Computing
in
Crystallography

C10053
10568

P1786
3158

BIBLIOTHEEK
TECHNISCHE UNIVERSITEIT
DELFT

```
    .    .    .    .    .    .    .    .    .    .    .  1.8  2.6  2.7  1.9    .    .    .    .
    .    .    .    .    .  2.4  2.9  2.4    .    .    .  3.1  4.8  5.1  3.6  1.7    .    .    .
    .    .    .    .  1.9  3.7  4.6  3.8  1.9    .    .  3.8  6.1  6.7  5.1  2.5    .    .    .
    .    .    .    .  1.9  4.1  5.1  4.2  2.2    .    .  3.3  5.6  6.2  4.8  2.5    .    .    .
    .    .    .    .    .  3.0  3.8  3.2  1.6    .    .  1.9  3.5  3.9  3.1  1.7    .    .    .
    .    .    .    .    .    .  1.7    .    .    .    .    .    .    .  1.5  1.5    .    .    .
    .    .  2.3  2.4  1.7    .    .    .    .    .    .    .    .    .    .  2.7  3.3  2.6    .
    .  3.1  4.8  4.9  3.3    .    .    .    .    .    .    .    .    .  1.9  4.3  5.4  4.5  2.4
    .  3.0  6.2  6.1  4.4  1.6    .    .    .    .    .    .    .    .  2.2  4.8  6.2  5.2  2.9
    .  3.6  5.6  5.8  3.9    .    .    .    .    .    .    .    .    .    .  3.6  4.8  4.2  2.5
    .  2.3  3.7  3.7  2.4    .    .    .    .    .    .    .    .    .    .    .  2.3  2.1    .
    .    .  1.6  1.6    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .
    .    .    .    .    .  2.2  2.8  2.4    .    .    .    .  2.5  2.7  1.7    .    .    .    .
    .    .    .    .  2.0  3.9  4.9  4.0  1.8    .    .  2.2  4.3  5.0  3.7    .    .    .    .
    .    .    .    .  2.1  4.6  5.7  4.6  2.1    .    .  2.7  5.1  6.0  4.7  2.1    .    .    .
    .    .    .    .    .  3.4  4.6  3.8  1.8    .    .  2.3  4.4  5.1  4.0  1.9    .    .    .
    .    .    .    .    .    .  2.3  2.1    .    .    .    .  2.5  3.0  2.3    .    .    .    .
    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .
    .  1.9  2.5  2.1    .    .    .    .    .    .    .    .    .    .    .    .  2.2  2.1    .
  1.6  3.3  4.5  4.4  2.9    .    .    .    .    .    .    .    .    .    .  2.9  4.1  3.9  2.5
  1.7  3.5  5.1  5.3  3.9  1.8    .    .    .    .    .    .    .    .    .  3.6  5.1  4.8  3.2
    .  2.7  4.1  4.5  3.5  1.7    .    .    .    .    .    .    .    .    .  3.3  4.3  4.1  2.7
    .    .  2.3  2.7  2.3    .    .    .    .    .    .    .    .    .    .  1.6  2.3  2.2    .
    .    .    .    .  1.6  1.5    .    .    .    .    .    .    .    .    .    .    .    .    .
    .    .    .    .  2.0  2.8  3.0  2.2    .    .    .    .  2.2  2.6  2.3    .    .    .    .
    .    .    .    .  2.6  4.2  4.7  3.9  2.2    .    .  2.2  3.8  4.6  4.0  2.4    .    .    .
    .    .    .    .  2.5  4.3  5.1  4.5  2.9    .    .  2.6  4.3  5.2  4.7  3.0    .    .    .
    .    .    .    .  1.6  3.1  3.9  3.6  2.5    .    .  2.2  3.4  4.1  3.7  2.5    .    .    .
    .    .    .    .    .  1.5  2.1  2.1  1.6    .    .    .  1.9  2.3  2.1    .    .    .    .
    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .    .
```

Computing in Crystallography

Proceedings of an
International Summer School on
Crystallographic Computing
held in Twente, The Netherlands
24 july - 1 August, 1978

H. Schenk
R. Olthof-Hazekamp
H. van Koningsveld
G. C. Bassi
eds

1786 3158

BIBLIOTHEEK DER TECHNISCHE HOGESCHOOL
78
012338
DOELENSTR.101
DELFT

DELFT UNIVERSITY PRESS

Published en distributed by
Delft University Press
Mijnbouwplein 11
2628 RT DELFT

Cover design:
B. W. van de Waal
Slant Plane Fourier Transform of quinazoline

Copyright © 1978 by Delft University Press.
No part of this book may be reproduced in any form by print, photoprint, microfiche or any other means without written permission from the publisher.

ISBN 90 6275 025 7

CONTENTS

PREFACE

This book consists of a series of papers on the basis of which the 1978 Summer School on Crystallographic Computing will be given. The idea to organize this school at the campus of the Twente University of Technology (Enschede, The Netherlands) originates from the Netherlands Discussion Group for Direct Methods and was supported by the Netherlands Organisation for Crystallography FOMRE, the Royal Netherlands Chemical Society KNCV, The International Union of Crystallography and its Commission on Crystallographic Computing and the Twente University of Technology.

The local organisation was taken care of by a committee consisting of H. Schenk (chairman), R. Olthof-Hazekamp, H. van Koningsveld (treasurer), R.A.G. de Graaff, C. Huiszoon, H. van der Meer, T.C. van Soest and A.L. Spek (members).

The scientific program was organised by a program committee consisting of G.C. Bassi, J.N.M. Bruggeman, J.L. Derissen, R.A.G. de Graaff, H. Schenk (chairman) and T.C. van Soest, under supervision of the Commission on Crystallographic Computing of the I.U.Cr. The School was focussed on the computational aspects of crystallographic research rather than on the description of crystallographic techniques. This book therefore gives a rather complete survey of the crystallographic computational facilities now available for large and small computers. The four topics of the School are:

1. Program-systems on large and mini-computers (chapters 1, 2 and 3 of this book)
2. Automatic data collection (chapter 4)
3. Molecular interactions (chapter 5)
4. Miscellaneous (chapter 6)

In six mornings the more than 20 subjects will be covered in introductory plenary lectures, followed by afternoon practical sessions of 3½ hours concerning the same subjects in classes of 15 students. The lecturers have the possibility to cover their subject in more detail and depth in these afternoon sessions. In this way the students are introduced to all subjects and get a more detailed knowledge of those six in which they are most interested. In addition, the school program features a computer program market, ad-hoc meetings and discussion sessions.

