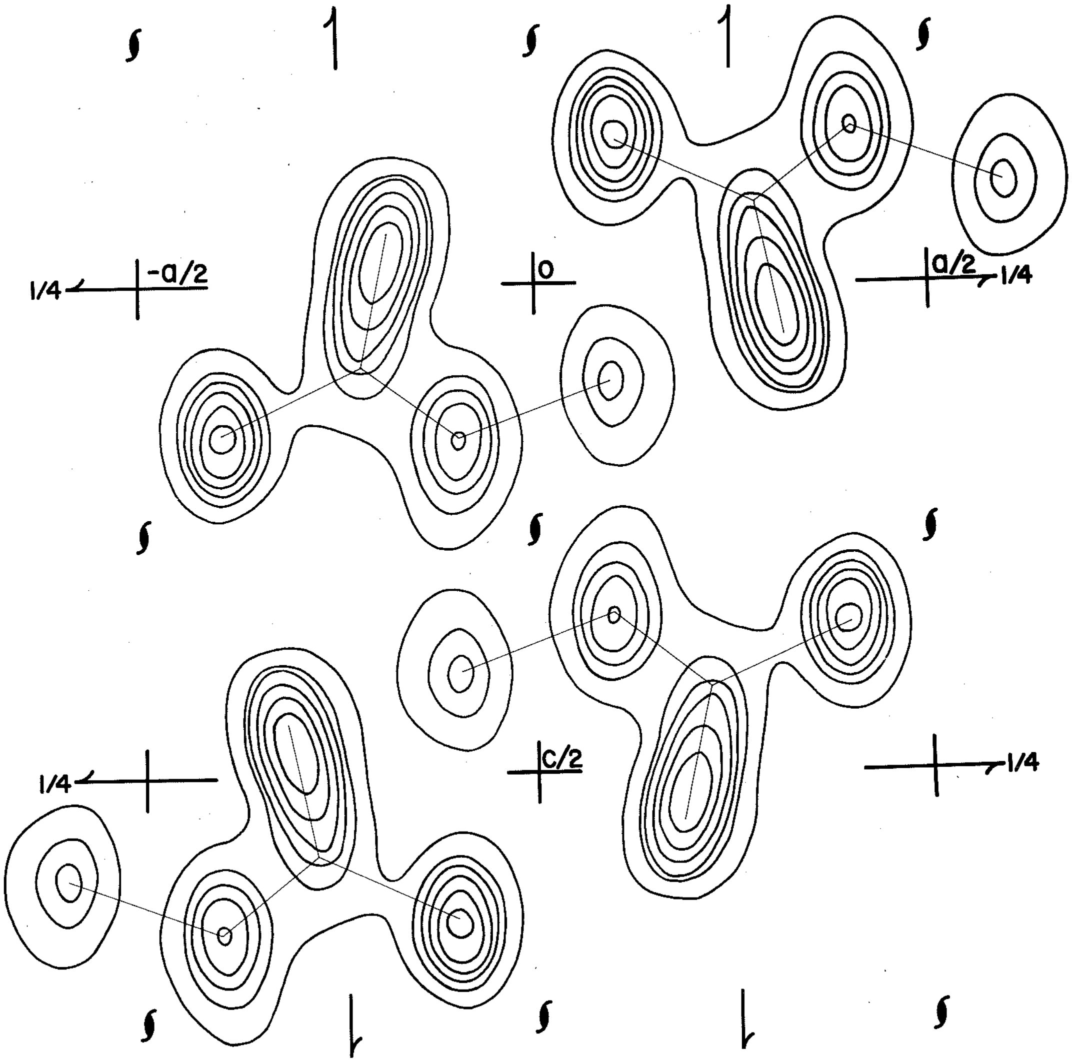


Figure 9. - Electron density projection of N-methylurea along the b-axis. Contours start at 2.7 e. Å$^{-2}$, increase by 1.4 e. Å$^{-2}$ over the first interval and by 0.7 e. Å$^{-2}$ over the remaining intervals.

undertaken. Figure 10 shows the infra-red spectra of N-methylurea, N,N'-dimethylurea and N,N-dimethylurea in pressed discs of potassium-bromide. The spectra were obtained on a Perkin-Elmer Model 21 double-beam recording spectrophotometer equipped with sodium-chloride optics. These spectra were intended as a preliminary to polarized IR studies of single crystals in order to confirm, if possible, the orientation of the C=O bond in N-methylurea. This work was terminated when the structure was solved in projection.

### iii) Three-dimensional refinement.

A y parameter of zero for the carbon atom $C_1$ was taken from the cork ball model; the other y parameters were computed from predicted interatomic distances. Although the derived set of x, y and z parameters for all the atoms formed a complete description of the structure relative to the origin, it was decided to express them relative to the origin defined by the expressions given in the International Tables (20) for the space group $P_{2_1 2_1 2_1}$. The translation components relating the two coordinate systems were determined in the following way: Four low-order reflections, one from each of the different parity groups, were used to test the choice of origin. All eight possible triple trigonometric products were computed for each reflection and were used to calculate eight "structure factors." Since a sine curve is equal to a cosine curve which has been shifted forward by a quarter cycle, the four groups of "structure factors" were examined to see what simple quarter-cycle shifts had to be applied to the parameter to give

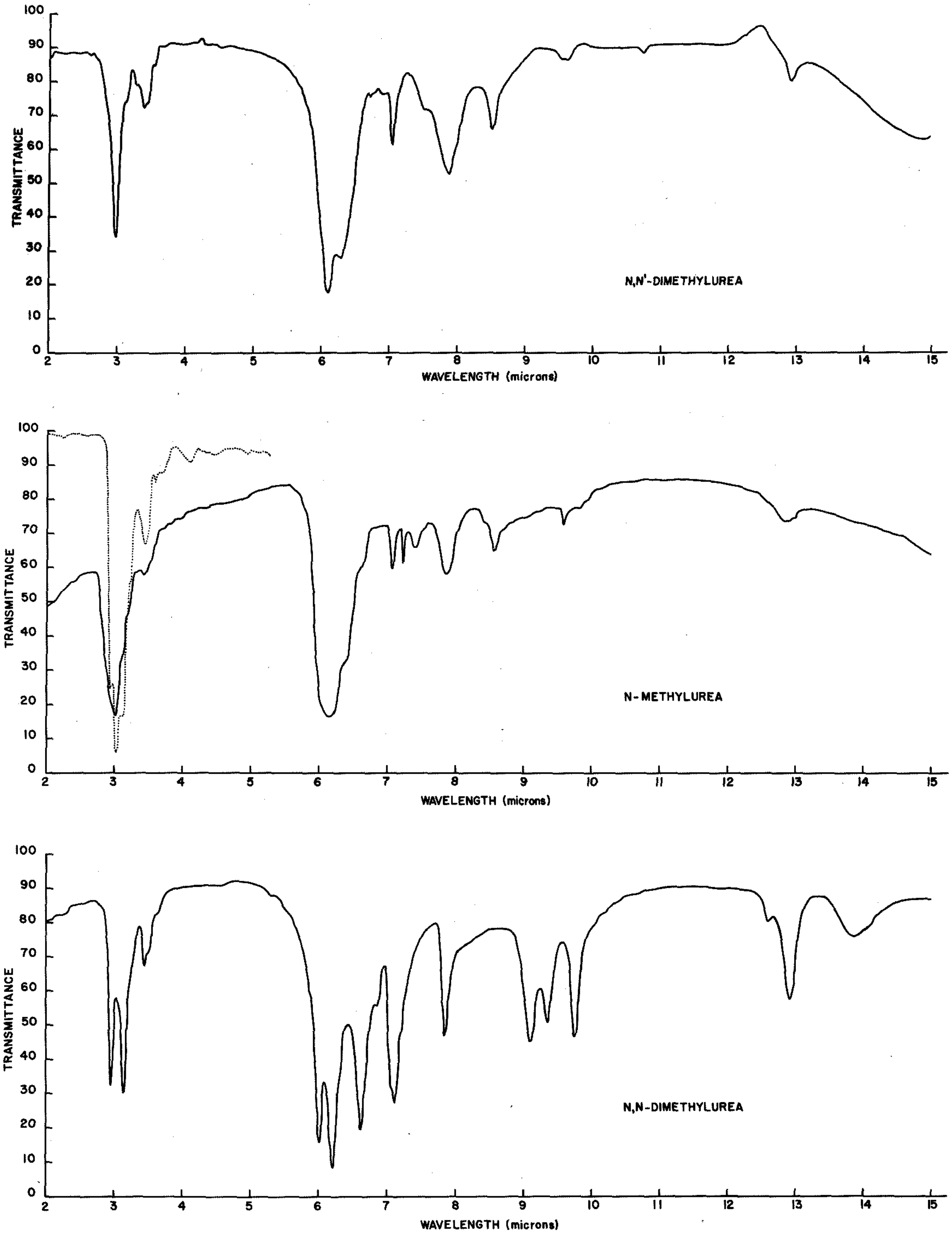


Figure 10. - Infra-red spectra of N-methylurea, N,N'-dimethylurea and N,N-dimethylurea.

the best fit for all four reflections. The set of parameters expressed relative to the new origin constituted the initial set of three-dimensional parameters used in the subsequent calculations. The hand computed structure factors and all of the intermediate calculations were also used during the debugging of the structure factor least-squares program for the Burroughs 220 computer described in Part II of this thesis and also in appendix I.

The refinement of the parameters progressed in stages, being dependent on the removal of program bugs from each of the major least-squares sections. The refinement, however, still followed the usual course, in that the positional parameters and the single isotropic temperature factor converged ahead of the anisotropic temperature factors. Because of the program debugging no definite figure can be given for the number of cycles that were needed for the complete refinement of the parameters.

Initially the layer-line scale factors were adjusted after each refinement cycle. This feature was soon removed, the individual layer-line scale factors being adjusted only a few times during the subsequent refinement. The least-squares refinement was started using unit weights for all reflections. This was later changed to $\sqrt{w} = \sqrt{w'}\,(0.2 + \sin\theta)/(1 + 0.01F_o)$, which was again changed to $\sqrt{w} = \sqrt{w'}/(0.48 + F_o)$, where $\sqrt{w'}$ is an external weight assigned to each reflection. The quantity being minimized was $\sum_q w_q(F_o^2 - F_c^2)_q^2$, and this last weighting scheme presumably reflects the pattern of reliability of the observed F's.

### iv) Hydrogen atom refinement.

Coordinates were assigned to the amino hydrogen atoms on the basis of a planar arrangement with distances and angles found in urea by Worsham, Levy and Peterson (21). The hydrogen atoms of the methyl group were placed in tetrahedral positions about the carbon atom with C-H distances of 1.0 Å and with two of the hydrogen atoms staggered with respect to the oxygen atom. Although the inclusion of the hydrogen atom contributions improved the agreement between calculated and observed structure factors, the improvement was less than had been anticipated, raising doubts as to the correctness of the assumed positions. Accordingly, an electron density map and a difference map were computed in the plane of the heavy atoms; these maps are shown in figures 11 and 12.

The difference map clearly showed the hydrogen atoms attached to the nitrogens. In the case of the methyl group, only one of the three hydrogens ($H_4$) should have appeared, because the other two are considerably out of the plane of the heavy atoms. However, as can be seen in figure 12, a second peak, labeled $H_{5,6}$, appeared. In order to find out what was going on around the methyl group, another difference map was computed, this time in the plane of the methyl hydrogen atoms. In this difference map as in the previous one the hydrogen atom contributions were included in the phase angles of the structure factors but were omitted from the amplitudes. The resulting map is shown in figure 13a. The assumed positions of the hydrogen atoms are shown by the small crosses. As can be seen, the peak heights associated with $H_5$ and $H_6$ are considerably less than that for $H_4$, and an additional peak of the same height as $H_5$ appeared opposite $H_4$.

In order to ascertain the effect of the inclusion of the hydrogen contributions to the phase angles of $F_{cal}$, another difference map was calculated with the hydrogen contributions to $F_{cal}$ omitted altogether. This synthesis is shown in figure 13b. The height of the peak opposite $H_4$ increased a little, while the height of the peak at $H_4$ decreased considerably. This map seems to indicate that the methyl group is not preferentially oriented. On the basis of this indication, it was decided to approximate the rotational torus of electron density by twelve quarter hydrogen atoms; that is, by twelve scattering centers each having 1/4 the scattering power of a hydrogen atom. The assumed coordinates of these fractional hydrogen atoms were computed in a manner analogous to the corresponding calculation for N,N'-dimethylurea.

A few more cycles of least-squares were run in which the methyl hydrogens were not allowed to move, using the weighting function $\sqrt{w} = \sqrt{w'}/(0.4 + 0.4764F_o + 0.1236F_o^2)$, which appropriately reflects the pattern of reliability of the observed $F^2$'s. Some improvement was obtained. The final R factor was 0.101. The R factors for each of

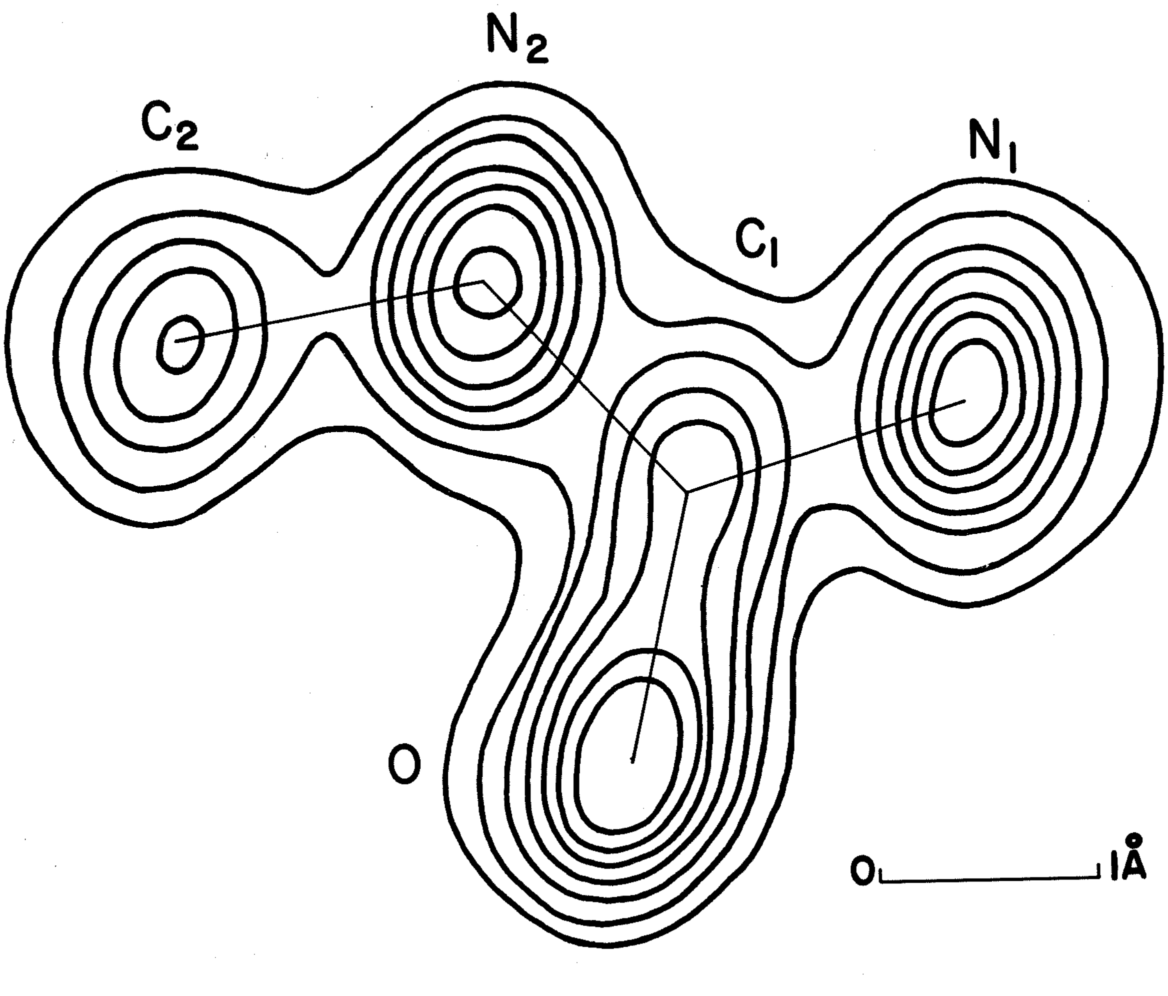


Figure 11. - The electron density section in the least-squares plane of N-methylurea. Contours at intervals of 1 e. Å$^{-3}$, beginning with 1 e. Å$^{-3}$.

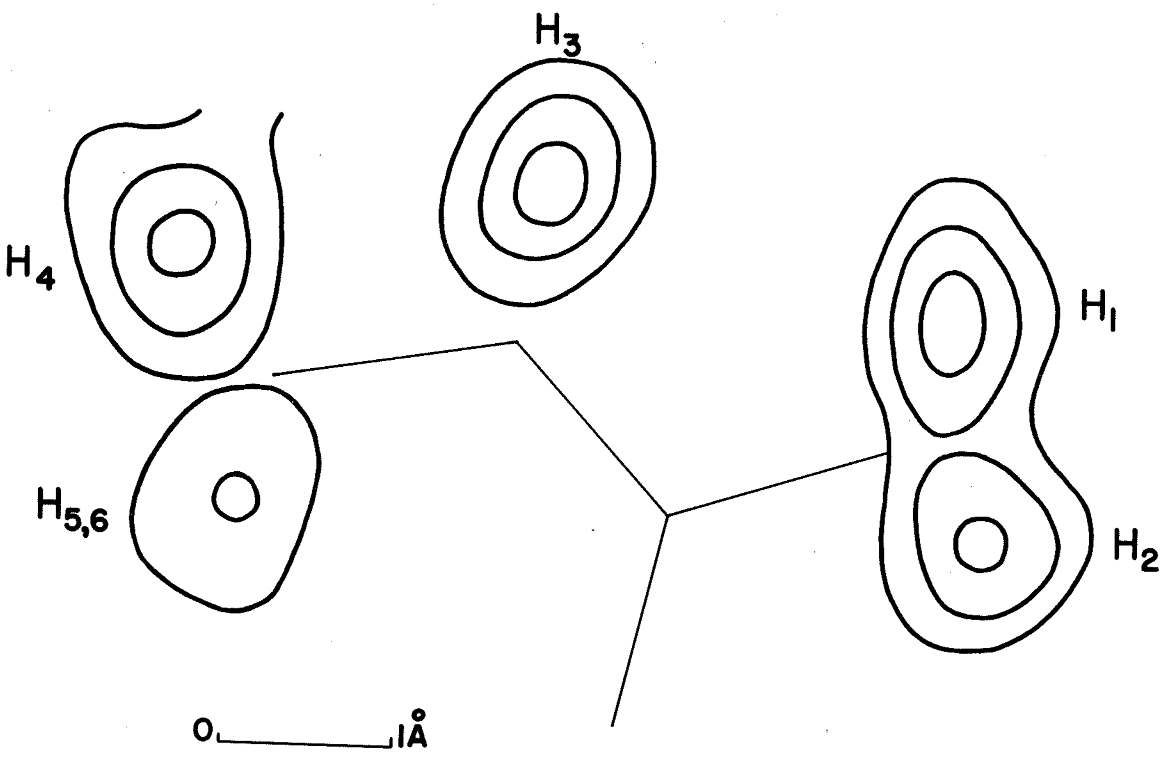


Figure 12. - The difference Fourier in the best plane of N-methylurea. Contours at intervals of 0.2 e. Å$^{-3}$, beginning with 0.2 e. Å$^{-3}$.

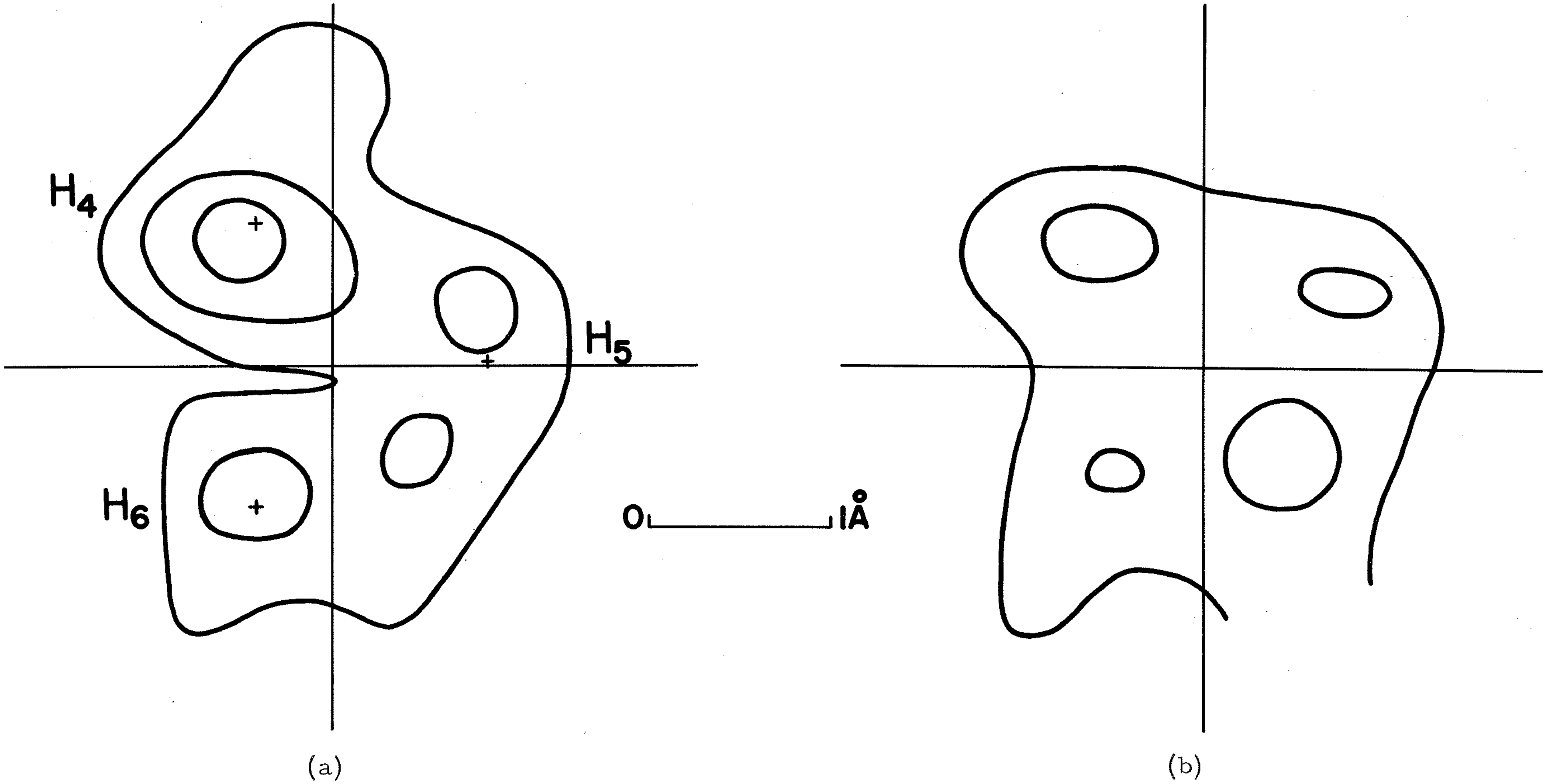


Figure 13.- Difference Fourier in the plane of the hydrogen atoms of the methyl group in N-methylurea. (a) Hydrogen atom contributions included in the phase angles of the structure factors but omitted from the amplitudes. (b) Hydrogen contributions to $F_{cal}$ omitted altogether. Contours at 0.2 e. $Å^{-3}$, beginning with 0.2 e. $Å^{-3}$.

the layers are: (O-layer) 0.071, (1) 0.097, (2) 0.073, (3) 0.140 and (4) 0.198. It did not seem that much more information could be obtained from the present set of data. The lack of resolution of $C_1$ in the electron density map shown in figure 11 is due to the fact that the resolution is poor along the z-axis since only layer lines 0-4 were available; the layer scaling constants were not determined experimentally except in a rough form which further complicated the refinement. Before any further work is done with this compound , it is felt that a complete three-dimensional set of correlated data should be collected.

Table 20 lists the final set of positional parameters and their standard errors. Table 21 lists the assumed coordinates for the methyl hydrogen atoms and table 22 gives the final structure factor list. The positional parameters in Angstrom units are listed in table 23. These were used to compute interatomic distances, bond angles and the best least-squares plane. The interatomic distances and angles are given in table 24. The equation for the least-squares plane is $0.3556X - 0.8080Y - 0.4698Z + 0.8292 = 0$. Further discussion of the structure will be given in a later section.

Table 20. - Final positional parameters and their standard errors for N-methylurea.

| Atom | x | y | z | σ(x) | σ(y) | σ(z) |
|---|---|---|---|---|---|---|
| O | 0.6842 | 0.1273 | 0.4472 | 0.0008 | 0.0008 | 0.0011 |
| $N_1$ | 0.8927 | 0.0769 | 0.6452 | 0.0010 | 0.0011 | 0.0015 |
| $N_2$ | 0.5958 | -0.0565 | 0.6560 | 0.0011 | 0.0011 | 0.0016 |
| $C_1$ | 0.7198 | 0.0581 | 0.5831 | 0.0011 | 9.0010 | 0.0018 |
| $C_2$ | 0.4051 | -0.0989 | 0.5967 | 0.0013 | 0.0014 | 0.0020 |
| $H_1$ | 0.915 | 0.035 | 0.773 | 0.008 | 0.008 | 0.010 |
| $H_2$ | 0.974 | 0.148 | 0.624 | 0.008 | 0.009 | 0.010 |
| $H_3$ | 0.622 | -0.118 | 0.735 | 0.008 | 0.008 | 0.009 |

Table 21. - Coordinates* assigned to methyl hydrogen atoms.

| | x | y | z |
|---|---|---|---|
| $H_4$ | 0.396 | -0.146 | 0.698 |
| $H_5$ | 0.410 | -0.076 | 0.698 |
| $H_6$ | 0.411 | -0.014 | 0.670 |
| $H_7$ | 0.399 | 0.025 | 0.622 |
| $H_8$ | 0.377 | 0.028 | 0.567 |
| $H_9$ | 0.352 | -0.002 | 0.519 |
| $H_{10}$ | 0.329 | -0.061 | 0.492 |
| $H_{11}$ | 0.315 | -0.131 | 0.492 |
| $H_{12}$ | 0.314 | -0.193 | 0.519 |
| $H_{13}$ | 0.326 | -0.231 | 0.567 |
| $H_{14}$ | 0.348 | -0.235 | 0.622 |
| $H_{15}$ | 0.373 | -0.203 | 0.670 |

*The methyl group is believed to be disordered and is represented here by 12 quarter hydrogen atoms.

Table 22. - Observed and calculated structure factors for N-methylurea. The five columns in each group contain the values, reading from left to right, of k, $10|F_o|$, $10|F_c|$, $10A_c$ and $10B_c$. Reflections indicated by an asterisk (*) were given zero weight in the final least-squares calculation.

| | | | | |
|---|---|---|---|---|
| | | 0k0 | | |
| 0 | ---* | 1602 | 1602 | 0 |
| 2 | 380* | 448 | 448 | 0 |
| 4 | 170 | 164 | -164 | 0 |
| 6 | 73 | 71 | -71 | 0 |
| 8 | 16 | 15 | -15 | 0 |
| | | 1k0 | | |
| 1 | 380* | 506 | 0 | -506 |
| 2 | 259 | 274 | 0 | 274 |
| 3 | 9 | 9 | 0 | 9 |
| 4 | 138 | 133 | 0 | 133 |
| 5 | 65 | 64 | 0 | 64 |
| 6 | 26 | 25 | 0 | 25 |
| 7 | 11 | 15 | 0 | 15 |
| 8 | 8 | 5 | 0 | -5 |
| | | 2k0 | | |
| 0 | 107 | 105 | -107 | 0 |
| 1 | 69 | 69 | -69 | 0 |
| 2 | 9 | 4 | -4 | 0 |
| 3 | 38 | 43 | 43 | 0 |
| 4 | 37 | 32 | 32 | 0 |
| 5 | 75 | 73 | 73 | 0 |
| 6 | <2 | 7 | 7 | 0 |
| 7 | 16 | 18 | 18 | 0 |
| 8 | <2 | 3 | 3 | 0 |
| | | 3k0 | | |
| 1 | 45 | 41 | 0 | -41 |
| 2 | 66 | 71 | 0 | 71 |
| 3 | 58 | 52 | 0 | -52 |
| 4 | 21 | 22 | 0 | -22 |
| 5 | 9 | 11 | 0 | -11 |
| 6 | 17 | 20 | 0 | -20 |
| 7 | 6 | 7 | 0 | 7 |
| 8 | 9 | 7 | 0 | 7 |
| | | 4k0 | | |
| 0 | 183 | 178 | -178 | 0 |
| 1 | 92 | 101 | 101 | 0 |
| 2 | 58 | 64 | -64 | 0 |
| 3 | 105 | 100 | 100 | 0 |
| 4 | 26 | 24 | 24 | 0 |
| 5 | 18 | 28 | 28 | 0 |
| 6 | 20 | 18 | 18 | 0 |
| 7 | 9 | 9 | 9 | 0 |
| | | 5k0 | | |
| 1 | 6 | 2 | 0 | -2 |
| 2 | 153 | 156 | 0 | -156 |
| 3 | 42 | 38 | 0 | -38 |
| 4 | 80 | 81 | 0 | -81 |
| 5 | 12 | 12 | 0 | -12 |
| 6 | 7 | 8 | 0 | -8 |
| 7 | 12 | 11 | 0 | 11 |
| | | 6k0 | | |
| 0 | 67 | 59 | -59 | 0 |
| 1 | 37 | 38 | -38 | 0 |
| 2 | 50 | 47 | -47 | 0 |
| 3 | 47 | 49 | -49 | 0 |
| 4 | 7 | 6 | -6 | 0 |
| 5 | 9 | 11 | -11 | 0 |
| 6 | 26 | 14 | 14 | 0 |
| | | 7k0 | | |
| 1 | 15 | 15 | 0 | 15 |
| 2 | <2 | 5 | 0 | 5 |
| 3 | 20 | 18 | 0 | 18 |
| | | 8k0 | | |
| 0 | <2 | 4 | 4 | 0 |
| 1 | 6 | 0 | 0 | 0 |
| | | 0k1 | | |
| 1 | 178* | 151 | 0 | -151 |
| 2 | 95 | 86 | -86 | 0 |
| 3 | 134 | 123 | 0 | -123 |
| 4 | <3 | 2 | -2 | 0 |
| 5 | 16 | 17 | 0 | 17 |
| 6 | 13 | 15 | 15 | 0 |
| 7 | <3 | 1 | 0 | -1 |
| 8 | <3 | 2 | 2 | 0 |
| | | 1k1 | | |
| 0 | 96 | 92 | 0 | 92 |
| 1 | 198 | 206 | 204 | -31 |
| 2 | 306 | 341 | -336 | 59 |
| 3 | 49 | 48 | -31 | -37 |
| 4 | 70 | 65 | -61 | 21 |
| 5 | 41 | 42 | -27 | -31 |
| 6 | 16 | 20 | 18 | -9 |
| 7 | 3 | 4 | 1 | -4 |
| 8 | 5 | 7 | 2 | -6 |
| | | 2k1 | | |
| 0 | 32 | 30 | 0 | -30 |
| 1 | 156 | 164 | -89 | 138 |
| 2 | <2 | 9 | -9 | 1 |
| 3 | 138 | 139 | 50 | 129 |
| 4 | 40 | 43 | -37 | 21 |
| 5 | 39 | 30 | 29 | 7 |
| 6 | 25 | 22 | -20 | -9 |
| 7 | <3 | 3 | -3 | -2 |
| 8 | 4 | 2 | -2 | 1 |
| | | 3k1 | | |
| 0 | 10 | 1 | 0 | -1 |
| 1 | 143 | 152 | -147 | -36 |
| 2 | 19 | 20 | 13 | -16 |
| 3 | 38 | 47 | 45 | -14 |
| 4 | 31 | 35 | 25 | -25 |
| 5 | 33 | 33 | 24 | 23 |
| 6 | 20 | 20 | 19 | 2 |
| 7 | 20 | 16 | -2 | 16 |
| 8 | 18 | 9 | 6 | 7 |
| | | 4k1 | | |
| 0 | 170 | 156 | 0 | 156 |
| 1 | 23 | 26 | -12 | -23 |
| 2 | 30 | 35 | 28 | 21 |
| 3 | 36 | 34 | -7 | -34 |
| 4 | 57 | 59 | 17 | -56 |
| 5 | 25 | 21 | 3 | -21 |
| 6 | 14 | 12 | 10 | -7 |
| 7 | 11 | 5 | 4 | -3 |
| | | 5k1 | | |
| 0 | 20 | 22 | 0 | -22 |
| 1 | 40 | 40 | 40 | -8 |
| 2 | 5 | 8 | 8 | 0 |
| 3 | 19 | 13 | -11 | -7 |
| 4 | 12 | 11 | -7 | 9 |
| 5 | 7 | 10 | - 9 | 4 |
| 6 | 11 | 5 | -5 | -1 |
| 7 | 3 | 1 | -1 | 0 |
| | | 6k1 | | |
| 0 | 56 | 56 | 0 | -56 |
| 1 | 28 | 32 | 22 | -23 |
| 2 | 21 | 25 | -2 | -24 |
| 3 | 16 | 16 | 9 | -14 |
| 4 | 10 | 9 | 0 | 9 |
| 5 | 17 | 15 | -5 | 14 |
| 6 | 10 | 11 | 0 | 11 |
| | | 7k1 | | |
| 0 | 14 | 15 | 0 | -15 |
| 1 | 30 | 26 | -23 | 12 |
| 2 | 9 | 7 | -2 | -7 |
| 3 | 15 | 15 | 2 | 15 |
| 4 | 14 | 10 | 10 | 2 |
| | | 8k1 | | |
| 0 | <3 | 2 | 0 | -2 |
| 1 | 5 | 2 | 1 | 2 |
| 2 | 11 | 8 | 7 | 3 |
| 3 | 8 | 3 | -3 | 2 |
| | | 0k2 | | |
| 0 | 148* | 165 | 165 | 0 |
| 1 | 397* | 478 | 0 | 478 |
| 2 | 45 | 40 | -40 | 0 |
| 3 | 131 | 121 | 0 | 121 |
| 4 | 77 | 76 | -76 | 0 |
| 5 | 40 | 38 | 0 | -38 |
| 7 | 50 | 47 | 0 | -47 |
| 8 | 8 | 8 | 8 | 0 |
| | | 1k2 | | |
| 0 | 113 | 102 | 102 | 0 |
| 1 | 214 | 224 | -215 | -61 |
| 2 | 172 | 173 | 140 | -102 |
| 3 | 200 | 194 | -187 | 51 |
| 4 | 73 | 63 | 52 | -36 |
| 5 | 37 | 33 | -28 | 18 |
| 6 | 22 | 22 | -21 | -5 |
| 7 | 10 | 8 | 7 | -4 |
| 8 | 26 | 23 | -23 | -2 |
| | | 2k2 | | |
| 0 | 167 | 180 | -180 | 0 |
| 1 | 152 | 167 | -152 | 71 |
| 2 | 138 | 137 | -46 | 129 |
| 3 | 128 | 125 | -114 | -52 |
| 4 | 57 | 58 | 28 | 51 |
| 5 | 20 | 23 | 8 | -22 |
| 6 | 7 | 10 | 9 | 2 |
| 7 | 11 | 14 | 7 | 12 |
| 8 | <3 | 4 | 3 | 2 |
| | | 3k2 | | |
| 0 | 86 | 71 | 71 | 0 |
| 1 | 148 | 148 | 71 | 130 |
| 2 | 108 | 110 | 33 | 105 |
| 3 | 51 | 52 | 52 | -6 |
| 4 | 13 | 13 | -5 | 12 |
| 5 | 23 | 27 | -7 | -26 |
| 6 | 12 | 14 | -4 | -13 |
| 7 | 14 | 12 | -11 | -4 |
| 8 | <2 | 2 | -1 | 2 |
| | | 4k2 | | |
| 0 | 53 | 53 | 53 | 0 |
| 1 | 83 | 91 | 38 | -83 |
| 2 | 29 | 28 | 28 | -5 |
| 3 | 17 | 22 | 19 | -12 |
| 4 | 34 | 38 | -3 | 38 |
| 5 | 25 | 23 | -11 | 21 |
| 6 | 32 | 34 | -9 | 32 |
| 7 | 10 | 11 | -3 | 11 |
| | | 5k2 | | |
| 0 | 36 | 30 | 30 | 0 |
| 1 | 33 | 33 | 33 | 3 |
| 2 | 34 | 25 | 12 | -22 |
| 3 | 73 | 67 | 66 | -12 |
| 4 | 6 | 5 | 3 | 4 |
| 5 | 45 | 40 | 40 | 2 |
| 6 | 12 | 10 | -1 | 10 |
| 7 | 10 | 8 | 6 | 6 |
| | | 6k2 | | |
| 0 | 24 | 24 | 24 | 0 |
| 1 | 70 | 62 | -9 | -61 |
| 2 | 17 | 17 | 5 | -17 |
| 3 | 6 | 7 | -7 | -3 |
| 4 | 23 | 21 | -4 | -20 |
| 5 | 23 | 15 | 0 | 15 |
| 6 | 12 | 9 | 1 | -9 |
| | | 7k2 | | |
| 0 | 36 | 31 | -31 | 0 |
| 1 | 22 | 17 | 0 | -17 |
| 2 | 17 | 14 | -14 | 4 |
| 3 | 9 | 7 | -2 | 7 |
| 4 | 12 | 9 | 5 | 8 |
| 5 | 10 | 8 | -5 | 6 |
| | | 8k2 | | |
| 0 | 14 | 14 | -14 | 0 |
| 1 | 16 | 14 | 4 | 14 |
| 2 | 2 | 1 | 0 | 1 |
| 3 | 12 | 8 | 8 | -1 |
| | | 0k3 | | |
| 1 | 75 | 56 | 0 | -56 |
| 2 | 111 | 92 | -92 | 0 |
| 3 | 53 | 47 | 0 | -47 |
| 4 | 34 | 33 | 33 | 0 |
| 5 | ---* | 8 | 0 | -8 |
| 6 | 43 | 49 | 49 | 0 |
| 7 | 5 | 5 | 0 | -5 |
| 8 | 8 | 8 | 8 | 0 |
| | | 1k3 | | |
| 0 | 9 | 1 | 0 | 1 |
| 1 | 13 | 18 | -11 | 14 |
| 2 | 54 | 55 | -51 | 20 |
| 3 | 61 | 62 | -58 | -22 |
| 4 | 10 | 15 | 10 | 11 |
| 5 | 53 | 53 | -13 | -52 |
| 6 | 9 | 14 | 9 | -11 |
| 7 | 15 | 15 | 10 | -12 |
| 8 | 6 | 5 | -4 | -3 |
| | | 2k3 | | |
| 0 | 99 | 74 | 0 | -74 |
| 1 | 108 | 104 | -96 | 39 |
| 2 | 15 | 18 | 11 | 14 |
| 3 | 52 | 53 | 52 | 11 |
| 4 | 65 | 55 | -36 | 42 |
| 5 | 31 | 37 | 28 | -24 |
| 6 | 22 | 24 | -23 | 8 |
| 7 | 8 | 11 | -9 | -7 |
| 8 | 5 | 3 | -1 | -3 |
| | | 3k3 | | |
| 0 | 36 | 39 | 0 | 39 |
| 1 | 47 | 46 | -35 | -30 |
| 2 | 32 | 37 | 33 | -17 |
| 3 | 30 | 30 | 27 | -13 |
| 4 | 34 | 36 | 0 | -36 |
| 5 | 26 | 27 | 13 | 23 |
| 6 | 12 | 14 | -13 | 7 |
| 7 | 14 | 14 | -5 | 14 |
| 8 | 9 | 13 | -7 | 10 |
| | | 4k3 | | |
| 0 | 52 | 38 | 0 | 38 |
| 1 | 9 | 7 | -6 | 1 |
| 2 | 30 | 30 | 30 | 3 |
| 3 | 22 | 26 | -26 | -1 |
| 4 | 22 | 19 | 16 | -12 |
| 5 | 5 | 8 | -2 | 8 |
| 6 | 8 | 4 | 3 | 3 |
| 7 | 14 | 12 | 9 | 8 |
| | | 5k3 | | |
| 0 | 22 | 20 | 0 | -20 |
| 1 | 28 | 30 | 23 | -20 |
| 2 | 14 | 12 | -2 | 12 |
| 3 | 18 | 18 | -3 | -18 |
| 4 | 24 | 25 | -7 | 24 |
| 5 | <3 | 3 | -1 | 3 |
| 6 | 4 | 2 | -2 | -1 |
| | | 6k3 | | |
| 0 | 3 | 7 | 0 | 7 |
| 1 | 18 | 16 | 15 | -6 |
| 2 | 9 | 11 | -9 | 5 |
| 3 | 15 | 15 | 15 | 0 |
| 4 | 2 | 4 | -4 | 2 |
| 5 | 10 | 7 | -1 | 7 |
| 6 | 3 | 4 | -3 | -3 |
| | | 7k3 | | |
| 0 | 17 | 14 | 0 | -14 |
| 1 | 15 | 12 | 2 | 12 |
| 2 | 18 | 15 | -14 | -6 |
| 3 | 23 | 15 | 4 | 14 |
| | | 8k3 | | |
| 0 | 6 | 3 | 0 | 3 |
| 1 | 7 | 5 | -2 | 4 |
| 2 | 9 | 8 | 8 | 1 |
| 3 | 4 | 5 | -1 | 5 |
| | | 0k4 | | |
| 0 | 130* | 148 | -148 | 0 |
| 1 | 79* | 85 | 0 | -85 |
| 2 | 83 | 94 | -94 | 0 |
| 3 | 46 | 44 | 0 | 44 |
| 4 | <3 | 6 | -6 | 0 |
| 5 | 18 | 20 | 0 | 20 |
| 6 | 25 | 30 | 30 | 0 |
| 7 | 10 | 10 | 0 | -10 |
| | | 1k4 | | |
| 0 | 101* | 114 | -114 | 0 |
| 1 | >55* | 50 | -11 | 49 |
| 2 | 70 | 72 | 9 | -71 |
| 3 | 51 | 39 | 15 | 36 |
| 4 | 64 | 65 | 54 | -36 |
| 5 | 32 | 29 | 29 | -1 |
| 6 | 4 | 5 | -2 | -5 |
| 7 | 17 | 13 | 8 | -11 |
| | | 2k4 | | |
| 0 | 36* | 38 | -38 | 0 |
| 1 | 43 | 28 | -28 | -2 |
| 2 | 26 | 12 | 5 | -11 |
| 3 | 74 | 53 | -30 | -44 |
| 4 | 49 | 37 | 18 | -33 |
| 5 | 14 | 11 | -7 | -9 |
| 6 | 21 | 11 | 0 | -11 |
| 7 | 19 | 16 | -4 | 15 |
| | | 3k4 | | |
| 0 | 85* | 64 | -64 | 0 |
| 1 | 34 | 35 | 25 | 24 |
| 2 | 23 | 39 | -35 | 17 |
| 3 | 3 | 3 | 2 | -2 |
| 4 | <7 | 9 | 7 | 6 |
| 5 | 24 | 23 | -23 | -4 |
| 6 | 18 | 13 | 13 | -1 |
| 7 | 9 | 7 | -7 | -1 |
| | | 4k4 | | |
| 0 | 36 | 37 | 37 | 0 |
| 1 | 42 | 39 | -1 | 39 |
| 2 | 57 | 49 | 20 | -45 |
| 3 | 17 | 12 | -11 | 3 |
| 4 | 6 | 5 | -5 | 2 |
| 5 | 20 | 16 | -14 | -7 |
| 6 | 19 | 20 | -12 | 16 |
| 7 | 9 | 8 | -5 | -6 |
| | | 5k4 | | |
| 0 | 19 | 23 | 23 | 0 |
| 1 | 15 | 20 | -19 | 6 |
| 2 | 27 | 22 | 5 | 21 |
| 3 | 15 | 13 | -13 | -1 |
| 4 | 34 | 29 | -7 | 28 |
| 5 | 7 | 6 | 5 | 2 |
| 6 | 11 | 11 | -3 | 10 |
| | | 6k4 | | |
| 0 | 14 | 25 | 25 | 0 |
| 1 | 7 | 8 | -1 | -8 |
| 2 | 24 | 17 | 13 | 10 |
| 3 | 18 | 13 | 7 | 11 |
| 4 | 8 | 2 | -2 | -1 |
| 5 | 10 | 10 | 9 | 5 |
| | | 7k4 | | |
| 1 | 15 | 9 | 2 | -9 |
| 2 | 8 | 4 | 4 | 1 |
| 3 | 6 | 3 | -2 | -3 |
| 4 | 4 | 2 | 2 | -1 |
| | | 8k4 | | |
| 1 | <2 | 1 | 1 | 0 |
| 2 | 9 | 9 | -3 | 9 |

Table 23. - Final coordinates given in Ångstrom units.

| Atom | X | Y | Z |
|---|---|---|---|
| O | 4.7140 | 0.8863 | 3.7792 |
| $N_1$ | 6.1504 | 0.5352 | 5.4517 |
| $N_2$ | 4.1052 | -0.3932 | 5.5434 |
| $C_1$ | 4.9591 | 0.4047 | 4.9274 |
| $C_2$ | 2.7910 | -0.6884 | 5.0419 |
| $H_1$ | 6.3067 | 0.2460 | 6.5314 |
| $H_2$ | 6.7086 | 1.0310 | 5.2747 |
| $H_3$ | 4.2882 | -0.8185 | 0.2074 |

The normalization equations are:

$X = 6.89\ x$

$Y = 6.96\ y$

$Z = 8.45\ z$

The inverse equations are:

$x = 0.14514\ X$

$y = 0.14368\ Y$

$z = 0.11834\ Z$

Table 24. - Interatomic distances and angles.

| | | | |
|---|---|---|---|
| $C_1$-O | 1.269 Å | O-$C_1$-$N_1$ | 120.0 |
| | | O-$C_1$-$N_2$ | 121.8 |
| $C_1$-$N_1$ | 1.308 | $N_1$-$C_1$-$N_2$ | 117.5 |
| $C_1$-$N_2$ | 1.321 | $C_1$-$N_2$-$C_2$ | 123.5 |
| $C_2$-$N_2$ | 1.437 | | |
| | | $C_1$-$N_2$-$H_3$ | 123.6 |
| $N_1$-$H_1$ | 1.13 | $C_2$-$N_2$-$H_3$ | 112.6 |
| $N_1$-$H_2$ | 0.77 | $C_1$-$N_1$-$H_1$ | 118.9 |
| $N_1$-$H_3$ | 0.81 | $C_1$-$N_1$-$H_2$ | 129.4 |
| | | $H_1$-$N_1$-$H_2$ | 106.6 |

Hydrogen Bonds:

| | | | | |
|---|---|---|---|---|
| $N_1$-$H_1$···O' | 2.980 | $C_1$-$N_1$···N' | 125.3 | Between molecules in adjacent chain |
| $N_1$-$H_2$···O'' | 2.969 | $C_1$-O···N'' | 129.0 | Between molecules in the same chain |
| $N_2$-$H_3$···O'' | 2.932 | O···$H''_1$-$N''_1$ | 174.1 | |

v) Temperature parameter discussion.

Table 25 lists the final anisotropic temperature parameters. The magnitudes and direction cosines of the principal axes of the vibration ellipsoids are given in table 26. The direction cosines are given relative to a Cartesian coordinate system coincident with the crystal axes.

Table 27 not only fails to substantiate the expectation of a predominantly out-of-plane libration but also indicates that something has gone badly astray. It is clear from table 26 that the magnitudes of the atomic displacements along the major axes of the heavy atom ellipsoids are too large. The major axis of each atom has oriented itself nearly parallel to the z-axis, as shown by the large values of $g_1(3)$.

It seemed likely that the individual layer line scale factors had run amuck, leading to overly large $B_{33}$ temperature fluctuations. A constant value of 0.02 was subtracted from the $B_{33}$ term of each atom and a new calculation was made of the magnitudes and direction cosines of the principal axes of the vibration ellipsoids. The major axes obtained from the new calculation were more nearly perpendicular to the least-squares plane as originally expected. However, the values of the atomic displacements along the semi-major and minor axes in a few cases became ridiculously small.

The difficulty cannot be blamed wholly on the individual layer line scale factors, because of the apparent involvement of the other

Table 25. - Final anisotropic temperature parameters. The temperature factors are in the form of:

$$T_i = \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}kl) \ .$$

| Atom | $B_{11}$ | $B_{22}$ | $B_{33}$ | $B_{12}$ | $B_{13}$ | $B_{23}$ |
|---|---|---|---|---|---|---|
| O | 0.0325 | 0.0287 | 0.0348 | -0.0038 | 0.0043 | 0.0094 |
| $N_1$ | 0.0218 | 0.0370 | 0.0434 | -0.0071 | -0.0053 | 0.0227 |
| $N_2$ | 0.0312 | 0.0281 | 0.0492 | -0.0036 | -0.0035 | 0.0085 |
| $C_1$ | 0.0279 | 0.0217 | 0.0469 | -0.0008 | 0.0047 | -0.0121 |
| $C_2$ | 0.0283 | 0.0361 | 0.0610 | -0.0216 | -0.0201 | 0.0377 |

Table 26. - Magnitudes and direction cosines of the principal axes of the vibration ellipsoids.

| Atom | Axis i | Bi | $g_i^{(1)}$ | $g_i^{(2)}$ | $g_i^{(3)}$ |
|---|---|---|---|---|---|
| O | 1 | 10.944 | 0.136 | 0.361 | 0.923 |
| | 2 | 6.635 | 0.875 | -0.481 | 0.059 |
| | 3 | 4.090 | 0.465 | 0.799 | -0.381 |
| $N_1$ | 1 | 15.993 | -0.147 | 0.529 | 0.836 |
| | 2 | 4.356 | 0.717 | -0.525 | 0.459 |
| | 3 | 3.368 | 0.682 | 0.666 | -0.302 |
| $N_2$ | 1 | 14.603 | -0.107 | 0.220 | 0.969 |
| | 2 | 6.060 | 0.899 | -0.395 | 0.189 |
| | 3 | 4.777 | 0.424 | 0.892 | -0.155 |
| $C_1$ | 1 | 14.336 | 0.119 | -0.270 | 0.955 |
| | 2 | 5.174 | 0.989 | 0.115 | -0.091 |
| | 3 | 3.389 | -0.085 | 0.956 | 0.281 |
| $C_2$ | 1 | 24.462 | -0.306 | 0.488 | 0.817 |
| | 2 | 4.094 | 0.828 | -0.287 | 0.481 |
| | 3 | 1.245 | 0.469 | 0.824 | -0.316 |

Table 27.

a. Angles between the thermal ellipsoids axes and the least-squares plane normal.

| Atom | major | semi-major | minor |
|---|---|---|---|
| O | 47.38 | 48.15 | 72.47 |
| $N_1$ | 29.25 | 62.36 | 81.13 |
| $N_2$ | 47.80 | 56.66 | 60.22 |
| $C_1$ | 79.16 | 72.46 | 20.81 |
| $C_2$ | 27.50 | 72.53 | 69.47 |

b. Angles between the thermal ellipsoid axes of the methyl group $C_2$ and the $N_2$-$C_2$ bond direction.

| Direction cosines of $N_2$-$C_2$ bond direction | Ellipsoid Axis | Acute Angle |
|---|---|---|
| $g(1) = 0.9144$ | major | 83.95 |
| $g(2) = 0.2054$ | semi-major | 29.97 |
| $g(3) = 0.3489$ | minor | 60.77 |

$B_{ij}$ terms. It is very probable that the lack of high order resolution along the c-axis is primarily responsible for the high uncertainties of the $B_{33}$, $B_{13}$ and $B_{23}$ terms. Although the lack of reflections with an $\ell$ index greater than 4 did not seem to seriously affect the refinement of the positional parameters, it certainly appears to have affected the anisotropic temperature parameter refinement.

## General Discussion of the Structural Results.

i) Intra-molecular features:

(a) Bond distances.

In table 28 are listed the observed bond lengths in urea and in several other closely related compounds. The table is divided into two parts; the first part lists the results of research included in this thesis and the published results of other compounds which are both closely related and accurately determined. The second group consists of a supplemental list of results of two types: (1) Those which were not thought to be of sufficient accuracy to be included in the first group. (2) Those from compounds which are not so closely related to urea as desired.

The compounds added to the first group are: Urea (x-ray diffraction, Vaughan and Donohue, 22; neutron diffraction, Worshan, Levy and Peterson 21) and biuret hydrate (Hughes, Yakel and Freeman, 23). Included in the second group are the following compounds: leucylprolylglycine (Leung and Marsh, 24), bis-biuret cadmium chloride (Cavalca, Nardelli and Fava, 25), potassium bis-biuret cuprate(II) tetrahydrate (Freeman, Smith and Taylor, 26), diacetylhydrazine (Shintani, 27), N-methyl acetamide (Katz and Post, 28), N-methylurea nitrate (Bryden, 29) and bis-(N-methylurea) cadmium chloride (Nardelli, Coghi and Azzoni, 30).

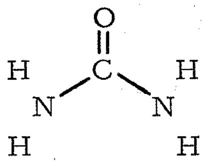
Table 28. - Bond lengths in urea and related compounds.

| I. Compound | C=O | $C_1$-N | $C_2$-N |
|---|---|---|---|
| Urea (x-ray) | 1.262 | 1.335 | |
| Urea (neutron) | 1.243 | 1.351 | |
| N-methylurea | 1.269 | 1.308 | 1.437 |
| | | 1.321 | |
| N,N'-dimethylurea | 1.262 | 1.318 | 1.449 |
| N,N-dimethylurea | 1.234 | 1.351 | 1.456 |
| | | 1.351 | 1.448 |
| Biuret hydrate | 1.255 | 1.332 | |
| | 1.246 | 1.361 | |
| Average of bond lengths: | 1.253 | 1.336 | 1.447 |
| Proposed by Pauling and Corey | 1.24 | 1.32 | 1.47 |
| | | | |
| II. Supplementary Compounds | | | |
| Leucylprolylglycine | 1.243 | 1.335 | 1.467 |
| | 1.207 | 1.331 | 1.457 |
| | | | 1.482 |
| $CdCl_2.(Biuret)_2$ | 1.26 | 1.35 | |
| | 1.23 | 1.34 | |
| | | 1.35 | |
| | | 1.35 | |
| $K_2$ $Cu(NHCONHCONH)_2.4H_2O$ | 1.25 | 1.34 | |
| | 1.27 | 1.33 | |
| Diacetyl hydrazine | 1.221 | 1.341 | |
| N-methyl acetamide | 1.236 | 1.290 | 1.465 |
| N-methylurea.$HNO_3$ | 1.28 | 1.29 | 1.44 |
| | | 1.30 | |
| $CdCl_2$.N-methylurea | 1.27 | 1.39 | 1.45 |
| | | 1.36 | |
| Average of bond lengths: | 1.247 | 1.336 | 1.460 |

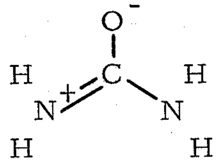
The bond lengths in N,N-dimethylurea are in excellent agreement with the neutron study of urea and the bond lengths in N-methylurea and N,N'-dimethylurea are in good agreement with the x-ray study of urea. Unfortunately the two sets of results are in only fair agreement with each other. In the opinion of this author, the discrepancies between the two sets of results should not be attributed to anything other than experimental uncertainties. The average values for the C=O, $C_1$-N and $C_2$-N bond lengths are 1.253 Å, 1.336 and 1.447 for the first group and 1.247, 1.336 and 1.460 for the second group. These averages are in good agreement with each other and with the values for the peptide linkage proposed by Pauling and Corey (31).

(b) Double bond character.

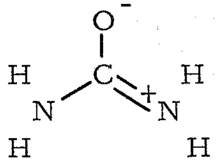
The values in table 28 show that the presence of the N-methyl groups does not change the interatomic distances within the urea nucleus by an amount greater than the apparent experimental error. Thus it appears that the relative contributions of the three resonance structures (a), (b) and (c) to the resonance hybrids are approximately the same for the methyl ureas as for urea itself.



(a)



(b)



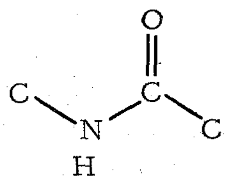
(c)

In order to estimate the relative contributions of these structures, the familiar equation of Pauling (32a) which relates bond length with bond number will be used:

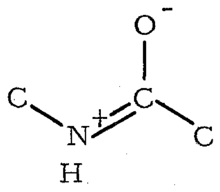
$$D_n = D_1 - (D_1 - D_2) \frac{1.84(n - 1)}{0.84n + 0.16}$$

It is first necessary to have reliable values for the lengths of carbon to oxygen and carbon to nitrogen single and double bonds. Satisfactory agreement with observed results is obtained if the values 1.166 for C=O and 1.240 for C=N are used, combined with the values 1.400 for C-O and 1.447 Å for C-N. The C=O value is the assumed pure C=O bond distance in $ClCOCl$, phosgene, reported by Robinson (33). This value is probably appropriate since not only is it accurate (the quoted standard deviation is 0.002 Å), but the electronegativity of chlorine is the same as that of nitrogen. The value 1.400 Å for C-O was also chosen so as to match the electronegativity of the environment; it is the value for the C-O bond in chloromethyl chloroformate reported by O'Gorman, Shand and Schomaker (34). The average of the $C_2$-N bond lengths, 1.447 Å, was taken as the average value of the C-N single bonds in the methyl and dimethyl ureas. The average observed C=O and $C_1$-N distances indicated that there is about a 40% contribution from structure (a), and that structures (b) and (c) contribute about 30% each to the resonance hybrid.

Table 29 gives a review of the double bond character of the C=O and $C_1$-N bonds in the peptide linkages and in the ureas. Corey and Pauling (31) estimated 60% and 40% contributions from the resonance structures (d) and (e) to the peptide linkage.



(d)



(e)

These values are compatible with the 40% and 30% values found for the ureas: The lower double bond character of the C=O bond in the ureas is explained by the additive contribution of both structures (b) and (c) to the resonance hybrid and hence to an additional lengthening of the urea C=O bond. The lower double bond character of the $C_1$-N bond distances in the ureas and their corresponding increased length is due to the competition between the electron pair donor nitrogen atoms for the acceptor group in structures (b) and (c); thus, the donor nitrogen atom in (e), having no competitor, makes a relatively larger contribution to the resonance hybrid of the peptide linkage than either (b) or (c) do in the ureas.

(c) Planarity of the molecules.

The contributions from the resonance structures (b) and (c) lead to the planarity of urea and its methyl derivatives. All three compounds investigated in this research have planar heavy atom configurations with an average deviation of an atom from the least-squares plane

Table 29. - Double bond character of the C=O and $C_1$-N bonds.

| Type of bond | C=O | $C_1$-N |
|---|---|---|
| Observed average distance | 1.253 | 1.336 |
| Proposed by Pauling and Corey (31) | 1.24(1.23)* | 1.32(1.33)* |
| Double bond character (peptide bond) | 60% | 40% |
| Double bond character: Urea (22) | 40% | 30% |
| Double bond character: This research | 40% | 30% |
| Calculated distances for urea | 1.27* | 1.36* |

*Calculated using the following values for the pure single and double bond lengths:

| | | | |
|---|---|---|---|
| C=O | 1.166 | C=N | 1.240 |
| C-O | 1.400 | C-N | 1.447 |

of 0.02 Å and a maximum deviation of 0.05 Å; the out-of-plane distances are listed in table 30.

(d) Configuration of the amide group.

Both N-methylurea and N,N'-dimethylurea have the trans-amide configuration discussed by Pauling, Corey and Branson (2); that is, the unsubstituted amide hydrogen atom is trans to the oxygen atom while the substituting group (in this case a methyl group) is cis. The preference of the amide group for the trans configuration, as shown by most linear peptides, was noted by Corey and Pauling (31). They listed only one exception to the trans configuration, namely, diketopiperazine (Corey 35) where the cis-configuration is demanded by the cyclic covalent bonding within the compound.

Another recently reported exception to the trans configuration is biuret; in this case the resultant configuration is not due to the geometry of the covalent bonding. An essentially planar cis-trans configuration was found for the two peptide linkages in biuret hydrate (Hughes, Yakel and Freeman 23) and also in bis-biuret cadmium chloride (Cavalca, Nardelli and Fava 25). (Two other possible planar configurations are the trans-trans and cis-cis arrangements.) It is the intra-molecular hydrogen bonding between one of the two $NH_2$ end groups and the oxygen atom of the second carbonyl group away from it which has apparently stabilized the cis-trans configuration. The trans-trans species, which requires two oxygen atoms to be positioned adjacent to one another, is unfavorable except when the molecule acts as

Table 30. - Distances of the heavy atoms from the least-squares plane.

| Atom | N-methyl-urea | N,N'-dimethyl-urea | N,N-dimethyl-urea |
|---|---|---|---|
| O | 0.014 | 0.000 | -0.012 |
| $N_1$ | 0.023 | 0.039 | 0.005 |
| $N_2$ | 0.003 | -0.039 | 0.046 |
| $C_1$ | -0.049 | 0.000 | -0.009 |
| $C_2$ | 0.009 | -0.018 | -0.007 |
| $C_3$ | --- | 0.018 | -0.024 |
| Average deviation: | 0.020 | 0.019 | 0.017 |
| Maximum deviation: | 0.049 | 0.039 | 0.046 |

a bidentate complexing agent as in bis-biuret copper (II) chloride (Freeman, Smith and Taylor 26). The cis-cis configuration is unfavorable because of the steric hindrance between the hydrogen atoms of the adjacent $NH_2$ groups; however, when two of these hydrogens have been removed by ionization, the molecule can be cis-cis and behave as a bidentate ligand as in potassium bis-biuret cuprate(II) tetrahydrate (Freeman, Smith and Taylor 26).

The fact that the cis configuration is found experimentally only if either a strong intra-hydrogen bond or the molecular geometry requires it suggests that the trans configuration is preferentially stabilized by both intra- and inter-molecular forces. Table 31 summarizes a qualitative evaluation of some statements about the relative stability of the cis and trans forms of the peptide linkage. Statements (3) to (5) are apposite to the argument for the existence of intramolecular stabilizing forces. Statement (3) appears to be equally significant to both the cis- and trans forms because an examination of molecular models indicated that there is very little steric hindrance which would prevent the side group from taking either of the two configurations. Statement (4) is certainly more significant than statement (5) because the $H\cdots O$ distance in the $NH\cdots O$ hydrogen bond would have to be about 2.77 Å, while the $H^{+}\cdots O$ distance would be about 2.28 Å for an eclipsed methyl hydrogen atom.

All of the inter-molecular statements, except statement (8),

Table 31. - Qualitative evaluation of some statements on the relative stability of the cis- and trans-amide configuration of the peptide linkage. (In making this evaluation a model of N-methyl formamide was considered.)

| | Statements | Evaluation | |
|---|---|---|---|
| | | cis | trans |
| 1) | Configuration stabilized by resonance | + | + |
| 2) | Experimentally the predominant form | | + |
| 3) | Methyl group is not sterically hindered | + | + |
| 4) | $\text{C-H}^{+}\cdots\text{O}$ attraction stabilized by hyper-conjugation | | + |
| 5) | Intra-molecular $\text{NH}\cdots\text{O}$ hydrogen bond stabilization | + | |
| 6) | Proton acceptor group is more accessible | | + |
| 7) | Proton donor group is more accessible | | + |
| 8) | More favorable for dimer formation | + | |
| 9) | Favorable for polymer formation | | + |

Key to evaluation: + Significant in a relative sense.

favor the trans-configuration; only in a moderately high pressure gas phase, where dimerization of the molecules becomes significant, would (8) become cogent. The dominance of the trans form in the solid, liquid and high pressure gas phases is largely due to the independent accessibility of the proton donor and proton acceptor groups in the trans form and accordingly to the fact that in this configuration it is possible to form long hydrogen-bonded chains.

(e) Orientations of the methyl group.

In each of the three compounds studied in this thesis, there are one or two methyl groups present. One of two situations exists whenever a methyl group is present: Either the hydrogen atoms of the methyl group occupy preferred orientations - that is, the rotation of the methyl group has been sufficiently restricted to produce three measurable electron density peaks in tetrahedral positions about the carbon atom; or they do not show any preferential orientation - that is, there is either a free-rotation about the heavy atom axis or a statistical distribution among two or more restricted orientations.

Only two restricted orientations seem likely for the N-methyl group located cis to the oxygen atom: either the eclipsed or the staggered configuration of the hydrogen atoms with respect to the carbonyl oxygen atom. The eclipsed configuration would result if an attractive force is operative, as, for example, a weak CH•••O intra-molecular hydrogen bond in which perhaps the hydrogen atom has become partially positive due to hyperconjugation (Pauling 32b). The staggered configuration would result from the repulsion between the electrons of the oxygen atom and the electron of the hydrogen atom. It is not clear which is more plausible; and if both are equally likely, then the potential function for the methyl group rotation would have six-fold symmetry and the methyl group would be expected to be statistically distributed among the two orientations.

The cis-methyl groups in N-methylurea and in N,N'-dimethylurea are not preferentially oriented as shown by the difference syntheses illustrated in figures 13b and 8, respectively. This implies that either the intra-molecular forces are not strong enough at room temperature to restrict the rotation of the methyl group or that there are two or more equally likely restricted orientations. A nearly eclipsed orientation was found for the cis-methyl group in N,N-dimethylurea; however, it appears that here the molecular packing along the chain axis contributes in part to the restriction of the orientation. The positions of the hydrogen atoms correspond to a rotation of the methyl group by about 10° from the eclipsed configuration (see figure 3b).

A molecular model (Corey and Pauling 10) of N,N-dimethylurea indicated that the trans-methyl group would most likely have a staggered orientation with respect to the amide hydrogen, $H_2$, which constitutes a formidable steric barrier to the rotation of the methyl group. The perfectly staggered position, however, is not found; the positions of the hydrogen atoms correspond to a rotation of the methyl group by about 30° from the staggered configuration (see figure 3a). Hydrogen atoms $H_3$ of the methyl group and $H_2$ of the amide group are in close van der Waals contact (2.00 Å). This configuration appears to result from the packing formed by the tetrahedral interlocking of atoms O and $H_7$ from one molecule with atoms $H_2$ and $H_3$ of the adjacent molecule related by the n-glide. This configuration is shown in figures 14 and 16, and a list of the interatomic distances for the atoms involved in the packing is given in table 32. The interlocking also explains the 10° deviation of the cis-methyl group from the eclipsed configuration.

ii) Inter-molecular features:

(a) The NH···O hydrogen bonds.

All of the NH···O hydrogen bond lengths observed in urea and its methyl derivatives are within the limits of the distances derived by Fuller (36); that is, $d = 2.96 \pm 0.09$ Å for the distance between the nitrogen atom of an amino ($NH_2$) group and the oxygen atom of a carbonyl group within a chain (rather than a ring). Table 33 lists for comparison the hydrogen bond distances found in these compounds. Throughout any further discussion involving these hydrogen bonds, it will be tacitly assumed that the relative strengths of all the hydrogen bonds are about equal. As Corey and Pauling (31) have pointed out, the precise relationship between the dimensions of the NH···O bonds in

Table 32. - Hydrogen-atom packing distances along the chain axis in N,N-dimethylurea.

| from / to | $H'_7$ | $H_2$ | $H_3$ | $H_4$ |
|---|---|---|---|---|
| O' | 2.28 | 2.04 | 2.49 | 3.74 |
| $H'_7$ | -- | 2.15 | 2.89 | 4.68 |
| $N_1$ | 2.85 | 1.04 | 2.54 | 3.36 |
| $H_2$ | 2.15 | -- | 2.00 | 3.18 |
| $H_3$ | 2.87 | 2.00 | -- | 1.91 |
| $H_4$ | 4.68 | 3.18 | 1.91 | -- |

Table 33. - Hydrogen bond distances found in urea and its methyl derivatives.

| Compound | Chain links | Cross links |
|---|---|---|
| Urea (neutron diffraction) | 2.99 | 3.03 |
| Urea (x-ray diffraction) | 2.99 | 3.04 |
| N-methylurea | 2.969<br>2.932 | 2.980 |
| N,N'-dimethylurea | 2.864 | |
| N,N-dimethylurea | 3.064 | 2.943 |
| (N-methyl acetamide | 2.825) | |

Table 33. - Hydrogen bond distances found in urea and its methyl derivatives.

| Compound | Chain links | Cross links |
|---|---|---|
| Urea (neutron diffraction) | 2.99 | 3.03 |
| Urea (x-ray diffraction) | 2.99 | 3.04 |
| N-methylurea | 2.969<br>2.932 | 2.980 |
| N,N'-dimethylurea | 2.864 | |
| N,N-dimethylurea | 3.064 | 2.943 |
| (N-methyl acetamide | 2.825) | |

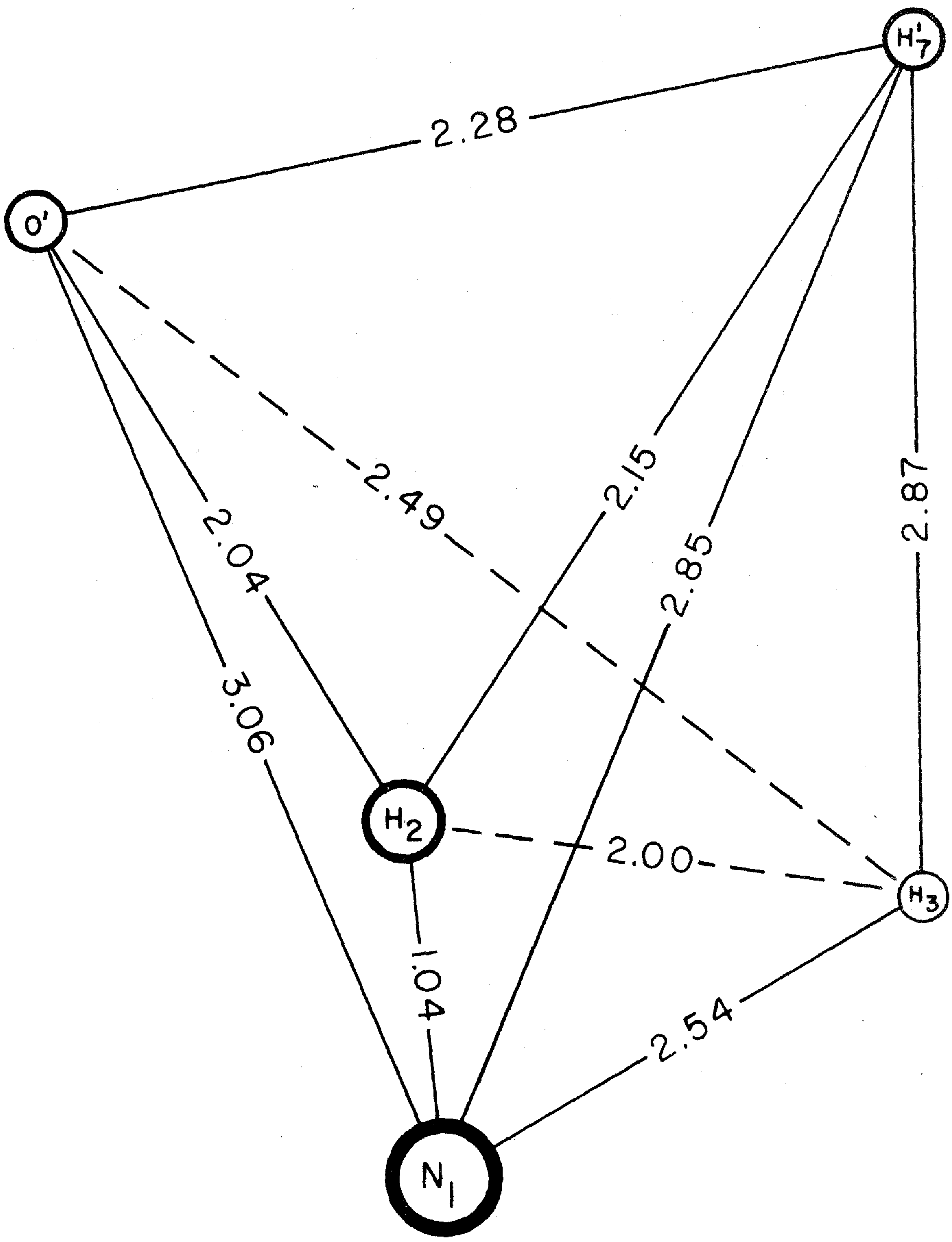


Figure 14. - The tetrahedrally interlocked hydrogen bond link along the chain axis of N,N-dimethylurea; some pertinent interatomic distances are shown.

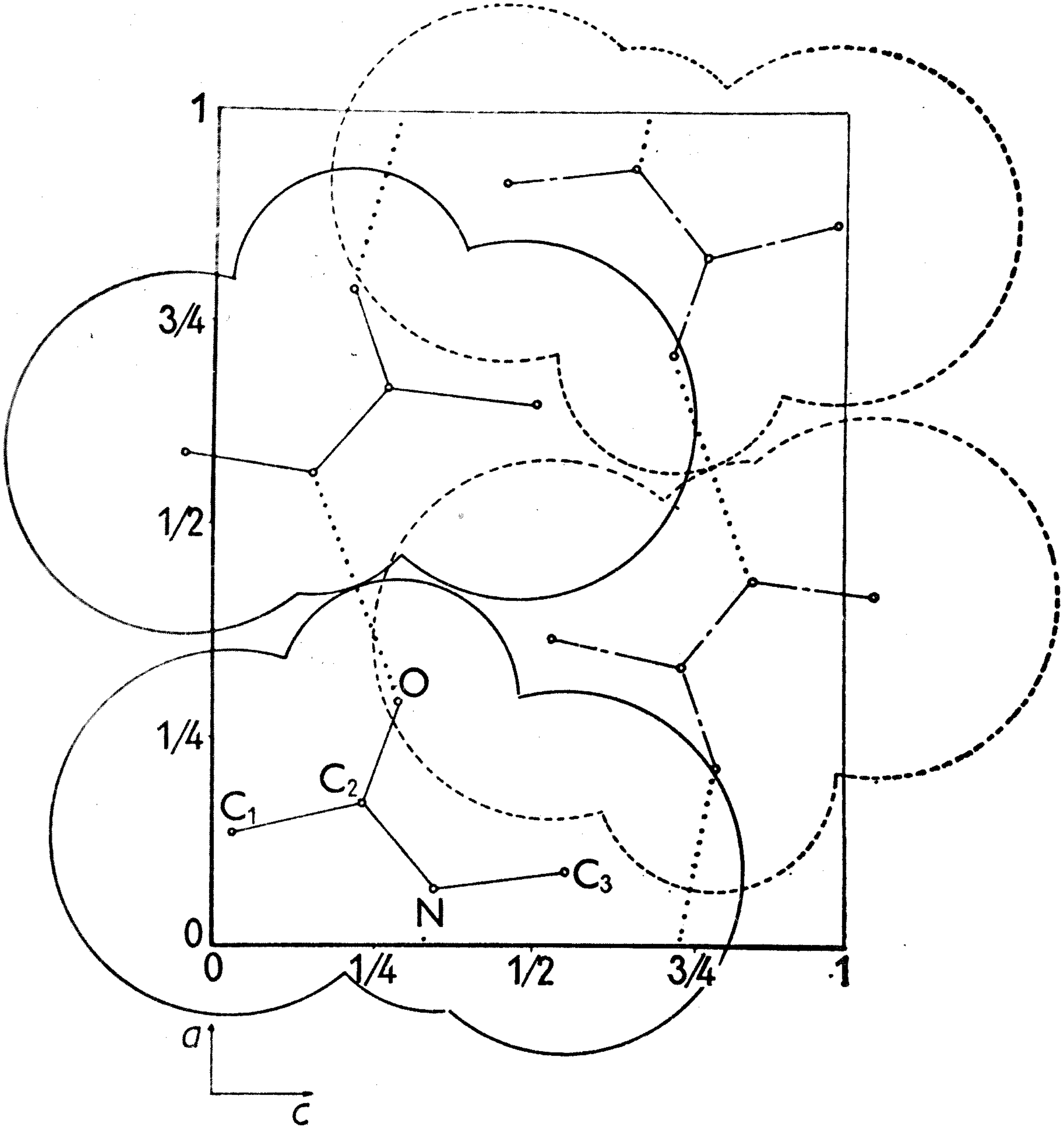


Figure 15. - A drawing of the trans-amide chain structure found in N-methyl acetamide (28). (Illustration from Acta Cryst.)

Figure 16. - A model of the trans-amide chain found in N,N-dimethylurea.

different compounds will always be complicated by the shapes of the molecules involved and the concessions which molecular packing must make to purely steric effects.

(b) Types of hydrogen bonded chains.

In each of the three substituted ureas investigated, the structure can be characterized, at least in part, by the type of hydrogen bond chain networks it contains. The essential characteristics of the hydrogen bonded molecular chains and the manner of their propagation depend on whether the links have the cis- or trans-amide configuration. Because of the importance of these chains, a brief review of some of the various types of chains formed by the cis- and trans-amide linkages will be given. The terms "cis" and "trans" refer to the position of the proton relative to the C-O bond in the HN-C=O links. The effect of combining the cis- and trans-amide links together into one molecule is included. The review will be used as an opportunity to introduce a system of nomenclature for describing multiple hydrogen bonded chains. The prefixes di-, tri-, etc., will be used to specify the number of each type of amide linkages connecting two molecules.

When there is only one trans-amide group in a molecule, it can form a trans-amide chain similar to the one formed in the low temperature modification (below 10°C) of N-methyl acetamide, illustrated in figure 15. In this case, the links are all coplanar, with the adjacent links alternatively having a 0 or 180° rotational alignment about the chain axis. In the high temperature modification, a disorder is introduced along the chain axis, which appears to be explained in terms of a statistical distribution of the molecules into the 0 and 180° orientations. This type of chain does not require the coplanarity of the amide links for its propagation as is shown by the trans-amide chain found in N,N-dimethylurea. The NH•••O hydrogen bond is directed approximately along one of the free electron pair orbitals of the oxygen atom; that is, directed along a direction approximately tetrahedral to the "banana bond" orbitals of the carbonyl double bond.

Two trans-amide groups in a molecule can result in the formation of ditrans-amide chains. Some variety exists for these types of chains depending on the geometric location of the two amide links in the parent molecule. One example of this chain is illustrated in figure 17 showing the two fused amide links in N,N'-dimethylurea forming a coplanar ditrans-amide chain. The same variant is easily spotted in urea (Vaughan and Donohue 22) in figure 18 and less easily in bis-biuret cadmium chloride (Cavalca, Nardelli and Fava 25) shown in figure 19. Coplanarity of all of the trans-amide units making up the chain appears to be the rule in the ditrans-amide chains, due, no doubt, to the tetrahedral hydrogen bonding to the two free pairs of electrons on the carbonyl oxygen. Like most rules, this one has its exception; the ditrans-amide chain in N-methylurea, illustrated with molecular models in figure 20, is a non-planar chain system (the acute angle between any two adjacent molecules is 56°).

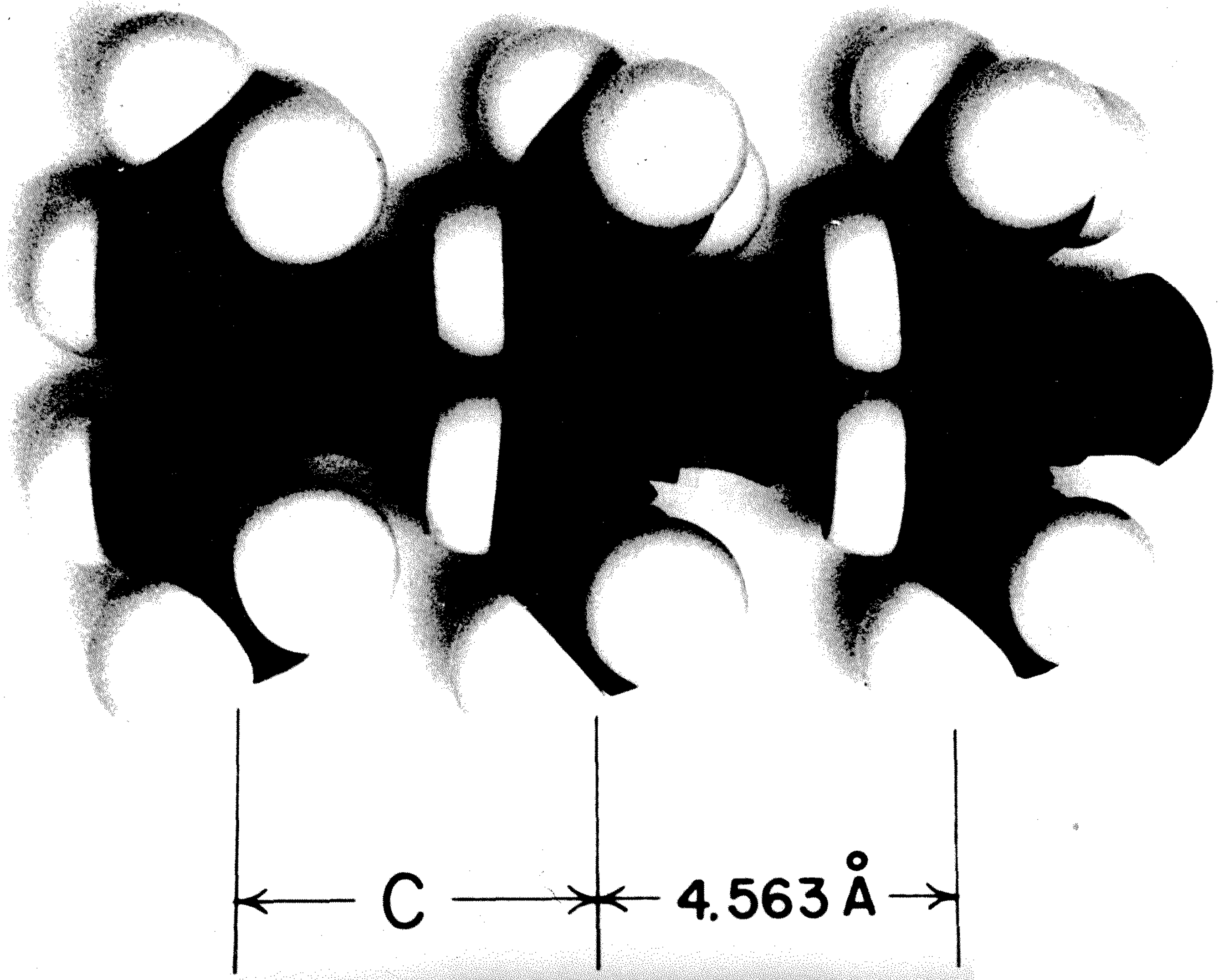


Figure 17. - A model of the ditrans-amide chain in N, N'-dimethylurea.