It is a great pleasure to acknowledge the generous support given to the School by the following sponsors:
Control Data, Digital Equipment, Dutch States Mines, Enraf-Nonius, International Business Machines Corporation, International Union of Crystallography, Koninklijke Nederlandse Chemische Vereniging (Royal Netherlands Chemical Society), Mathematical Centre Amsterdam, Mathematical Centre Leiden, Philips, Syntex Analytical Instruments, Twentse Hogeschoolfonds, Twente University of Technology, Unilever.

The School edition of the Proceedings was sponsored exclusively by Digital Equipment Corporation.

The editors also gratefully acknowledge the cooperation of the authors to fulfill the sharp time limits for submitting their manuscripts. The editors also wish to thank the Delft University Press, which offered to produce this edition of the Proceedings of the School.

Apart from the bodies mentioned above the editors are indebted to many individuals for their valuable assistence in realizing the School, in particular to Drs B. van Geelen, B.O. Loopstra, D.P. den Os, W.P. Pijper and J.N. King.

Amsterdam, June 1978

H. Schenk
R. Olthof-Hazekamp
H. van Koningsveld
G.C. Bassi

CHAPTER 1

Program systems on large computers

PROGRAM SYSTEMS

James M. Stewart*

Summary

The purpose of this presentation is to delineate some methods and practices related to the programming of digital computers which will give rise to program systems having the desired properties delineated.
There are a few clearly identifiable categories, each of which will aid in systematization of programs. They are Documentation, Input-Output, Crystallographic Generality, and Accommodation to a Generalized Turing Machine. Each category when carefully applied to a program will cause the results to be more systematic. When all categories are applied together the result will be a highly systematic program.

Introduction

The dictionary definition of "system" is: "anything formed of parts placed together to make a regular and connected whole working as one machine" or "a method of organization; a customary plan; a method of procedure; a regular method or order; a method or scheme of classification." From the point of view of crystallographic computing any or all of these definitions have merit but it is the first which has the greatest appeal to the working scientist who is carrying out crystallographic studies.

For the purpose of structure determination from the view of the research scientist, the "parts" are the familiar calculations required to carry the raw diffraction data to the final refined structure and the interpretation of and display of that structure in an instructive manner.

Documentation (figure I)

Documentation is the most important aspect of preparing an application program. If it is lacking, all other programming effort will be lost. An author unwilling to document his programs would be well advised to abandon his efforts. He will be either the only person ever to make use of the programs or he will become the one all others who use the programs depend upon for preparation of data. Good documentation consists of two major categories; instructions for use of the programs and comments on the program code.

---

*University of Maryland U.S.A.

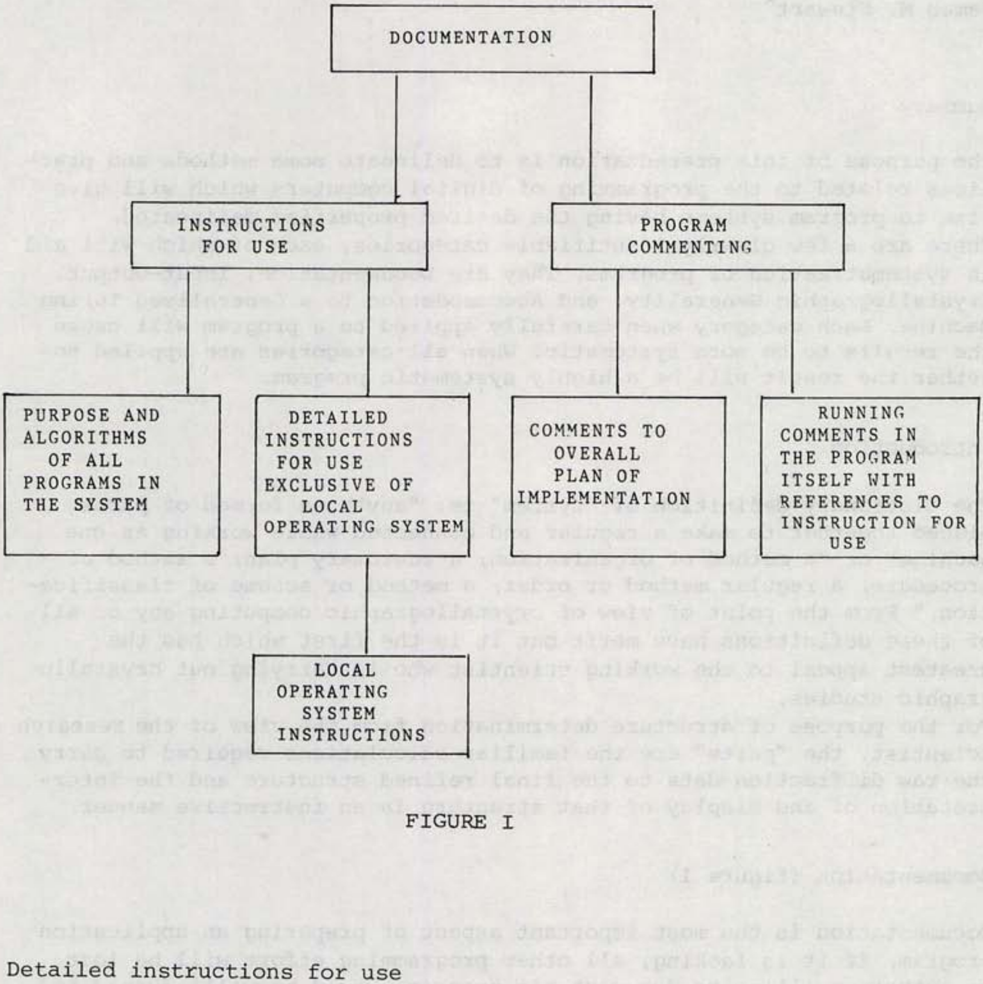


FIGURE I

Detailed instructions for use

The detailed instructions for the use of programs may be divided into three categories. The first and most important are the detailed instructions concerning the necessary data and "keys" for causing a given calculation to be accomplished. The form in which the data and keys are to be supplied, must be carefully delineated. This may be done in many systematic ways. The two most common are through formatted and non-formatted (or free-formatted) input lines.