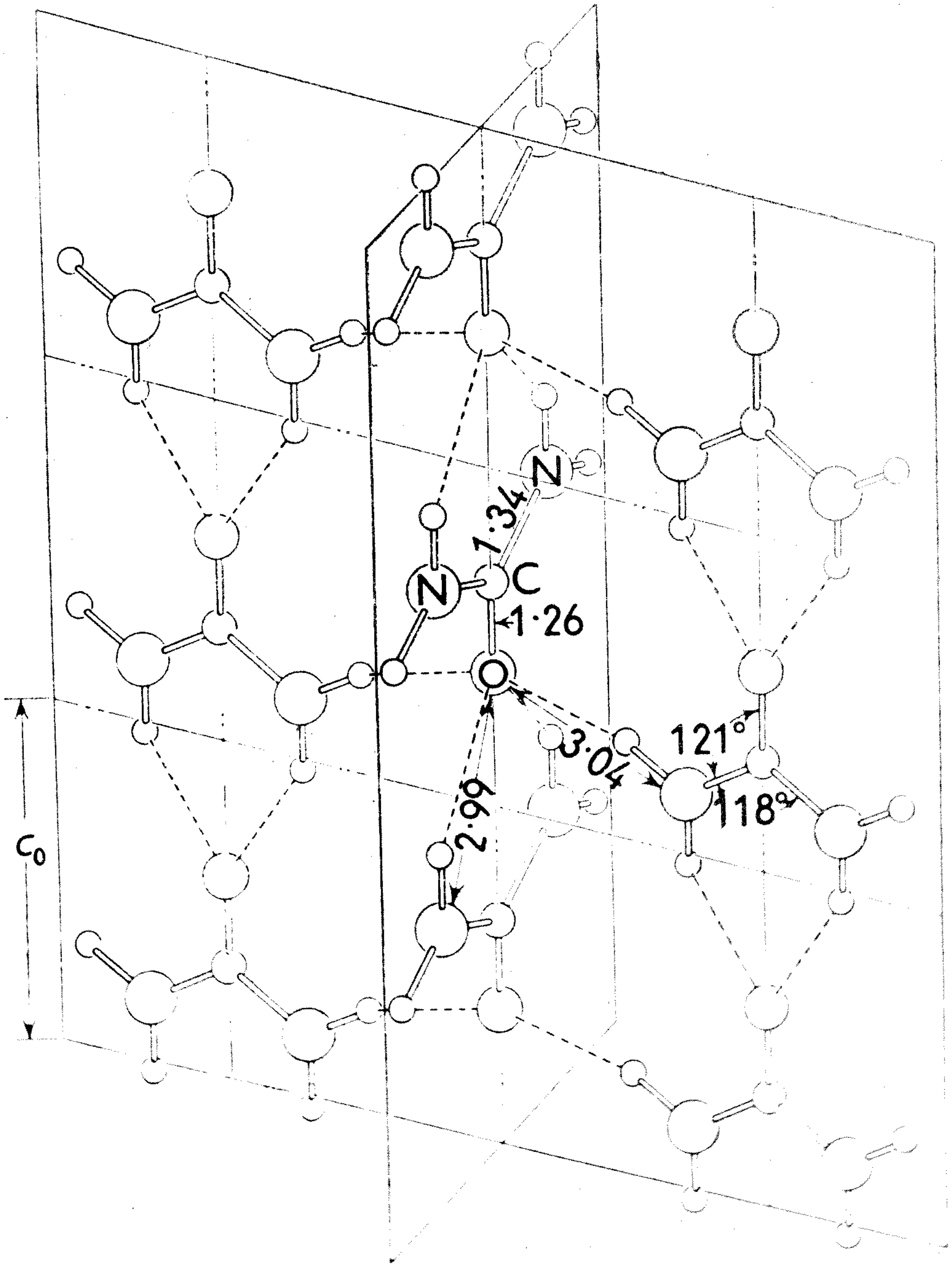


Figure 18. - A view of the crystal structure of urea (22), showing the ditrans-amide chains running parallel to the c-axis with two cis-amide chains, running through each molecule, cross linking the ditrans-amide chains. (Illustration from Acta Cryst.)

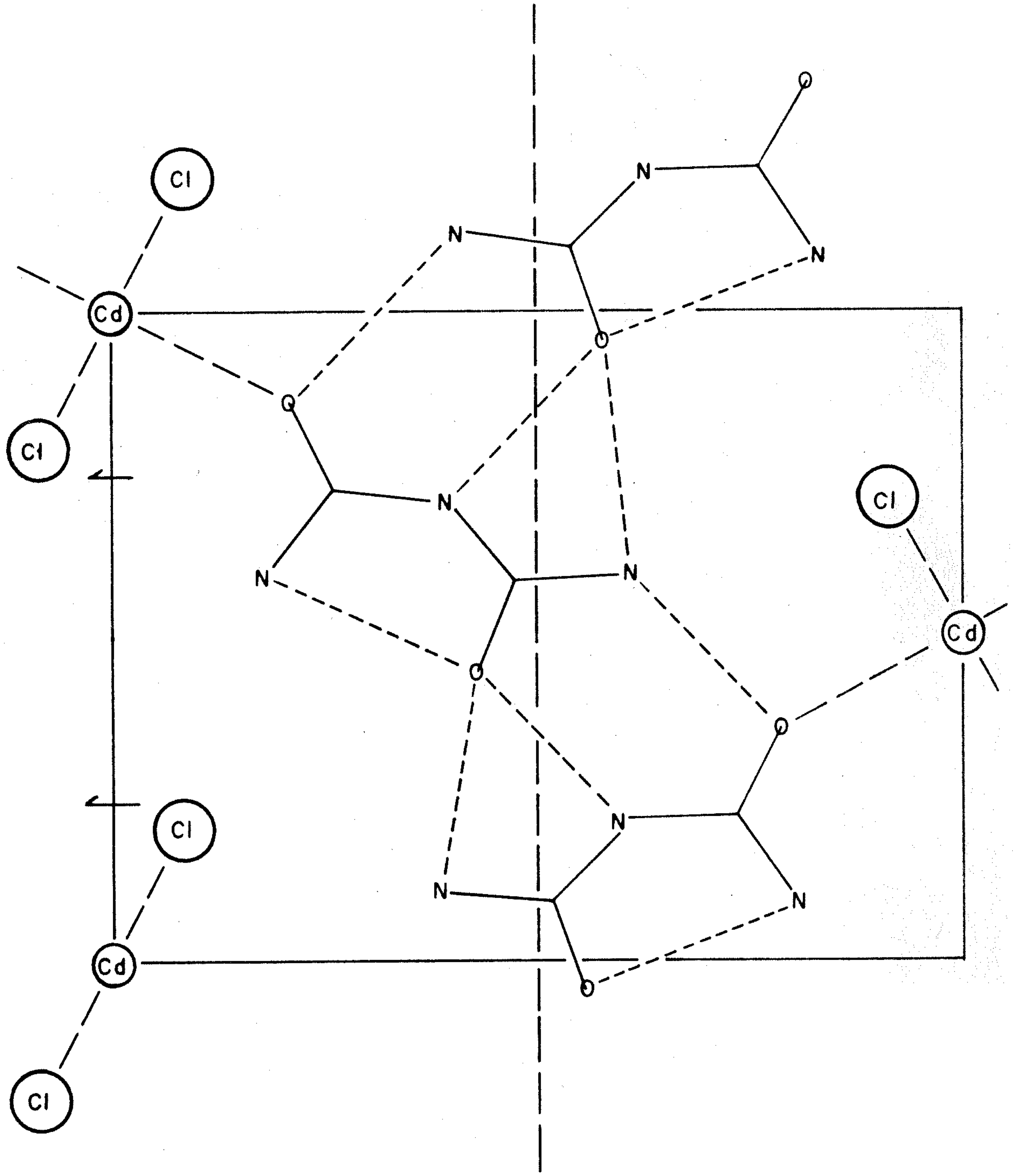
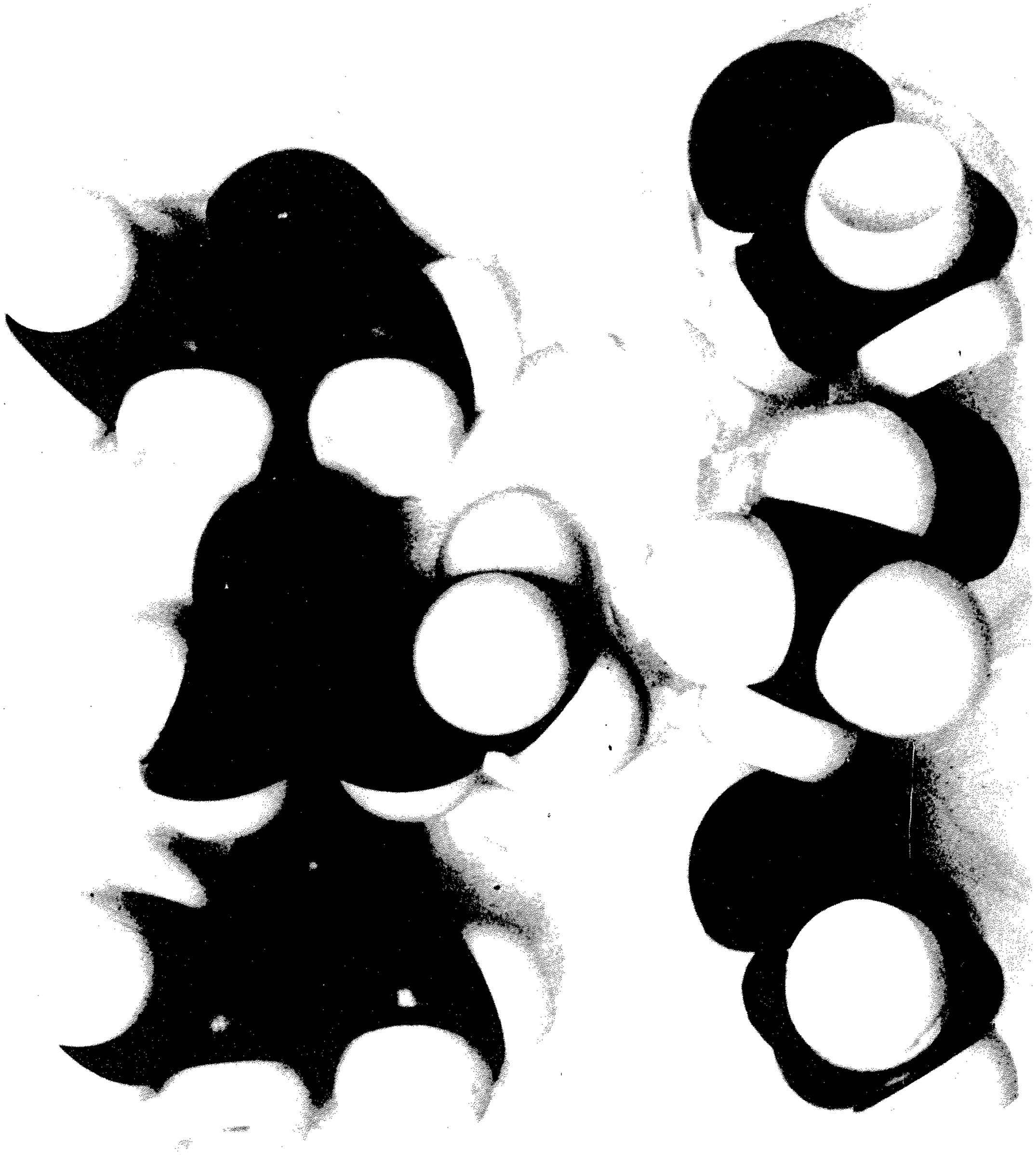


Figure 19. - An illustration of the ditrans-, biscis D-amide chains in *bis*-biuret cadmium chloride (25).



(a) (b)

Figure 20. - Two views of the ditrans-amide chain in N-methylurea showing the non-planar propagation of the chain; (a) a broad side view, (b) an edge view.

Another variant, in which the amide groups are separated in the molecule, is the type of ditrans-amide chain formed in anhydrous diacetyl hydrazine (Shintani 27); it is illustrated in figure 21. The pleated-sheet configuration of polypeptide chains (Pauling and Corey 37) is an example of a polytrans-amide chain.

The most common hydrogen-bond configuration for the cis-amide groups is the formation of a centric dimer. Figure 22 shows the cisD-amide dimer formed in 1-methyl thymine (Hoogsteen 38). (The term cisD-amide is coined to denote this type of dimer.) Figure 23 shows a photograph of the cisD-amide dimer found in N,N-dimethylurea. Because of the cyclic nature of the hydrogen bonding within the dimer, no chain structures are possible with only one cis-amide link per molecule. Chain structures, which contain the cisD-amide linkage as a unit, exist when there are two or more cis-amide linkages in each molecule. (The number of times the cisD-amide unit is replicated on one molecule will be denoted by the prefixes bis-, tris, and tetrakis, etc.)

Thymine hydrate (Gerdil 39), illustrated in figure 24, is an example of a compound having a biscisD-amide chain structure. (Note that in this example it is the two fused cis-amide groups of the urea nucleus which have been used to propagate the chain.) Another example of a compound having biscisD-amide chains in it is diketopiperazine (Corey 35; Degeilh and Marsh 40); it is illustrated in figure 25.

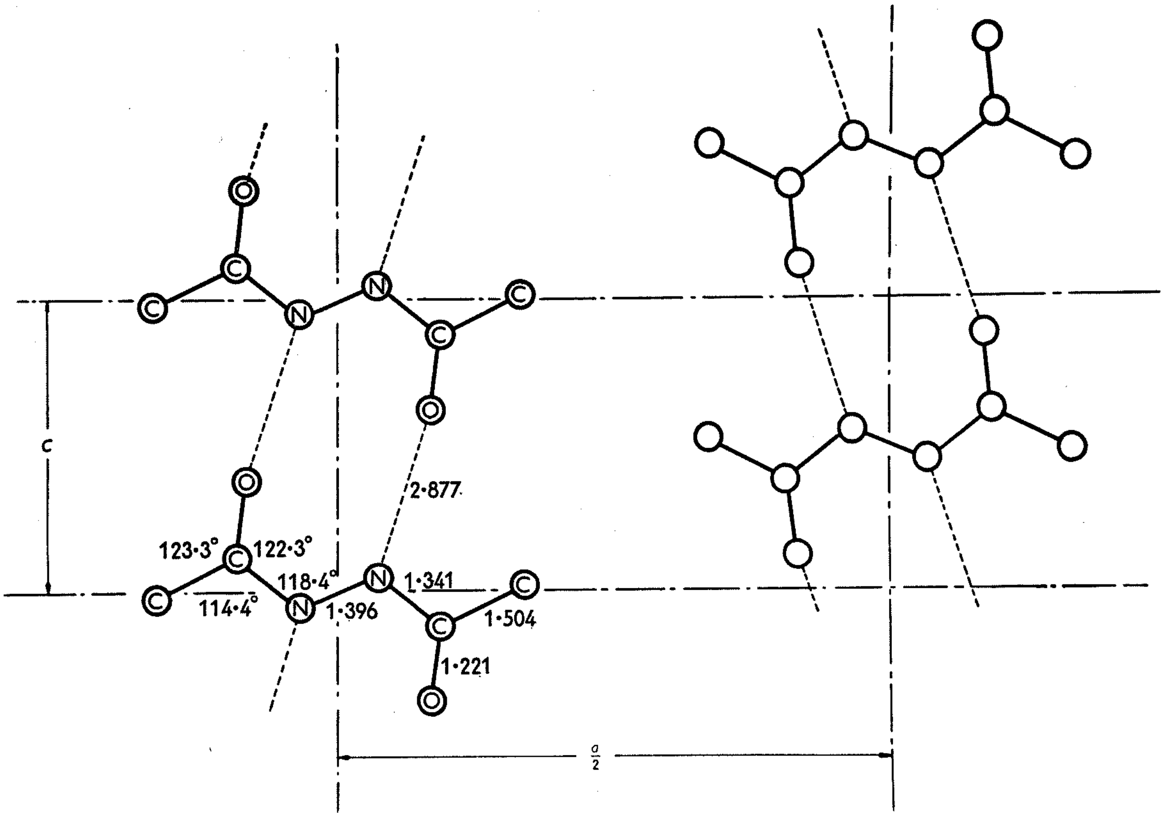


Figure 21. - An illustration showing the ditrans-amide chain in anhydrous diacetyl hydrazine (27). (Illustration from Acta Cryst.)

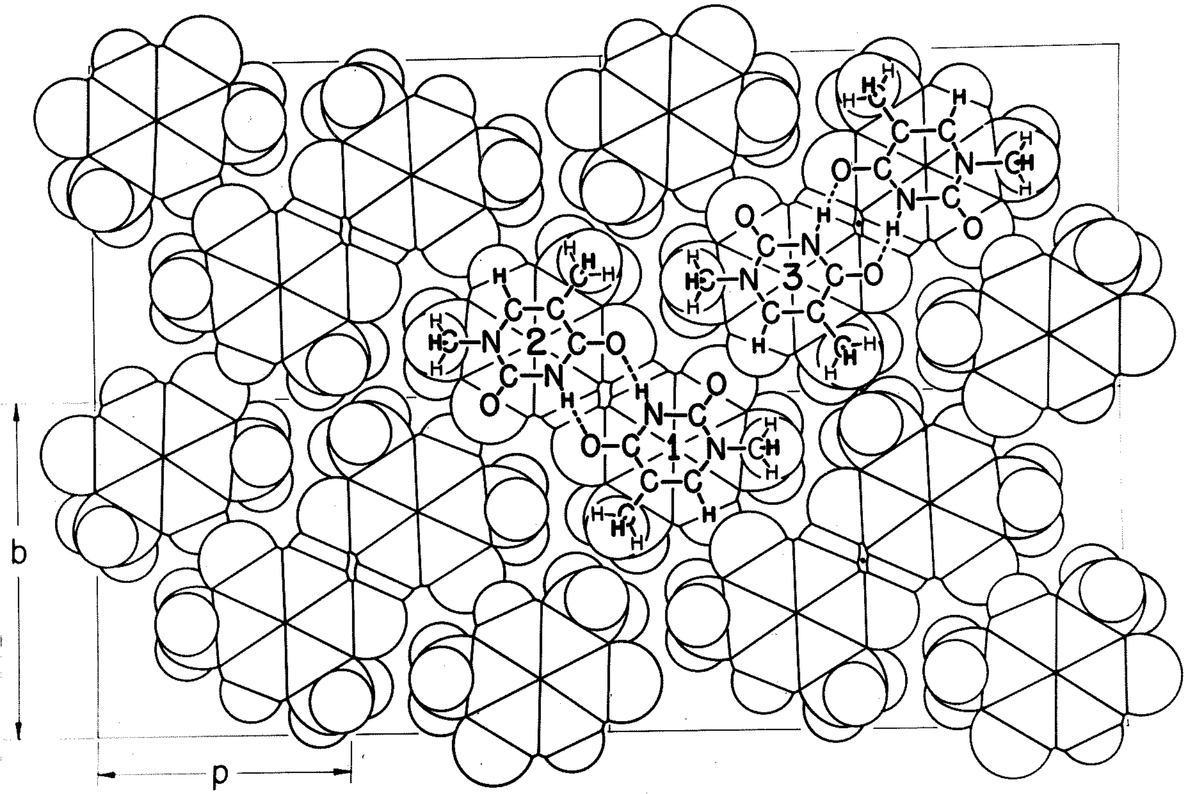


Figure 22. - An illustration of the cisD-amide dimer in 1-methyl thymine (38).

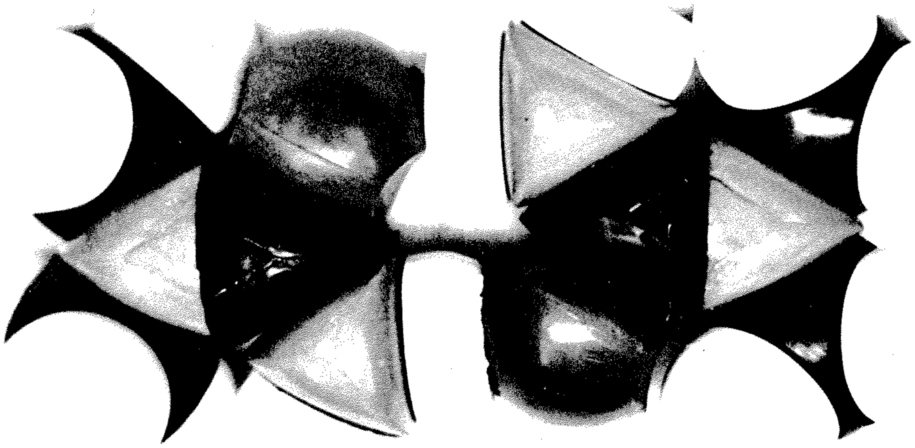


Figure 23. - Models showing the cisD-amide dimers found in N,N-dimethylurea.

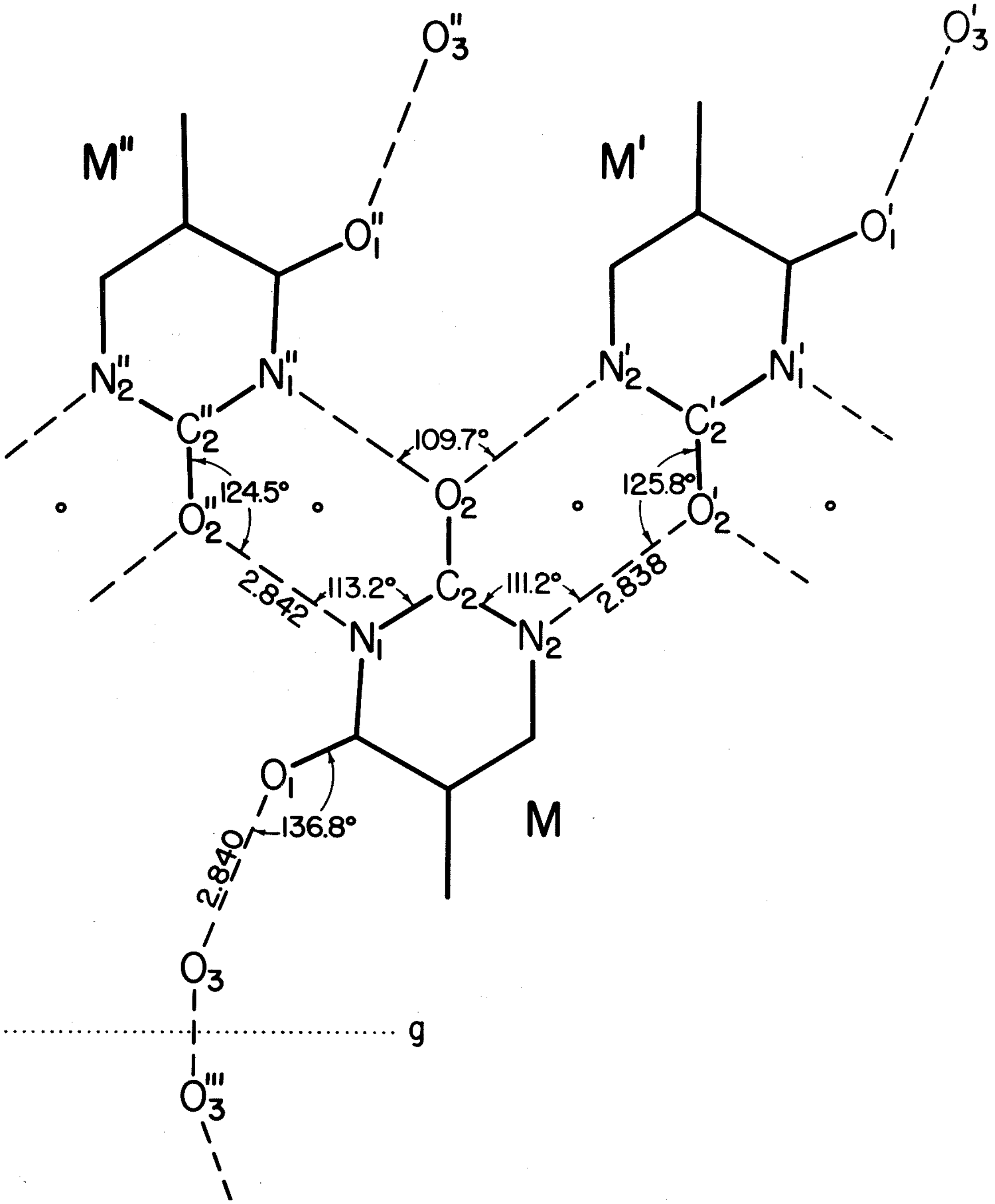


Figure 24. - An illustration of the biscisD-amide chain in thymine hydrate (39).

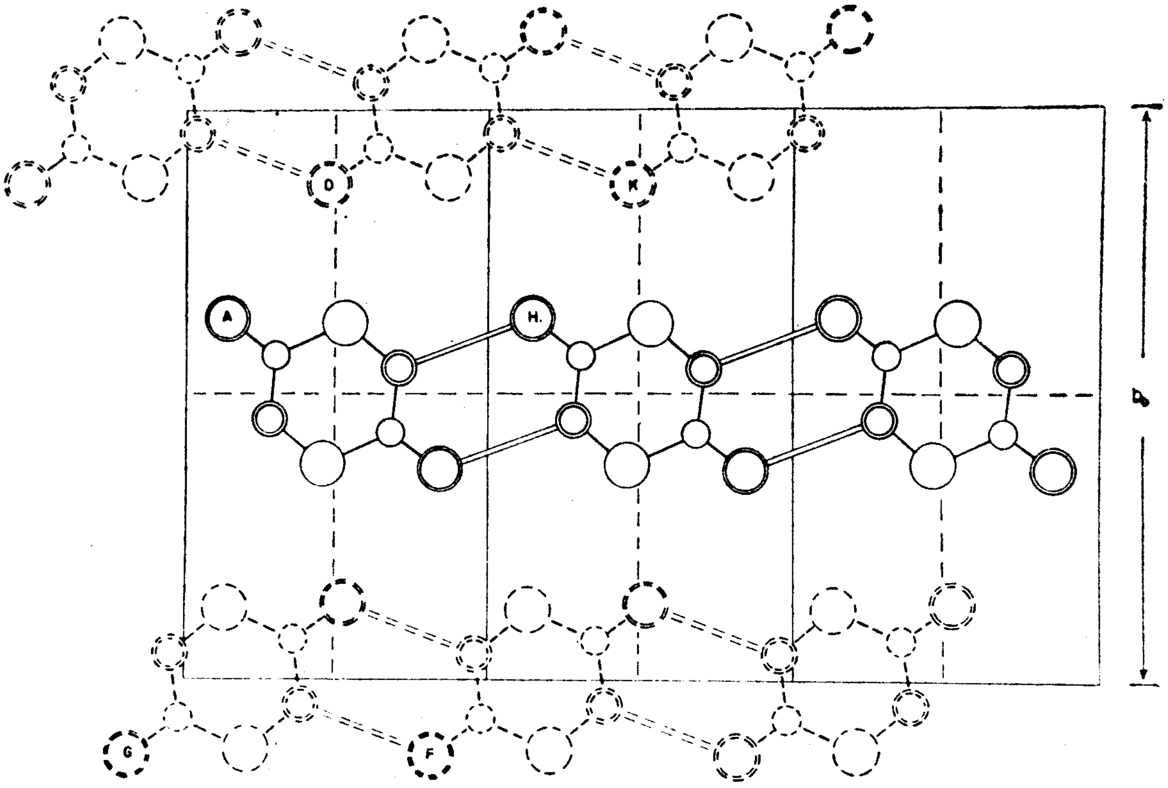


Figure 25. - An illustration of the biscisD-amide chain in diketopiperazine (35). (Illustration from J. Am. Chem. Soc. with permission of the author.)

This same type of chain is easily spotted in cyanuric acid (Wiebenga and Moerman 41), shown in figure 26. Newman and Badger (42) called the biscisD-amide hydrogen bonds in this compound "type B."

In addition to the biscisD-amide hydrogen bond in cyanuric acid, there are the NH•••O hydrogen bonds denoted as "type A" by Newman and Badger, in which the cis-amide groups do not form centric dimers. We shall call these hydrogen bond chains the biscis-amide type, since the chain requires two cis-amide links per molecule for propagation. Figure 27 shows the structure of uracil (Parry 43). Both the cisD-amide dimers and the biscis-amide chains are easily seen in the illustration.

Digressing a bit, we notice that if the $\mathrm{O{=}C{-}\underset{H}{C}{=}\underset{H}{C}{-}}$ portion in uracil is replaced by $\mathrm{O{=}C{-}\underset{H}{N}{-}\underset{H_2}{C}{-}}$ , then the resultant hydrogen bond system would be a triscisD-amide planar network. The short CH•••O distance (3.28 Å) in the pseudo biscisD-amide chain of uracil suggests strongly that the carbon atom is hydrogen bonded through the proton to the oxygen atom. Allowing 1 Å for the C-H bond distance, the distance between the hydrogen atom and the oxygen atom is 2.28 Å, 0.3 Å shorter than the 2.6 Å distance predicted from the van der Waals radii given by Pauling (32c). The other CH•••O distance (3.19 Å) is even shorter, being 0.4 Å less than the predicted distance. A single-crystal polarized IR study of the CH frequencies in this compound might establish the validity of the CH•••O hydrogen bond conjecture.

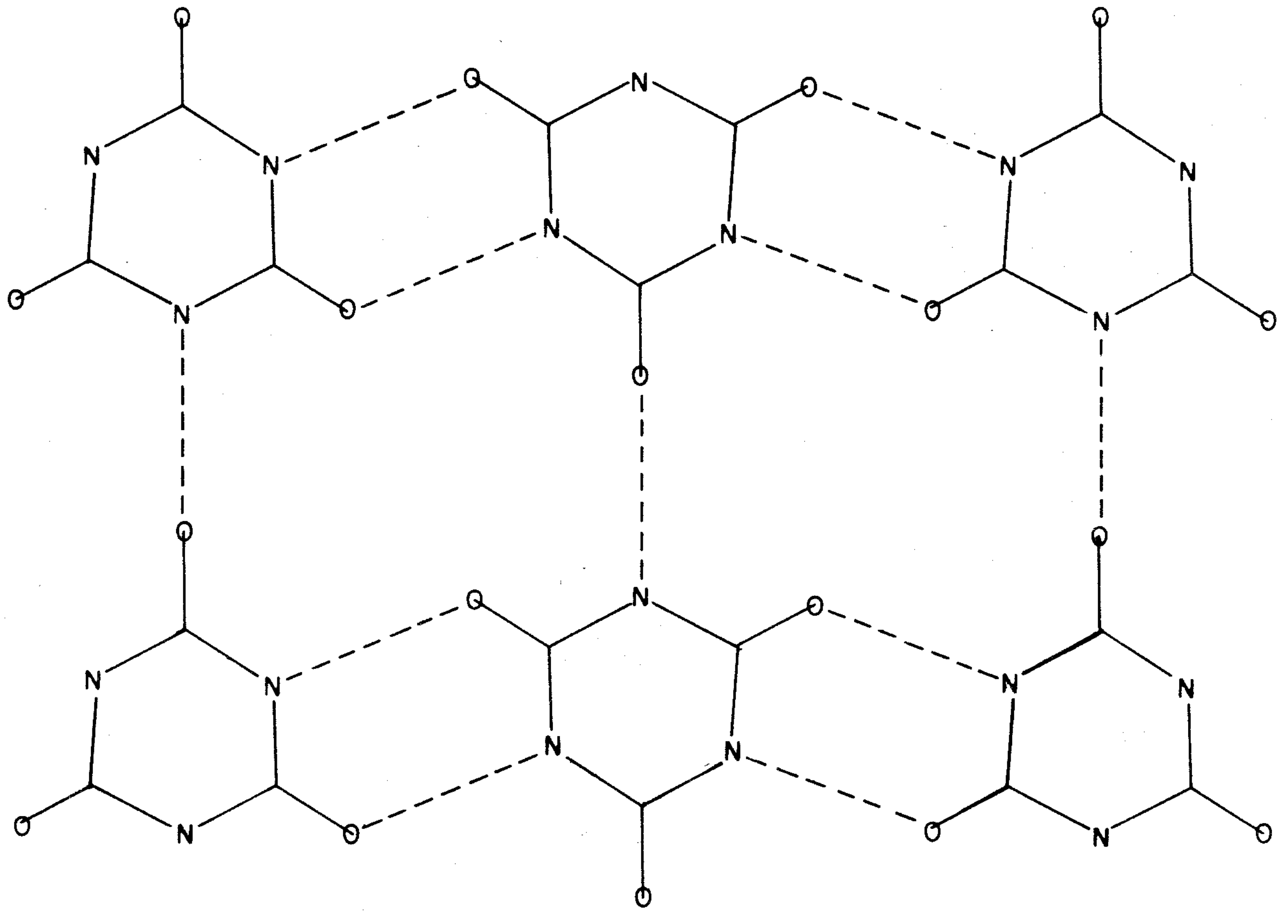


Figure 26. - An illustration of the structure of cyanuric acid (42). Both the biscisD-amide chains and the biscis-amide chains are shown.

THE CRYSTAL STRUCTURE OF URACIL

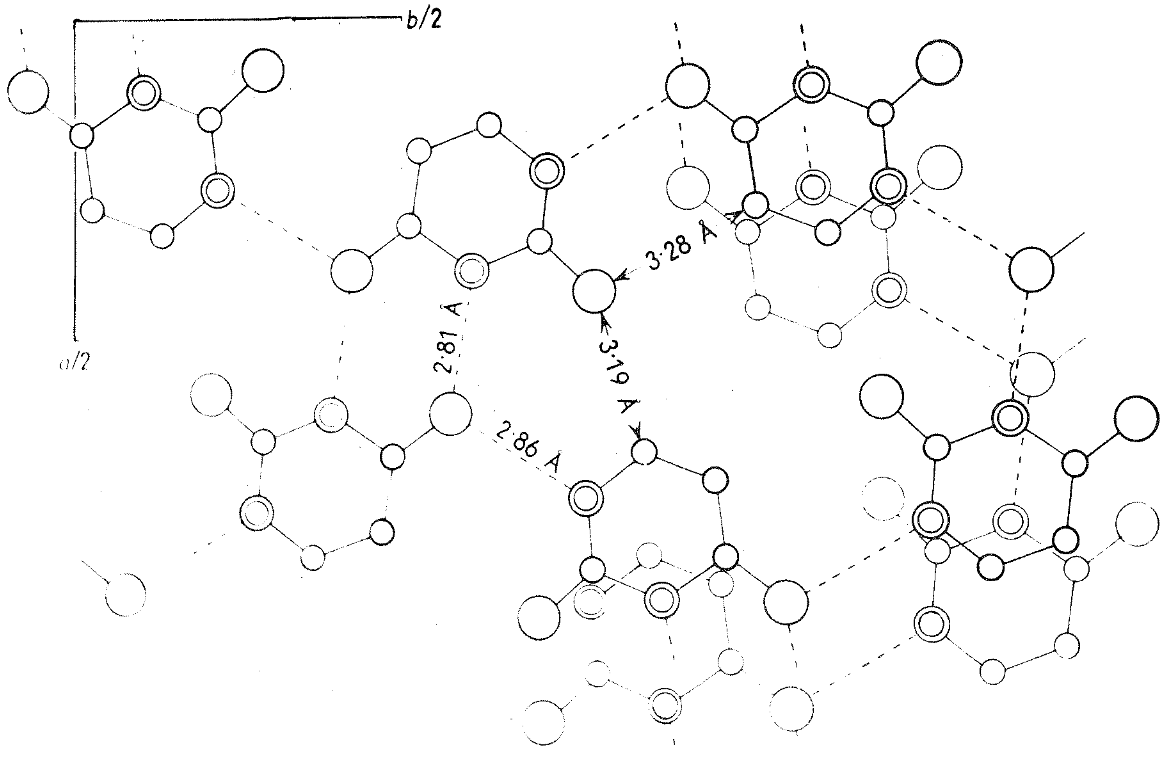


Figure 27. - A drawing showing the hydrogen bonding network in uracil (43). The cisD-amide dimers and the biscis-amide chains are shown. (Illustration from Acta Cryst.)

Table 34. - Summary of hydrogen bond type chains.

| | Type of chain | Example found in: | Figure |
|---|---|---|---|
| A. | Mono-amide type. | | |
| | (1) Trans-amide | N-methyl acetamide | 15 |
| | (2) CisD-amide (dimer) | 1-methyl thymine | 22 |
| | (3) Cis-amide | N-methyl urea (cross link) | 28 |
| B. | Poly-amide type. | | |
| | (1) Ditrans-amide | N,N'-dimethylurea | 17 |
| | (2) BiscisD-amide | diketopiperazine | 25 |
| | (3) Biscis-amide | uracil | 27 |
| | (4) Polytrans-amide | pleated-sheet (polypeptide) | -- |
| C. | Mixed-amide types. | | |
| | (1) Ditrans, biscisD-amide | bis-biuret cadmium chloride | 19 |

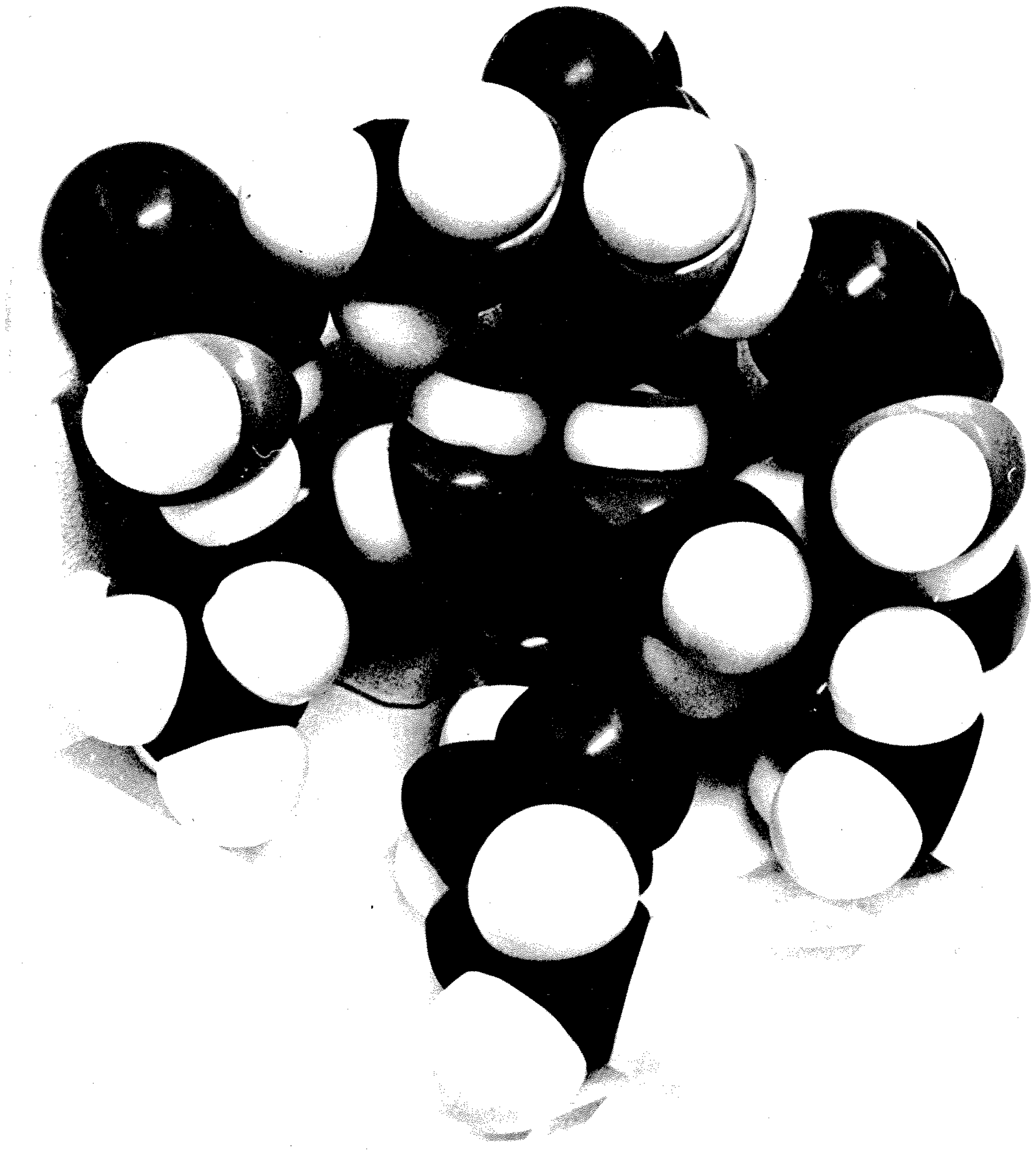


Figure 28. - A model of the structure of N-methylurea.