It has become customary in computing to use either cards or some kind of keyboard machine to accept data. Under these conditions, one "line" or one "card" constitutes the smallest collection of information acted upon by the machine. Under the control of formatted lines, each item of data in the line must conform rigidly to the previously programmed format. Under the free-format mode of input, some mark of punctuation is used to separate the data items on a line.

Regardless of the method of data input, the detailed instructions of the program system will contain all the information required to supply the requisite data and keys to a given program.

Local operating system instructions

A special category of detailed instructions must be recognized by the programmer preparing application programs. These are the detailed instructions which must be given to "run a job" on any given machine. The formulation of these instructions was in general outside the programmers control. The operating system of any computer is, however, based on exactly the same principles as is the system which he is preparing. If the application program is to be used on many different machines, very little use, or better, no use will be made of "special" features of a local operating system.

Under this method of preparing documentation there will be an addendum to the detailed instructions for the crystallographic system for each different machine and local operating system on which it is utilized.

Explanation of purpose and algorithms

Once documentation of the instructions has been prepared, many knowledgeable people will be able to make use of the "system" which has been programmed. However, another important part of the documentation procedure is a description of the purpose and structure of the algorithms actually programmed. This will be a great help to less experienced users and will assure experienced users that the calculations performed are, in fact, the ones the user wishes to have performed. This phase of the documentation is a brief textbook with references to the literature as well as to the detailed instructions. In addition, sample input data are shown for common calculations. In general examples of use are the most powerful means of explaining the relationship between input data and calculations to be performed.

Program documentation

The process of documenting the programs themselves is complicated somewhat by the method of programming. Whatever the method, be it "machine language" (e.g. assembler), FORTRAN, ALGOL, PASCAL, RATFOR, or any other, there will be a provision for "commenting" the coded instructions. If the programmer has any hope that others, besides himself, or even he himself will later need to modify or correct the code, there must be documentation as to the purpose and meaning of the code.

Running comments

Every line of code deserves to be considered for commenting. Moreover, the comment must not be fatuous, simply echoing what is obvious from the code.
For example: PRINT 5, A, B, C, should not have the comment:
C PRINT A LINE
but rather should have a comment such as:
C PRINT DIRECT CELL CONSTANTS BEFORE REFINEMENT.

In addition to the line by line comments, there should be a general statement at the front of each "routine" (program, subroutine, etc.) which gives the purpose and scope of the calculation carried out by the routine. This statement should also have a glossary which relates the symbols used in the programs to the scientific data which the symbols represent. Variables which are local and temporary should be clearly distinguished from those which are global and used by other routines.

Overall plan of implementation of a program system

If the program system is small, the overall plan of implementation may be unnecessary. However, if it consists of many different routines which may be used either together or separately, then it is necessary to provide comments on how all the parts of the system fit together. On all the common machines and operating systems, the programs which accept the codes for the application systems accept them as a "symbolic deck". These decks are then operated upon by the local "compiler" (assembler, FØRTRAN processor, etc.) and from this operation a "relocatable element" will be produced. The exact name of these "relocatable elements" will vary from machine to machine, but in general they constitute small portions of the overall scientific calculation in machine code with all the addresses of the code relative to the same starting address. When all of the "relocatable elements" of a large system of application programs are to be used, it is necessary to have some method of describing how this is to be done. The description of this plan is an important feature of the documentation of a program system.

A very difficult problem in documentation is caused by the lack of standards for "loaders". The problem may be compared in many respects to that of the documentation of the local operating system of which the loader is a part.

Systematic treatment of input and output (figure II)

As described above under "Detailed instructions for use", the most important aspect of creating a program system, is the ability to make use of the system by means of line input. The problem of documenting and hence of using the program system will be simplified if the structures of all input lines are carefully formatted under a set of easily learned and remembered conventions. Similar data should always be prepared in the same way. This means that the program system will have a subroutine for treating "raw data" and a clearly defined method of "tagging" the data so that it is clear what measurements or other information the numbers represent. Data accepted as "line input" will be transformed and stored, where possible and necessary, in a structured mass storage file of some kind (e.g. disc). This mass storage file will obviate the need to supply the same information over and over again through the line input device.

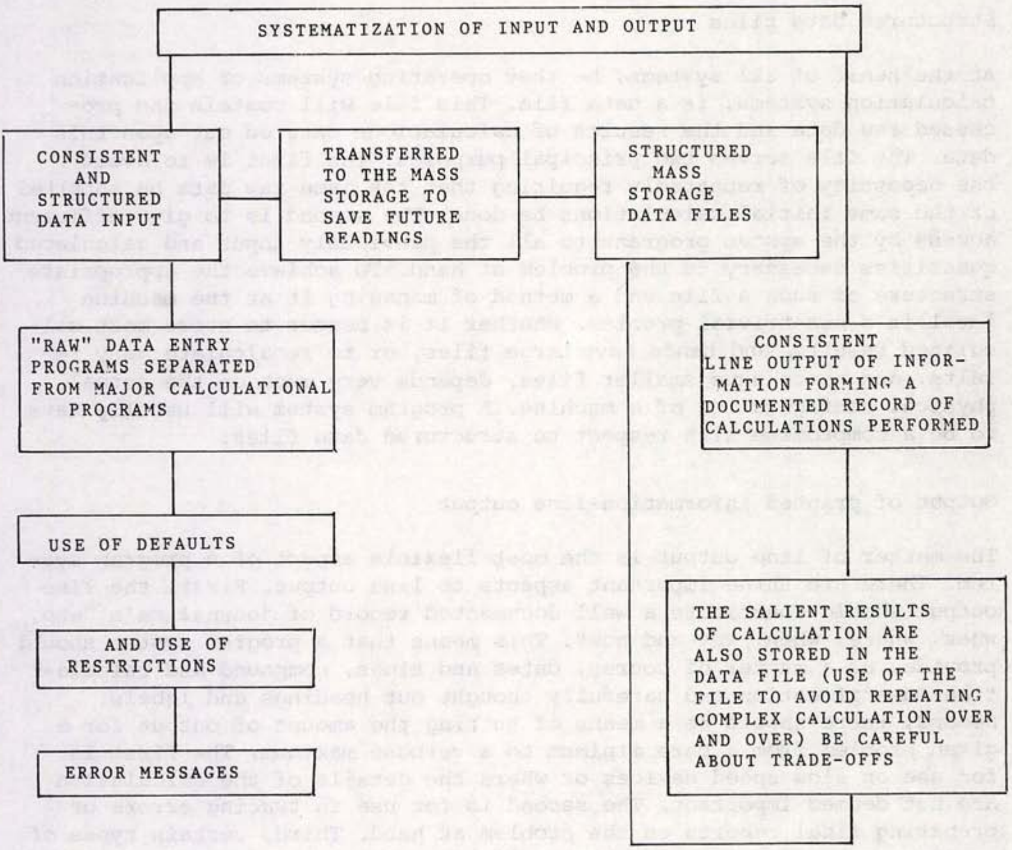


FIGURE II

Raw data entry

Special routines may be set up in a program system to serve for the screening and conversion of "raw data" into a structured mass storage file. These programs can be separated from the major calculational programs of the system and will serve to set "default" values, protect against physically unreasonable or inconsistent raw data and to produce useful error diagnostic messages concerning detected irregularities. This "buffering" of the line input from the "working" programs should help mitigate wasted machine time, aid in making the documentation of the programs clearer to novice users, and prevent careless mistakes.

Structured data files

At the heart of all systems, be they operating systems or application calculation systems, is a data file. This file will contain the processed raw data and the results of calculations carried out upon this data. The file serves two principal purposes. The first is to avoid the necessity of repeatedly requiring that the same raw data be supplied or the same initial calculations be done. The second is to give efficient access by the system programs to all the previously input and calculated quantities necessary to the problem at hand. To achieve the appropriate structure of such a file and a method of managing it at the machine level is a non-trivial problem. Whether it is better to store most calculated results, and hence have large files, or to recalculate many results, and hence have smaller files, depends very much on the actual physical configuration of a machine. A program system will usually have to be a compromise with respect to structured data files.

Output of printed information-line output

The matter of line output is the most flexible aspect of a program system. There are three important aspects to line output. First, the line output should constitute a well documented record of journalism's "who, what, where, when, why and how". This means that a program system should provide, as a matter of course, dates and times, compound and calculation identification and carefully thought out headings and labels. Second, there should be a means of setting the amount of output for a given problem from a bare minimum to a verbose maximum. The first is for use on slow speed devices or where the details of the calculation are not deemed important. The second is for use in tracing errors or preparing final reports on the problem at hand. Third, certain types of error messages should be highly stylized, as clear as possible in delineating their cause, and never suppressed. This feature can save the user of a program system quite a bit of frustration and unnecessary effort.

Generality of calculations (figure III)

A program system should carry out (in the greatest generality possible), the calculations of the discipline for which it is created. The programmer should seek algorithms which will not restrict the calculation in any way, yet will still make physical sense. For crystallographic calculations, generality would be applied especially to symmetry, problem size and measurement techniques.

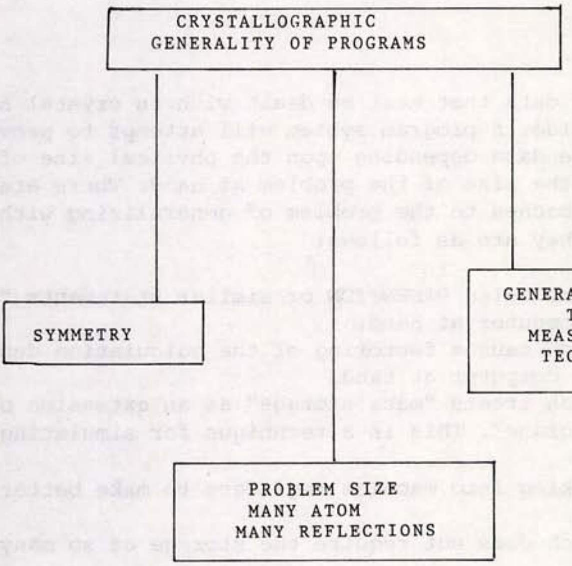


FIGURE III

Symmetry

The very nature of crystallographic symmetry is systematic and therefore leads directly to a means of treating symmetry in general. The International Tables for Crystallography Vol. I give the general equivalent positions of all 230 space groups in their "standard settings". For digital computers it is usually more useful to store the symmetry information in the form of rotation matrices and translation vectors than in the "algebraic" notation of the International Tables. Programs have been written for interpreting either the space group symbol or the "algebraic" notation and for forming the corresponding matrices.

In crystallographic system programs, there are two principal approaches which can be made to the general coding of calculations with respect to symmetry. One is to code all calculations as if they were either centric triclinic or acentric triclinic. In this case the stored symmetry operators may be used to generate as needed, the symmetrically related reflections in reciprocal space or points in direct space. The other approach is to use the "factored" form of the structure factor or Fourier expressions found in the second half of the International Tables Vol. I. Both methods have merit and can be used to generalize programs with respect to symmetry. The first method may require more calculation time in the program: the second may require more time and complexity in the program.

Problem size

The range of amounts of data that must be dealt with in crystal structure analysis is very wide. A program system will attempt to provide means of "factoring" the data depending upon the physical size of the computer available and the size of the problem at hand. There are five somewhat different approaches to the problem of generalizing with respect to problem size. They are as follows:

1) Recompile the programs using DIMENSION or similar statements "tuned" to the problem and computer at hand.
2) Use an algorithm which causes factoring of the calculation depending upon the size of the computer at hand.
3) Use an algorithm which treats "mass storage" as an extension of "immediate access storage". This is a technique for simulating "virtual memory".
4) Use "bit by bit" packing into machine registers to make better use of available storage.
5) Use an algorithm which does not require the storage of so many intermediate results.

"Dimension" adjustment

This method of storage management is of limited use especially for "small" machines. It also requires that much time be spent in program modification on a problem by problem bases.

Factoring of the problem

For crystallographic calculations, the technique of factoring is often used. Structure factors, for example, may be summed in parts. Electron density calculations have always been summed by factoring methods because the process saves space and enhances the speed of the calculations.