A single cis-amide link per molecule can form a cis-amide chain. The chain propagation is from one cis-amide link through an NH···O hydrogen bond to another cis-amide link which is oriented perpendicular to the first; hence, the chain propagates along a zig-zag course. Cis-amide chains cross link the ditrans-amide chains in N-methylurea in planes perpendicular to the ditrans-amide chain axes, see figure 28. Urea, figure 18, has two cis-amide chains propagating through each molecule. In both of these examples the C=O bonds in two adjacent amide links of the cis-amide chains have opposite orientations; that is, the adjacent ditrans-amide chains are inverted. The cis $C_1$-N···O' angle in urea is 129.3°, and in N-methylurea it is 125.3°.

iii) Description of the molecular packing.

A description of the molecular packing in urea and its methyl substituted derivatives will now be given. Urea has been included for purposes of comparison. In this connection, one should note the basic similarity in packing between N-methylurea and urea, while N,N'-dimethylurea retains only the ditrans-amide chain character of urea and N,N-dimethylurea is completely different.

(a) Urea (Vaughan and Donohue 22).

The packing found in urea is illustrated in figure 18. Planar ditrans-amide chains are packed perpendicular to four adjacent inverted chains with each molecule cross-linked to the four inverted chains via two cis-amide chain systems.

(b) N-methylurea.

A view of the molecular packing in crystals of N-methyl urea is shown in figure 28. The kinked ditrans-amide chains are easily seen in two edge views and in a broad-side view. A cis-amide chain is seen propagating from the molecule in the lower foreground to a molecule in the center ditrans-amide chain which in turn is seen cross-linked to a molecule in the chain on the left. On the right the methyl group is shown packed into the fold of the ditrans-amide chain. Except for the presence of the methyl group which has introduced the kinks into the ditrans-amide chains, this structure is basically similar to urea. The angle between the b-axis and the line of intersection of the molecular least-squares plane with the ab-plane is 66.2°; this indicates that the ditrans-amide chains are twisted by only 23.8° from the orientation presumed by Gordon on the basis of his magnetic anisotropy measurements.

(c) N,N'-dimethylurea.

The packing of the ditrans-amide chains in N,N'-dimethylurea is basically different from urea because all of the chains have the same orientation; that is, the C=O bonds all point in the same direction along the c-axis. Figure 29 shows a packing view of N,N'-dimethylurea along the c-axis. The angle between two molecules related by a diamond glide plane is 73.9°. As can be seen from the figure the packing is governed by the manner in which the methyl groups are best accommodated into interstitial positions along the chain axis.

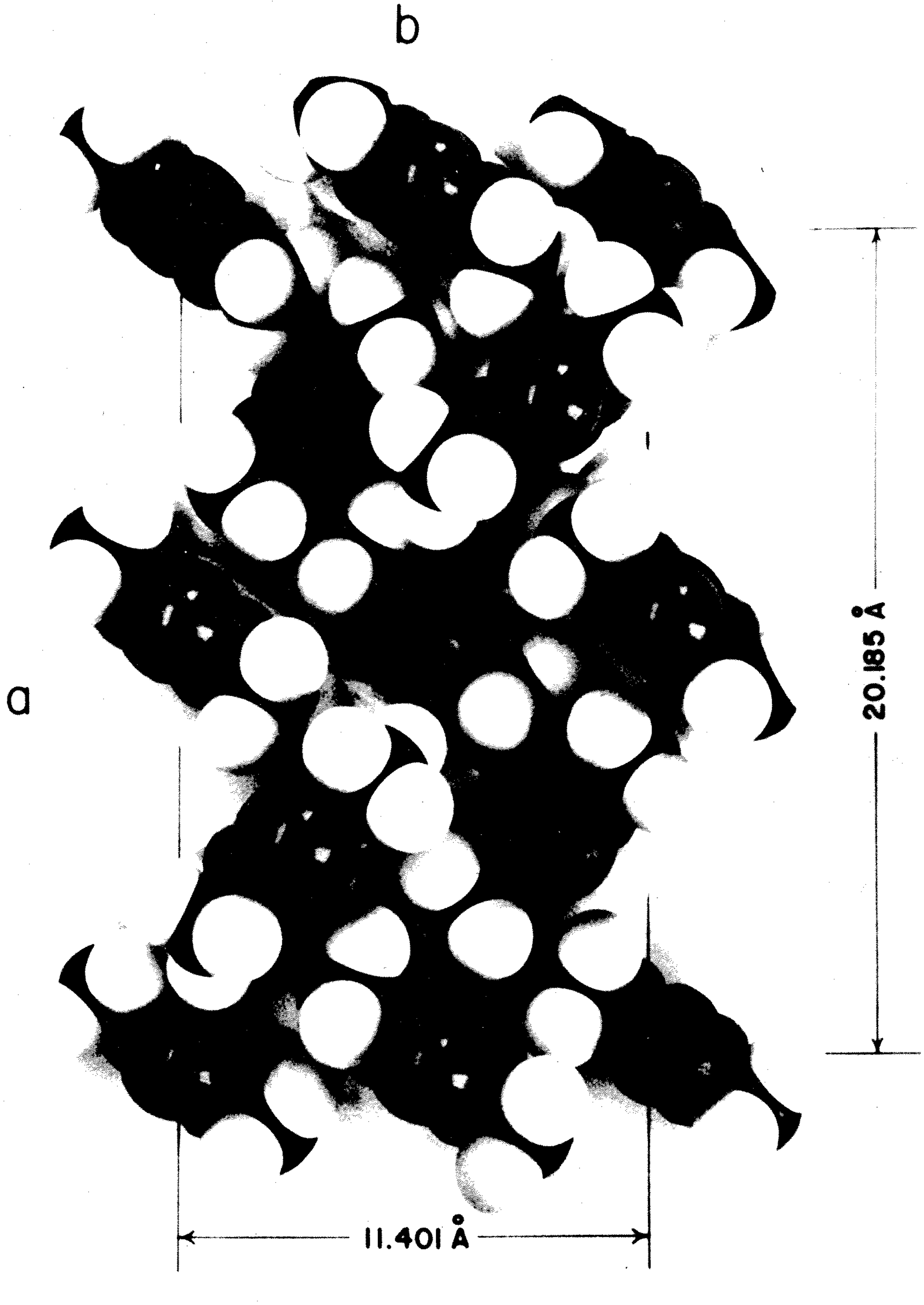


Figure 29. - A model of the structure of N,N'-dimethylurea viewed along [001] showing the contents of one unit cell.

(d) N,N-dimethylurea.

The trans-amide chains in N,N-dimethylurea pack into parallel arrays in which each chain is cross-linked to two other chains by formation of cisD-amide dimers. These cross links thus fuse two arrays of chains into sheets having a polar center and a non-polar outer covering of methyl groups. These sheets are in turn stacked back-to-back; figure 30 shows this packing in projection on the [010] zone.

iv) A rationalization of the melting point anomaly.

A clew to the mechanism of melting of these hydrogen-bonded chain structures is given by our knowledge of the structural changes undergone by N-methyl acetamide during its solid phase transformation (Katz and Post 28, hereafter KP). In the low-temperature form, adjacent molecules along the chain axes are related by a two fold screw-axis, while in the high-temperature form the relation is statistically either a screw axis or a unit-cell translation. KP infer that the disorder is present in each chain although they have not shown that the disorder is not a random packing of ordered chains. This point could easily be resolved if it could be shown that the phase transformation is reversible in a single crystal, in which case the disorder with the individual molecules randomly oriented would be more probable.

If we assume that the latter is true, it would appear that as the temperature is increased there is an increased molecular libration about the chain axes which increases the effective chain volume and weakens

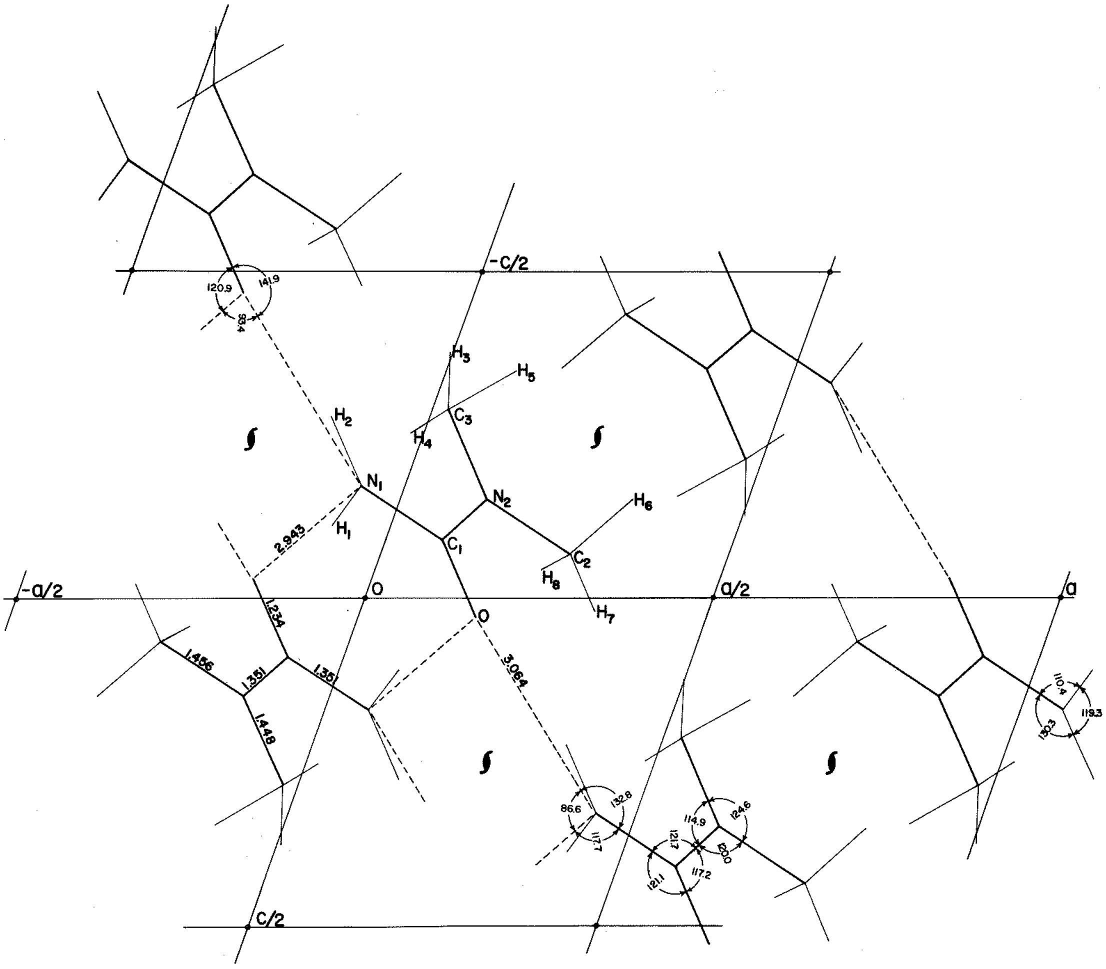


Figure 30. - A drawing of the structure of N,N-dimethylurea projected on [010] ; some bond distances and angles are included.

the intra-molecular interactions, which include van der Waals forces and any hydrogen bonds present. The crystal melts when the chain is fragmented and each fragment has achieved a mean-free volume within which it is as free as a man in a moving crowd. Thus more efficiently packed crystals require higher temperatures to increase the mean-free volume to a point at which each fragment acts independently of its neighbors. The presence of hydrogen bonds in the crystal affects the melting point by increasing the effective chain length of the fragments; as in the case of the normal hydrocarbons, the longer the chain length the higher will be the melting point. The average chain length of the individual fragments is undoubtedly a dynamic property in the liquid state.

The trans-amide chain formation of N-methyl acetamide is the dominant feature in its two crystal modifications. The libration of one molecule about its chain axis must be paralleled by a similar libration in the adjacent chains; thus the molecules in the high-temperature phase move about the chain axis like the cogs in a timing gear of a clock moving by the ratchet gear. The crystal melts when enough additional energy has been added to cleave the chains into short fragments and to disengage the ratchet wheel system.

Trimethylurea (melting at 75.5°C) probably has a trans-amide chain structure. Its melting point is higher than that of N-methyl acetamide (28°C) because of the increased molecular weight. On the other hand, the still higher melting point of N,N'-dimethylurea (106°C) is due to the formation of the ditrans-amide chains which means that the fragments have a longer average chain length and more energy is needed to produce the fragments. N-methylurea also contains the ditrans-amide chains. It has a lower molecular weight than N,N'-dimethylurea and appears to have available within the crystal most of the mean-free volume needed for molecular libration about the chain axis; these facts would imply a low melting point relative to N,N'-dimethylurea. However, the formation of the additional cis-amide chain cross links would tend to raise the melting point. The result is a compromise; the crystal melts at 102°C, only slightly lower than N,N'-dimethylurea. Urea, like N-methylurea, has most of the mean-free volume needed for molecular libration already built in. However, it has two cis-amide chains forming cross links between the ditrans-amide chains. These predominate over the loss of the second methyl group, resulting in a melting point of 132°C.

N,N-dimethylurea contains the trans-amide chain structure with each molecule cross linked to form cisD-amide dimers. The cross linking of the trans-amide chains has produced an interwoven two-dimensional network of hydrogen bonds; each molecule is attached to the network by four hydrogen bonds. The molecules are efficiently packed into the crystal, leaving little free space available for large molecular librations. The librations of the molecule about its librational axes are damped by the cisD-amide dimer formation, by the interlocked hydrogen atom packing at the site of the trans-amide hydrogen bond link and by the denseness of the methyl-to-methyl packing. All of these features probably contribute to the fact that the melting point of N,N-dimethylurea (melting at 182°C) is higher than any of the other compounds so far considered.

# PART II

# THE DETERMINATION AND USE OF CRYSTALLOGRAPHIC PARAMETERS

## Introduction

The purpose of a crystallographic investigation is to obtain a complete description of the atomic arrangement of the crystal under examination. This inquiry ultimately results in a set of experimentally determined parameters (positional and thermal) which abstractly describe the scattering power, location, and vibrational motion of each atom within any unit cell of a periodic assembly of cells. However complete this description may be, within the sophistication of the method, the set of parameters still needs to be translated into more perceptible quantities, such as interatomic distances, bond angles, planarity of groups of atoms, magnitudes and direction of the principal axes of the vibrational ellipsoids of each atom and the rigid body representation of the thermal displacements and vibrations.

The first three sections in this part are devoted to a description of one of the methods used for the determination of the parameter set, namely, the method of least-squares. The remaining five sections recount several methods for the conversion of the parameter set into other more perceptible quantities. Original work by the author is given in section ii; the other sections have been added to complement this work. These additions have been purposely selected for presentation as a collected work. Presently many of these results are to be found only in diverse locations throughout the original literature.

i) General discussion of the least-squares method.

We start with a set of observables, $|F_o|^2$, and a function $|F(\lambda)|^2$, presumed to adequately predict the observed quantities to within the accuracy of the experiment provided, of course, that the parameters $\lambda_i$ are known. We define the residue between the observed and calculated values as $v_q = |F(\lambda)|^2 - |F_o|^2$. The method of least-squares is employed to obtain the best set of $\lambda_i$ by requiring that the quantity $\sum_q w(q)v_q^2$ be a minimum, where $w(q)$ is a weighting function. (Hereafter $w$ will be written for $w(q)$.) The transcendental nature of the function $|F(\lambda)|^2$, however, prohibits the immediate application of the method of least-squares.

A circuitous route (Whittaker and Robinson 1) is used to obtain the parameter set $\left\{\lambda_i\right\}$ from an approximate set $\left\{\zeta_i\right\}$. We assume that $\lambda_i = \zeta_i + \Delta\zeta_i$, and the residual equation becomes $v_q = |F(\zeta + \Delta\zeta)|^2 - |F_o|^2$. We now expand $|F(\zeta + \Delta\zeta)|^2$ in a Taylor series about the parameters $\zeta_i$, neglecting all derivatives greater than the first. This gives a residual equation linearized in terms of the $\Delta\zeta_i$ thus:

$$vq = \Delta\zeta_1\left(\frac{\partial|F|^2}{\partial\zeta_1}\right) + \cdots + \Delta\zeta_n\left(\frac{\partial|F|^2}{\partial\zeta_n}\right) - \Delta G ,$$

where $\Delta G = |F_o|^2 - |F(\zeta)|^2$

and $\left(\frac{\partial|F|^2}{\partial\zeta}\right) = \left(\frac{\partial A^2}{\partial\zeta}\right) + \left(\frac{\partial B^2}{\partial\zeta}\right)$ etc.

This equation is linear in $\Delta\zeta_i$ and can be used to set up the least-squares normal equations:

$$\sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_1}\right)^2 \Delta\zeta_1 + \cdots + \sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_1}\right)\left(\frac{\partial|F|^2}{\partial\zeta_n}\right)\Delta\zeta_n - \sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_1}\right)\Delta G = 0$$

$$\sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_n}\right)\left(\frac{\partial|F|^2}{\partial\zeta_1}\right)\Delta\zeta_1 + \cdots + \sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_n}\right)^2 \Delta\zeta_n - \sum_q w\left(\frac{\partial|F|^2}{\partial\zeta_n}\right)\Delta G = 0$$

The parameters generally encountered in crystallographic problems are of three types: the positional parameters, which locate, in terms of fractional coordinates, the positions of the atoms in the unit cell; the isotropic or anisotropic temperature parameters, which modify the scattering function for an atom at rest to take into account temperature-induced vibrations; and the population parameters, which take into account point position multiplicities less than the space group general position multiplicity or denote partial occupancy of an atomic site. The population parameters can also be used to obtain information concerning the over-all scale factor needed to place the observed quantities $|F_o|^2$ on an absolute scale.

The large number of parameters involved frequently requires that some simplifying assumptions be made in order to reduce the size of the normal equation matrix. The usual procedure is to assume that large sections of the matrix contain zeros as elements, the explicit calculation of these terms being omitted. For machine calculations the off-diagonal terms $\left(\frac{\partial|F|^2}{\partial \zeta_i}\right)\left(\frac{\partial|F|^2}{\partial \zeta_j}\right)$ commonly omitted are those in which:

1) i and j refer to parameters of different atoms.
2) i refers to a positional parameter of an atom and j refers to a temperature or population parameter of the same atom or vice versa.
3) i and j refer to positional parameters of the same atom and $i \neq j$. (This is a poor approximation for non-orthogonal space groups.)

This set of assumptions reduces the complexity of the normal equations matrix so that positional shifts are computed from the diagonal terms only. (If assumption 3 is not made, then the shifts are computed from a 3x3 matrix for each atom.) The thermal and population parameter shifts are obtained by solving a 7x7 matrix for each atom. The overall scale factor correction may be assumed to be the average of the individual population parameter shifts, and can then be used to reduce the 7x7 matrices to 6x6 matrices which are solved for the 6 thermal parameter shifts for each atom.

If we define M to be the matrix of elements

$m_{ij} = \sum_q w\left(\frac{\partial|F|^2}{\partial \zeta_i}\right)\left(\frac{\partial|F|^2}{\partial \zeta_j}\right)$ of the least-squares normal equations and we further define $n_{ij}$ as the elements of $M^{-1}$, then the standard deviation $\sigma(\zeta_i)$ of a parameter $\zeta_i$ is equal to $[n_{ii} \sum_q w(\Delta G)^2/(r-s)]^{1/2}$, where r-s is the excess of data over the number of variables. If the assumptions (1-3) are made, then the complete matrix M will not be available. The standard deviation of a positional parameter is $s(x_i) = [(1/m_{ii}) \sum_q w(\Delta G)^2/(r-s)]^{1/2}$ and for a thermal parameter it is given as $s(B_{ij}) = [t_{ij} \sum_q w(\Delta G)^2/(r-s)]^{1/2}$, where $t_{ij}$ are the elements in the inverse matrix of the corresponding 7x7 matrix. As has been suggested by Templeton (2), s(x) is written in place of $\sigma(x)$ to indicate that the standard deviations have been obtained by an approximation of $n_{ij}$.

### ii) Structure factor and least-squares calculation for orthorhombic systems.

In the following paper, accepted for publication in Acta Crystallographica, expressions employable for the calculation of structure factors and their derivatives for any orthorhombic space group are developed. The resultant set of expressions was used as the basis of the program for the Burroughs 220 computer described in appendix I.

# Structure Factor and Least-Squares Calculation for Orthorhombic Systems with Anisotrpic Vibrations *

By Albert Hybl and Richard E. Marsh
Gates and Crellin Laboratories of Chemistry
California Institute of Technology
Pasadena, California, U.S.A.

## Abstract

A set of expressions is presented for calculating structure factors and least-squares coefficients for orthorhombic structures containing atoms with anisotropic temperature factors. These expressions are analogous to those previously derived by Rollett and Davies (1955) for monoclinic space groups.

## Introduction

The increased availability of high-speed digital computers for crystallographic calculations has made practical more accurate determinations of atomic parameters with the inclusion of anisotropic thermal vibrations of the atoms in the analysis. This increase in computer utilization has been paralleled by an increase in the use of the least-squares method (Hughes, 1941) for the refinement of the various parameters. Three features make the least-squares method particularly suitable for machine computation. First,

---

*Contribution No. 2629 from the Gates and Crellin Laboratories of Chemistry.

any desired weighting function may be prescribed. Second, totals for the normal equations can be accumulated in concurrence with the calculation of the structure factors. Third, the procedure is iterative.

The programming of digital computers for structure factors and least-squares calculations has followed two courses. The first course makes use of no lattice symmetry; by using redundant parameters and calculating symmetry equivalent reflections one can obtain results consistent with any space group (Sparks, Prosen, Kruse, and Trueblood, 1956). Proponents of this method consider the ease of coding to justify the computing inefficiency encountered in symmetric structures. Following the other course are those who program the computer for specific crystal classes (see, for example, Lavine and Rollett, 1956), feeling that the saving of computing time warrants the additional efforts invested in programming.

With this latter course in mind, Rollett and Davies (1955), hereafter RD, have derived a set of expressions that can be used to calculate structure factors and least-squares coefficients for any monoclinic space group. It is the purpose of the present paper to present an analogous, though somewhat more complicated, set of expressions that are applicable to all orthorhombic symmetries. In developing these expressions we make use of the formulations of Trueblood (1956) and the International Tables (1952).

## General (hk$\ell$) structure factor

We shall start with the expression for the scattering factor for a vibrating atom:

$$f_1 = f_o \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}k\ell),$$

where $f_o$ is the scattering factor for the atom at rest. As explained by RD, the orthorhombic symmetry gives rise, in general, to three additional orientations of the vibrational ellipsoid; and we write the corresponding scattering factors:

$$f_2 = f_o \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk - B_{13}h\ell - B_{23}k\ell)$$

$$f_3 = f_o \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 - B_{12}hk + B_{13}h\ell - B_{23}k\ell)$$

$$f_4 = f_o \exp -(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 - B_{12}hk - B_{13}h\ell + B_{23}k\ell)$$

RD also showed that the structure factor expression for any orthorhombic space group contains the term $(f_1 + f_2 + f_3 + f_4)$ and three terms in which two of $f_2$, $f_3$ and $f_4$ are negative. Accordingly we define, in Table 1, the four functions $E_1$, $E_2$, $E_3$ and $E_4$ and their derivatives with respect to the temperature coefficients. (For the isotropic case, where $f_1 = f_2 = f_3 = f_4 = f_o \exp(-B \sin^2 \theta/\lambda^2)$, $E_1 = 4f_1$ and $E_2 = E_3 = E_4 = 0$.) We see that each derivative is, except for a multiplicative constant, equal to one of the E's.

Table 1

| Definitions | $\frac{\partial E}{\partial B_{11}}$ | $\frac{\partial E}{\partial B_{22}}$ | $\frac{\partial E}{\partial B_{33}}$ | $\frac{\partial E}{\partial B_{12}}$ | $\frac{\partial E}{\partial B_{13}}$ | $\frac{\partial E}{\partial B_{23}}$ | $\frac{\partial E}{\partial B}$ |
|---|---|---|---|---|---|---|---|
| $f_1 + f_2 + f_3 + f_4 = E_1$ | $E_1$ | $E_1$ | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_1$ |
| $f_1 + f_2 - f_3 - f_4 = E_2$ | $E_2$ | $E_2$ | $E_2$ | $E_1$ | $E_4$ | $E_3$ | 0 |
| $f_1 - f_2 + f_3 - f_4 = E_3$ | $E_3$ | $E_3$ | $E_3$ | $E_4$ | $E_1$ | $E_2$ | 0 |
| $f_1 - f_2 - f_3 + f_4 = E_4$ | $E_4$ | $E_4$ | $E_4$ | $E_3$ | $E_2$ | $E_1$ | 0 |
| Multiplicative constants obtained from the differentiation | $-h^2$ | $-k^2$ | $-\ell^2$ | $-hk$ | $-h\ell$ | $-k\ell$ | $-\frac{\sin^2\theta}{\lambda^2}$ |

We next introduce the trigonometric portions of the scattering factors. There are eight possible combinations of triple products of sines and cosines, which we call $T_1 - T_8$; these, together with their products, P, with the various E's, are defined in Table 2. The resulting P's are now combined in eight different ways (see Table 3) to form the pertinent coefficients $S_1 - S_8$. With the exception of certain classes of reflections in space groups Fdd2 and Fddd, the structure factor and all the parameter derivatives for any reflection in the orthorhombic system contain as the principal factor one or another of these S's.

The cascade of definitions culminates in Table 4, which, in conjunction with Tables 5 and 6, provides the prescription for choosing the S functions appropriate to any class of reflections in any orthorhombic space group.

We now introduce the relations $|F|^2 = \rho_c^2 (A^2 + B^2)$, $\partial |F|^2/\partial \zeta = 2\rho_c^2 (A\partial A/\partial\zeta + B\partial B/\partial\zeta)$ and $A = \sum_i \rho_i A_i$, $B = \sum_i \rho_i B_i$. Here $\rho_c$ reflects the space-group multiplicity. For acentric primitive space groups $\rho_c$ has the value unity; a center of symmetry introduces a factor of two; the non-primitive lattices A, C and I contribute a factor of two; while a face-centered lattice F, a factor of four. We also introduce $\rho_i$ as a population parameter, which, if other than unity, indicates either partial occupancy of an atomic site or an atom in a special position having multiplicity less than that of the space group. By including $\rho_i$ in

Table 2

| Definition | Product with $E_1$ | $E_2$ | $E_3$ | $E_4$ |
|---|---|---|---|---|
| $\cos 2\pi hx \cdot \cos 2\pi ky \cdot \cos 2\pi \ell z = T_1$ | $P_1$ | $P_9$ | $P_{17}$ | $P_{25}$ |
| $\cos 2\pi hx \cdot \cos 2\pi ky \cdot \sin 2\pi \ell z = T_2$ | $P_2$ | $P_{10}$ | $P_{18}$ | $P_{26}$ |
| $\cos 2\pi hx \cdot \sin 2\pi ky \cdot \cos 2\pi \ell z = T_3$ | $P_3$ | $P_{11}$ | $P_{19}$ | $P_{27}$ |
| $\cos 2\pi hx \cdot \sin 2\pi ky \cdot \sin 2\pi \ell z = T_4$ | $P_4$ | $P_{12}$ | $P_{20}$ | $P_{28}$ |
| $\sin 2\pi hx \cdot \cos 2\pi ky \cdot \cos 2\pi \ell z = T_5$ | $P_5$ | $P_{13}$ | $P_{21}$ | $P_{29}$ |
| $\sin 2\pi hx \cdot \cos 2\pi ky \cdot \sin 2\pi \ell z = T_6$ | $P_6$ | $P_{14}$ | $P_{22}$ | $P_{30}$ |
| $\sin 2\pi hx \cdot \sin 2\pi ky \cdot \cos 2\pi \ell z = T_7$ | $P_7$ | $P_{15}$ | $P_{23}$ | $P_{31}$ |
| $\sin 2\pi hx \cdot \sin 2\pi ky \cdot \sin 2\pi \ell z = T_8$ | $P_8$ | $P_{16}$ | $P_{24}$ | $P_{32}$ |

Table 3

$S_1 = P_1 - P_{15} - P_{22} - P_{28}$

$S_2 = P_2 - P_{16} + P_{21} + P_{27}$

$S_3 = P_3 + P_{13} - P_{24} + P_{26}$

$S_4 = P_4 + P_{14} + P_{23} - P_{25}$

$S_5 = P_5 + P_{11} + P_{18} - P_{32}$

$S_6 = P_6 + P_{12} - P_{17} + P_{31}$

$S_7 = P_7 - P_9 + P_{20} + P_{30}$

$S_8 = P_8 - P_{10} - P_{19} - P_{29}$

Table 4

| Multiplicative Constants | Functions | Group* | | | | Type I | Type II |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | | |
| 1 | $A_i$ | $+S_1$ | $-S_4$ | $-S_7$ | $-S_6$ | ------ | $\partial B/\partial z$ |
| $2\pi\rho ih$ | $\partial A/\partial x$ | $-S_5$ | $+S_8$ | $-S_3$ | $-S_2$ | $-\partial B/\partial B_{23}$ | $-\partial B/\partial B_{13}$ |
| $2\pi\rho ik$ | $\partial A/\partial y$ | $-S_3$ | $-S_2$ | $-S_5$ | $+S_8$ | $-\partial B/\partial B_{13}$ | $-\partial B/\partial B_{23}$ |
| $2\pi\rho i\ell$ | $\partial A/\partial z$ | $-S_2$ | $-S_3$ | $+S_8$ | $-S_5$ | $-\partial B/\partial B_{12}$ | $-B_i$ |
| $-hk\rho i$ | $\partial A/\partial B_{12}$ | $-S_7$ | $-S_6$ | $+S_1$ | $-S_4$ | $\partial B/\partial z$ | ------ |
| $-h\ell\rho i$ | $\partial A/\partial B_{13}$ | $-S_6$ | $-S_7$ | $-S_4$ | $+S_1$ | $\partial B/\partial y$ | $\partial B/\partial x$ |
| $-k\ell\rho i$ | $\partial A/\partial B_{23}$ | $-S_4$ | $+S_1$ | $-S_6$ | $-S_7$ | $\partial B/\partial x$ | $\partial B/\partial y$ |
| 1 | ------ | $+S_8$ | $-S_5$ | $-S_2$ | $-S_3$ | $-B_i$ | $-\partial B/\partial B_{12}$ |
| $-h^2\rho i$ | $\partial A/\partial B_{11}$ | same as for $A_i$ & $B_i$ | | | | | |
| $-k^2\rho i$ | $\partial A/\partial B_{22}$ | " | | | | | |
| $-\ell^2\rho i$ | $\partial A/\partial B_{33}$ | " | | | | | |
| $-\rho i \sin^2\theta/\lambda^2$ | $\partial A/\partial B$ | " | | | | | |
| 1 | $\partial A/\partial\rho i$ | " | | | | | |

*$V_i$ or $W_i$ replace $S_i$ whenever specified by Table 5.

Type III: The imaginary part of the expressions are all zero.

Table 5

| Space Group | | Class | | Planes | Group Key |
|---|---|---|---|---|---|
| No. | Symbol | Type | $\rho_c$ | | |
| 16 | P222 | I | 1 | All planes | 1 |
| 17 | $P222_1$ | I | 1 | $\ell$ even | 1 |
| | | | | $\ell$ odd | 2 |
| 18 | $P2_12_12$ | I | 1 | (h+k) even | 1 |
| | | | | (h+k) odd | 3 |
| 19 | $P2_12_12_1$ | I | 1 | (h+k) even, (k+$\ell$) even | 1 |
| | | | | (h+k) even, (k+$\ell$) odd | 2 |
| | | | | (h+k) odd, (k+$\ell$) odd | 3 |
| | | | | (h+k) odd, (k+$\ell$) even | 4 |
| 20 | $C222_1$ | I | 2 | (h+k) even, $\ell$ even | 1 |
| | | | | (h+k) even, $\ell$ odd | 2 |
| 21 | C222 | I | 2 | (h+k) even | 1 |
| 22 | F222 | I | 4 | (h+k) even and (k+$\ell$) even | 1 |
| 23 | I222 | I | 2 | (h+k+$\ell$) even | 1 |
| 24 | $I2_12_12_1$ | I | 2 | h even and k even and $\ell$ even | 1 |
| | | | | h even and k odd and $\ell$ odd | 4 |
| | | | | h odd and k even and $\ell$ odd | 3 |
| | | | | h odd and k odd and $\ell$ even | 2 |
| 25 | Pmm2 | II | 1 | All planes | 1 |
| 26 | $Pmc2_1$ | II | 1 | $\ell$ even | 1 |
| | | | | $\ell$ odd | 2 |
| 27 | Pcc2 | II | 1 | $\ell$ even | 1 |
| | | | | $\ell$ odd | 3 |
| 28 | Pma2 | II | 1 | h even | 1 |
| | | | | h odd | 3 |
| 29 | $Pca2_1$ | II | 1 | h even and $\ell$ even | 1 |
| | | | | h even and $\ell$ odd | 4 |
| | | | | h odd and $\ell$ even | 3 |
| | | | | h odd and $\ell$ odd | 2 |

Table 5 (continued)

| Space Group | | Class | | Planes | Group Key |
|---|---|---|---|---|---|
| No. | Symbol | Type | $\rho_c$ | | |
| 30 | Pnc2 | II | 1 | $(k+\ell)$ even | 1 |
| | | | | $(k+\ell)$ odd | 3 |
| 31 | $Pmn2_1$ | II | 1 | $(h+\ell)$ even | 1 |
| | | | | $(h+\ell)$ odd | 2 |
| 32 | Pba2 | II | 1 | $(h+k)$ even | 1 |
| | | | | $(h+k)$ odd | 3 |
| 33 | $Pna2_1$ | II | 1 | $(h+k)$ even and $\ell$ even | 1 |
| | | | | $(h+k)$ even and $\ell$ odd | 4 |
| | | | | $(h+k)$ odd and $\ell$ even | 3 |
| | | | | $(h+k)$ odd and $\ell$ odd | 2 |
| 34 | Pnn2 | II | 1 | $(h+k+\ell)$ even | 1 |
| | | | | $(h+k+\ell)$ odd | 3 |
| 35 | Cmm2 | II | 2 | $(h+k)$ even | 1 |
| 36 | $Cmc2_1$ | II | 2 | $(h+k)$ even and $\ell$ even | 1 |
| | | | | $(h+k)$ even and $\ell$ odd | 2 |
| 37 | Ccc2 | II | 2 | $(h+k)$ even and $\ell$ even | 1 |
| | | | | $(h+k)$ even and $\ell$ odd | 3 |
| 38 | Amm2 | II | 2 | $(k+\ell)$ even | 1 |
| 39 | Abm2 | II | 2 | $(k+\ell)$ even, $k$ even | 1 |
| | | | | $(k+\ell)$ even, $k$ odd | 3 |
| 40 | Ama2 | II | 2 | $(k+\ell)$ even, and $h$ even | 1 |
| | | | | $(k+\ell)$ even, and $h$ odd | 3 |
| 41 | Aba2 | II | 2 | $(k+\ell)$ even, $(h+k)$ even | 1 |
| | | | | $(k+\ell)$ even, $(h+k)$ odd | 3 |
| 42 | Fmm2 | II | 4 | $(h+k)$ even, $(k+\ell)$ even | 1 |
| 43 | Fdd2 | II | 4 | $h+k+\ell = 4n$, $(h+\ell)$ even, $(k+\ell)$ even | 1 |
| | | | | $h+k+\ell = 4n+1$, $(h+\ell)$ even, $(k+\ell)$ even | 1* |
| | | | | $h+k+\ell = 4n+2$, $(h+\ell)$ even, $(k+\ell)$ even | 3 |
| | | | | $h+k+\ell = 4n+3$, $(h+\ell)$ even, $(k+\ell)$ even | 3* |

Table 5 (continued)

| Space Group No. | Space Group Symbol | Class Type | Class $\rho_c$ | Planes | Group Key |
|---|---|---|---|---|---|
| 44 | Imm2 | II | 2 | $(h+k+\ell)$ even | 1 |
| 45 | Iba2 | II | 2 | $(h+k+\ell)$ even and $\ell$ even | 1 |
| | | | | $(h+k+\ell)$ even and $\ell$ odd | 3 |
| 46 | Iam2 | II | 2 | $(h+k+\ell)$ even and h even | 1 |
| | | | | $(h+k+\ell)$ even and h odd | 3 |
| 47 | Pmmm | III | 2 | All planes | 1 |
| 48 | Pnnn | III | 2 | $(h+k)$ even and $(k+\ell)$ even | 1 |
| | | | | $(h+k)$ even and $(k+\ell)$ odd | 3 |
| | | | | $(h+k)$ odd and $(k+\ell)$ even | 2 |
| | | | | $(h+k)$ odd and $(k+\ell)$ odd | 4 |
| 49 | Pccm | III | 2 | $\ell$ even | 1 |
| | | | | $\ell$ odd | 3 |
| 50 | Pbam | III | 2 | h even and k even | 1 |
| | | | | h even and k odd | 4 |
| | | | | h odd and k even | 2 |
| | | | | h odd and k odd | 3 |
| 51 | Pmma | III | 2 | h even | 1 |
| | | | | h odd | 4 |
| 52 | Pnna | III | 2 | h even and $(k+\ell)$ even | 1 |
| | | | | h even and $(k+\ell)$ odd | 3 |
| | | | | h odd and $(k+\ell)$ even | 2 |
| | | | | h odd and $(k+\ell)$ odd | 4 |
| 53 | Pmna | III | 2 | $(h+\ell)$ even | 1 |
| | | | | $(h+\ell)$ odd | 2 |
| 54 | Pcca | III | 2 | h even and $\ell$ even | 1 |
| | | | | h even and $\ell$ odd | 3 |
| | | | | h odd and $\ell$ even | 4 |
| | | | | h odd and $\ell$ odd | 2 |
| 55 | Pbam | III | 2 | $(h+k)$ even | 1 |
| | | | | $(h+k)$ odd | 3 |

Table 5 (continued)

| Space Group | | Class | | Planes | Group Key |
|---|---|---|---|---|---|
| No. | Symbol | Type | $\rho_c$ | | |
| 56 | Pccn | III | 2 | (h+k) even and (h+ℓ) even | 1 |
| | | | | (h+k) even and (h+ℓ) odd | 3 |
| | | | | (h+k) odd and (h+ℓ) even | 2 |
| | | | | (h+k) odd and (h+ℓ) odd | 4 |
| 57 | Pbcm | III | 2 | k even and ℓ even | 1 |
| | | | | k even and ℓ odd | 2 |
| | | | | k odd and ℓ even | 3 |
| | | | | k odd and ℓ odd | 4 |
| 58 | Pnnm | III | 2 | (h+k+ℓ) even | 1 |
| | | | | (h+k+ℓ) odd | 3 |
| 59 | Pmmn | III | 2 | h even and k even | 1 |
| | | | | h even and k odd | 2 |
| | | | | h odd and k even | 4 |
| | | | | h odd and k odd | 3 |
| 60 | Pbcn | III | 2 | (h+k) even and ℓ even | 1 |
| | | | | (h+k) even and ℓ odd | 2 |
| | | | | (h+k) odd and ℓ even | 4 |
| | | | | (h+k) odd and ℓ odd | 3 |
| 61 | Pbca | III | 2 | (h+k) even and (k+ℓ) even | 1 |
| | | | | (h+k) even and (k+ℓ) odd | 2 |
| | | | | (h+k) odd and (k+ℓ) even | 4 |
| | | | | (h+k) odd and (k+ℓ) odd | 3 |
| 62 | Pnma | III | 2 | (h+ℓ) even and k even | 1 |
| | | | | (h+ℓ) even and k odd | 3 |
| | | | | (h+ℓ) odd and k even | 4 |
| | | | | (h+ℓ) odd and k odd | 2 |
| 63 | Cmcm | III | 4 | (h+k) even and ℓ even | 1 |
| | | | | (h+k) even and ℓ odd | 2 |
| 64 | Cmca | III | 4 | (h+k) even and (k+ℓ) even | 1 |
| | | | | (h+k) even and (k+ℓ) odd | 2 |
| 65 | Cmmm | III | 4 | (h+k) even | 1 |

Table 5 (continued)

| Space Group | | Class | | Planes | Group Key |
|---|---|---|---|---|---|
| No. | Symbol | Type | $\rho_c$ | | |
| 66 | Cccm | III | 4 | $(h+k)$ even and $\ell$ even | 1 |
| | | | | $(h+k)$ even and $\ell$ odd | 3 |
| 67 | Cmma | III | 4 | $(h+k)$ even and $h$ even | 1 |
| | | | | $(h+k)$ even and $h$ odd | 2 |
| 68 | Ccca | III | 4 | $(h+k)$ even and $k$ even and $\ell$ even | 1 |
| | | | | $(h+k)$ even and $k$ even and $\ell$ odd | 3 |
| | | | | $(h+k)$ even and $k$ odd and $\ell$ even | 4 |
| | | | | $(h+k)$ even and $k$ odd and $\ell$ odd | 2 |
| 69 | Fmmm | III | 8 | $(h+k)$ even and $(k+\ell)$ even | 1 |
| 70 | Fddd | III | 8 | $h+k = 4n$, $\ell+h = 4n$, $k+\ell = 4n$ | 1 |
| | | | | $h+k = 4n$, $\ell+h = 4n+2$, $k+\ell = 4n+2$ | 3 |
| | | | | $h+k = 4n+2$, $\ell+h = 4n+2$, $k+\ell = 4n$ | 2 |
| | | | | $h+k = 4n+2$, $\ell+h = 4n$, $k+\ell = 4n+2$ | 4 |
| | | | | $h+k = 4n$, $\ell+h = 4n+2$, $k+\ell = 4n$ | 1** |
| | | | | $h+k = 4n$, $\ell+h = 4n$, $k+\ell = 4n+2$ | 3** |
| | | | | $h+k = 4n+2$, $\ell+h = 4n$, $k+\ell = 4n$ | 2** |
| | | | | $h+k = 4n+2$, $\ell+h = 4n+2$, $k+\ell = 4n+2$ | 4** |
| 71 | Immm | III | 4 | $(h+k+\ell)$ even | 1 |
| 72 | Ibam | III | 4 | $(h+k+\ell)$ even, $\ell$ even | 1 |
| | | | | $(h+k+\ell)$ even, $\ell$ odd | 3 |
| 73 | Ibca | III | 4 | $(h+k+\ell)$ even, $h$ even, $k$ even | 1 |
| | | | | $(h+k+\ell)$ even, $h$ even, $k$ odd | 4 |
| | | | | $(h+k+\ell)$ even, $h$ odd, $k$ even | 3 |
| | | | | $(h+k+\ell)$ even, $h$ odd, $k$ odd | 2 |
| 74 | Imma | III | 4 | $(h+k+\ell)$ even, $k$ even | 1 |
| | | | | $(h+k+\ell)$ even, $k$ odd | 2 |

---

*Use $V_{1-8}$ in place of $S_{1-8}$.