Simulating "virtual memory"

Many machines have operating systems which provide a feature called "virtual memory". Under these operating systems, no limit, or a very high limit is placed on the size of arrays which may be specified by users. This system can be simulated on any machine by treating a small region of immediate access memory as a buffer and using a mass storage device as the actual region of storage. The method can be costly in time for those calculations which "skip around" in the data (e.g. phase determinations). This is because "random access" will require many input output operations as each buffer of data is written to and read from the mass storage devices. On the other hand, calculations which pass through the data in a systematic fashion (e.g. Fourier transforms) will not be slowed much by a "virtual memory" system since each datum is treated in turn from the first to the last.

Packing

Packing capitalizes on the fact that for many calculations the range of magnitude of numbers used is very small relative to the range provided for in a computer word. For example, the magnitude of h, k, or l for crystal structures will be 0 to 63 which can be represented as 000000 to 111111, in binary. A machine with words of 32 bits can thus "pack" h, k, l, and a sign indicator in one word rather than four. Packing is often used in direct phasing methods in crystallography. There are of course overheads in time, when packing and unpacking must be done. But the use of a "virtual memory" algorithm as described above, will be even slower than packing when random access is used.

"Overlaying" of programs

Another important method of simulating "virtual memory" while leaving as much space as possible for data is the use of program overlaying. In this method, the calculation to be performed is broken down into as many "significant" steps as possible. To use crystallographic least-squares as an illustration: First, load atom parameters and clear the matrix; second, calculate constraints for special positions and groups; third, calculate the structure factors and derivatives and build the matrix; fourth, invert the matrix and apply the shifts to the parameters; fifth, write the new parameters to the file. In this case, the program for each step in the process is brought into the computer in turn. This is completely analogous to the case where data is buffered. This method simply buffers the program instead.

Use of other algorithms

Changing the way the calculation is done may seem like an obvious or trivial method for dealing with the problem of saving time and space, but it is one much used. The common example in crystallographic computing is the use of diagonal or block diagonal calculations instead of full matrix calculations in least-squares refinement. In this case there are debateable trade-offs in time and results but the method works and is the only one practical for solving large problems on small machines.

Another important aspect of this method of dealing with application programs is the possibility of discovering completely new algorithms which will simply not require the storage of intermediate results or the amount of time the current algorithms require.

Summary of space husbanding methods

In practice, a combination of all of the methods mentioned in the foregoing paragraphs is necessary for the successful preparation of a large application program system. In what follows, an attempt will be made to discuss the constraints imposed by the variety of machines and operating systems available in the market place.

Range of calculations performed

An application program system will, in general, have a very wide range of different calculations that may be carried out. For crystal structure analysis, the following partial list indicates those often included:

1) Diffractometer or film raw data to integrated intensity data (X-ray or neutron; powder or single crystal; small or large molecule)
2) Cell reduction
3) Reflection sorting and averaging
4) Absorption corrections
5) Calculation of estimated normalized structure factors, E
6) Various "direct" methods of estimating phases
7) Patterson search methods
8) "Other" methods of structure solution
9) Translation functions
10) Fourier Transforms
11) Structure factor
12) Structure factor least-squares
13) Constrained refinement least-squares
14) "Other" refinement methods - differential synthesis etc.
15) Bond lengths and angles with errors
16) Least-squares planes and lines
17) Thermal motion analysis
18) Graphic display of structure
19) Data preparation for publication
20) Electron density distributions and bonding
21) Data and information archiving

Accommodation to a Turing machine (figure IV)

Most of the computers which are available at the present time are very similar in their general architecture. These machines consist of some sort of "memory" device and of various registers for interpreting "instructions" and manipulating data. In addition, machines must have some means of receiving data for storage and calculation and displaying these data and the results of the manipulations of these data. Thus it is, in theory, possible to write programs which may be used on any and all Turing machines.

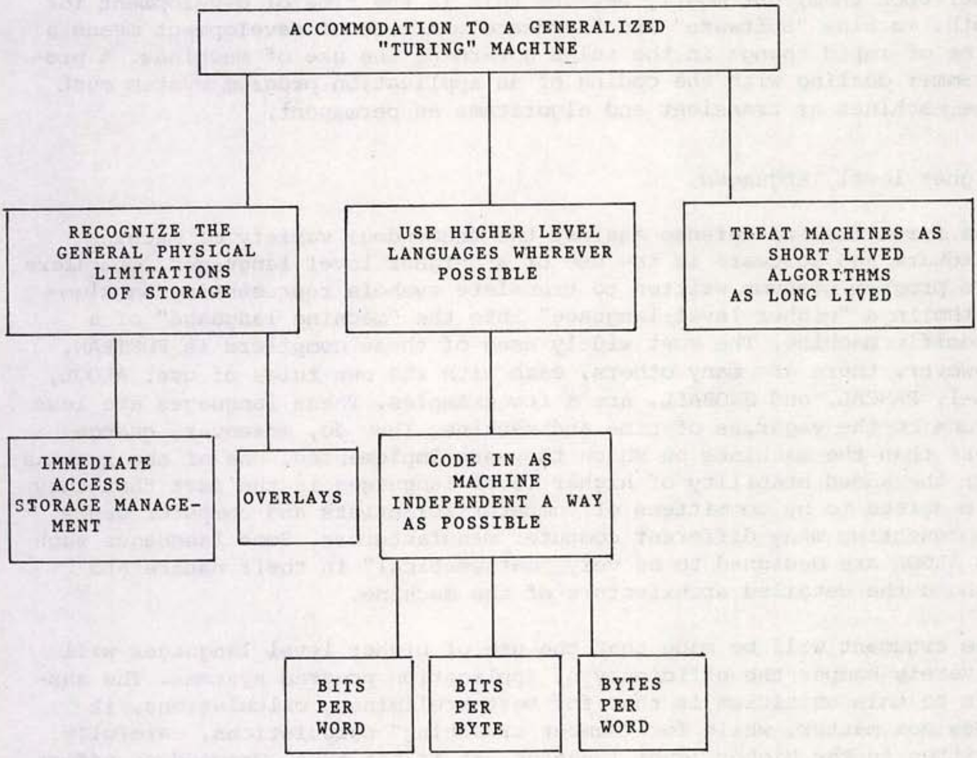


FIGURE IV

The truth of the matter, however, is that the details of the structure of each machine are different so that in practice, the use of every machine is governed by a different set of detailed rules. In order to be able to make a program system that will run on many machines, it is necessary to study carefully the rules for many different machines. During this study one must look for a method which will avoid the specific details while responding to the common basis for computing.