**Use $W_{1-8}$ in place of $S_{1-8}$.

Table 6

$V_1 = 0.5\,(S_1-S_2-S_7-S_8)$

$V_2 = 0.5\,(S_1+S_2+S_7-S_8)$

$V_3 = 0.5\,(S_3-S_4+S_5+S_6)$

$V_4 = 0.5\,(S_3+S_4-S_5+S_6)$

$V_5 = 0.5\,(S_3+S_4+S_5-S_6)$

$V_6 = 0.5\,(-S_3+S_4+S_5+S_6)$

$V_7 = 0.5\,(-S_1-S_2+S_7-S_8)$

$V_8 = 0.5\,(S_1-S_2+S_7+S_8)$

$W_1 = 0.5\,(S_1-S_4+S_6-S_7)$

$W_2 = 0.5\,(S_2+S_3-S_5-S_8)$

$W_3 = 0.5\,(S_2+S_3+S_5+S_8)$

$W_4 = 0.5\,(-S_1+S_4+S_6-S_7)$

$W_5 = 0.5\,(-S_2+S_3+S_5-S_8)$

$W_6 = 0.5\,(S_1+S_4+S_6+S_7)$

$W_7 = 0.5\,(-S_1-S_4+S_6+S_7)$

$W_8 = 0.5\,(-S_2-S_3-S_5-S_8)$

the least-squares treatment, one can obtain information concerning either the over-all scale factor or the degree of occupancy of the site.

We suggest that when using these tables one start by finding the appropriate space group listing in Table 5. From there one obtains the value of $\rho_c$, the type designation and the group key as a function of the index parity tests. The left-hand column of Table 4 contains the constant that multiplies both parts (A and B) of the relation whose A term is listed under "Functions." The value of $A_i$ and the derivatives of A with respect to the various positional and temperature-factor parameters $(\partial A/\partial \zeta)$ can then be selected according to which of the four groups the particular reflection falls into. The terms involving B, however, require that we first separate the space groups into the three crystal classes 222-$D_2$ (denoted Type I in Tables 4 and 5), <u>mm</u>2-$C_{2v}$ (Type II), and <u>mmm</u>-$D_{2h}$ (Type III). The last two columns of Table 4 provide the key for choosing the correct functions of B for space groups of Type I and II; since those of Type III are centrosymmetric, no B term is necessary (provided the origin of coordinates is chosen at the center of symmetry, as is customary).

## Application

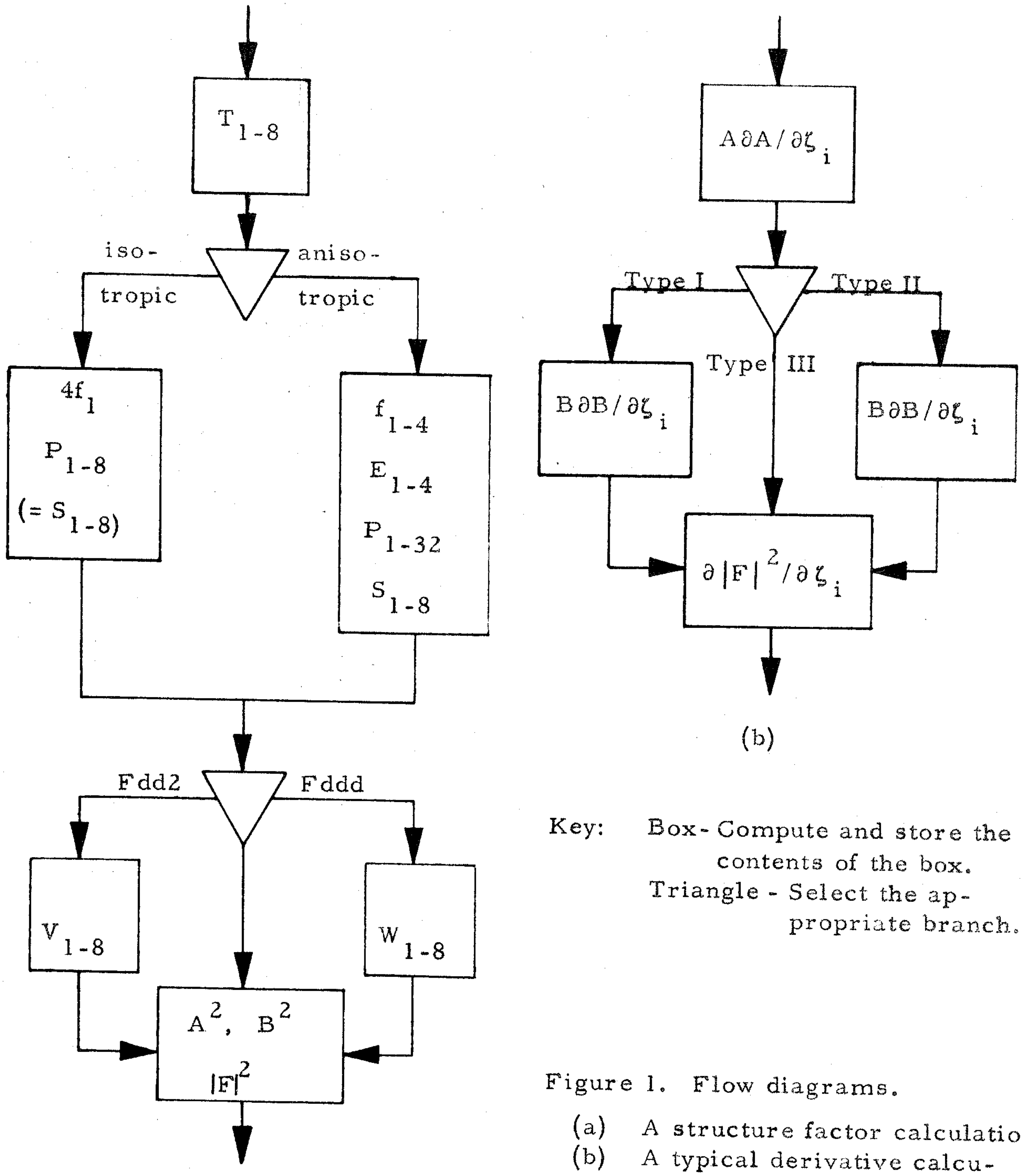
As an example illustrating the use of Tables 4 and 5, let us consider the reflections with <u>k</u> and <u>$\ell$</u> odd in space group Abm2-$C_{2v}^{15}$. From Table 5 we find that this space group is Type II and the reflections are in group 3; accordingly, $A_i = (-S_7)_i$, $B_i = -(+S_8)_i$, and $\partial |F|^2/\partial B_{13i} = -2(2^2)h\ell \rho_i \left\{ A(-S_4)_i + B\,[-(-S_3)_i\,] \right\}$.

Certain classes of reflections in space groups Fdd2 and Fddd require separate treatment. For them it is convenient to use additional functions $V_{1-8}$ or $W_{1-8}$, as defined in Table 6, which are substituted for the corresponding S functions and used as prescribed in Table 4.

A suggested procedure for calculating structure factors and derivatives is illustrated in Figure 1. For each atom i, the trigonometric triple products $(T_{1-8})_i$ are first calculated. The intermediate products $(P_1\text{-}P_{32})_i$ are then formed, and these are combined to give the coefficients $(S_1\text{-}S_8)_i$. (For an isotropic atom, only $P_1\text{-}P_8$ need be calculated since $P_1\text{-}P_8 = S_1\text{-}S_8$.) Space group and index parity tests are then made, and if necessary the additional functions $(V_1\text{-}V_8)_i$ or $(W_1\text{-}W_8)_i$ are calculated. The correct structure factor and derivative terms are then selected in accordance with Table 4.

We have used these orthorhombic expressions and also the monoclinic expressions of RD as the bases for two separate structure factor and least-squares programs for the Burroughs 220 computer. This computer has an access time of approximately 100 microseconds; a complete structure factor least-squares calculation, including the collecting of 7 x 7 matrices involving scale and temperature-factor derivatives for each atom, takes approximately 0.25 seconds per atom reflection. Isotropic atoms, for which many of the calculations can be by-passed, require less than half this time.

One of the authors (A.H.) wishes to acknowledge the tenure of a Shell Fellowship.



(a)

(b)

Key: Box - Compute and store the contents of the box.
Triangle - Select the appropriate branch.

Figure 1. Flow diagrams.

(a) A structure factor calculation.
(b) A typical derivative calculation.

# References

International Tables for X-ray Crystallography, Vol. I (1952) Birmingham: Kynoch Press.

Hughes, E. W. (1941) J. Am. Chem. Soc., 63, 1737.

Lavine, L. R. and Rollett, J. S., (1956). Acta Cryst. 9, 269.

Rollett, J. S. and Davies, D. R., (1955). Acta Cryst. 8, 125.

Sparks, R. A., Prosen, R. J., Kruse, F. H. and Trueblood, K. N., (1956). Acta Cryst. 9, 350.

Trueblood, K. N. (1956). Acta Cryst. 9, 359.

iii) Computer approximations frequently employed.

As in many other computer applications, the crystallographic least-squares programs require some basic function subroutines. The approximating expressions employed depend on one or more of the following considerations: The accuracy desired, the storage space available and the machine time needed to execute the calculations. In general, some compromise between use of space-filling look-up tables and of time-consuming polynomial approximations is employed. The expressions used in the function subroutines for the Burroughs 220 orthorhombic least-squares program (see appendix I) are listed in the following paragraphs.

The values for sine or cosine $2\pi hy$ are computed by using a linear interpolation of a 51-word look-up table. The arguments of the trigonometric functions are first reduced to the appropriate argument of the corresponding cosine function in the interval $0 \leqq 0.nmx \leqq 0.5$; the relations employed are:

$$\sin, \cos 2\pi hy = \cos 2\pi\, 0.nmx \doteq A(0.nm) + 0.00x\, B(0.nm) .$$

The values of A(0.nm) and B(0.nm) are obtained by requiring

$$\int_{0.nm}^{0.nm+0.01} [A(0.nm) + (\zeta - 0.nm)\, B(0.nm) - \cos 2\pi\zeta\,]^2\, d\zeta$$

be a minimum. The two normal equations obtained by differentiation of the above integral with respect to A and B, respectively, are

$$\int_{0.nm}^{0.nm+0.01} [A(0.nm) + (\zeta - 0.nm)\, B(0.nm) - \cos 2\pi\zeta ]\, d\zeta = 0$$

$$\int_{0.nm}^{0.nm+0.01} (\zeta - 0.nm)\, [A(0.nm) + (\zeta - 0.nm)\, B(0.nm) - \cos 2\pi\zeta ]\, d\zeta = 0$$

These are solved for $A(0.nm)$ and $B(0.nm)$ in each of the 51 intervals.

The exponential function $e^{-x}$ is computed using the expression given by Hastings (4):

$$e^{-x} \doteq 1/(1 + 0.2507213\, x + 0.0292732\, x^2 + 0.0038278\, x^3)^4 .$$

The maximum absolute error for $e^{-x}$ in the interval $0 \leqq x < \infty$ is $\pm$ 0.0002.

The scattering factor values are computed using a quadratic interpolation formula within ten regions of reciprocal space for each atom type. The quadratic expression is

$$f(0.nx) \doteq a(n) + 0.0x\, b(n) + (0.0x)^2\, c(n) ,$$

where $0.nx$ is equal to $\sin\theta$ for Mo radiation ($\lambda = 0.7107$), and $a(n)$, $b(n)$ and $c(n)$ are the coefficients for the best quadratic approximation to the form factor in the interval $0.n \leqq 0.nx < 0.(n+1)$ for the atom kind in question. These can be determined by a least-squares fit of published scattering values (Webb 5).

Square roots are extracted by use of Newton's formula $\sqrt{A} \doteq x_{n+1} = 0.5(A/x_n + x_n)$ , where $x_o = 1$. The iterations are terminated when the first five significant digits of $x_n$ and $x_{n+1}$ are identical.

iv) Orthogonalized coordinates and analytic geometry.

It is convenient and sometimes necessary to orthogonalize the fractional unit-cell coordinates into a Cartesian system. When this has been done, the resultant coordinates can be easily used to make a variety of different geometric calculations, such as interatomic distances, bond angles, best planes, and such. Defined below is the orthogonalization matrix D which transforms the fractional crystal coordinates x, y and z into a Cartesian set X, Y and Z. The Y and Z axes in this representation are coincident with the b and c* crystal axes, respectively, while X is in the ab plane perpendicular to bc*.

$$D = \begin{pmatrix} a \sin\gamma & 0 & c(\cos\beta - \cos\alpha\cos\gamma)/\sin\gamma \\ a\cos\gamma & b & c\cos\alpha \\ 0 & 0 & V/ab\sin\gamma \end{pmatrix}$$

thus $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = D \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = D^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ , where

$$D^{-1} = \begin{pmatrix} 1/(a\sin\gamma) & 0 & -bc(\cos\beta - \cos\alpha\cos\gamma)/V\sin\gamma \\ -\cos\gamma/b\sin\gamma & 1/b & -ac(\cos\gamma\cos\beta - \cos\alpha)/V\sin\gamma \\ 0 & 0 & ab\sin\gamma/V \end{pmatrix}$$

and $V = abc(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{1/2} = |D|$.

The interatomic distance between atoms $P_1(X_1Y_1Z_1)$ and $P_2(X_2Y_2Z_2)$ is $d_1 = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2}$. Further, if we define $d_2$ as the interatomic distance $P_1$ to $P_3$ and $d_3$ as the interatomic distance $P_2$ to $P_3$, then the bond angle $\theta$ between atoms $P_1$-$P_2$-$P_3$ is given by

$$\theta = \cos^{-1}[(d_1^2 + d_2^2 - d_3^2)/2d_1d_2] \quad , \text{ where } 0 \leq \theta \leq 180°.$$

Several other useful analytic relations can be found in any text on the subject; one that has been very useful to the author is "Solid Analytic Geometry" by J. Olmsted (6).

v) Least-squares planes.

It is sometimes desired to fit a least-squares plane to a set of atoms found to be approximately coplanar. This problem has been rigorously treated by Schomaker, Waser, Marsh and Bergman (7).

We define the plane as $m_1X + m_2Y + m_3Z - d = 0$, where $m_1$, $m_2$ and $m_3$ are the direction cosines of the plane normal and $d$ is the distance from the plane to the origin. A Cartesian coordinate system is assumed. The distance $d_k$ from the plane to an atom $P_k(X_kY_kZ_k)$ is then given by $d_k = m_1X_k + m_2Y_k + m_3Z_k - d$. To find the best plane we require that $\sum_k w_kd_k^2 = \sum_k w_k(m_1X_k + m_2Y_k + m_3Z_k - d)^2$ be a minimum subject, as we shall see, to one side condition. The problem of the plane is essentially equivalent to the problem of finding the principal plane of a set of mass points with the least inertia. To find the principal planes of inertia we introduce the side condition that $m_1^2 + m_2^2 + m_3^2 = 1$ by use of the Lagrange multiplier $\lambda$. Thus we define

$$F = \sum_k w_k(m_1X_k + m_2Y_k + m_3Z_k - d)^2 - \lambda (m_1^2 + m_2^2 + m_3^2)$$

The normal equations resulting from the differentiation of the above expression are

$$\frac{1}{2}\left(\frac{\partial F}{\partial m_1}\right) = \sum_k w_kX_k(m_1X_k + m_2Y_k + m_3Z_k - d) - \lambda m_1 = 0$$

$$\frac{1}{2}\left(\frac{\partial F}{\partial m_2}\right) = \sum_k w_kY_k(m_1X_k + m_2Y_k + m_3Z_k - d) - \lambda m_2 = 0$$

$$\frac{1}{2}\left(\frac{\partial F}{\partial m_3}\right) = \sum_k w_kZ_k(m_1X_k + m_2Y_k + m_3Z_k - d) = \lambda m_3 = 0$$

and

$$-\frac{1}{2}\left(\frac{\partial F}{\partial d}\right) = \sum_k w_k(m_1X_k + m_2Y_k + m_3Z_k - d) = 0$$

The last equation can be readily solved for d:

$$d = \frac{\sum_k w_k(m_1X_k + m_2Y_k + m_3Z_k)}{\sum_k w_k} = m_1\bar{X} + m_2\bar{Y} + m_3\bar{Z} \quad \text{, where}$$

$\bar{X} = \sum_k w_kX_k / \sum_k w_k$ etc. Thus if we make the transformations $x_k = X_k - \bar{X}$, $y_k = Y_k - \bar{Y}$ and $z_k = Z_k - \bar{Z}$, we eliminate d from the first three normal equations which are reduced to

$$(\sum_k w_k x_k^2 - \lambda)m_1 + \sum_k w_k x_k y_k m_2 + \sum_k w_k x_k z_k m_3 = 0$$

$$\sum_k w_k x_k y_k m_1 + (\sum_k w_k y_k^2 - \lambda)m_2 + \sum_k w_k y_k z_k m_3 = 0$$

$$\sum_k w_k x_k z_k m_1 + \sum_k w_k y_k z_k m_2 + (\sum_k w_k z_k^2 - \lambda)m_3 = 0$$

We write this in the matrix form

$$\begin{pmatrix} A_{11} - \lambda & A_{12} & A_{13} \\ A_{12} & A_{22} - \lambda & A_{23} \\ A_{13} & A_{23} & A_{33} - \lambda \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \underline{A} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = 0$$

In order to solve this secular equation, we first expand it to $-\lambda^3 + a\lambda^2 + b\lambda + c = 0$, where $a = (A_{11} + A_{22} + A_{33})$, $b = -(A_{11}A_{22} + A_{11}A_{33} + A_{22}A_{33}) + (A_{12}^2 + A_{13}^2 + A_{23}^2)$ and $c = A_{11}A_{22}A_{33} + 2A_{12}A_{13}A_{23} - (A_{11}A_{23}^2 + A_{22}A_{13}^2 + A_{33}A_{12}^2)$.

The cubic equation can, of course, be solved directly, but this method is tedious. Since $\lambda$ is a sum of squares, all solutions of the cubic must be positive and since $\underline{A}$ is a symmetric matrix all three solutions are real. The desired solution is the least of these. Blow (7) has shown that an approximate value to the desired solution $\lambda^{(1)}$ is $\lambda_o^{(1)} = [-b - (b^2 - 4ac)^{1/2}] / 2a$. It is clear that this represents a solution provided by a parabola which approximates to the cubic at small $\lambda$. The form of $\lambda_o^{(1)}$ gives the smaller of the two solutions.

This solution may be refined by Newton-Raphson method (Scarborough 9). The n+1 st refined value is obtained from the n th value by evaluating:

$$\lambda_{n+1}^{(1)} = \lambda_n - \frac{(-\lambda_n^3 + a\lambda_n^2 + b\lambda_n + c)}{(-3\lambda_n^2 + 2a\lambda_n + b)} .$$

Because the slope of the cubic at $\lambda_o^{(1)}$ must be steeper than that of the quadratic, refinement by the Newton-Raphson method will always lead to the smallest solution.

Values of $m_1$, $m_2$ and $m_3$ are obtained from substitution of $\lambda^{(1)}$ into matrix $\underline{A}$ and normalizing so that $m_1^2 + m_2^2 + m_3^2 = 1$.

vi) Orthogonalization and rotation of temperature parameter axis.

(a) Orthogonalization.

By least-squares analysis it is possible to determine the scattering factor of an atom in the form:

$$f = f_o T = f_o \exp^-(B_{11}h^2 + B_{22}k^2 + B_{33}\ell^2 + B_{12}hk + B_{13}h\ell + B_{23}k\ell) .$$

To make use of the constants $B_{11}$, $B_{22}$, etc., it is convenient to first transform them to orthogonalized reciprocal axes. Further, it is appropriate to demand that these Cartesian axes be identical to the axes obtained by orthogonalization of the positional parameters.

We wish the Y and Z axes to be coincident with the b and c* crystal axes respectively, while X is in the ab plane perpendicular to bc*. Consequently:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} a^* \sin\beta^* & b^*(\cos\gamma^* - \cos\alpha^* \cos\beta^*)/\sin\beta^* & 0 \\ 0 & V^*/a^*c^* \sin\beta^* & 0 \\ a^* \cos\beta^* & b^* \cos\alpha^* & c^* \end{pmatrix} \begin{pmatrix} h \\ k \\ \ell \end{pmatrix}$$

We define the 3x3 orthogonalization matrix as $\underline{Q}$. Table 7 gives several relations between the real and reciprocal cell which are useful for the explicit evaluation of the elements of $\underline{Q}$. A discussion of the geometry of oblique cells and their reciprocals is given in Buerger (10). An example for the monoclinic system is given on page 26 of part I of this thesis.

Since we will need $\underline{Q}^{-1}$ later in our discussion, we shall compute it now:

$$\underline{Q}^{-1} = \begin{pmatrix} 1/a^*\sin\beta^* & -b^*c^*(\cos\gamma^* - \cos\alpha^*\cos\beta^*)/V^*\sin\beta^* & 0 \\ 0 & a^*c^*\sin\beta^*/V^* & 0 \\ -\cos\beta^*/c^*\sin\beta^* & -a^*b^*(\cos\beta^*\cos\gamma^* - \cos\alpha^*)/V^*\sin\beta^* & 1/c^* \end{pmatrix}$$

Thus $\begin{pmatrix} h \\ k \\ \ell \end{pmatrix} = Q^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$ .

The temperature factor T expressed in matrix form becomes:

$$-\ln T = (hk\ell) \begin{pmatrix} B_{11} & B_{12}/2 & B_{13}/2 \\ B_{12}/2 & B_{22} & B_{23}/2 \\ B_{13}/2 & B_{23}/2 & B_{33} \end{pmatrix} \begin{pmatrix} h \\ k \\ \ell \end{pmatrix} = (hk\ell)\, \underline{B}_{ij} \begin{pmatrix} h \\ k \\ \ell \end{pmatrix}.$$

Table 7. - Real and reciprocal cell relations computed from $a, b, c, \cos\alpha$, $\cos\beta$ and $\cos\gamma$.

$\sin\alpha = (1 - \cos^2\alpha)^{1/2}$; likewise for $\sin\beta$ and $\sin\gamma$

$V = abc(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)^{1/2}$

$a^* = bc\sin\alpha/V$

$b^* = ac\sin\beta/V$

$c^* = ab\sin\gamma/V$

$\cos\alpha^* = (\cos\beta\cos\gamma - \cos\alpha)/\sin\beta\sin\gamma$

$\cos\beta^* = (\cos\alpha\cos\gamma - \cos\beta)/\sin\alpha\sin\gamma$

$\cos\gamma^* = (\cos\alpha\cos\beta - \cos\gamma)/\sin\alpha\sin\beta$

$\sin\alpha^* = (1 - \cos^2\alpha^*)^{1/2}$; likewise for $\sin\beta^*$ and $\sin\gamma^*$

$V^* = 1/V$

The orthogonalization transforms the equation as follows:

$$-\ln T = (hk\ell)\,\underline{B}_{ij}\begin{pmatrix}h\\k\\\ell\end{pmatrix} = (XYZ)\,(\underline{Q}^{-1})^{T}\,\underline{B}_{ij}\,Q^{-1}\begin{pmatrix}X\\Y\\Z\end{pmatrix} = (XYZ)\,\underline{B}'_{ij}\begin{pmatrix}X\\Y\\Z\end{pmatrix},$$

where $(\underline{Q}^{-1})^{T}$ is the transpose of $Q^{-1}$. The $\underline{B}'_{ij}$ matrix is then the orthogonalized temperature parameter matrix.

(b) Rotation of the orthogonalized temperature parameter ellipsoid axes.

It is sometimes convenient to express the temperature parameters with respect to orthogonal molecular axes different from the orthogonal axes described above. If we define the molecular axes as $P_1$, $P_2$ and $P_3$ and let $\ell_1$, $m_1$ and $n_1$ be the direction cosines of the $P_1$ axis relative to the X, Y and Z axes respectively, then

$$\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix} = \begin{pmatrix}\ell_1 & m_1 & n_1\\ \ell_2 & m_2 & n_2\\ \ell_3 & m_3 & n_3\end{pmatrix}\begin{pmatrix}X\\Y\\Z\end{pmatrix} = \underline{R}\begin{pmatrix}X\\Y\\Z\end{pmatrix} \quad \text{hence} \begin{pmatrix}X\\Y\\Z\end{pmatrix} = \underline{R}^{-1}\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix}.$$

Utilizing these transformations gives:

$$-\ln T = (XYZ)\,\underline{B}'_{ij}\begin{pmatrix}X\\Y\\Z\end{pmatrix} = (P_1P_2P_3)\,\underline{R}\,\underline{B}'_{ij}\,\underline{R}^{-1}\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix} = (P_1P_2P_3)\,\underline{B}''_{ij}\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix}.$$

(c) Rescaling of the $\underline{B}_{ij}$ matrix elements.

The elements of the $\underline{B}_{ij}$ matrices discussed above are in units of $\sin^2\theta/\lambda^2$. It is desirable at times that they be in units of $\mathring{A}^2$ which is more readily interpreted in terms of mean square displacements. The rescaling may be written:

$$-\ln T = (P_1P_2P_3)\,\underline{S}\,\underline{B}''_{ij}\,\underline{S}\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix} = (P_1P_2P_3)\,\underline{U}_{ij}\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix} \quad , \text{ where}$$

$$\underline{S} = \underline{S}^T = \begin{pmatrix}1/\pi\sqrt{2} & 0 & 0\\ 0 & 1/\pi\sqrt{2} & 0\\ 0 & 0 & 1/\pi\sqrt{2}\end{pmatrix} .$$

vii) Magnitudes and principal axes of temperature ellipsoids.

The anisotropic temperature factor of a given atom represented in an orthogonalized form is:

$-\ln T = (XYZ)\,\underline{B}'_{ij}\begin{pmatrix}X\\Y\\Z\end{pmatrix}$ . The $B'_{ij} = B'_{ji}$ are the components of a symmetric tensor describing the temperature effect. Waser (11) has given a rigorous method of obtaining the principal axes from the $\underline{B}'_{ij}$ tensor. We shall follow his method as it applies to the orthogonalized system defined in section vi. He considered the surface for which the temperature factor is constant; that is $(g_1g_2g_3)\,\underline{B}'_{ij}\begin{pmatrix}g_1\\g_2\\g_3\end{pmatrix} = B = \text{const.}$, where we have replaced X, Y and Z by $g_1$, $g_2$ and $g_3$. Owing to the physical nature of the tensor $\underline{B}'_{ij}$ the quadratic form is positive definite and this equation represents an ellipsoid.

For the principal ellipsoid directions only, the surface of the vibrational ellipsoid is normal to the radius vector $q = g_1 b_1 + g_2 b_2 + g_3 b_3$, where $b_1$, $b_2$ and $b_3$ are the unit vectors in the directions X, Y and Z, respectively. The problem of finding the principal axes of this ellipsoid is equivalent to finding the extreme values of $q^2 = g_1{}^2 + g_2{}^2 + g_3{}^2$ subject to the side condition $(g_1 g_2 g_3)\ \underline{B}'_{ij} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} = B$. Using Lagrange's method, we introduce one multiplier and consider the vector equation:

$$\text{grad}\left[ (g_1 g_2 g_3) \left(\underline{I} - \frac{1}{\lambda}\ \underline{B}'_{ij}\right) \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \right] = 0,$$

where $\underline{I}$ is the identity matrix. This may be re-written:

$$\begin{pmatrix} (B'_{11} - \lambda) & B'_{12} & B'_{13} \\ B'_{12} & (B'_{22} - \lambda) & B'_{23} \\ B'_{13} & B'_{23} & (B'_{33} - \lambda) \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} = 0$$

The secular equation represents a cubic equation in $\lambda$ which has three real, positive solutions (since the quadratic form is positive definite). We label these as $\lambda^{(1)}$, $\lambda^{(2)}$ and $\lambda^{(3)}$. For each $\lambda^{(k)}$, the secular equation is used to find the ratios of the $g_i$ which are then normalized to direction cosines by requiring $g_1{}^2 + g_2{}^2 + g_3{}^2 = 1$. The lengths of the semi-axes of the ellipsoid are $(2\pi^2/\lambda^{(k)})^{1/2}$, which are the inverse of the root-mean square displacements of the atom along the principal direction. The displacements in B units $(2 \sin \theta/\lambda)^2$ are given by $4\lambda^{(k)}$.

The values of $\lambda^{(k)}$ can be obtained from the cubic either by an exact solution or by a numerical method similar to the one given for the least-squares plane (section v).

### viii) Rigid body analysis of atomic thermal motions.

Cruickshank (12) has developed a method for the determining the anisotropic rigid-body translational and rotational vibration tensors of molecules in crystals from the vibration tensors of individual atoms. He assumes that symmetric tensors $\underline{U}^r{}_{ij}$ have been found for each atom r(xyz) such that

$$\bar{u}^2 = (P_1 P_2 P_3)\ \underline{U}^r{}_{ij} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}$$ is the mean square amplitude of vibration of atom r in the direction specified by the unit vector $\underline{P} = (P_1 P_2 P_3)$. (See section vi for the derivation of the $\underline{U}^r{}_{ij}$ tensor.) He also assumes that the axes are orthogonal, and defined by the molecule rather than by the crystal.

If there are N atoms in the molecular unit, the problem is to interpret the N sets of $U^r{}_{ij}$ in terms of the rigid-body and internal vibrations of the molecule. As a rule the contributions from the rigid-body vibrations will be much larger than the contributions from the internal vibrations. The simplest hypothesis that can be made about anisotropic rigid-body vibrations is to suppose that the motion of a molecule can be expressed in terms of two symmetric tensors, each with six independent components, one, $\underline{T}_{ij}$, giving the translational vibrations of the mass center and the other, $\underline{W}_{ij}$, the angular oscillations about the center.

Thus for rigid-body vibrations we assume that each of the atomic $\underline{U}^{r}_{ij}$ tensors can be expressed as

$\underline{P}_{1j}\,\underline{U}^{r}_{ij}\,\underline{P}_{i1} = \underline{P}_{1j}\,\underline{T}_{ij}\,\underline{P}_{i1} + (\underline{P}\wedge\underline{r})_{1j}\,\underline{W}_{ij}\,(\underline{P}\wedge\underline{r})_{i1}$. The problem is now to find the $\underline{T}_{ij}$ and $\underline{W}_{ij}$ tensors, given N $U^{r}_{ij}$ tensors.

We now expand the right side of the above equation and collect terms in $P_1^2$, $P_2^2$, $P_3^2$, $P_1P_2$, $P_1P_3$ and $P_2P_3$. By matching coefficients we obtain six equations relating the $U^{r}_{ij}$ to the $T_{ij}$ and $W_{ij}$, thus:

$$U^{r}_{11} = T_{11} + W_{22}z^2 - 2W_{23}yz + W_{33}y^2$$

$$U^{r}_{22} = T_{22} + W_{11}z^2 - 2W_{13}xz + W_{33}x^2$$

$$U^{r}_{33} = T_{33} + W_{11}y^2 - 2W_{12}xy + W_{22}x^2$$

$$U^{r}_{12} = T_{12} - W_{12}z^2 + W_{13}yz + W_{23}xz - W_{33}xy$$

$$U^{r}_{13} = T_{13} + W_{12}yz - W_{13}y^2 + W_{23}xy - W_{22}xz$$

$$U^{r}_{23} = T_{23} + W_{12}xz + W_{13}xy - W_{23}x^2 - W_{11}yz$$

Since the number of observed $U^{r}_{ij}$ will be greater than the number of independent $T_{ij}$ and $W_{ij}$, the method of least-squares should be used to best determine the $T_{ij}$ and $W_{ij}$. Table 8 gives the derivative relations of the $U_{ij}$ with respect to the $T_{ij}$ and $W_{ij}$. The twelfth-order

Table 8. - Derivative relations of $U_{ij}$ with respect to $T_{k\ell}$ and $W_{k\ell}$.

| $\partial U^r_{ij}$ / $\partial \zeta_{k\ell}$ | $U_{11}$ | $U_{22}$ | $U_{33}$ | $U_{12}$ | $U_{13}$ | $U_{23}$ |
|---|---|---|---|---|---|---|
| $T_{11}$ | 1 | 0 | 0 | 0 | 0 | 0 |
| $T_{22}$ | 0 | 1 | 0 | 0 | 0 | 0 |
| $T_{33}$ | 0 | 0 | 1 | 0 | 0 | 0 |
| $T_{12}$ | 0 | 0 | 0 | 1 | 0 | 0 |
| $T_{13}$ | 0 | 0 | 0 | 0 | 1 | 0 |
| $T_{23}$ | 0 | 0 | 0 | 0 | 0 | 1 |
| $W_{11}$ | 0 | $z^2$ | $y^2$ | 0 | 0 | $-yz$ |
| $W_{22}$ | $z^2$ | 0 | $x^2$ | 0 | $-xy$ | 0 |
| $W_{33}$ | $y^2$ | $x^2$ | 0 | $-xy$ | 0 | 0 |
| $W_{12}$ | 0 | 0 | $-2xy$ | $-z^2$ | $yz$ | $xz$ |
| $W_{13}$ | 0 | $-2xz$ | 0 | $yz$ | $-y^2$ | $xy$ |
| $W_{23}$ | $-2yz$ | 0 | 0 | $xz$ | $xy$ | $-x^2$ |

normal equations are shown in table 9. The elements have yet to be summed over all atoms. The inverse of the normal equations matrix gives the desired $T_{ij}$ and $W_{ij}$.

The accuracy of the determination of the $T_{ij}$ and $W_{ij}$ may be estimated using the standard least-squares formula, namely $\sigma^2(T_{pp} \text{ or } W_{pp}) = C_{pp}\,\sigma^2(U)$, where $C_{pp}$ is the appropriate diagonal element of inverse matrix and $\sigma^2(U)$ is estimated as $\sigma^2(U) = \sum_n (U^r_{ij} \text{ obs.} - U^r_{ij} \text{ calc.})^2/t$, where t is the difference between the total number of $U^r_{ij}$ obs. and the number of parameters determined.

Table 9.- Normal equation matrix elements. The weighted sum is accumulated for each of these elements over all atoms in the rigid body.

| i<br>j | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 | 0 | 0 | $z^2$ |
| 3 | 0 | 0 | 1 | 0 | 0 | 0 | $y^2$ |
| 4 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 1 | $-yz$ |
| 7 | 0 | $z^2$ | $y^2$ | 0 | 0 | $-yz$ | $z^4+y^4+y^2z^2$ |
| 8 | $z^2$ | 0 | $x^2$ | 0 | $-xz$ | 0 | $x^2y^2$ |
| 9 | $y^2$ | $x^2$ | 0 | $-xy$ | 0 | 0 | $x^2z^2$ |
| 10 | 0 | 0 | $-2xy$ | $-z^2$ | $yz$ | $xz$ | $-2xy^3-xyz^2$ |
| 11 | 0 | $-2xz$ | 0 | $yz$ | $-y^2$ | $xy$ | $-2xy^3-xy^2z$ |
| 12 | $-2yz$ | 0 | 0 | $xz$ | $xy$ | $-x^2$ | $+x^2yz$ |

| i<br>j | 8 | 9 | 10 | 11 |
|---|---|---|---|---|
| 1 | $z^2$ | $y^2$ | 0 | 0 |
| 2 | 0 | $x^2$ | 0 | $-2xz$ |
| 3 | $x^2$ | 0 | $-2xy$ | 0 |
| 4 | 0 | $-xy$ | $-z^2$ | $yz$ |
| 5 | $-xz$ | 0 | $yz$ | $-y^2$ |
| 6 | 0 | 0 | $xz$ | $xy$ |
| 7 | $x^2y^2$ | $x^2z^2$ | $-2xy^3-xyz^2$ | $-2xz^3-xy^2z$ |
| 8 | $z^4+x^4+x^2z^2$ | $y^2z^2$ | $-2x^3y-xyz^2$ | $xy^2z$ |
| 9 | $y^2z^2$ | $y^4+x^4+x^2y^2$ | $xyz^2$ | $-2x^3z-xy^2z$ |
| 10 | $-2x^3y-xyz^2$ | $xyz^2$ | $4x^2y^2+z^4+y^2z^2+x^2y^2$ | $-yz^3-y^3z+x^2yz$ |
| 11 | $xy^2z$ | $-2x^3z-xy^2z$ | $-yz^3-y^3z+x^2zy$ | $4x^2y^2+y^2z^2+y^4+x^2y^2$ |
| 12 | $-2yz^3-x^2yz$ | $-2y^3z-x^2yz$ | $-xy^3+xy^2z-x^3z$ | $xyz^2-xy^3-x^3y$ |

Table 9. - continued

| i<br>j | 12 | Parameter column | | Right hand column |
|---|---|---|---|---|
| 1 | $-2yz$ | $T_{11}$ | | $U_{11}$ |
| 2 | $0$ | $T_{22}$ | | $U_{22}$ |
| 3 | $0$ | $T_{33}$ | | $U_{33}$ |
| 4 | $xz$ | $T_{12}$ | | $U_{12}$ |
| 5 | $xy$ | $T_{13}$ | | $U_{13}$ |
| 6 | $-x^2$ | $T_{23}$ | $=$ | $U_{23}$ |
| 7 | $x^2yz$ | $W_{11}$ | | $z^2U_{22}+y^2U_{33}-yzU_{23}$ |
| 8 | $-2yz^3-x^2yz$ | $W_{22}$ | | $z^2U_{11}+x^2U_{33}-xzU_{13}$ |
| 9 | $-2y^3z-x^2yz$ | $W_{33}$ | | $y^2U_{11}+x^2U_{22}-xyU_{12}$ |
| 10 | $-xz^3+xy^2z-x^3z$ | $W_{12}$ | | $-2xyU_{33}-z^2U_{12}+yzU_{13}+xzU_{23}$ |
| 11 | $xyz^2-xy^3-x^3y$ | $W_{13}$ | | $-2xyU_{22}+yzU_{12}-y^2U_{13}+xyU_{23}$ |
| 12 | $4y^2z^2+x^2z^2+x^2y^2+x^4$ | $W_{23}$ | | $-2yzU_{11}+xzU_{12}+xyU_{13}-x^2U_{23}$ |

# References

## Part I

1. L. Pauling and R. B. Corey, J. Am. Chem. Soc., 71, 5349 (1950).

2. L. Pauling, R. B. Corey and H. R. Branson, Proc. Nat. Acad. Sci., 37, 205-211 (1951).

3. L. Pauling and R. B. Corey, Proc. Nat. Acad. Sci., 37, 235-285 (1951).

4. J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore, Nature (London), 185, 422-427 (1960).

5. Groth, Chemische Krystallographie, Vol. 3, pp. 548-551 (1910). Leipzig: Verlag von Wilhelm Engelmann.

6. A. J. Wilson, Nature (London), 150, 151-152 (1942).

7. R. McWeeny, Acta. Cryst., 4, 513-519 (1951).

8. J. Berghuis, I. J. Haanappel, M. Potters, B. O. Loopstra, C. H. MacGillavry and A. L. Veenendaal, Acta Cryst., 8, 478-483 (1955).

9. J. A. Hoerni and J. A. Ibers, Acta Cryst., 7, 744-746 (1954).

10. R. B. Corey and L. Pauling, Rev. Sci. Instrum., 24, 621-627 (1953).

11. R. A. Pasternak, Acta Cryst., 9, 341-349 (1956).

12. E. W. Hughes, J. Am. Chem. Soc., 63, 1737-1752 (1941).

13. J. S. Rollett and D. R. Davies, Acta Cryst., 8, 125-128 (1955).

14. H. Mark, Ber., 57B, 1820-1827 (1924).

15. D. H. Templeton, Acta Cryst., 12, 771-773 (1959).

16. G. Mez, Z. Krystallogr., 35, 242-271 (1902).

17. R. B. Corey and R. W. G. Wyckoff, Z. Krystallogr., 85, 132-142 (1933).

18. D. A. Gordon, Ph. D. Thesis, California Institute of Technology, Part II, pp. 102-119 (1958).

19. D. A. Gordon, J. Phys. Chem., 64, 273-275 (1960).

20. International Tables for x-Ray Crystallography, Vol. I. (1952). Birmingham: Kynoch Press, p. 105.

21. J. E. Worsham, Jr., H. A. Levy and S. W. Peterson, Acta Cryst., 10, 319-323 (1957).

22. P. Vaughan and J. Donohue, Acta Cryst., 5, 530-535 (1952).

23. E. W. Hughes, H. L. Yakel and H. C. Freeman, Acta Cryst., (in press); private communication from Hughes.

24. Y. C. Leung and R. E. Marsh, Acta Cryst., 11, 17-31 (1958).

25. L. Cavalca, M. Nardelli and G. Fava, Acta Cryst., 13, 594-600 (1960).

26. H. C. Freeman, J. E. W. L. Smith and J. C. Taylor, Nature (London), 184, 707-710 (1959).

27. R. Shintani, Acta Cryst., 13, 609-618 (1960).

28. J. L. Katz and B. Post, Acta Cryst., 13, 624-628 (1960).

29. J. H. Bryden, Acta Cryst., 10, 714 (1957).

30. M. Nardelli, L. Coghi and G. Azzoni, Gazz. Chim. Ital., 88, 235-247 (1958).

31. L. Pauling and R. B. Corey, Proc. Roy. Soc., B 141, 10 (1953).

32. L. Pauling, The Nature of the Chemical Bond, (1960), 3rd ed. Ithaca: Cornell University Press.
    (a) pp. 232-239;
    (b) pp. 308-309;
    (c) p. 260.

33. G. W. Robinson, J. Chem. Phys., 21, 1741-1745 (1953).

34. J. M. O'Gorman, W. Shand and V. Schomaker, J. Am. Chem. Soc., 72, 4222-4228 (1950).

35. R. B. Corey, J. Am. Chem. Soc., 60, 1598-1604 (1938).

36. W. Fuller, J. Phys. Chem., 63, 1705-1717 (1959).

37. L. Pauling and R. Corey, Proc. Nat. Acad. Sci., 37, 251-256 (1951).

38. K. Hoogsteen (1960), private communication.

39. R. Gerdil, Acta Cryst., (in press); private communication from R. Marsh.

40. R. Degeilh and R. E. Marsh, Acta Cryst., 12, 1007-1014 (1959).

41. E. H. Wiebenga and N. F. Moerman, J. Am. Chem. Soc., 74, 6156-6157 (1952).

42. R. Newman and R. M. Badger, J. Am. Chem. Soc., 74, 3545-3548 (1952).

43. G. S. Parry, Acta Cryst., 7, 313-320 (1954).

## References

### Part II

1. E. Whittaker and G. Robinson, *The Calculus of Observation*, 4th ed., (1954), pp. 209-259. Blackie & Son Limited; Glasgow.

2. D. H. Templeton, *Acta Cryst.*, *12*, 771-773 (1959).

3. E. W. Hughes, *J. Am. Chem. Soc.*, *63*, 1737-1752 (1941).

4. C. Hastings, Jr., *Approximations for Digital Computers*, (1955), pp. 125-200. Princeton University Press: Princeton.

5. Ned C. Webb, "Calculation of Quadratic f-Curve Coefficients," pp. 1-4 (1961); private communication (1960).

6. J. M. H. Olmsted, *Solid Analytic Geometry*, (1947). New York: London : D. Appleton-Century Company, Inc.

7. V. Schomaker, J. Waser, R. E. Marsh and G. Bergman, *Acta Cryst.*, *12*, 600-604 (1959).

8. D. M. Blow, *Acta Cryst.*, *13*, 168 (1960).

9. J. B. Scarborough, *Numerical Mathematical Analysis*, (1950), pp. 192-198. Baltimore: The Johns Hopkins Press.

10. M. J. Buerger, *X-Ray Crystallography*, (1953), pp. 347-363. New York: John Wiley & Sons, Inc.

11. J. Waser, *Acta Cryst.*, *8*, 731 (1955).

12. D. W. J. Cruickshank, *Acta Cryst.*, *9*, 754-756 (1956).

# APPENDICES

# Appendix I. Instructions for the use of the orthorhombic least-squares and structure-factor program for the Burroughs 220 system.

Contents: Page

## Introduction.

This appendix is added as a guide for those persons who wish to use the orthorhombic least-squares and structure factor program written by the author for the Burroughs 220 system. The method used to compute the structure factor and its derivatives is described in part II of this thesis. The input and output used in this program has been written specifically for the peripheral equipment available at the California Institute of Technology; they are a photo-electric reader, a supervisory printer, a high speed paper tape punch, and a high speed electro-static printer.

The organization of the appendix is such as to be most useful to the reader while operating the computer. For this reason the operating instructions are given first followed by supplements A-C inclusive. Supplements D-H inclusive describe the preparation of the data tape. Since the data tape need be prepared before a problem can be run on the computer, these supplements will be of little use during the computation and have been placed last. A section describing the output has been added to aid in interpreting the results.

Operating Instructions:

1.0) If ON-line print out is desired, set up the SPO unit. (See suppl. B.)

2.0 Select and set up desired out put route (see suppl. C).

2.1) Use either route 4, 5 or 6 for LS&SF problem.

2.2) Use route 1 if a new program tape is to be prepared.

3.0) Set "Hold PZT to Zero" switch in the up position; all other switches should be in the normal (down) position.

4.0) Read the program tape into the computer:

4.1) Enter 04 into (rC:62).

4.2) Place computer into execute phase.

4.3) Press start.

5.0) If a new program tape is to be prepared change (rP) to 0200 and then press start.

5.1) If the program is to be used on a problem, set the "Hold PZT to Zero" switch in the normal (down) position.

6.0) Set the program control switches to desired position. (See suppl. A.)

7.0) Place the data reel on the optical reader (read in on low speed).

7.1) If it is desired that the computer stop before the parameter shifts are printed (to change output route), then turn on the S to P switch and enter 1531 into (rS).

8.0) Press start. The computer now has control of the problem.

9.0) If the computer has stopped because of (rS) = 1531, i.e. an S to P halt; then

9.1) Change the output route.

9.2) Press start.

10.0) The computer has finished a cycle of refinement when (rC:06) = 002244.

10.1) To iterate with the new parameters, reset the output route.

10.2) Change any of the key words or parameters desired.

10.3) Position data tape to the start of data. (Bypass the old parameters, etc.)

10.4) Press start to iterate.

11.0) To punch out a new parameter tape for future use.

11.1) Make any changes in the parameters or the constants table desired via the key board.

11.2) Set up output route 1 on the tape punch unit 0. Feed enough blank tape to act as the leader. (See also suppl. C.)

11.3) Load 1441 in (rP) and set the computer into the fetch cycle; press start.

11.4) After the computer has come to a halt, (rC) = 9999 00 1334, feed more blank tape and splice the new parameter tape onto the data tape in place of the old set.

11.5) To recycle from this point, clear the registers and set (rP) = 2000. Follow the instructions starting with (10.1).

12.0) To enter the bond distance section at any time.

12.1) The value of $d^2$, the upper limit on the printout of the distances, is in cell (2458). The program contains $d_{min}$ = 5 Å. To change this, enter in (L:2458) $(dddd)^2$ 30 0591 via (rS) and (rD). (Note: do not change the 06 portion, since this is a command!)

12.2) Set (rP) = 2398, set the fetch light on and press start.

12.3) The computer will stop ready to recycle when it has finished printing the bond distances. Proceed from instruction (10.0).

Supplement A: Program Control Switch Selections.

| PCS | State | Remarks |
|---|---|---|
| 1 | OFF | Omit printout of k, $F_o$, $F_c$, $\Delta F$, q. |
| | ON | Print k, $F_o$, $F_c$, $\Delta F$, q. |
| 2 | OFF | Omit printout of A and B. |
| | ON | Acentric Structures: Print A and B. |
| | | Centric Structures: A and B are not printed. |
| | | (Note: If PCS 1 is OFF then PCS 2 must also be OFF.) |
| 3 | OFF | The same scale factor is used on all data. |
| | ON | Layer scale factors are to be used and layer statistics are to be collected. (Note the program is written for a max. of 9 layers in $\ell$.) |
| 4 | OFF | Sub-totals after each $h\ell$ group are not printed. |
| | ON | Sub-totals after each $h\ell$ group are printed. |
| 5 | OFF | The structure is acentric. |
| | ON | The structure is centric. |
| 6 | OFF | Enter least squares. |
| | ON | Omit least squares. |
| 7 | OFF | The structure is of type I or III (see suppl. F). |
| | ON | The structure is of type II (see suppl. F). |
| 8 | OFF | Only reflections with q = to 0 or 1 are used in LS. |
| | ON | In addition to reflections with q = 0, 1 add those with q = 2 whenever $F_c > F_{min.}$ $(F_o)$ to the LS. |
| 9 | | Irrelevant. |
| 0 | OFF | Omit interatomic distance calculation. |
| | ON | Compute interatomic distances. |

*Store $K_\ell$ in cells $(2490 + \ell)$ in floating point notation.

Supplement B: Set up of SPO Unit.

1) The paper should be brought into the machine so that the perforated edge is on the right and the printed face showing.

2) Press line feed until the type bar is set to print on the first line of a new page.

3) Set the form-out gear.

4) Place the zero suppress switch in the up position and the map memory switches in the down position.

5) Set format switch to space.

6) For OFF-line printing of punched tape use Local operation.

7) For ON-line printing use Remote.

8) Press line feed until the type bar is in the numbered area at the lower left half of a page. (The computer gives an initial form-out instruction.)

Supplement C: A Review of Output Routes:

Commands: PWR ± Onni 06 aaaa. This program uses punch unit zero!

SPO ± OnnO 09 aaaa.

Definition of output routes:

Route

(1) A PWR command to give a punched tape on unit zero.
(2) A PWR command to print on the Telewriter.
(3) A PWR command to print on the high speed printer.
(4) A SPO command to give a punched paper tape on unit zero.
(5) A SPO command to print on the Telewriter.
(6) A SPO command to print on the high speed printer.

Definition of important unit switches:

Switch

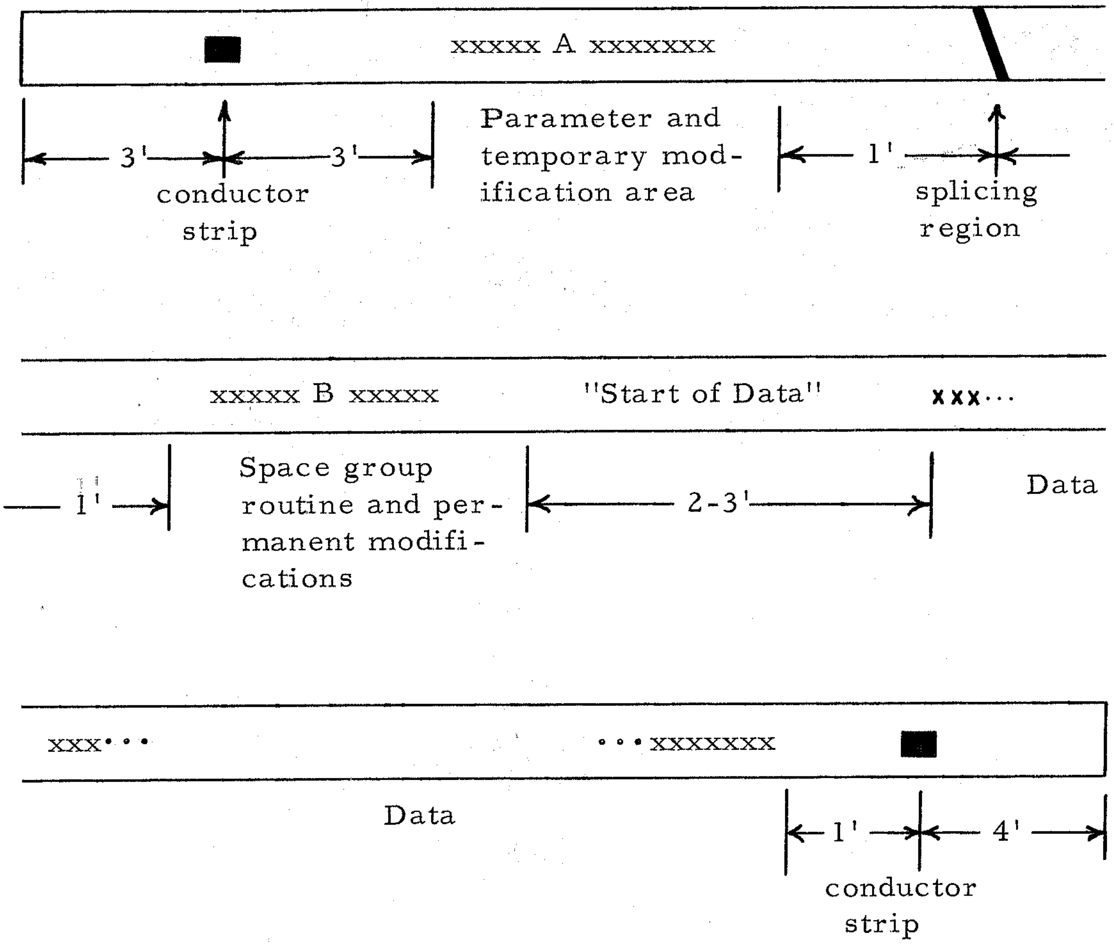
(a) SPO-PTW switch on paper punch zero.
(b) SPO switch on Telewriter.
(c) 0 unit designation toggle on Telewriter.
(d) Unit designation switch on high speed printer.

Positioning of switches:

| Route | Switch | Positioning | Remarks |
|---|---|---|---|
| (1) | (a) | up | |
| | (b) | i | i denotes irrelevancy. |
| | (c) | down | |
| | (d) | i | |
| (2) | (a) | i | |
| | (b) | i | |
| | (c) | up | |
| | (d) | i | |

| Route | Switch | Positioning | Remarks |
|---|---|---|---|
| (3) | (a) | i | |
| | (b) | i | The paper punch unit must not have the same designation! |
| | (c) | down | |
| | (d) | to 0 | |
| (4) | (a) | down | |
| | (b) | down | |
| | (c) | i | |
| | (d) | i | |
| (5) | (a) | i | |
| | (b) | up | |
| | (c) | i | |
| | (d) | i | |
| (6) | (a) | up | |
| | (b) | down | |
| | (c) | i | |
| | (d) | to SPO | |

Supplement D: Lay out of Data Reel.



1.0) The initial and final conductor strips are used to control the stop of rewind and the automatic initiation of rewind operations, respectively.

2.0) The details of the parameter part A are given in suppl. E. Temporary modifications are defined as those changes to the main program that are to be used only for one or a few cycles of refinement.

3.0) The splicing area is suggested as the easiest way of changing parameters.

4.0) Section B is for the space group changes (*i.e.*, changes from $P2_12_12_1$ to any of the other orthorhombic space groups). See suppl. F and permanent modifications: defined as those changes to the main program that are to be made every time the following set of data is to be used; as an example, the use of different f-curves (see suppl. H).

5.0) "Start of Data" is the point to which Operating Instructions 10.3 is making reference - the position of the data for start of the next iteration.

6.0) The details of the data part are given in suppl. G.

Supplement E: Parameter Part of Data Reel.

The following is an example of how to enter the table of necessary constants. Part of this table, i. e., the constant in location (0118), must precede the atomic parameters. sff xxxxxxxx denotes a floating point number with sign.

| Location | Constant | Remarks |
|---|---|---|
| | 6 0000 04 0100 | This changes the read command to the base location of the constants table. |
| (0100) | s. xxxx xx xxxx | A x 10° where $\sin^2\theta/\lambda^2 = Ah^2 + Bk^2 + C\ell^2$. |
| 1) | s. xxxx xx xxxx | B x 10° |
| 2) | s. xxxx xx xxxx | C x 10° |
| 3) | 0 0000 00 0000 | |
| 4) | 0 0000 00 0000 | |
| 5) | 0 0000 00 0000 | |
| 6) | 0 xxxx xx xxxx | Lambda x $10^{-2}$ for the radiation used to obtain the data. |
| 7) | 0 ffxx xx xxxx | $\rho_c$ from table at end of this section. |
| 8) | 0 0000 00.9999 | Sin $\theta$ (max.) x $10^{-6}$. |
| 9) | 0 0000 00.0000 | Sin $\theta$ (min.) x $10^{-6}$. |
| (0110) | 0 aaaa aa aaa# | Atom type code numbers needed for this problem, listed starting from the left. Any extra digits between the last a and # are filled with zeros. The f-curves already incorporated are : 1=Carbon, 2=Nitrogen, 3=Oxygen, 4=Hydrogen and 5=Chloride. # is the number of a's $\neq 0$. |
| 1) | 0 ffxx xx xxxx | The scale factor. |

| Location | Constant | Remarks |
|---|---|---|
| (0112) | s ffxx xx xxxx | P $\sqrt{w} = \sqrt{w_e}\,\dfrac{(P + Q\sin\theta)}{(R + SF_o + TF_o^2)}$, where $\sqrt{w_e}$ is an external weight and $F_o$ is the scaled value, i. e., it has been put on the absolute scale. |
| 3) | s ffxx xx xxxx | Q |
| 4) | s ffxx xx xxxx | R |
| 5) | s ffxx xx xxxx | S |
| 6) | s ffxx xx xxxx | T |
| 7) | s ffxx xx xxxx | A shift factor to be applied to all shifts. |
| 8) | 0 0000 00 aaaa | The base location into which the first atomic parameter is to be stored; denoted as $B_p$. aaaa must be greater than 2549. The parameters will occupy 30 N consecutive storage location , where N = the number of atoms. |
| 9) | 0 0000 00 aaaa | The base location for the storage of the scale and temperature LS coefficients; denoted as $B_t$. $B_t$ must be greater than or equal to $B_p$ + 30 N. |
| (0120) | 0 dd00 00 0000 | This constant designates the location of the decimal point in the list of $F_o$'s. The following table gives values of dd:<br>$F_o \times 10^{-x}$ x = 8 7 6 5 etc.<br>dd = 58 57 56 55 etc.<br>An example: the make up of an $F_o$ word is + $\sqrt{\text{www}}$ q FFFF kk. x refers to the location of the decimal point in $F_o$, thus FF.FF is $F_o \times 10^{-6}$ and dd = 56. |

(Note: The location between the end of the program in cell (2549) and $B_p$ are not cleared after end of a cycle of refinement and the start of the next. Therefore, this is a good place to store any program modification.)

| | | |
|---|---|---|
| 6 | 6 0000 04 2473 | Changes the read in command. |
| (2473) | 0 xxxx xx xxxx | $a \times 10^{-3}$ The unit cell length of the *a*-axis. |
| 4) | 0 xxxx xx xxxx | $b \times 10^{-3}$ The unit cell length of the *b*-axis. |
| 5) | 0 xxxx xx xxxx | $c \times 10^{-3}$ The unit cell length of the *c*-axis. |

Any temporary modification of the main program should be spliced onto the data tape following the above set of constants.

Table of $\rho_c$ Values:

| Type of Lattice | Value of $\rho_c$ in Crystal Class: | | |
|---|---|---|---|
| | 222-$D_2$ | mm2-$C_{2v}$ | mmm-$D_{2h}$ |
| P | 1 | 1 | 2 |
| A | 2 | 2 | 4 |
| C | 2 | 2 | 4 |
| I | 2 | 2 | 4 |
| F | 4 | 4 | 8 |

The following is an example of how to enter atomic parameters:

6 0000 30 1281 — This command precedes the 1st atom; it places the parameter read-in under computer control.

| For each isotropic Atom | | | For each anisotropic Atom |
|---|---|---|---|
| 7 0070 03 0000 | Change read command (for iso-atoms only) | | |
| 3 iiii ia bcde | Key word for atom | | 3 iiii ia bcde |
| 0 0000.xx xxxx | $\rho_z$ | | 0 00xx.xx xxxx |
| ± 0000.xx xxxx | $X_i$ | | ± 0000.xx xxxx |
| ± 0000.yy yyyy | $Y_i$ | | ± 0000.yy yyyy |
| ± 0000.zz zzzz | $Z_i$ | | ± 0000.zz zzzz |
| 0 0000.xx xxxx | $\rho_i$ | | 0 0000.xx xxxx |
| 0 00BB.BB BBBB | $B_i$ | or $B_{11i}$ | 0 00BB.BB BBBB |
| | | $B_{22i}$ | 0 00BB.BB BBBB |
| | | $B_{33i}$ | 0 00BB.BB BBBB |
| | | $B_{12i}$ | 0 00BB.BB BBBB |
| | | $B_{13i}$ | 0 00BB.BB BBBB |
| | | $B_{23i}$ | 0 00BB.BB BBBB |
| etc. | | | etc. |

6 0000 30 2000 — This command follows the last atom and transfers control back to the computer.

i = { In the write up: Denotes the atom number.
On punched tape example: Irrelevant digit.

$\rho_z$ : The multiplicative factor relating the atom's point position multiplicity to the space group multiplicity; that is, it is equal to unity or 1/2, 1/4, 1/8, etc.

$\rho_i$ : The atom population parameter used in least-squares; it is less than one only if there is a statistical absence.

Meaning of key-word symbols:

a = 0 : Delta $\rho_i$ is to be used to adjust the overall scale factor.

1 : Delta $\rho_i$ is to be used to adjust $\rho_i$, the population parameter. The correction is on $f_i$.

2 : Ignore delta $\rho_i$. (It is not used in any average or to adjust $\rho_i$.)

b = 0 : Calculate an isotropic temperature factor for this atom in the SF.

1 : Calculate the anisotropic temperature factor for this atom in the SF.

c = 0 : Include the atom in the SF.

1 : Exclude this atom from the SF.

d = 0 0 : Exclude this atom from all least-squares.

1 : Exclude this atom from scale and temperature factor LS.

2 : Calculate isotropic temperature and scale factor LS shifts.

3 : Calculate anisotropic temperature and scale factor shifts.

e = Atom type code number. Digits 1-9 incl. can be used. e can never be zero!

Abbreviations:

SF = Structure factor

LS = Least-squares

Supplement F: Space Group Modifications.

This program can be used for the space group $P2_12_12_1$ without modification; all other space groups require that a space group modification routine be part of the data tape. See supplement D for placement of this modification on the data tape with respect to the parameters of other data.

Memory locations (0654-0689) inclusive are reserved for the space group modifications. The program will enter the subroutine via a BUN 0654 command in location (2353). This program uses the equations developed in part II of this thesis; therefore the space group subroutine should be written so that the parity tests indicated in table 5 are made.

It is suggested that the first thing the routine tests is the positioning of program control switches 5 and 7. The settings for these switches are listed in the following table:

| Space group type | Setting of PCS 5 | 7 |
|---|---|---|
| I | OFF | OFF |
| II | OFF | ON |
| III | ON | OFF |

One should code these tests in such a way that if the switches are not in the correct position the computer will stop with a "flag" Halt in the (rC) register. By following this halt with a BUN 0654, the problem can be continued by pressing "START" after the program control switches have been set to the correct positions. Since it is not necessary to make this test more than once, successfully, per problem, the BUN command in location (2353) should be changed to bypass further testing of these control switches. (See the examples at the end of this supplement.)

Table 5 in part II of this thesis gives the parity tests that need to be made for each space group and the corresponding group key. There are up to four possible exits from the space group subroutine; these correspond to the four possible group keys. The contents of location (0127) is + hh00 kk 00ℓℓ. One uses this location to obtain the indices for the parity test. The contents of this location should not be destroyed or changed in any way since it will be used again by the program. The locations (0150-0160) are free for any intermediate calculations necessary for the parity tests.

The following is a list of the group key exits:

| Group Key | Exit to Location |
|---|---|
| (1) | 1082 |
| (2) | 1080 |
| (3) | 1078 |
| (4) | 1076 |

Two examples have been added to this supplement as a guide for the preparation of other space group modification. The first example is the space group parity test which is permanently included in this program, that is, for $P_{2_1 2_1 2_1}$. The other is for space group $F_{dd2}$.

For the two special classes of reflections within space groups $F_{dd2}$ and $F_{ddd}$, it is necessary to allow for the additional calculation of $V_i$ or $W_i$ respectively. To do this, one should add to the beginning of the parity test routine the following two commands: + 6212 40 STB 1231 and + 6201 26 IFL 1231. Then whenever a reflection is encountered which has an * or ** following its group key in table 5, add the command + 6229 26 1231 before the appropriate exit is made. (See example 2 for the space group $F_{dd2}$.)

The following is a list of useful constants, which are available for use in making the parity test. Do not destroy these numbers since the program will be using them later.

| Location | Contents |
|---|---|
| (0040) | + 5000 00 0000 |
| (1070) | + 0050 00 5000 |
| (1071) | + 0100 01 0001 |

Example 1 for Space Group $P_{2_1 2_1 2_1}$.

| Location | S | Variant | Order | | Address | Comments |
|---|---|---|---|---|---|---|
| | | | Num | Alpha. | | |
| | | | | | | |
| | 6 | | 04 | | 0654 | |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 0654 | | 5000 | 38 | BCS | 0668 | |
| 5 | | 7000 | 38 | BCS | 0668 | |
| 6 | | 0103 | 26 | IFL | 2353 | |
| 7 | | | 10 | CAD | 0127 | (rA) = + hh00 kk 00ℓℓ |
| 8 | | 2 | 45 | CLR | 0000 | |
| 9 | | 1 | 49 | SLT | 0004 | |
| 0660 | | | 12 | ADD | 0127 | = (h+k) 0 (k+ℓ) 00ℓ |
| 1 | | | 40 | STA | 0150 | |
| 2 | | | 10 | CAD | 1071 | = + 0100 01 0001 |
| 3 | | | 17 | EXT | 0150 | |
| 4 | | 6600 | 36 | BFA | 1082 | Group (1): h+k=2n and k+ℓ=2n |
| 5 | | 2200 | 36 | BFA | 1080 | Group (2): h+k=2n but k+ℓ=2n=1 |
| 6 | | 6200 | 36 | BFA | 1076 | Group (4): h+k=2n+1 while k+ℓ=2n |
| 7 | | | 30 | BUN | 1078 | Group (3): h+k=2n+1 and k+ℓ=2n+1 |
| 8 | | 5555 | 00 | HLT | 0668 | Error! Reset PCS's 5 and 7 |
| 9 | | | 30 | BUN | 0654 | |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

Example 2 for Space Group $F_{dd2}$.

| Location | S | Variant | Order |  | Address | Comments |
|---|---|---|---|---|---|---|
|  |  |  | Num. | Alpha. |  |  |
|  | 6 |  | 04 |  | 0654 |  |
|  |  |  |  |  |  |  |
| ~~0~~ |  |  |  |  |  |  |
| ~~1~~ |  |  |  |  |  |  |
| ~~2~~ |  |  |  |  |  |  |
| ~~3~~ |  |  |  |  |  |  |
| 065 4 |  | 5000 | 38 | BCS | 0656 |  |
| 5 |  | 7000 | 38 | BCS | 0657 |  |
| 6 |  | 5555 | 00 | HLT | 0656 | Error! Reset PCS's 5 and 7. (Note 1) |
| 7 |  | 0104 | 26 | IFL | 2353 |  |
| 8 |  | 6212 | 40 | STB | 1231 |  |
| 9 |  | 6201 | 26 | IFL | 1231 |  |
| 066 0 |  |  | 10 | CAD | 0127 | $(rA) = +\ hh00\ kk\ 00\ell\ell$ |
| 1 |  | 2 | 45 | CLR | 0000 |  |
| 2 |  | 1 | 49 | SLT | 0004 |  |
| 3 |  |  | 12 | ADD | 0127 | $= (h+k)\ (k+\ell)\ \ell$ |
| 4 |  |  | 40 | STA | 0150 |  |
| 5 |  | 1 | 49 | SLT | 0004 |  |
| 6 |  |  | 12 | ADD | 0127 | $= (h+k+\ell)\ (k+\ell)\ \ell$ |
| 7 |  |  | 40 | STA | 0151 |  |
| 8 |  |  | 14 | MUL | 0040 | $= (h+k+\ell)/2$ etc. |
| 9 |  |  | 40 | STA | 0152 |  |
| 067 0 |  |  | 12 | ADD | 1070 | $= ((h+k+\ell)/2 + 1/2)$ |
| 1 |  |  | 40 | STA | 0153 |  |
| 2 |  |  | 10 | CAD | 1071 | = + 0100 01 0001 |
| 3 |  |  | 17 | EXT | 0150 |  |
| 4 |  | 6600 | 36 | BFA | 0676 |  |
| 5 |  | 5555 | 00 | HLT | 0675 | Error! $h+k \neq 2n$ or $k+\ell \neq 2n$ . |
| 6 |  |  | 10 | CAD | 1071 | (rA) = +0100 01 0001 |
| 7 |  |  | 17 | EXT | 0127 |  |
| 8 |  | 0000 | 36 | BFA | 0680 | Branch if $h, k, \ell$ are all even |
| 9 |  | 6229 | 26 | IFL | 1231 | Include calc. of $V_{1-8}$ |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

| Location | S | Variant | Order | | Address | Comments |
|---|---|---|---|---|---|---|
| | | | Num | Alpha | | |
| | | | | | | |
| | | | | | | |
| 068 0 | | | 10 | CAD | 1071 | (rA) = + 0100 01 0001 |
| 1 | | | 17 | EXT | 0151 | |
| 2 | | 2201 | 36 | BFA | 0687 | Branch if $h+k+\ell \equiv 2n+1$;($4n+1$ or $4n+3$) |
| 3 | | | 10 | CAD | 1071 | (rA) = + 0100 01 0001 |
| 4 | | | 17 | EXT | 0152 | |
| 5 | | 2201 | 36 | BFA | 1078 | Group (3): $h+k+\ell = 4n+2$ |
| 6 | | | 30 | BUN | 1082 | Group (1): $h+k+\ell \equiv 4n$ |
| 7 | | | 17 | EXT | 0153 | |
| 8 | | 2201 | 36 | BFA | 1082 | Group (1)*: $h+k+\ell \equiv 4n+1$ |
| 9 | | | 30 | BUN | 1078 | Group (3)*: $h+k+\ell = 4n+3$ |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | Notes: | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

1. Because of the limited space the HLT in location (0656) is not followed by a BUN 0654; therefore the program control switches must be correctly set and (rP) can be set to 0654 before continuing. Do not clear! (The machine should be in the remember mode!)

## Supplement G: Index and $F_o$ Input Format.

All index and flag words are negative! An index word is defined as those words that contain the indices h and $\ell$. The data are entered in groups of constant h and $\ell$.

All words which contain $F_o$ are positive! These words also contain the k index as well as three digits of an external weight and one digit for a key to the observational quality of the reflection.

The following is an example of the data format:

| | |
|---|---|
| 1 0000 00 hh$\ell\ell$ | An index word |
| 0 $\sqrt{.www}$q FF FFkk | A reflection word, whose indices are the h and $\ell$ of the preceding index word and the kk contained in the same word as the reflection $F_o$ value. |
| 0 $\sqrt{.www}$q FF FFkk | |
| . | |
| . | |
| . | |
| 1 9999 99 9999 | A flag word denoting the end of data. |
| 6 0000 30 0838 | A command transferring control back to the computer. Note: If there are less than 99 punched words preceding the command, then change 0838 to 2007.<br>A conductor strip should be placed one foot beyond the transfer control command and four feet from the end of tape. |

### Table of q Code

| q | Implies | Remarks |
|---|---|---|
| 0 | $F_c = F_o$ | $\sqrt{.www}$ is irrelevant (it may be punched as 000); the external weight is taken as = 1.000. |
| 1 | $F_c = F_o$ | External weight taken as $\sqrt{.www}$ ( 000 means zero). |

| q | Implies | Remarks |
|---|---|---|
| 2 | $F_c < F_o$ | External weight taken as $\sqrt{.www}$ (see also program control switch number 8). |
| 3 | $F_c > F_o$ | External weight taken as zero. |
| 4 | $F_o = ?$ | External weight taken as zero. |
| >4 | (no meaning) | External weight taken as zero. |

Supplement H: Entry of f-Curves.

This program allows for up to nine atom types. Each atom type occupies 30 consecutive storage locations. The following table lists the atom type code number and the corresponding base storage location; that is, the location where the first of the thirty words are stored.

| Code number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Base location | 0300 | 0331 | 0362 | 0393 | 0424 | 0455 | 0486 | 0517 | 0548 |
| f-Curve | C | N | O | H | $Cl^-$ | - | - | - | - |

The formulae used in the computer program are:

$\sin \theta M_o = (\sin \theta/\lambda)\, \lambda M_o = 0.nxxxx$ , where $\lambda M_o = 0.71069$ and $\lambda$ is the wave length of the radiation used to collect the data. Let $\Delta$ be defined as $10(\sin \theta M_o - 0.n0000) = 10(0.0xxxx)$ , then $f(hk\ell) = A(n) + B(n)\,\Delta + C(n)\,\Delta^2$.

Thus, each atom type has ten A's, ten B's and ten C's. These are the thirty numbers referred to above and the order in which they are stored is: A(0), A(1),... B(0),... C(0),... C(9). Once the values of the A's etc. are determined, they must be converted to floating point notation. They are punched in the following form:

| | |
|---|---|
| 6 0000 04 aaaa | aaaa is one of the base locations referred to above and reference to this atom is denoted in the computer by its code number. |
| s ffxx xx xxxx | A(0) |
| s ffxx xx xxxx | A(1) |
| . | |
| . | |
| . | |
| s ff xx xxxx | C(9) Other f-curves are entered in an analogous manner. |

6 0000 30 2006 After the last set of f-curves, this command must be used.

Several f-curves are permanently included in this program. They are:

| Code Number | Type | Source |
|---|---|---|
| 1 | C | An average of Berghuis et al. (1955) and Hoerni and Ibers (1954). |
| 2 | N | Same as above. |
| 3 | O | Same as above. |
| 4 | H | McWeeny (1951). |
| 5 | $Cl^-$ | Berghuis et al. (1955). |

## Output

The output of this program is in three parts. The first part is the print out of the structure factor lists, the second is an output of the least-squares results, and the third is a print out of interatomic distances.

The following is an example of the output from the first part:

```
hh kk ll    |Fo|      |Fc|      ΔF        A        B      q   Remarks

0  0  0     0.00    384.49  -384.49   384.49    0.00    4  *
      4     9.43     10.41  -  1.48    10.91    0.00    1
      8     3.86      3.27     0.59  -  3.27    0.00    1
     12     8.37      8.21     0.16     8.21    0.00    1

2  2  0    92.11     97.69  -  5.57    97.69    0.00    1

                           etc.
```

The above is the maximum print out for a non-centrosymmetric space group. The A and B terms are omitted for centric space groups and $F_o$ and $F_c$ replace $|F_o|$ and $|F_c|$ respectively. The A and B terms may be omitted by placing program control switch 2 <u>OFF</u>. When program control switches 1 and 2 are both <u>OFF</u>, all of the structure factor print out is omitted.

The first four rows in the LS print out are:

| | | | |
|---|---|---|---|
| $\Sigma\lvert kF_o\rvert$ | $\Sigma\lvert F_c\rvert$ | $\Sigma\lvert\Delta F\rvert$ | (floating point) |
| $\Sigma w\lvert G_o\rvert^2$ | $\Sigma w\lvert G_c\rvert^2$ | $\Sigma w\lvert\Delta G_o\rvert^2$ | " |
| $\Sigma\lvert kF_o\rvert^2$ | $\Sigma\lvert F_c\rvert^2$ | $\Sigma\lvert kF_o - F_c\rvert^2$ | " |
| $n$ | $n_w$ | | |

In these equations $G_o = |kF_o|^2$, $G_c = |F_c|^2$ and $\Delta G = |kF_o|^2 - |F_c|^2$; n is the number of reflections for which q = 0 or 1, and $n_w$ = the number of reflections included in the least squares with $\sqrt{w} \neq 0$.

The next information printed is the parameter groups. Within each group are all the parameters pertinent to one atom. The atom groups are printed in the same sequence in which the atoms were entered. The first line of each group consists of: #-00000xxxxx ffxxxxxxxx , where # is the number of the atom; that is, if it were the first atom in the list, then # = 1. -00000xxxxx is the key word for the atom except that the 3 in the sign position has been translated as a minus. ffxxxxxxxx is the value of $\rho_z$ for that atom in floating point notation.

The remaining 5 or 10 lines, each contain four columns. The first column is a list of the old parameters, the second is the shift times the shift factor, the third is the new parameter and the last is the standard deviation of that parameter. Going down one column, the sequence of parameters is x, y, z, $\rho_i$, B if the atom is isotropic or $B_{11}$ if it is anisotropic; if the atom is isotropic no further parameters are printed out while if it is anisotropic the next parameter is $B_{22}$, then $B_{33}$, $B_{12}$, $B_{13}$ and finally $B_{23}$. If the third and fourth columns of rho of i are blank, then the value of delta rho of i is either being used to adjust the overall scale factor or is to be ignored. In the latter case, the value of delta rho i will be zero identically.

Following the atomic parameter print out, will be the print out of the scale factor adjustment. When only one scale factor is used, there will be one line of print out; it will consist of the old scale factor, the correction factor and the new scale factor, all printed in floating point notation. The correction factor is computed in the following way: correction factor $= 1/(1 + \frac{1}{N} \sum_i \Delta\rho_i)$, where the $\Delta\rho_i$ are the values obtained from the least-squares times the shift factor and N is the number of atoms for which $\Delta\rho_i$ was to be included in the average as indicated by the individual key words.