Machines and algorithms

The basis of all attempts to build an application program system will be the algorithms which represent the calculations required by the discipline. These algorithms will, in general, have a very long life. In most cases where the theoretical understanding of the subject is as advanced and as sophisticated as it is in diffraction, the life of the algorithms will be "from now on". Machines and their operating systems on the other hand, are very short lived partly because of the wear and tear upon them, but mostly because this is the time of development for both, machine "software" and "hardware". A time of development means a time of rapid change in the rules governing the use of machines. A programmer dealing with the coding of an application program system must see machines as transient and algorithms as permanent.

Higher level languages

The first line of defense against the tremendous variety of machine hardware and software is the use of a "higher level language". Compilers are program systems written to translate symbols representing an algorithm in a "higher level language" into the "machine language" of a specific machine. The most widely used of these compilers is FORTRAN. However, there are many others, each with its own rules of use. ALGOL, PL-1, PASCAL, and SNOBALL, are a few examples. These languages are less prone to the vagaries of time and machine. They do, moreover, change less than the machines on which they are implemented. One of the reasons for the added stability of higher level languages is the fact that they are agreed to by committees of computer scientists and computer users representing many different computer manufacturers. Some languages such as ALGOL are designed to be very "mathematical" in their nature and ignore the detailed architecture of the machine.

The argument will be made that the use of higher level languages will severely hamper the efficiency of application program systems. The answer to this criticism is that for most preliminary calculations, it does not matter, while for "number crunching" calculations, carefully written in the higher level language, it is not true. Tremendous effort by compiler programmers has gone into solving the problem by writing compiler "optimizers". These optimizers automate the work of making efficient machine code from the higher language.

The problems arising from different machine architecture

The use of higher level languages to prepare program systems is not a panacea. It surely mitigates the problem of transportability of the codes but it leaves unanswered several key problems. These problems have to do with the structure of the registers of the machines. As the development of computers has gone on, there have been many different ways of storing characters and quantities. The three most important ways for purposes of crystallographic calculations is as characters, as integers, and as reals. A comparison of these three will serve to illustrate the problem.

Many machines have an architecture in which the basic storage register is called a word. The word consists of a vector of bits. The bit is a binary zero or one. Common machines vary in word size from 16 to 60 bits per word. This means that of these registers are used for signed integers, the one machine will be limited to integers of value ± 32767 while the other might be capable of handling ± 10^{18}. Thus one may see that the FORTRAN statement I=J+L is very limited on the one machine when compared to the other. Account must be taken of this fact when the coding is being done for a transportable program system. A similar kind of restriction but over a wider range of values applies to the REAL numbers. In this case, the danger is that precision or accuracy will be lost during large calculations on a "small word" machine. This can sometimes be helped by using double precision where two or more words of the machine are treated as one longer word. This is usually costly in time and space but must be planned for if transportability of systems is an objective.

Finally the problem of "strings" of characters present the most difficult challenge. Often a word on a given machine is broken into bit strings called bytes. The number of bits per byte will determine the number of bytes per word. Programs which hope to be general are very difficult to write so that the wretched confusion which exists concerning these "bytes", "character codes", or "internal codes" can be dealt with. The simplest method is to store one character (byte) per word. However, the waste of space on a machine which will hold 10 characters (bytes) per word seems unconscionable in the light of all the effort to make efficient use of immediate access memory.

The general solution for transportable program systems is to move and store all variables as REAL in an appropriate precision for the machine at hand; then use a suitable packing subroutine for character string manipulation which is specific for that machine. This puts another constraint on the systems programmer. A possible solution to this difficult problem of transportability will be shown in the presentation on the RATMAC preprocessor to be presented later in the School.

Machine independent codes

The best application program systems will be coded in as machine independent a way as possible. The algorithms should be written so as to be independent of the bits per word, bits per byte and bytes per word of machines.

Input-output

The other area of higher level languages which is ill-defined is "input-output". For this reason transportable program systems will use very "primative" subroutines to isolate the actual "READ" and "WRITE" statements. The basis of this restriction is again to be found in the architecture of the machines. The customary number of words which constitute a "record" will vary from machine to machine. Therefore if input-output is to be efficient and is to make efficient use of mass storage devices, the lenght of "records" in a program must be carefully controlled.

To illustrate the problem relative to FORTRAN, consider the following statement: WRITE (10) (A(J), J=1,MAX). The local operating system and the implementation of FORTRAN which is used on it will usually have a value for maximum "physical" record size. Often this number will be around 256 or 512 words depending upon the machine. If the application program system writes records with MAX set equal to 513 on the 512 words/record machine, two actual write's will occur for every programmed WRITE. These double WRITE's will degrade the input-output performance of the application program system since all READ's and WRITE's of file "10" will be double read's and write's in actuality. To avoid this problem the application systems programmer will treat MAX as a global control variable which is set for each different machine and operating system.

A general method for dealing with input-output

Since input-output is so machine specific, it is useful in an application program system to carry out all actual read's and write's in subroutines of the system itself. By this means, the actual interaction with any given machine is confined to one place in the whole system and the burden of finding and changing input-output statements is reduced to the minimum. This minimum procedure can then be carried out in a high level language or machine language as is expedient.

Overview

The definition of "system" given at the beginning was "anything formed of parts placed together to make a regular and connected whole working as one machine". In what followed the "parts" singled out for attention in a program consisted of subroutines usually written in a "high level" computer language. These subroutines would fall into two major classes. First, would be the service subroutines. These deal with the problems of immediate access storage management; input-output; bit, byte and word moving; and straight forward counting or reckoning with real numbers. Second would be all the subroutines, grouped by algorithms treated. This second group would be dependent on the first group.

By applying this simple structure and by taking care to isolate the ill-defined features of common computing machines, the working scientist may hope to build an application subsystem of any operating system for any available digital computer.

Finally, and most important, the whole system must be thoroughly and rigorously documented.

THE NRC SYSTEM OF CRYSTALLOGRAPHIC COMPUTER PROGRAMS

F.R. Ahmed*

Summary

The NRC system is a generalized and self-consistent set of programs for the determination and refinement of crystal structures, excluding the proteins. It was developed by crystallographers at the National Research Council of Canada, and has been in constant use for over a decade. Descriptive write-ups and program listings in FORTRAN IV are available at no charge. This system is operable on a medium-size computer with 32K words of core memory, floating point hardware, a disk or three magnetic tape units, a card reader, a card punch, and a printer. Though initially designed for an IBM 360 model 50 with the Basic Operating System, it has been used with Operating System 360, and with the IBM Time Sharing System. It should be applicable, with very minor changes, to other computer makes and models.