For those cases where individual layer scale factors are used, the output following the parameter will consist of four groups of information. Each group will have one to four columns in it and each row, when there is more than one, will correspond to the layer lines $\ell = 0, 1, \ldots$ max. (in this program the maximum number of layers is 9). All numbers will be in floating point notation. The information in the first group is:

$$\Sigma|F_o| \qquad \Sigma|F_c| \qquad \Sigma|\Delta F| \qquad \text{Ratio } \Sigma|F_c|/\Sigma|F_o|$$

The next group has two columns. The first column lists the old layer line scale factors and the second column lists the new layer line scale factors. The new layer line values are obtained by multiplying each old scale factor by the least-squares correction factor. The third group contains one number; it is the least-squares correction factor. This factor is the same as described in the preceding paragraph. The last group tabulates the following columns of data:

$$\Sigma w|F_o|^2 \qquad \Sigma w|F_c|^2 \qquad \text{Ratio: } (\Sigma w|F_c|^2/\Sigma w|F_o|^2)^{1/2}$$

The third output part is a list of all interatomic distances less than 5 Å. The new parameters are used in these calculations. These calculations will be made only if PCS number 0 is on. The format of the output is in fixed point. One line corresponds to one interatomic distance calculation and the printout will contain the parameter group number of both atoms and the distance. The smallest atom group number is always listed first.

## Appendix II. Lorentz-Polarization Factor program for Burroughs 220 Computer.

Program Synopsis: This will make Lorentz-Polarization corrections for equi-inclination, antiequi-inclination (zero layer) or general inclination Weissenberg intensity data or for oscillation-rotation intensity data. Simultaneous correction of two observations of the same reflection is possible and the ratio of the corrected values of the first to the second is computed. Thus estimates of a reflection from films taken around two different axes are each corrected for the Lorentz-Polarization factor respective to its axis of rotation and a scaling factor between them is computed.

Range of Variable: The maximum value that any index may take is ± 99.

Computation Method:

References:

(1) Sparks and Trueblood *et al.*, *Acta Cryst.*, 9, 350-358 (1956).

(2) Sparks and Trueblood *et al.*, *Acta Cryst.*, 10, 88 (1957).

(3) Cullity, *Elements of X-Ray Diffraction,* addison-Wesley, Inc., (1956), pp. 124-129.

Equations Employed:

$$\mathrm{Sin}^2\theta = Ah^2 + Bk^2 + C\ell^2 + Dh\ell + Ehk + Fk\ell$$

$\mathrm{Sin}^2\theta'$ is obtained by setting the layer-line index for the axis of rotation equal to zero and then evaluating the above sum, *i.e.*, $\sin^2\theta$.

(1) Equi-inclination Weissenberg ($\sin^2\mu = \sin^2 v_\perp = \lambda^2 h_i^2 / 4a_i^2$):

$$\frac{T_i}{L_p} = \frac{\sqrt{\sin^2\theta - \sin^4\theta}}{1-2(\sin^2\theta - \sin^4\theta)} \sqrt{1 - \frac{\lambda^2 h_i^2}{4a_i^2 \sin^2\theta}}$$

(2) Antiequi-inclination Weissenberg of the zero layer ($\sin^2\mu$ = same as (1)):

$$\frac{T_i}{L_p} = \frac{\sqrt{\sin^2\theta - \sin^4\theta}}{1-2(\sin^2\theta' - \sin^4\theta')} \sqrt{1 - \frac{\lambda^2 h_i^2}{4a_i^2 \sin^2\theta}}$$

Note: $h_i \neq 0$ but is equal to the corresponding equi-inclination layer line; only in the evaluation of $\sin^2\theta'$ is $h_i = 0$.

(3) General inclination ($\sin^2\mu$ is a free parameter):

$$\frac{T_i}{L_p} = \frac{\sqrt{\sin^2\theta - \sin^4\theta}}{1-2(\sin^2\theta - \sin^4\theta)} \sqrt{1 - \frac{\sin^2\mu}{\sin^2\theta}}$$

(4) Rotation-oscillation ($\sin^2\mu = \sin^2 v_\perp / 4 = \lambda^2 h_i^2 / a_i^2$):

$$\frac{T_i}{L_p} = \frac{\sqrt{\sin^2\theta - \sin^4\theta}}{1-2(\sin^2\theta - \sin^4\theta)} \sqrt{1 - \frac{\lambda^2 h_i^2}{a_i^2 \sin^2\theta}}$$

where in each case $h_i$ is the layer-line index of rotation $i$ whose repeat distance is $a_i$ and $\lambda$ is the wave length of the radiation used; $\mu$ is the Weissenberg inclination angle.

Unit Requirement:

(1) Burroughs Datatron 220 system with 5000 words of storage.

(2) One paper tape reader.

(3) One paper tape punch unit and one on-line supervisory printer.

(4) A supervisory printer for off-line printing of results.

Operating Instructions:

(a) Set PSLZ and PZT switches in down (normal) position.

(b) Read in program tape.

(c) Set output:

  (1) Type one only: Set both the SPO switch on the printer and the PTW-SPO switch on the punch No. 0 to down position. (Output on punch.)

  (2) Type both one and two: Set both the SPO switch on the printer and the PTW-SPO switch on the punch in the up position. Set both MM and ZS to down.

(d) Set Program Control Switches

PCS No. 1 - OFF: Input has form + IIII.I IIII.I
ON: Input has form + IIII.W IIII.W

PCS No. 2 - OFF: Print F
ON: Print $F^2$

PCS No. 3 - OFF: Output type one
ON: Output both type one and two

PCS No. 4 - OFF: Both values of I are independent estimations from the <u>same</u> rotation axis for the same reflection.

| | | |
|---|---|---|
| | ON: | Data either consists of one estimation, $(\text{IIII.W})_2$, per reflection or two estimations about two axes, $a_1$ and $a_2$. |
| PCS No. 5 - | OFF: | Print $T_i/L_p$ |
| | ON: | Omit printing $T_i/L_p$ |
| PCS No. 6 - | OFF: | Print $\sin^2\theta$ ($\sin\theta$) |
| | ON: | Omit $\sin^2\theta$ print out |
| PCS No. 7 - | OFF: | Only one observation per reflection. It is entered in $(\text{IIII.W})_2$ position on the data tapes (Set PCS #4 ON). |
| | ON: | Two observations per reflection |
| PCS No. 8 - | OFF: | Print $F_1^2/F_2^2(F_1/F_2)$ |
| | ON: | Omit $F_1^2/F_2^2$ print out |
| PCS No. 9 - | OFF: | When $\sin^2\theta \geq 1$ print $h_3$ and $\sin^2\theta$ |
| | ON: | When $\sin^2\theta \geq 1$ skip to next reflection |
| PCS No. 0 - | OFF: | Print $\sin^2\theta$ |
| | ON: | Print $\sin\theta$ |

(e) Read in data tape(s):

(1) The parameter part reads in and all pre-program modifications are made. This makes the (4000-4500) area available for data.

(2) The first data part reads in after all pre-program modifications have been made. If more than 4500 words of data are to be processed, the input must be in sections.

(3) Once the data are started reading into the computer all further input is computer controlled (except for changing spools of input data).

Data Input:

(a) Parameter part of input tape.

6 0000 04 PRB 4500

| | | | |
|---|---|---|---|
| 4500 | s ffxx xx xxxx | A | where $\sin^2\theta = Ah^2+Bk^2$ |
| 1 | | B | $+C\ell^2+Dh\ell+Ekh$ |
| 2 | | C | $+Fk\ell$ |
| 3 | | D | |
| 4 | | E | |
| 5 | | F | |
| 6 | | $(\lambda/a_1)^2$ | |
| 7 | | $(\lambda/a_2)^2$ | |
| 8 | 0 $iiiz_1y_1z_2\ y_2h_1h_2h_3$ | | |
| | 6 0000 30 BUN 4000 | transfer control to pre-program routine. | |

Cells (4500-4507) inclusive are entered in floating point notation.

In cells (4506 and 4507), $\lambda$ is the wave length of the radiation employed; $a_1$ and $a_2$ are the unit-cell repeat distances whose layers are indexed by $h_1$ and $h_2$ respectively (see data of input tape). If Control Switch No. 7 is off only (4507), $(\lambda^2/a_2^2)$ need be entered. If Control Switch No. 4 is off, only (4506), $(\lambda^2/a_1^2)$ need be entered.

This program is written so that any permutation of the crystallographic indices h k $\ell$ can be entered as $h_1\ h_2\ h_3$. The crystallographic indices are coded in the following manner: h = 1, k = 2 and $\ell$ = 3. These designations are entered into the 03 partial field of the word in location (4508). If, for example, a set of data were punched with $h_1 = k$, $h_2 = h$ and $h_3 = \ell$, the $h_1h_2h_3$ in (L:4508:03) would be 213.

The contents of (4508:74) = $z_1y_1\ z_2y_2$ are defined thus:

$z_1y_1$ –

(00 denotes axis i = 1 is a rotation axis.

(01 denotes axis i = 1 is an equi-inclination axis .

(02 denotes axis i = 1 is a general inclination axis and (4506) contains $\sin^2 \mu$. (See the description under General Inclination).

$z_2y_2$ –

(00 denotes axis i = 2 is a rotation axis.

(01 denotes axis i = 2 is an equi-inclination axis.

(02 denotes axis i = 2 is a general inclination axis and (4507) contains $\sin^2 \mu$. (See description under General Inclination).

( 11 implies that the $(\text{IIII.W})_2$ data are the corresponding zero layer reflections taken from antiequi-inclination films around axis i = 1.

<u>General Inclination:</u>

If more than one layer is to be treated by the general inclination $L_p$ factors, then the data must be entered in sections, where each section has $\sin^2 \mu$ = const. Preceding each data section the new value of $\sin^2 \mu$ must be entered. The following is a method for making the change where both axes are general inclination axes.

6 0000 $\underline{04}$ 4506

4506 s ffxx xx xxxx new $\sin^2 \mu$ from axis i = 1

4507 s ffxx xx xxxx new $\sin^2 \mu$ from axis i = 2

(in floating point notation)

↓

Next data section
see data part of input.

Computer Stops Due to Errors:

(1) If (4508 $\pm$ 1) = - then control is transferred to HLT in cell (4127).

(2) If (4508:07) = 0 then control is transferred to HLT in cell (4130).

(3) If (4508:52) $\neq$ 00, 01 or 02 then control is transferred to HLT in cell (4150).

(4) If (4508:72) $\neq$ 00, 01, 02 or 11 then control is transferred to HLT in cell (4159).

(5) If words (4500-4505) = 0 then control is transferred to HLT in cell (4109).

(6) If words (4501-4505) = 0 then control is transferred to HLT in cell (4111).

(7) If $\Sigma h_i$ (i = 1, 2, 3) $\neq$ 6 then control is transferred to HLT in cell (4181).

(b) Data part of input tape.

Data are read into cells (0000-4499) inclusive. Each data word will be one of two types. An index word (always negative) designates two constant indices and the starting value of the third running index. The other type of input word is the intensity word (always positive). An example of a portion of the data input tape is given below.

| | | |
|---|---|---|
| | 6 0000 04 0000 | |
| 0000 | $1\ P \pm h_1 h_1 + h_2 h_2 \pm h_3 h_3$ | an index word |
| 1 | $0(\mathrm{IIII.W})_1\ (\mathrm{IIII.W})_2$ | an intensity word with $h_3$ as running index. The next intensity word has $\pm(\lvert h_3 \rvert + P)$ as its index. |
| 2 | $0(\mathrm{IIII.W})_1\ (\mathrm{IIII.W})_2$ | |
| ⋮ | ⋮ | |
| ≤ 4499 | 1 8000 00 0000 | an index word with P = 8 |
| | 6 0000 30 4550 | branch to start of calculation |

⋮

intermediate loads, all are the same as above

⋮

| | |
|---|---|
| 6 0000 04 0000 | |
| $1\ P \pm h_1 h_1 \pm h_2 h_2 \pm h_3 h_3$ | final load |
| $0\ (\mathrm{IIII.W})_1\ (\mathrm{IIII.W})_2$ | |
| ⋮ | |
| 1 9000 00 0000 | an index word with P = 9 |
| 6 0000 30 4550 | |

The index word is detected by the computer as a negative number whereas the intensity word is always a positive number. The digit position denoted by P has the following significance.

$P = 9$; End of totality of data: Computer HLT in cell (4645)

$P = 8$: End of partial data: Read more data from tape reader

$P = 0$: Constant by which $h_3$ is increased = 1

$1 \leq P \leq 7$: Constant by which $h_3$ is increased = P (Note $P = 1$ has same effect as $P = 0$)

$\pm h_1 h_1$: This is the layer-line index of intensity data denoted by $(\text{IIII.W})_1$.

The actual crystallographic axis is recorded in the computer via the key word in cell (4508:81) by $h_1$.

$\pm h_2 h_2$: This is the layer-line index of intensity data denoted by $(\text{IIII.W})_2$. The actual crystallographic axis is recorded in the computer by $(4508:91) = h_2$. (If Control Switch No. 4 is OFF, then an exception to the above exists. In this case, $(\text{IIII.W})_2$ is an independent estimation from axis $h_1$ and $(4508:91) = h_2$ denotes the common index for the set of I's that follow on the data tape.)

If Control Switch No. 7 is OFF then $(4508:81) = h_1$ denotes the common index that intensity data that follow have.

$\pm h_3 h_3$: This denotes the common row line smallest index, for example if for some $h_1 h_2$ group the smallest positive index was +2 then $\pm h_3 h_3 = 002$. The computer will increase this running index by P after each intensity word is processed; $\pm h_3$ is increased to $\pm(h_3+P)$, for example if $\pm h_3 = -1$, $P = 0$ or $1$, then $\pm(h_3+1) = -2$ and $\pm h_3 = +2$, then $\pm(h_3+1) = +3$. Therefore, an intensity word must be entered for each $h_3$ between the smallest and largest values to be computed. Note that if such an intensity is absent due to space group extinctions or unobserved, a word of zeros should be entered for the intensity word. If a large gap exists within an $h_1h_2$ group, an alternative is to punch a new index word for the upper group of intensities.

Whenever the sign of $h_1$, $h_2$ or $h_3$ changes, a new index word must be punched.

The Intensity Word (always positive):

(IIII.W)$_1$ is the observed raw intensity taken around axis $a_1$ and with layer-line index $h_1$ whose true crystallographic index is recorded in the computer by (4508:81) = $h_1$. The digit W may have the following significance. If Program Control Switch No. 1 is ON, it denotes an estimate of the observational quality of the measured reflection. Any scheme or code for the ten bits is left the crystallographer. W is handled for output purposes only (see output). If Program Control Switch No. 1 is OFF, then W = .I (that is W denotes the least significant digit of the preceding intensity. If Program Control Switch No. 7 is OFF then (IIII.W)$_1$ = 0.

(IIII.W)$_2$ is the observed raw intensity data obtained from around axis $a_2$ with layer-line index $h_2$ (except if Program Control Switch No. 4 is OFF). Its true crystallographic axis is recorded in the computer by (4508:91) = $h_2$. W has the same significance as discussed under (IIII.W)$_1$. If Program Control Switch No. 4 is OFF then (IIII.W)$_2$

is an independent observation of the intensity taken around axis $a_1$ with layer-line index $h_1$. In this case $\underline{+}h_2h_2$ denotes the common row line whose running index is $\underline{+}h_3h_3$.

The scaling of the index word within the computer is taken as IIII.W. If there is only one observation per reflection, then enter $(\text{IIII.W})_1 = 0$ and $(\text{IIII.W})_2$ will denote the observed reflection and is treated as above. In this case $\underline{+}h_1h_1$ denotes the common row line index with $\underline{+}h_3h_3$ as the running index. Also set $(\lambda/a_1)^2 = 0$ and $z_1y_1 = .00$.

At the end of each memory load, which cannot exceed 4500 words, the command 6 0000 30 4550 is placed to start the data processing.

<u>Type One Output:</u> SPO or high speed printer (SPO)

The following is an example of the output (maximum print out):

$\underline{+}h_1h_1$ $\underline{+}h_2h_2$

$\underline{+}h_3h_3$ $\quad$ $\text{Sin}^2\,\theta$ $\quad$ $T_1/Lp$ $\quad$ $T_2/Lp$ $\quad$ $F_1^{\ 2}$ $\quad$ $F_2^{\ 2}$ $\quad$ $W_1W_2$ $\quad$ $F_1^{\ 2}/F_2^{\ 2}$

$\underline{+}(h_3+1)$ $\quad$ 0.xxxx $\quad$ 0.xxxx $\quad$ 0.xxxx $\quad$ xxxx.x $\quad$ xxxx.x $\quad$ WW $\quad$ xxx.xxx

.

.

.

$\underline{+}(h_3+k)$ $\quad$ s ffxx xx xxxx

$\underline{+}h_1h_1$ $\underline{+}h_2h_2$

$\underline{+}h_3h_3$ $\quad$ etc.

If $\sin^2\theta \geq 1$, then only $\underline{+}h_3$ and $\sin^2\theta$ in floating point notation are printed. In this case $\sin^2\theta$ is printed without regard to the position of Program Control Switch No. 6; however, PCS No. 9 ON will suppress this print out.

If $F_1^2/F_2^2$ is too large, i. e., the ratio is outside the range shown above, the ratio is not printed.

Print out modifications: see operating instructions.

Type Two Output: On Punch (Output suitable for further data processing.)

Whenever Program Control Switch No. 3 is ON, the following output is obtained on punch unit No. 0.

1 $P \pm h_1h_1 \pm h_2h_2 \pm h_3h_3$

0 $(F^2F^2F^2F^2.W)_1$ $(F^2F^2F^2F^2.W)_2$

.
.
.

1 P000 00 0000 P = 8 or 9 (see input description)

6 0000 00 0000

.
.

etc.

If Program Control Switch No. 1 is OFF, W = the least significant digit of the preceding $F^2$. This output will have the same form as the data part of the input tapes except that $F^2$ (or F) replaces I. F's are punched as FF.FFW whenever Control Switch No. 2 is OFF. The data tape will have the same length as the input tape and the data may be entered into consecutive storage locations as specified by the next data handling program.

## Appendix III. Molecular Plane Fourier Program.

Synopsis: This program calculates the 3-dimensional Fourier summation for a plane in any orientation relative to the crystallographic axes. A square mesh of arbitrary size is ruled on the plane about a point $\bar{x}, \bar{y}, \bar{z}$.

### Operating Instructions:

(1) Read program in with hold PZT to zero switch up. After the program has been read in and (rC) = 9995000705 switch the "Hold PZT to zero" to its normal (down) position.

(2) Put the data reel on the optical tape reader. Set up the desired format. (This program uses SPO output.) Set the necessary PCS's ON; then press START.

### To Punch Out a New Program Tape:

(1) Read program in with "Hold PZT to zero" up. Leave this switch up during program punchout.

(2) Set tape punch unit designation to 0 and for PTW output.

(3) Set to "FETCH" with (rP) = 0200 and then press START.

Program Control Switch Options:

| PCS # | State | Remarks |
|---|---|---|
| 1 | OFF | For monoclinic, orthorhombic, tetragonal and cubic structures |
| | ON | For triclinic, trigonal, hexagonal (not yet programmed) |
| 2 | OFF | Print out X, ρ (XYZ) |
| | ON | " " X, Y, Z and ρ(XYZ) |
| 3 | OFF | Punch out XYZ 0 of mesh points in 2000 word blocks. Computer halt at end of each 2000 word to give operator time to feed blank tape. Press start to continue. |
| | ON | Omit print out of point mesh. |
| 4 | OFF | Fourier point mesh will fit in memory, so store it in sequence. |
| | ON | (PCS # 3 OFF). Point mesh not stored in sequence, but punch out sequence of 2000 word block. This allows the calculation to be done in parts. |
| 5 | OFF | For acentric structures |
| | ON | For centric structures |

Preparation of Data Tape:

(1) Space group tape if other than $P2_12_12_1$. (See "Other space groups" under "Equations used." See also examples at end of this write-up.)

(2) Constants: The constants needed by the program are defined and their manner of entry illustrated.

| | | | | |
|---|---|---|---|---|
| | 6 0000 04 | 0380 | Change (rC) to read into (0380) etc. | |
| 0380 | 00xx.xx | xxxx | $\bar{X} \times 10^{-4}$ | Coordinates of Centroid (Cartesian) |
| 1 | 00xx.xx | xxxx | $\bar{Y} \times 10^{-4}$ | |
| 2 | 00xx.xx | xxxx | $\bar{Z} \times 10^{-4}$ | |
| 3 | 000x.xx | xxxx | $A \times 10^{-4}$ | Eq. const. for L.S. Plane $AX + BY + CZ = D$ |
| 4 | '' | | $B \times 10^{-4}$ | |
| 5 | '' | | $C \times 10^{-4}$ | |
| 6 | '' | | $D \times 10^{-4}$ | |
| 7 | 00xx.xx | xxxx | $a \times 10^{-4}$ | Real cell constants |
| 8 | | | $b \times 10^{-4}$ | |
| 9 | | | $c \times 10^{-4}$ | |
| 0390 | | | $\cos \alpha \times 10^{-4}$ | |
| 1 | | | $\cos \beta \times 10^{-4}$ | |
| 2 | | | $\cos \gamma \times 10^{-4}$ | |
| 3 | | | $R \times 10^{-4}$ | value of mesh interval |
| 4 | | | $R_{max} \times 10^{-4}$ | max. length of mesh |
| 5 | 0000 00 | 0000 | empty cells | |
| 6 | '' | | '' | |
| 7 | '' | | '' | |
| 8 | '' | | '' | |
| 9 | '' | | '' | |
| 0400 | 00xx xx | xxxx | $2/V_c \times 10^{-1}$ | give final result value of $e/A^3$ |
| 1 | ff00 00 | 0000 | (see table) Data scaling const. | |
| 2 | 0000 00 | aaaa | B(XYZ) | address of first X |
| | 6 30 | 0620 | (Note B(XYZ) $\geq$ 0780 required) | |

## Table of Values of ff:

| Value ff | Scale of Data |
|---|---|
| 55 | A. AAA |
| 56 | AA. AA |
| 57 | AAA. A |
| 58 | AAAA . |

etc.

## Reflections:

A. Multiplicities

(1) F(000): Enter A(000) with 1/2 its true value.

(2) All others: Enter at full value.

B. Format: Punch data tape in the following manner:

±hh0±kk0±ℓℓ

+ ±AAAA±BBBB

.
.
.

etc.

.
.
.

3 0000 00 0000 final data word

6 0000 30 0479 " command word

12"

■ Place conductor strip here

4'

(Note the constants part must precede the data part.)

Output:

On the first page the following numbers are printed in floating point notation:

$-R(B^2+C^2)^{1/2}$
$RAB/(B^2+C^2)^{1/2}$
$RC/(B^2+C^2)^{1/2}$
$RAC/(B^2+C^2)^{1/2}$
$-RB/(B^2+C^2)^{1/2}$
$1/a$
$1/b$
$-c \cos \beta$
$1/c\,(1 - \cos^2\beta)^{1/2}$

Starting with the second page, the results of the Fourier calculation are printed out in the following manner:

PCS#2 OFF: $X$ $\rho(XYZ)$ $C_R$

$\pm 0.\text{xxxxxx}$ $\pm\text{xx.xx}$

etc.

PCS#2 ON: $X$ $Y$ $Z$ $\rho(XYZ)$ $C_R$

$\pm 0.\text{xxxxxx}$ $\pm 0.\text{xxxxxx}$ $\pm 0.\text{xxxxxx}$ $\pm\text{xx.xx}$

etc.

This format has 10 lines per group and 5 groups per page. $C_R$ denotes carriage return action by the typewriter. For each point with coordinates (X Y Z), there correspond a pair of integers $N_x$, $N_{yz}$, which are defined later. These pairs of numbers are essentially the grid coordinates of the L.S. plane. The output sequence starts with $+N_x$, $+N_{yz}$. The second value of $\rho(XYZ)$ corresponds to $N_x$, $N_{yz}-1$. $N_{yz}$ is decreased by one until it is equal to $-N_{max}$ at which time $N_x$ is decreased by one for the first time and $N_{yz}$ is reset to $+N_{max}$. The print out terminates when the value of $\rho(XYZ)$ is printed for $N_x = -N_{max}$ and $N_{yz} = -N_{max}$.

## Equations Used:

The electron density equations for P1 and P$\bar{1}$ are employed. For P1 the equation is:

$$\begin{aligned}\rho(XYZ) = \sum_q \Big\{ & A(hk\ell)\cos 2\pi(hX+kY+\ell Z) + B(hk\ell)\sin 2\pi(hX+kY+\ell Z) \\ & + A(\bar{h}k\ell)\cos 2\pi(-hX+kY+\ell Z) + B(\bar{h}k\ell)\sin 2\pi(-hX+kY+\ell Z) \\ & + A(h\bar{k}\ell)\cos 2\pi(hX-kY+\ell Z) + B(h\bar{k}\ell)\sin 2\pi(hX-kY+\ell Z) \\ & + A(hk\bar{\ell})\cos 2\pi(hX+kY-\ell Z) + B(hk\bar{\ell})\sin 2\pi(hX+kY-\ell Z)\Big\}\end{aligned}$$

For P$\bar{1}$ the equation used does not have any sine terms and $F(hk\ell)$ replaces $A(hk\ell)$ in the above expression .

The program will generate $A(\bar{h}k\ell)$, $A(h\bar{k}\ell)\cdots B(\bar{h}k\ell)\cdots$ etc. for space group $P_{2_1 2_1 2_1}$ using the formulation of the International Table for Crystallography (1952).

## Other Space Group:

Memory location (0300)-(0379) incl. are reserved for space group modifications. The program will STP 0300 then BUN 0301. The modification must be written in the form of a subroutine. On the basis of an index parity test the various values of $A(q)$ and $B(q)$ must be generated for use in the above formula. Storage location assignments are:

| | | | |
|---|---|---|---|
| $A(hk\ell)$ | (0406) | $B(hk\ell)$ | (0410) |
| $A(\bar{h}k\ell)$ | (0407) | $B(\bar{h}k\ell)$ | (0411) |
| $A(h\bar{k}\ell)$ | (0408) | $B(h\bar{k}\ell)$ | (0412) |
| $A(hk\bar{\ell})$ | (0409) | $B(hk\bar{\ell})$ | (0413) |

At the end of this write-up are two examples of space group subroutines.

Calculation of X, Y, Z.

The following method is used to calculate the coordinates of a square point grid on the molecular plane. Let $\bar{x}$, $\bar{y}$, $\bar{z}$ be the centroid of the group of atoms lying in the desired plane, obtained say from a least squares process. Further, let $\bar{x}$, $\bar{y}$, $\bar{z}$ be in the Cartesian orthogonal system. The equation of the plane is $Ax + By + Cz = D$, where A, B and C are the direction cosines of the plane and D is the distance to the origin. We take as one of the two axes in the plane the line $\Lambda_1$ formed by the intersection of the L.S. plane with the plane $x = \bar{x}$. (Note that the LS plane cannot be parallel to the $x = \bar{x}$ plane. The program will have to be modified if the plane is parallel or nearly parallel to it.)

The direction numbers of the line $\Lambda_1$ are: $\ell_1 = 0$, $m_1 = C$ and $n_1 = -B$. The direction cosines obtained by normalization are: $\lambda_1 = 0$, $\mu_1 = C(B^2 + C^2)^{1/2}$ and $\nu_1 = -B/(B^2 + C^2)^{1/2}$. We need a second line $\Lambda_2$ perpendicular to $\Lambda_1$. The direction numbers of $\Lambda_2$ must be $\ell_2 = B\nu_1 - C\mu_1$, $m_2 = -A\nu_1$ and $n_2 = A\mu_1$ which we can also express thus: $\ell_2 = -(B^2 + C^2)^{1/2}$, $m_2 = AB/(B^2 + C^2)^{1/2}$ and $n_2 = AC/(B^2 + C^2)^{1/2}$. The direction numbersin this case are also the direction cosines since $\ell^2 + m^2 + n^2 = (B^2 + C^2) + A^2(B^2/(B^2 + C^2)) + A^2(C^2/(B^2 + C^2)) = A^2 + B^2 + C^2 = 1$.

If we take R to be the desired interval of the square grid, then the following set of equations produces the point mesh:

$$x = -N_x R(B^2 + C^2)^{1/2} \qquad + \bar{x}$$

$$y = N_x RAB/(B^2 + C^2)^{1/2} + N_{yz} RC/(B^2 + C^2)^{1/2} \qquad + \bar{y}$$

$$z = N_x RAC/(B^2 + C^2)^{1/2} - N_{yz} RB/(B^2 + C^2)^{1/2} \qquad + \bar{z} \quad ,$$

where $N_x$ and $N_{yz}$ define the grid points relative to $\bar{x}$, $\bar{y}$, $\bar{z}$ as origin.

The square grid generated by this program is:

$$N_{max} \geq N_x \geq -N_{max}$$

$$N_{max} \geq N_{yz} \geq -N_{max}$$

The maximum value of $N_x$ = the maximum value of $N_{yz} \equiv N_{max}$, where $N_{max}$ is defined as the integral part of $\left[ \frac{R_{max}}{R} + 1 \right]$.

The Cartesian x, y, z of the point mesh are transformed into the fractional coordinate system of the monoclinic system by:

$X = (1/a)(x - Zc \cos \beta)$ (Note Z not z is used here.)

$Y = (1/b)\, y$

$Z = z/c(1 - \cos^2 \beta)^{1/2}$ .

Example 1.

Space Group $P_{2_1 2_1 2_1} - D_2^4$ No. 19

Origin midway between 3 pairs of non-intersecting screw axes as defined in the International Tables.

$|F(hk\ell)| = |F(\bar{h}k\ell)| = |F(h\bar{k}\ell)| = |F(hk\bar{\ell})| = |F(\bar{h}\bar{k}\bar{\ell})|$

$h + k = 2n$ $A(\bar{h}k\ell) = A(hk\ell)$ and $B(\bar{h}k\ell) = -B(hk\ell)$
$k + \ell = 2n$ $A(h\bar{k}\ell) = A(hk\ell)$ $B(h\bar{k}\ell) = -B(hk\ell)$
$A(hk\bar{\ell}) = A(hk\ell)$ $B(hk\bar{\ell}) = -B(hk\ell)$

$h + k = 2n$ $A(\bar{h}k\ell) = A(hk\ell)$ and $B(\bar{h}k\ell) = -B(hk\ell)$
$k + \ell = 2n + 1$ $A(h\bar{k}\ell) = -A(hk\ell)$ $B(h\bar{k}\ell) = B(hk\ell)$
$A(hk\bar{\ell}) = -A(hk\ell)$ $B(hk\bar{\ell}) = B(hk\ell)$

$h + k + 2n + 1$ $A(\bar{h}k\ell) = -A(hk\ell)$ and $B(\bar{h}k\ell) = B(hk\ell)$
$k + \ell = 2n$ $A(h\bar{k}\ell) = A(hk\ell)$ $B(h\bar{k}\ell) = -B(hk\ell)$
$A(hk\bar{\ell}) = -A(hk\ell)$ $B(hk\bar{\ell}) = B(hk\ell)$

$h + k = 2n + 1$ $A(\bar{h}k\ell) = -A(hk\ell)$ and $B(\bar{h}k\ell) = B(hk\ell)$
$k + \ell = 2n + 1$ $A(h\bar{k}\ell) = -A(hk\ell)$ $B(h\bar{k}\ell) = B(hk\ell)$
$A(hk\bar{\ell}) = A(hk\ell)$ $B(hk\bar{\ell}) = -B(hk\ell)$

| Location | S | Variant | Order |  | Address | Comments |
|---|---|---|---|---|---|---|
|  |  |  | Num. | Alpha |  |  |
|  |  |  |  |  |  |  |
|  | 6 |  | 04 |  | 0300 |  |
| $030^0$ |  |  | 30 | BUN | "P" |  |
| 1 |  | 4400 | 28 | DLB | 0484 | (rB) = $B_D$ |
| 2 | 1 | 1 | 10 | CAA | 0200 | (rA) = hh oo kk oo $\ell\ell$ |
| 3 |  |  | 48 | SRA | 0004 | = oooo hh oo kk |
| 4 | 1 | 1 | 12 | ADA | 0200 | = hh o(h+k) o(k+$\ell$) |
| 5 |  |  | 40 | STA | 0420 |  |
| 6 |  |  | 10 | CAD | 0429 | (rA) = 01 0001 0001 |
| 7 |  |  | 17 | EXT | 0420 |  |
| 8 |  | 0600 | 36 | BFA | 0344 | Branch if h+k, k+$\ell$ are e, e |
| 9 |  | 0200 | 36 | BFA | 0333 | " " " " o, e |
| $031^0$ |  | 6200 | 36 | BFA | 0322 | " " " " e, o |
| 1 |  |  | 11 | CSU | 0406 | (rA) = -A(hk$\ell$) |
| 2 |  |  | 40 | STA | 0407 |  |
| 3 |  |  | 40 | STA | 0408 |  |
| 4 |  |  | 10 | CAD | 0406 | = A(hk$\ell$) |
| 5 |  |  | 40 | STA | 0409 |  |
| 6 |  |  | 10 | CAD | 0410 | = B(hk$\ell$) |
| 7 |  |  | 40 | STA | 0411 |  |
| 8 |  |  | 40 | STA | 0412 |  |
| 9 |  |  | 11 | CSU | 0410 | = -B(hk$\ell$) |
| $032^0$ |  |  | 40 | STA | 0413 |  |
| 1 |  |  | 30 | BUN | 0300 |  |
| 2 |  |  | 11 | CSU | 0406 | = -A(hk$\ell$) |
| 3 |  |  | 40 | STA | 0408 |  |
| 4 |  |  | 40 | STA | 0409 |  |
| 5 |  |  | 10 | CAD | 0406 | = A(hk$\ell$) |
| 6 |  |  | 40 | STA | 0407 |  |
| 7 |  |  | 10 | CAD | 0410 | = B(hk$\ell$) |
| 8 |  |  | 40 | STA | 0412 |  |
| 9 |  |  | 40 | STA | 0413 |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

| Location | S | Variant | Order |  | Address | Comments |
|---|---|---|---|---|---|---|
|  |  |  | Num. | Alpha. |  |  |
| 0330 |  |  | 11 | CSU | 0410 | (rA) = $-B(hk\ell)$ |
| 1 |  |  | 40 | STA | 0411 |  |
| 2 |  |  | 30 | BUN | 0300 |  |
| 3 |  |  | 11 | CSU | 0406 | = $-A(hk\ell)$ |
| 4 |  |  | 40 | STA | 0407 |  |
| 5 |  |  | 40 | STA | 0409 |  |
| 6 |  |  | 10 | CAD | 0406 | = $A(hk\ell)$ |
| 7 |  |  | 40 | STA | 0408 |  |
| 8 |  |  | 10 | CAD | 0410 | = $B(hk\ell)$ |
| 9 |  |  | 40 | STA | 0411 |  |
| 0340 |  |  | 40 | STA | 0413 |  |
| 1 |  |  | 11 | CSU | 0410 | = $-B(hk\ell)$ |
| 2 |  |  | 40 | STA | 0412 |  |
| 3 |  |  | 30 | BUN | 0300 |  |
| 4 |  |  | 10 | CAD | 0406 | = $A(hk\ell)$ |
| 5 |  |  | 40 | STA | 0407 |  |
| 6 |  |  | 40 | STA | 0408 |  |
| 7 |  |  | 40 | STA | 0409 |  |
| 8 |  |  | 11 | CSU | 0410 | = $-B(hk\ell)$ |
| 9 |  |  | 40 | STA | 0411 |  |
| 0350 |  |  | 40 | STA | 0412 |  |
| 1 |  |  | 40 | STA | 0413 |  |
| 2 |  |  | 30 | BUN | 0300 |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |

Example 2.

Space Group $P2_{1/n} - C^{5}_{2h}$ No. 14 b as unique axis

Origin at $\bar{1}$ .

$h + k + \ell = 2n$ $\quad F(h\bar{k}\ell) = F(hk\ell)$

$F(hk\bar{\ell}) = F(\bar{h}k\ell)$

$h + k + \ell = 2n + 1$ $\quad F(h\bar{k}\ell) = -F(hk\ell)$

$F(hk\bar{\ell}) = -F(hk\bar{\ell})$

| Location | S | Variant | Order |  | Address | Comments |
|---|---|---|---|---|---|---|
|  |  |  | Num | Alpha. |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  | 6 |  | 04 |  | 0300 |  |
| 0300 |  |  | 30 | BUN | "P" |  |
| 1 |  | 4400 | 28 | DLB | 0484 | $(rB) \equiv B_D$ |
| 2 |  | 1 | 10 | CAA | 0200 | $(rA) \equiv$ hh oo kk oo $\ell\ell$ |
| 3 |  |  | 48 | SRA | 0004 |  |
| 4 |  | 1 | 12 | ADA | 0200 | $(rA{:}03) \equiv k + \ell$ $(rA{:}63) \equiv h + k$ |
| 5 |  |  | 48 | SRA | 0004 |  |
| 6 |  | 1 | 12 | ADA | 0200 | $(rA{:}03) \equiv h + k + \ell$ |
| 7 |  |  | 40 | STA | 0420 |  |
| 8 |  | 4412 | 40 | STB | 0560 |  |
| 9 |  | + | 10 | CAD | 0403 |  |
| 0 |  | 1 | 33 | BSA | 0321 | Branch if sign of h is negative |
| 1 |  |  | 10 | CAD | 0429 | = 01 0001 0001 |
| 2 |  |  | 17 | EXT | 0420 |  |
| 3 |  | 4202 | 26 | IFL | 0560 |  |
| 4 |  | 0201 | 36 | BFA | 0318 | Branch if $h + k + \ell \equiv 2n + 1$ |
| 5 |  |  | 10 | CAD | 0406 |  |
| 6 |  |  | 40 | STA | 0408 |  |
| 7 |  |  | 30 | BUN | 0300 |  |
| 8 |  |  | 11 | CSU | 0406 |  |
| 9 |  |  | 40 | STA | 0408 |  |
| 0320 |  |  | 30 | BUN | 0300 |  |
| 1 |  | 0 | 43 | LSA | 0000 | Change sign h to + |
| 2 |  | + | 40 | STA | 0403 |  |
| 3 |  |  | 10 | CAD | 0406 |  |
| 4 |  |  | 40 | STA | 0407 |  |
| 5 |  | 4203 | 26 | IFL | 0560 |  |
| 6 |  |  | 10 | CAD | 0429 |  |
| 7 |  |  | 17 | EXT | 0420 |  |
| 8 |  | 0201 | 36 | BFA | 0332 | Branch if $h + k + \ell \equiv 2n + 1$ |
| 9 |  |  | 10 | CAD | 0407 |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

| Location | S | Variant | Order | | Address | Comments |
|---|---|---|---|---|---|---|
| | | | Num. | Alpha. | | |
| $033^0$ | | | 40 | STA | 0409 | |
| 1 | | | 30 | BUN | 0300 | |
| 2 | | | 11 | CSU | 0407 | |
| 3 | | | 40 | STA | 0409 | |
| 4 | | | 30 | BUN | 0300 | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| 0 | | | | | | |
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |
| 6 | | | | | | |
| 7 | | | | | | |
| 8 | | | | | | |
| 9 | | | | | | |
| | 6 | | 04 | | 0556 | |
| | | 4400 | 28 | | 0560 | |
| | 6 | | 04 | | 0568 | |
| | | 0002 | 21 | | 0557 | |

# PROPOSITIONS

1. In a good many crystal structure investigations, particularly of organic crystals, the whole of the molecule can be formulated as a rigid body; that is, models of one or more plausible configurations can be constructed, say, using the Corey and Pauling atomic models. For each such model a set of positional coordinates can be computed relative to a set of molecular axes. Only six parameters (three translational components and the three Eulerian angles shown in figure 1) are required to define the transformation from the molecular axes to the fractional coordinate system of the unit cell.

I propose that the preliminary manipulations of the more plausible trial models can be greatly expedited by use and refinement of the six rigid body parameters.

In many instances the failure of the usual least-squares process during the early stages of the application of the method can be attributed to the convergence of the positional parameters to false minima in which the atoms appear to be completely oblivious of reasonable interatomic distances and angles. I argue, therefore, that during these early stages, the model must be refined in such a way that the interatomic distances and angles remain invariant from one iteration to the next, thus avoiding all false minima in which these invariances are not observed. The refinement of the rigid body parameters will preserve the interatomic relationships of the trial model. In the last part of this proposition I have developed the least-squares equations which allow for the refinement of the rigid body parameters. These equations make use of the structure factor and derivative expressions already derived for use in the usual application of the least-squares.

When the correctness of the trial model has been substantiated and no further improvement is obtainable from the rigid body least-squares, the present least-square methods can be used to continue the refinement. Assuming that the perversity encountered during a refinement is proportional to the number of independent parameters being refined, one can see that it will be constant from structure to structure using only the six rigid body parameters while it is proportional to N/2 times the same constant if the present least-squares process is used (where N is the number of atoms). Thus it appears possible to greatly simplify a problem by use of a rigid body least-squares method. This suggests another plausible application of the method.

I think that it could be used for problems in which the number of parameters has become too large to handle by usual methods, such as, in large polypeptides, crystalline globular proteins or other large molecular substances. The individual amino-acids or other structural units can be treated as rigid bodies; thus only 6R parameters need be refined.

In addition, it seems feasible that the Monte Carlo method could be used to produce a structure-probability map of the six dimensional volume defined by the limits of the six rigid body parameters. The structure-probability map would have local maximums in regions where the sextet of transformation parameters give an acceptable packing of the trial model into the unit cell. That is, the six parameters would be obtained by a random number generator and the trial model would then be transformed into the unit cell. The resultant molecular packing would be tested to see that no intermolecular distances are too short with respect to the requirements of the unit cell dimensions and the space group symmetry. In practice, a computer program could be written to seek and print out sextets of parameters which give acceptable packing. Each of these could be tested to see that they make chemical sense. Those passing the last test could then serve as a starting point for the rigid body least-squares.

In the remainder of this proposition I develop in a general form the rigid body least-squares equations.

Let us consider the atom i of a rigid body trial structure; its coordinates with respect to a set of Cartesian molecular axes are $\underline{U}_i = (U_i V_i W_i)$. We shall consider the transformation of these coordinates into the fractional coordinate system of the unit cell as a two step process. The first step is a general rigid body rotation and translation to give new Cartesian coordinate positions $\underline{X}_i = (X_i Y_i Z_i)$ which are related to the unit cell positions $\underline{x} = (x_i y_i z_i)$ by the usual orthogonalization relations (see section iv of part II of this thesis). The second step is the inverse of the orthogonalization of the unit cell coordinates.

We write the first step in matrix form: $\underline{X}_i = \underline{T}\,\underline{I} + \underline{R}\,\underline{U}_i$, where $\underline{I}$ is the identity matrix, T is a 3x3 diagonal matrix whose only non-zero components are $t_{11}$, $t_{22}$ and $t_{33}$ and

$$\underline{R} = \begin{pmatrix} \cos\phi\cos\theta\cos\psi - \sin\phi\sin\psi & -\cos\phi\cos\theta\sin\psi - \sin\phi\cos\psi & \cos\phi\sin\theta \\ \sin\phi\cos\theta\cos\psi + \cos\phi\sin\psi & -\sin\phi\cos\theta\sin\psi + \cos\phi\cos\psi & \sin\phi\sin\theta \\ -\sin\theta\cos\psi & \sin\theta\sin\psi & \cos\theta \end{pmatrix}$$

Step two can be written: $\underline{x}_i = \underline{D}^{-1}\underline{X}_i$, where $\underline{D}^{-1}$ is as defined on page 156 . Thus $t_{11}$, $t_{22}$ and $t_{33}$ are the three translational components and $\phi$, $\theta$ and $\psi$ are the three Eulerian angles. These are the six rigid body parameters for which we wish to find the best values. We denote this best set by use of an asterisk (*); that is, $t^*_{11}$ is the best value of $t_{11}$, etc. Hence we write:

$$t^*_{11} = t_{11} + e_{11} \qquad \phi^* = \phi + e(\phi)$$

$$t^*_{22} = t_{22} + e_{22} \qquad \theta^* = \theta + e(\theta)$$

$$t^*_{33} = t_{33} + e_{33} \qquad \psi^* = \psi + e(\psi)$$

Let us now write the residual equations from which the least-square normal equations are derivable.

$$V_q = \sum_{i=1}^{N} \left[ \left( \frac{\partial |F|^2}{\partial_{x_i}} \right) e(x_i) + \left( \frac{\partial |F|^2}{\partial_{y_i}} \right) e(y_i) + \left( \frac{\partial |F|^2}{\partial_{z_i}} \right) e(z_i) \right] - \Delta G$$

We rewrite this as

$$V_q = \sum_{i=1}^{N} \left[ \left( \frac{\partial |F|^2}{\partial_{x_i}} \right) \left( \frac{\partial |F|^2}{\partial_{y_i}} \right) \left( \frac{\partial |F|^2}{\partial_{z_i}} \right) \right] e\,(\underline{x}_i) - \Delta G, \text{ where } e(\underline{x}_i) = \begin{vmatrix} e(x_i) \\ e(y_i) \\ e(z_i) \end{vmatrix}$$

The problem is to replace the $e(x_i)$, $e(y_i)$ and $e(z_i)$ by $e_{11}$, $e_{22}$, $e_{33}$, $e(\phi)$, $e(\theta)$ and $e(\psi)$.

We assume that the step one transformation equations contain the best rigid body parameters, and make the substitutions $t^*_{11} = t_{11} + e_{11}$ etc. Now we expand the transformation equations in a Taylor series neglecting all terms in which e is of an order greater than one. The result is: $\underline{X}^*_i = \underline{T}\,\underline{I} + \underline{R}\,\underline{U}_i + \underline{E}\,\underline{I} + \underline{P}_i\,\underline{e}$, where $\underline{E}$ is a diagonal matrix whose only non-zero elements are $e_{11}$, $e_{22}$ and $e_{33}$,

$$\underline{P}_1 = \begin{vmatrix} (\partial X_i/\partial\phi) & (\partial X_i/\partial\theta) & (\partial X_i/\partial\psi) \\ (\partial Y_i/\partial\phi) & (\partial Y_i/\partial\theta) & (\partial Y_i/\partial\psi) \\ (\partial Z_i/\partial\phi) & (\partial Z_i/\partial\theta) & (\partial Z_i/\partial\psi) \end{vmatrix} \quad \text{and} \quad \underline{e} = \begin{vmatrix} e(\phi) \\ e(\theta) \\ e(\psi) \end{vmatrix}$$

Table 1 lists the matrix elements of $\underline{P}_i$ explicitly. Since $\underline{T}\,\underline{I} + \underline{R}\,\underline{U}_i = \underline{X}_i$, we rewrite this equation in the form: $e(\underline{X}_i) = \underline{X}^*_i - \underline{X}_i = \underline{E}\,\underline{I} + \underline{P}_i\,\underline{e}$. But $e(\underline{X}_i)$ is the orthogonalized value of $e(\underline{x}_i)$. These are related by $e(\underline{x}_i) = \underline{D}^{-1} e(\underline{X}_i) = D^{-1}(\underline{E}\;\underline{I} + \underline{P}_i\;e)$. The residual equations are now written:

$$V_q = \sum_{i=1}^{N} [(\partial|F|^2/\partial x_i)(\partial|F|^2/\partial y_i)(\partial|F|^2/\partial z_i)] \; \underline{D}^{-1}(\underline{E}\,\underline{I} + \underline{P}_i\,\underline{e}) - \Delta G .$$

This equation can now be used in the usual way (2) to set up the six normal equations which are used to refine the six rigid body parameters.

## References

1. A. R. Edmonds, Angular Momentum in Quantum Mechanics, (1957), p. 7, Princeton: Princeton University Press.
2. E. Whittaker and G. Robinson, The Calculus of Observation, (1954), 4th ed., pp. 209-259.

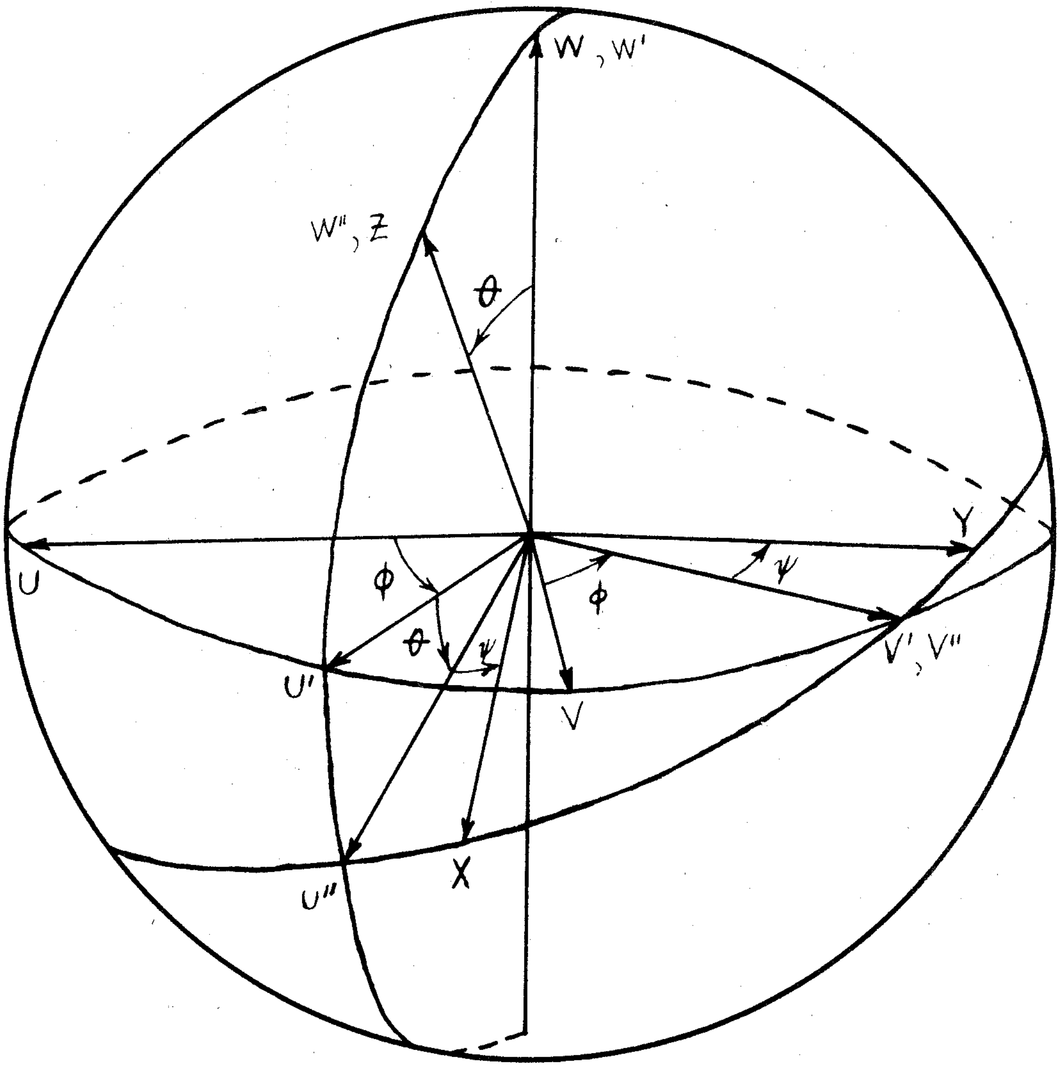


Figure 1. - The Eulerian angles (1).

1) Rotation $\phi$ ($0 \leqq \phi < 2\pi$) about W takes U to U' etc.

2) Rotation $\theta$ ($0 \leqq \theta \leq \pi$) about V' takes U' to U'' etc.

3) Rotation $\psi$ ($0 \leqq \psi < 2\pi$) about W'' takes U'' to X etc.

<u>Table 1.</u> The matrix elements of $\underline{P}_i$.

----------------

$$(\partial X_i/\partial\phi) = (-\sin\phi\cos\theta\cos\psi-\cos\theta\sin\psi)U_i + (\sin\phi\cos\theta\sin\psi-\cos\phi\cos\psi)V_i + (-\sin\phi\sin\theta)W_i$$

$$(\partial X_i/\partial\theta) = (-\cos\phi\sin\theta\cos\psi)U_i+(\cos\phi\sin\theta\sin\psi)V_i+(\cos\phi\cos\theta)W_i$$

$$(\partial X_i/\partial\psi) = (-\cos\phi\cos\theta\sin\psi-\sin\phi\cos\psi)U_i+(-\cos\phi\cos\theta\cos\psi+\sin\phi\sin\psi)V_i$$

----------------

$$(\partial Y_i/\partial\phi) = (\cos\phi\cos\theta\cos\psi-\sin\phi\sin\psi)U_i+(-\cos\phi\cos\theta\sin\psi-\sin\phi\cos\psi)V_i + (\cos\phi\sin\theta)W_i$$

$$(\partial Y_i/\partial\theta) = (-\sin\phi\sin\theta\cos\psi)U_i+(\sin\phi\sin\theta\sin\psi)V_i+(\sin\phi\cos\theta)W_i$$

$$(\partial Y_i/\partial\psi) = (-\sin\phi\cos\theta\sin\psi+\cos\phi\cos\psi)U_i+(-\sin\phi\cos\theta\cos\psi-\cos\phi\sin\psi)V_i$$

----------------

$$(\partial Z_i/\partial\phi) = 0$$

$$(\partial Z_i/\partial\theta) = (-\cos\theta\cos\psi)U_i+(\cos\theta\sin\psi)V_i+(-\sin\theta)W_i$$

$$(\partial Z_i/\partial\psi) = (\sin\theta\sin\psi)U_i+(\sin\theta\cos\psi)W_i$$

----------------

2. For three years I watched the freshmen in their traditional struggle through Swift's "A System of Qualitative Analysis for a Representative Group of the Chemical Elements" (1957). The majority of the students are able to follow the procedures quite well until they get into the aluminum-chromium group (T.O. 13, page 261). At this point several students do not find any Al or Zn when they may have one of these in their sample. It may be that they have lost the Zn to the Pb group or perhaps they inadvertently removed some of the precipitates along with the benzoic acid while dissolving the benzoate precipitates as directed in procedure 72, page 264. How this situation arises is worth further study; however, I wish to propose an alternative procedure for the detection and the estimation of Al and Zn.

I propose to start with the first precipitate under TO 16 on page 286. The following outline details the proposed system:

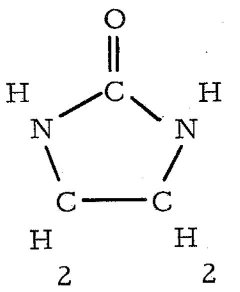
Precipitates: $H_2SiO_3$, $SnO_2 \cdot {}_xH_2O$, $PbCO_3$, $CuCO_3$, $Al(OH)_3$, $ZnCO_3$

Add $HNO_3$ (6N) until $CO_2$ ceases to be evolved, dilute to 6 ml.

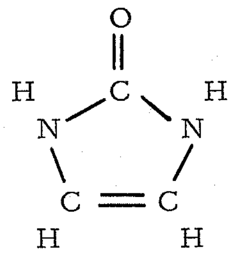
| | |
|---|---|
| Residue: $H_2SiO_3$, $SnO_2 \cdot {}_xH_2O$<br>Treat for $H_2SiO_3$ | Solution: $Pb^{++}$, $Cu^{++}$, $Al^{+++}$, $Zn^{++}$, $NO_3^-$<br>Add: $I^-$ and cool in ice. |
| Residue: CuI ($K_{sp} = 4 \times 10^{-12}$)<br>$PbI_2$ ($K_{sp} = 1 \times 10^{-8}$)<br>Discard or use to estimate for Cu and Pb. | Solution: $Al^{+++}$, $Zn^{++}$, $I_3^-$, $I^-$, $NO_3^-$<br>Remove $I_3^-$ with $SO_3^{=}$ or extract with $CCl_4$ and tritrate for $I_2$.<br>Add: $NH_4OH$ |

| Residue: $Al(OH)_3$ | Solution: $Zn(NH_3)_4^{++}$ |
|---|---|
| To aluminon test procedure 75. | Titrate with $HS^-$ (see procedure 76)<br>(Confirmation: procedure 77) |

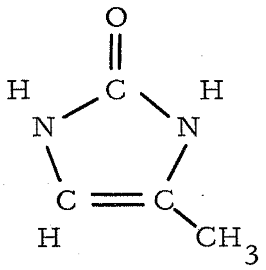
3. In the introduction to part I of this thesis, I have stated that the information obtained from the comparative crystallography of the methyl-substituted ureas can be used to deduce the trial model structures of related compounds. I, therefore, feel somewhat obligated to use the information to predict the structure of some compounds in which I happen to have become interested. I propose that biscisD-amide chains are the major structural feature in the following three compounds:



(I) mp. 131°C
2-keto-1,3-diazolidine (cyclo-ethyleneurea)



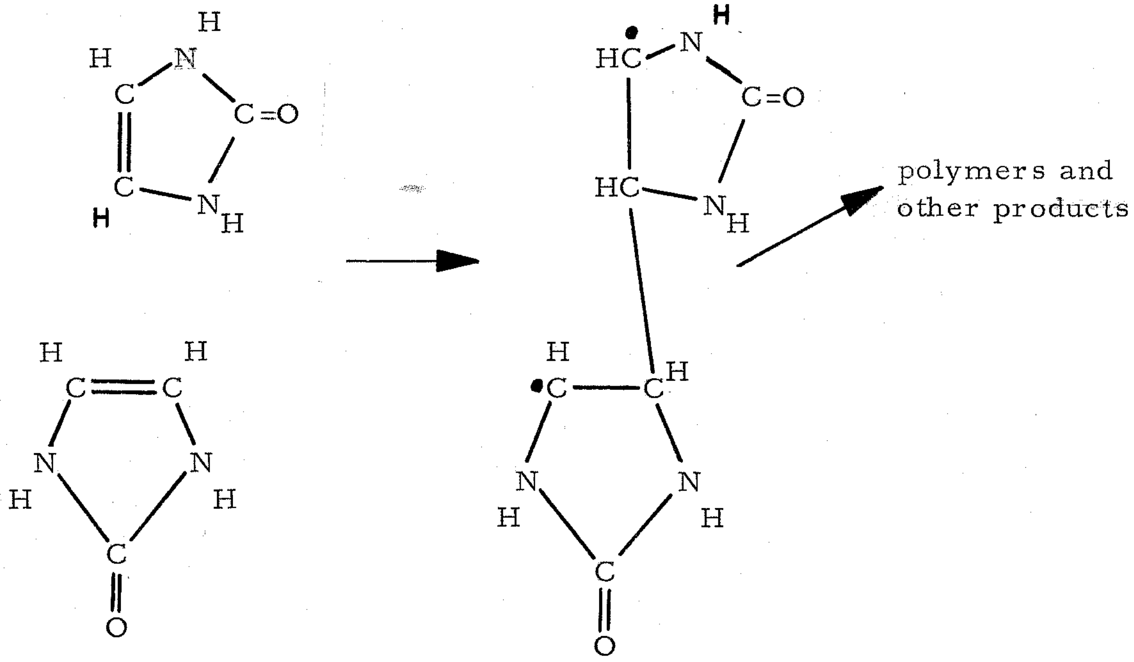
(II) mp. 250°C dec.
2-keto-1,3-diazol-4-ine (2-imidazolone)



(III) mp. 121°C
4-methyl-2-keto-1,3-diazol-4-ine

i) Introduction.

The structures of (I) and (II) are proposed in detail in this proposition. Note also the anomalous differences in the melting points of these compounds. The fact that (II) decomposes before melting indicates that some solid state reaction occurs. The packing of the molecules in the crystals of (II) suggests that a free-radical polymerization may be occurring; that is,



ii) The structure of cyclo-ethyleneurea (I) is proposed.

Although I was able to predict the nature of the hydrogen bond chains, I could not predict the packing of the chains without knowledge of the unit cell dimensions and space group. Since the unit cell information was not available in the literature, I synthesized about 5 grams of the compound using the procedure reported by Fisher and Koch (1). Their procedure is to place stoichiometric portions of pure ethylenediamine and diethyl carbonate into a sealed tube. The tube is then heated for six hours in an oil bath at 180°C.

In my experiments I used a stainless steel reaction bomb of about 30 ml capacity instead of the sealed tubes. (I wish to thank Dr. Robert L. Poynter of the JPL laboratories for the loan of the reaction bomb used for my work.) The product was recrystallized from ethanol until colorless crystals were obtained. One of the crystals was mounted parallel to its c-axis. This crystal was used to obtain a rotation photograph, zero and first layer (hk0 and hk1) Weissenberg photographs plus ($h0\ell$) and ($0k\ell$) precession photographs. These showed that cyclo-ethyleneurea is orthorhombic with a = 11.86 Å, b = 13.42 Å, c = 9.92 Å; space group Fddd:($D_{2h}^{24}$ - No. 70); Z = 16; density (x-rays) = 1.45 g $cm^{-3}$. The density by floatation was found to be about 1.425-1.430 g $cm^{-3}$; however, because of the flaky nature of the crystals used for the floatation experiments the accuracy was very poor. In view of this, I think the two density observations are in fair agreement.

The structure of one chain was assumed to be planar except for the ethylene hydrogens. A drawing of a few links of the chain was prepared by using predicted bond distances, and assuming a cross ring N to N distance of 2.256 Å and assuming a hydrogen bonding distance of 2.90 Å. This showed that the most probable orientation for the chain axis must be coincident with the ac face diagonal. (I have taken the origin of the space group Fddd to be at a center of symmetry.) From these deductions, proposed structural coordinates were obtained. I decided to calculate the structure factors for a few reflections because I was anxious to test the structure.

The first calculation showed poor agreement for the (0k0) reflections. These reflections are quite sensitive to the assumed length of the hydrogen bonds. These could be improved by shortening the hydrogen bond to about 2.75 Å. A second structure factor calculation was made using the coordinates obtained from the new model; the proposed coordinates are given in table 1. The results of the structure factor calculations are given in table 2. The proposed coordinates are probably correct to within 0.15 Å. Visual comparison with the x-ray photographs indicates good agreement. There are a few cases where the relative magnitudes of the observed intensity and the computed structure factor appear reversed. These are probably due to minor inaccuracies in the proposed atomic coordinates, and a refinement of the structure would, I am sure, remove these inconsistencies.

Table 1. - Proposed coordinates* for cyclo-ethyleneurea.

| Atom | x | y | z | |
|---|---|---|---|---|
| $C_1$ | -0.125 | 0.035 | 0.125 | a special position |
| N | -0.052 | 0.089 | 0.052 | |
| O | -0.125 | -0.058 | 0.125 | a special position |
| $C_2$ | -0.075 | 0.194 | 0.075 | |

*Since the molecular two-fold axis is utilized by the space group, it is only necessary to give the coordinates of one asymmetric portion.

Table 2. - Calculated Structure Factors for cyclo-Ethyleneurea Using the Proposed Coordinates given in Table 1.

| | $F_c$ |
|---|---|
| 0 k 0 | |
| 0 | 640.73 |
| 4 - | 23.45 |
| 8 - | 131.14 |
| 12 - | 5.99 |
| 16 - | 3.33 |
| 0 k 2 | |
| 2 - | 134.61 |
| 6 - | 32.02 |
| 10 - | 0.88 |
| 0 k 4 | |
| 0 - | 125.82 |
| 4 - | 54.15 |
| 8 | 47.75 |
| 0 k 6 | |
| 2 - | 89.98 |
| 6 - | 5.52 |
| 0 k 8 | |
| 0 - | 22.72 |
| 4 | 30.89 |
| 2 k 0 | |
| 2 | 141.82 |
| 6 | 33.08 |
| 10 | 0.87 |
| 14 | 9.31 |
| 4 k 0 | |
| 0 - | 145.80 |
| 4 - | 61.41 |
| 8 | 52.28 |
| 12 | 32.01 |
| 16 - | 5.71 |
| 6 k 0 | |
| 2 | 116.01 |
| 6 | 7.74 |
| 10 - | 48.61 |
| 14 | 1.69 |

| | $F_c$ |
|---|---|
| 12 k 0 | |
| 0 - | 28.49 |
| 4 | 4.12 |
| 1 k 1 | |
| 1 - | 86.53 |
| 3 - | 43.98 |
| 5 | 18.68 |
| 7 - | 39.87 |
| 9 - | 7.64 |
| 11 | 23.82 |
| 13 | 1.46 |
| 15 | 1.06 |
| 3 k 1 | |
| 1 | 173.58 |
| 3 - | 77.04 |
| 5 - | 5.19 |
| 7 - | 31.66 |
| 9 - | 40.91 |
| 11 | 12.09 |
| 5 k 1 | |
| 1 | 104.98 |
| 3 | 66.19 |
| 5 | 24.72 |
| 7 | 21.79 |
| 9 - | 37.32 |
| 7 k 1 | |
| 1 - | 28.32 |
| 3 - | 1.45 |
| 5 - | 4.75 |
| 7 | 35.39 |
| 9 - | 1.49 |
| 9 k 1 | |
| 1 | 16.50 |
| 3 | 20.22 |
| 5 - | 13.22 |
| 7 - | 0.61 |

Table 2. (continued)

| | $F_c$ |
|---|---|
| 8 k 0 | |
| 0 - | 30.98 |
| 4 | 41.33 |
| 8 | 8.82 |
| | |
| 10 k 0 | |
| 2 - | 34.12 |
| 6 - | 23.51 |
| 10 | 22.85 |

| 11 k 1 | $F_c$ |
|---|---|
| 1 - | 28.42 |
| 3 | 8.11 |
| 5 | 10.80 |
| | |
| 13 k 1 | |
| 1 - | 12.01 |
| 3 - | 12.29 |
| | |
| 15 k 1 | |
| 1 | 4.66 |

iii) The structure of 2-imidazolone (II) is proposed.

Fenton and Wilks (2) synthesized (II) from urea and dihydroxymaleic acid. They reported that the crystals are tetragonal, probably holohedral with c/a = 0.7063. The forms present were: a 100 quite small and not found on all crystals, p 111 large. The angle ap (100:111) is 60° and pp'' $(111:\bar{1}\bar{1}1)$ is 90°. The crystals are uniaxial and negative.

I have assumed that the molecule lies in a mirror plane with the C=O bond coincident with a $4_2$ axis. The packing of two molecules along this axis would give a c-axis repeat distance of 14.2 Å . A value of 5.00 Å was obtained for the a-axis by using the postulated distance for the c-axis, and assuming that the repeat unit in the biscisD-amide chain is 7.1 Å long and that the goniometric axial ratio corresponds to the expression: c/4a = 0.71. The density calculated from the derived cell dimensions is 1.51 g $cm^{-3}$ for four molecules per cell. This value is higher than the density for <u>cyclo</u>-ethyleneurea and is compatible with the corresponding higher melting point. The trial coordinates listed in table 3 were obtained from a drawing of the trial model, made by using predicted bond distances and angles.

The NH···O hydrogen bond distance between adjacent molecules in the biscisD-amide chain is taken to be 2.8 Å. Viewed down the c-axis, one would see layers of parallel chains. The direction of propagation of the chains in two neighboring layers is perpendicular and oriented at 45° with respect to the a and b-axes. The space group for this trial structure is $P_{4_2/n}\ {}^{2}{}_{1/c}\ 2/m\ (D^{16}_{4h}$ - No 138).

Table 3. - Proposed coordinates for 2-imidazolone. The origin is taken to be at a center of symmetry.

| Atom | x | y | z |
|---|---|---|---|
| O | -0.250 | 0.250 | -0.027 |
| $C_1$ | -0.250 | 0.250 | 0.061 |
| N | -0.093 | 0.093 | 0.110 |
| $C_2$ | -0.156 | 0.156 | 0.206 |

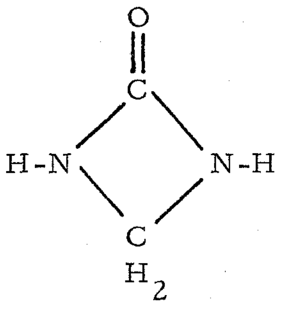
## References

1. V. E. Fisher and H. Kock, Annalen. (1886). 232, 222-228.

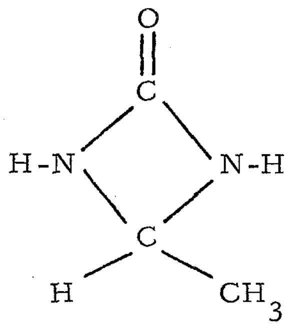
2. H. J. Fenton and W. A. R. Wilks, J. Chem. Soc. (1909). 95, 1329-1334.

4. Synthetic methods for the preparation of several related derivatives of substituted methyleneurea are proposed. A ring closure method is proposed as an alternative technique for the synthesis of substituted uretediones. A method for the preparation of 1-methyl-2-imidazolidone is also appended.

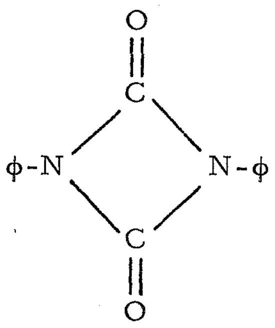
i) The synthesis of methyleneurea (I) was reported by Hemmelmayr (1) and later by Dixon and Taylor (2). Schiff (3) reported the synthesis of related monoethylideneurea (II).



(I)



(II)



(III)

I repeated the Dixon and Taylor synthesis. A powder x-ray diagram of the product indicated that it is either micro crystalline or an amorphous polymer. Taylor (4) and Kadowaki (5) have already postulated the polymeric nature of this material using chemical arguments. It appears that there are no confirmed syntheses of substituted methyleneureas.

An examination of the literature for other compounds with the same basic ring structure disclosed that the compound diphenyluretedione (III) had been synthesized by Hoffman (6). He obtained this product from the dimerization of phenylisocyanate in the presence of pyridine. The basic ring configuration was confirmed by x-ray analysis by Brown (7). Although the interatomic distances in this structure determination require further substantiation (8), the basic configuration reported by Brown is undoubtedly correct. Thus I feel that N, N'-substituted derivatives of (I) should also exist and be reasonably stable at room temperature.

In the following discussion two different routes for the synthesis of substituted methyleneureas are put forward. The first is a direct attack via ring closure techniques. The second is by reduction of the known substituted uretediones.

### ii) Synthesis of substituted methyleneureas by ring closure.

The ring closure reaction step is necessarily competitive with a polymerization reaction. It is, therefore, important to give some consideration to the steric configuration of the reacting system so as to maximize the desired ring closure while at the same time decreasing the opportunity for polymerization.

The intermediate formation of N, N, N'-trisubstituted ureas in which the cis N, N'-urea positions contain unreactive substituents (say, for example, methyl groups), while the trans position contains the reactive substituent, represents one avenue of approach. A one step synthesis of the desired product may possibly be accomplished by reacting N,N'-dimethylurea and ethylal, $EtOCH_2OEt$. The first step in the reaction between these two compounds would give the desired urea intermediate. The trans $-CH_2OEt$ substituent has the advantage that there is reason to hope for the formation of an intermolecular hydrogen bond similar to the type found in biuret (see figure 19 on page 108). This feature should favor the elimination of EtOH with subsequent ring closure. Other candidates as a suitable trans substituent are $CH_2Br$ and $-CH_2SH$ groups, where in the latter case the elimination of $H_2S$ could be enhanced by use of mercuric compounds.

There is some question as to the stability for the product in the reaction mixture. A partial answer to this problem can be obtained by consideration of a ring closure procedure for the preparation of diphenyluretedione (III). I think diphenylurea and diethylcarbonate may react by ring closure to give (III). The isolation of (III) from the products of the reaction, indicating that the ring closure does occur, would strengthen the hope for the success of the method for the preparation of the substituted methyleneureas. A continued failure to isolate the corresponding methyleneurea would probably indicate an instability of the product.

The reactions of $RNHCH_2NHR$ with diethyl carbonate or phosgene look attractive as another avenue of attack.

iii) Synthesis of substituted methyleneureas by reduction of the corresponding uretedione.

It is cogent to consider the reduction of substituted uretediones to give the corresponding substituted methyleneureas. Some chemical reagents that might work are diborane dihydride and $LiAlH_4$. Both catalytic and electrolytic hydrogenation methods are possible approaches.

iv) Synthesis of 1-methyl-2-imidazolidone.

I propose the following synthetic route for the preparation of 1-methyl-2-imidazolidone. Heat stoichiometric quantities of pure $CH_3NHCH_2CH_2NH_2$ and diethyl carbonate in a sealed reaction bomb for 6 hours at 180° C.

## References

1. F. v. Hemmelmayr, Montashefte fur Chemie. (1891). 12, 89-97.
2. A. E. Dixon and J. Taylor, J. Chem. Soc., (1916), 109, 1244-1262.
3. v. H. Schiff, Annalen. (1869). 151, 186-213.
4. J. Taylor, J. Chem. Soc., (1918), 113, 238-248.
5. H. Kadowaki, Bull. Chem. Soc. Japan. (1936), 11, 248-261.
6. A. Hofmann, Annalen. (1871), Suppl. 1, 57-58.
7. C. J. Brown, J. Chem. Soc., (1955), 2931-2936.
8. W. J. Takei, Ph. D. Thesis, California Institute of Technology, (1957), pp. 59-67.

5. I propose a method for manipulating a set of normal equations for which the determinant of the equations is of an order less than the number of parameters being refined.

i) Introduction.

I wish first to explain how such a situation can arise (1). Assume that a set of observed quantities $F_o$ is adequately represented by a non-linear function $F(\lambda)$, where $\lambda$ denotes the set of parameters $\{\lambda_i\}$. Suppose further that there are other sets of parameters, say $\{\lambda_i + a\}$ (where $a$ is any arbitrary constant but is the same constant for all $\lambda_i$), which equally well represents the observables. This means that the $\lambda_i$ are not linearly independent. Assume that the linear relationship is unknown. Further, suppose that for purposes of other physical considerations any one of the equally satisfactory parameter sets suffices as a basis for making significant inferences.

Thus we wish to refine by least-squares methods an approximate set of parameters $\{\zeta_i\}$ so as to make it converge to any one of the $\{\lambda_i\}$ sets. Let us assume $\lambda_i = \zeta_i + \epsilon_i$ and by Taylor series expansion a residual equation linear in the $\epsilon_i$ is obtained which in turn is used to give the normal equations: $\underline{M}\ \underline{\epsilon} = \underline{Q}$. Because of the linear dependence of the parameters, the determinant of $\underline{M}$ is equal to zero. The question now is how do we proceed?

ii) A method of solution is proposed.

Assume that there are m normal equations and that the order of $\underline{M}$ is n, where m = n + 1. The geometric representation of the normal equation matrix is a surface in m-dimensional $\epsilon$ space, which must look like an elliptic cylinder, since all minors of the diagonal elements of $\underline{M}$ are symmetric positive definite. Let $\underline{N}$ be the minor consisting of the first n by n elements of $\underline{M}$ . We first solve for the line of centers of the cylinder. By matrix methods we obtain:

$$\text{I} \qquad \left|\begin{array}{c} \epsilon_1 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_n \end{array}\right| = \underline{N}^{-1} \left|\begin{array}{cc} Q_1 & -M_{1m} \\ \cdot & \\ \cdot & \\ \cdot & \\ Q_1 & -M_{nm} \end{array}\right| \begin{pmatrix} 1 \\ \\ \epsilon_m \end{pmatrix} = \left|\begin{array}{cc} q_1 & -a_{1m} \\ \cdot & \\ \cdot & \\ \cdot & \\ q_n & -a_{nm} \end{array}\right| \begin{pmatrix} 1 \\ \\ \epsilon_m \end{pmatrix}$$

And the equation of the line in parametric form becomes:

$\epsilon_i = q_i - a_{im}\,\epsilon_m$ for i = 1 to n

$\epsilon_m = \epsilon_m$ , where $\epsilon_m$ is the parameter.

Any point along this line defines a theoretically acceptable set of $\epsilon_i$ shift values. For example, taking $\epsilon_m = 0$ is particularly convenient, since it permits us to ignore the column of terms $-a_{1m}$ to $-a_{nm}$. Another choice, where a digital computer is used to accumulate the elements of $\underline{M}$, would be to take $\underline{\epsilon} = \underline{M}^{-1}\,\underline{Q}$. (It is unlikely that the determinant of $\underline{M}$ would be identically zero, because of the inherent digital limitation of computers in general.) The objection to this method is that the set of $\epsilon_i$ could be inconveniently large. However, the fact that this choice might be reasonably acceptable could mask and prevent the detection of an inherent singularity in the system of equations.

As we shall see, the best method for choosing $\epsilon_m$ is to demand that the m-dimensional distance from the origin to the cylinder axis be a minimum. That is, we want $\sum_{i=1}^{m} (\epsilon_i)^2$ to be a minimum. This requires that

$\sum_{i=1}^{m} \epsilon_i(\partial\epsilon_i/\partial\epsilon_m) = 0 = \epsilon_m - \sum_{i=1}^{n} a_{im}\,\epsilon_i$ . The optimum value of $\epsilon_m$ is given by

$$\epsilon_m = \frac{\sum_{i=1}^{m} a_{im} q_i}{1 + \sum_{i=1}^{n} a_{im}^2} \quad , \text{ since}$$

$$0 = \epsilon_m - \sum_{i=1}^{n} a_{im}\epsilon_i = \epsilon_m - \sum_{i=1}^{n} a_{im}(q_i - a_{im}\epsilon_m) = \epsilon_m\left(1 + \sum_{i=1}^{n} a_{im}^2\right) - \sum_{i=1}^{n} a_{im} q_i$$

Once the value of $\epsilon_m$ is available, the other shift values are easily computed using the parametric equation of the line of centers; that is, $\epsilon_i = q_i - a_{im}\,\epsilon_m$.

iii) Statistical significance.

The values for $\sigma(\epsilon_i)$ for i = 1 to n are computed in the usual way from knowing the value of the elements, denoted $n_{ij}$, of $\underline{N}^{-1}$ and the sum of the squared differences $(F_o - F(\lambda))^2$, denoted as $\Delta^2$. Thus we have $\sigma(\epsilon_i) = [n_{ij}\ \Sigma\Delta^2/(r-s)]^{1/2}$, where (r-s) is the excess of data over parameters determined.

The value of $\sigma(\epsilon_m)$ is obtained from the expression $\epsilon_m = \sum_{i=1}^{n} a_{im}\epsilon_i$ by use of the addition rule of variance (2): $\sigma^2(\zeta_1 + \zeta_2 + \cdots \zeta_n) = \sigma^2(\zeta_1) + \sigma^2(\zeta_2) + \cdots + \sigma^2(\zeta_n)$ and the rule that $\sigma(a\zeta + b)$ is equal to $|a|\ \sigma(\zeta)$. With these we find that

$$\sigma(\epsilon_m) = [\sum_{i=1}^{n} a_{im}{}^2\sigma^2(\epsilon_i)]^{1/2}.$$

iv) Other suggestions.

Quite often the normal equations are solved assuming that the off-diagonal terms are zero. I suggest that the geometric orientation of the ellipsoid be determined for a series of related problems for both the complete matrix and the matrix obtained by assuming that some elements are zero. These then can be examined to determine if there exists a mean or systematic shift factor relating the exact shift to the approximate shift. (For example, it might be found that the exact shift is always less than the approximate shift by about a factor of 3/4.) The knowledge and use of such a factor could then expedite the refinement of the parameters.

## References

1. D. H. Templeton, Acta Cryst., 12, 771-773 (1959).
2. H. Cramer, Mathematical Methods of Statistics, (1957). Princeton: Princeton University Press, pp. 179-180.

6. A proof of a linear dependence among the least-squares shifts and among the parameters themselves in polar space groups is proposed.

i) Proof of linear dependence of parameters in polar space groups.

We have a function $G(x)$ where $x$ denotes the set of parameters $\{x_i\}$. We also know that $G(x) = G(x + a)$, where $a$ is any arbitrary constant. Let $\zeta(a)$ denote the set $\{x_i + a\}$. We can write the equation of the dependence in parametric form. Since $\zeta_i(a) = x_i + a$, we have $\zeta_m(a) = x_m + a$. Let us take $a = \mu - x_m$, where $\mu$ is another arbitrary constant. Thus we have $\zeta_i(a) = \zeta_i(\mu) = x_i - x_m + \mu$ and $\zeta_m(\mu) = \mu$. Thus we see that we have a linear dependence of the $\zeta_i$ for $i = 1$ to $m - 1$ on $\zeta_m$ in a parametric form.

ii) Linear dependence of the shift terms.

Let a best parameter set be denoted as $\lambda = \{\lambda_i\}$, where $\lambda_i = x_i + \epsilon_i$. We also have $\lambda_i(a) = x_i + \epsilon_i + a$. Let $a = \mu - \epsilon_m$, then we have $\lambda_i(a) = \lambda_i(\mu) = x_i + \epsilon_i - \epsilon_m + \mu$ and $\lambda_m(\mu) = x_i + \mu$. If we let $\Delta_i = \epsilon_i - \epsilon_m + \mu$, and $\Delta_m = \mu$, we see that we have a linear dependence of the $\Delta_i$ for $i = 1$ to $m-1$ on the $\Delta_m$ in a parametric form. This linear dependence implies that the least-squares equations $\underline{M}\,\underline{\Delta} = \underline{Q}$ must be singular and the determinant of M, denoted $|M|$, is equal to zero.