1. Introduction

The main part of the NRC system was developed by Ahmed, Hall, Pippy and Huber (1, 2), and was described briefly by Ahmed (3). All the programs are written in FORTRAN IV (4), and all except one are tailored to fit into a 32K word (32 bits each) core memory. The computing system is assumed to have a disk or three magnetic tape units, a card reader, a card punch, and a printer. Each program is a separate entity with its own subprograms, and is carefully documented with details of the mathematical procedure, input/output formats, and a list of the arrays whose sizes may be changed to reduce or expand the memory requirements. Well-planned data sheets are also available for the main programs. For each compound under study, the system generates a standardized data file that is utilized and sometimes updated by the programs. This file, therefore, serves as

*Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada K1A 0R6

the common link between the programs of the system, and its format has not been changed since the first program was written. This facilitates the execution of several programs in succession, in any desired order, by means of the job control cards. The programs offer an adequate degree of flexibility, and include various useful options.

In addition to program listings and write-ups, card images of the FORTRAN statements on magnetic tape are also available, provided the user supplies his own magnetic tape.

2. The standard data file

The use of a standard data file in the NRC system is intended to facilitate the transfer of data from one program to another, and to avoid as much as possible wasteful recalculation of the same quantities that are needed at various stages. For each crystal structure, therefore, a standard file is generated at the data reduction stage and made available to the other programs. When revised atomic parameters are calculated, the file is automatically updated.

The format of the file is basically a simplified version of that devised by Cruickshank et al. (5) for the KDF-9 system. It starts with an identification record, followed by a directory of the lists in the file, then the data lists in ascending order. These lists are: (1) unit cell; (2) symmetry; (3) form factors; (4) atomic parameters; (5) cos and sin $(2\pi x)$; (6) planes. The unit-cell list holds the cell parameters and the tensors for calculating $\sin^2\Theta$, bond lengths, and orthogonal coordinates in Å. The symmetry list contains the equivalent positions of the space group and its different symmetries. The tables of $f_i(\sin\Theta/\lambda)$ are stored in the form-factor list. Atomic positional and thermal parameters and their e.s.d.'s are stored in list 4; there may be one or two of these lists with different serial numbers. The planes list has one record for each permissible reflexion, containing h, k, l, obs/unobs status, multiplicity, parity, order indicator (to specify which indices have changed), $\sin^2\Theta$, $\sqrt{w}$, F_o, F_c (and A_o, B_o, A_c, B_c if needed), total intensity count, and the interpolated f values.

3. Available programs

The programs in the NRC system which are available for distribution are listed below, with a brief description of their purpose, main features, and options incorporated in them.

3.1. NRC-1: Goniostat settings

It generates the indices (hkl) and calculates the goniostat settings (χ, ϕ, $2\Theta_1$, $2\Theta_2$) of the possible non-equivalent reflexions within a selected sphere (or part of a sphere) in reciprocal space, for use with the G.E. manual, or the card-operated Picker, diffractometer. Either one reciprocal axis must be at $\chi = 90°$, or two reciprocal axes must be at $\chi = 0°$. The scan range can be increased in stages as 2Θ increases.

3.2. NRC-2A: Picker diffractometer data processing

It reads the data cards produced by the Picker diffractometer (rejecting the master cards), calculates the net integrated intensities, scales them, and tests whether they are observed above threshold. Duplicate measurements of a given reflexion are averaged if the difference is small, otherwise the higher count is taken as the true count. The data are then sorted in any specified order of $|h|$, $|k|$, $|l|$, and the results output on the printer, on magnetic tape or disk, or on cards. The program also checks that there are no missing permissible reflexions within the reciprocal space specified, and notifies the user of any such omissions.

3.3. NRC-2: Data reduction and file generation

This program generates the standard data file needed by the other programs. During this process, it calculates $\sin^2\Theta$, interpolates on the scattering factor curves, derives $|F_o|$ from the measured intensities, assigns weights to F_o, and applies a sharpening function to F_o^2. It can also apply absorption corrections as a function of ϕ or 2Θ. The input data may be the intensities (either on cards or from the output file of NRC-2A), or the structure amplitudes (on cards). The Lorentz and polarization correction may be calculated for goniostat or Weissenberg geometry.

3.4. NRC-3: Absorption correction

It calculates and applies the absorption corrections to the intensities and structure amplitudes, using the m-point Gaussian integration formula, as described by Busing and Levy (6). It is applicable only to the 3-circle goniostat geometry, and to crystals with planar faces. Intensity data are accepted from cards or from the standard data file. The number of points for the Gaussian integration may be different in the three spatial directions. The crystal is described to the program by the coordinates of its corners relative to any suitable Cartesian system with the origin within the crystal. Each face and edge is defined by stating on the input cards the serial numbers of 3 or 2 corners as appropriate. This eliminates from the program any search for faces or edges. Preliminary short tests may be executed prior to the full run, in order to ensure that the crystal orientation and measurements are properly described on the input cards, and to test the accuracy of integration as the number of points m_1, m_2, m_3 are varied. These tests are usually conducted on a few of the axial reflexions at $\chi = 90°$, measured at different values of ϕ.

3.5. NRC-4: Symbolic addition procedure

This is a very powerful program for estimating the phases for centrosymmetric structures using the symbolic addition procedure of Karle and Karle (7). It was written by Hall and partially revised by Ahmed (8). The program is applicable only to triclinic, monoclinic and orthorhombic systems. It is divided into four parts which perform the following steps.

I. To estimate the temperature factor B and scale K from the linear Wilson plot of

$$\log_e\left[\ \langle \sum_{\underset{\sim}{h}} m_{\underset{\sim}{h}} \varepsilon_{\underset{\sim}{h}} \sum^{N} f_{\underset{\sim}{h}}^{\,2} \rangle_s \Big/ \langle \sum_{\underset{\sim}{h}} m_{\underset{\sim}{h}} |F_{\underset{\sim}{h}}|^2 \rangle \right]_s \quad \text{versus} \quad \langle s^2 \rangle$$

where $s = \sin\Theta/\lambda$, m is the multiplicity, ε is an integer ≥ 1 which is dependent on the space group symmetry and represents the number of identical contributions to $F_{\underset{\sim}{h}}$ by the symmetrically equivalent atoms, N is the number of atoms in the full unit cell, and Σ is over all the permissible reflexions in the asymmetric unit including those accidentally unobserved. The parameters of the straight line through the points of the plot are derived for best fit by the regression line procedure.

II. To calculate the normalized structure amplitudes $|E_h|$ and their statistical distribution, then sort those with $|E| \geq$ E-minimum in descending order of $|E|$, where

$$|E_{\underset{\sim}{h}}| = [|F_{\underset{\sim}{h}}|^2/(\varepsilon_{\underset{\sim}{h}} \overset{N}{\Sigma} f_{\underset{\sim}{h}}{}^2)]^{\frac{1}{2}}$$

and E-minimum is a specified lower limit, usually 1.2-1.5. The number of reflexions above this limit should be about 10 x the number of non-hydrogen atoms in the asymmetric unit cell.

III. To generate the sigma-2 triples. For each reflexion h with $E_h \geq$ E-minimum, a list is assembled of the pairs of reflexions k and h-k, provided that $|E_{\underset{\sim}{k}}|$ and $|E_{\underset{\sim}{h}-\underset{\sim}{k}}|$ are both higher than a specified optimum limit, usually 1.8-2.0. The latter limit is intended to reduce the amount of computation, and is chosen to provide about 200-400 reflexions above that limit. During this search, the program derives the symmetrically equivalent reflexions, their phase relationships to the input set, and includes them in the search. The output is produced on the printer and on magnetic tape or disk. It may contain either a non-redundant set of triples for computer operation, or a redundant set that would be useful for manual application of the symbolic addition. The latter option, however, is recommended only for a small data set. This search for the triples can be very time-consuming, and considerable care has been taken in the program logic to speed up this operation.

IV. To select the origin-defining reflexions, assign the symbols as needed, and apply the sigma-2 relationship

$$s(E_{\underset{\sim}{h}}) \approx s\underset{\underset{\sim}{k}}{\Sigma}(E_{\underset{\sim}{k}} E_{\underset{\sim}{h}-\underset{\sim}{k}})$$

in order to determine the signs (s) of the reflexions. This is carried out in two main stages: first, for the 200-400 reflexions with $|E| \geq$ E-optimum, to determine the values of the assumed symbols; second, for the weaker reflexions. At the end, it sorts the signed reflexions and outputs them in the standard format for the Fourier program. The origin-defining reflexions may be selected by the program or by the user. In the former case, the three origin-defining reflexions are chosen from the top 10-15 $|E|$'s, to be linearly independent, and to have the largest number of sigma-2 triples. If the starting phases are input from cards, the starting set can include any number of reflexions. The symbolic phases, on the other hand, have to be assigned by the program as needed, to a maximum of four symbols. During this operation, all the multiple indications which might lead to the true or relative values of the assumed symbols and symbol products are accumulated, and at the end are analysed for consistency. The probabilities of the phase indications

$$P_+(E_{\underset{\sim}{h}}) = \tfrac{1}{2} + \tfrac{1}{2}\tanh(\sigma_3\sigma_2^{-3/2}\Sigma_{\underset{\sim}{k}} E_{\underset{\sim}{k}} E_{\underset{\sim}{h}-\underset{\sim}{k}}),$$

where $\sigma_n = \sum_{i=1}^{N} Z_i^{\,n}$ and Z is the atomic number, are calculated throughout this process, and examined before acceptance of the phase indications. The program allows for these acceptance limits on the probability to be set high at the early stages of phasing and to be lowered gradually in steps. The printed output from this program provides full information about the sequence of phase determination, and lists the multiple indications, so that the user may make a more critical evaluation of the true values of the assumed symbols. Both the numeric and symbolic phases are available on the output tape, so that the user may change the values of the symbols if he so chooses.

3.6. NRC-5: Direct phase determination

This program determines the phases for non-centrosymmetric structures by repetitive application of the tangent or weighted tangent formula

$$\tan\phi_{\underset{\sim}{h}} \approx \frac{\Sigma_{\underset{\sim}{k}} w_{\underset{\sim}{k}} w_{\underset{\sim}{h}-\underset{\sim}{k}} |E_{\underset{\sim}{k}} E_{\underset{\sim}{h}-\underset{\sim}{k}}| \sin(\phi_{\underset{\sim}{k}}+\phi_{\underset{\sim}{h}-\underset{\sim}{k}})}{\Sigma_{\underset{\sim}{k}} w_{\underset{\sim}{k}} w_{\underset{\sim}{h}-\underset{\sim}{k}} |E_{\underset{\sim}{k}} E_{\underset{\sim}{h}-\underset{\sim}{k}}| \cos(\phi_{\underset{\sim}{k}}+\phi_{\underset{\sim}{h}-\underset{\sim}{k}})} = \frac{S_{\underset{\sim}{h}}}{C_{\underset{\sim}{h}}}$$

where $w_{\underset{\sim}{h}} = \tanh(\alpha_{\underset{\sim}{h}}/2)$, $\alpha_{\underset{\sim}{h}} = 2\sigma_3\sigma_2^{-3/2}|E_{\underset{\sim}{h}}|(S_{\underset{\sim}{h}}^{\,2}+C_{\underset{\sim}{h}}^{\,2})^{\frac{1}{2}}$.

It may also be applied to centrosymmetric structures as a special case. The program consists of five parts, to be executed consecutively. Parts I, II, and III perform the first three steps described in NRC-4, except that parts I and II include an option to estimate and employ anisotropic thermal parameters in the calculation of $|E_{\underset{\sim}{h}}|$ [Maslen(9)], and part III utilizes a more general expression for the phase relationships among the equivalent reflexions, of the form

$$\phi(\bar{h}kl,\ h\bar{k}l,\ \text{or}\ hk\bar{l}) = (0\ \text{or}\ \pi) \pm \phi(hkl)$$

depending on the space group and the parity of the indices. Part IV refines and extends the structure factor phases, in discrete cycles for decreasing |E|-threshold and with a few iterations in each cycle. Each of these iterations involves one full sweep through the relevant sigma-2 triples. Part V is to be used only when a molecular fragment is known and it is necessary to select those reflexions with |E| above |E|-minimum and $|F_c|$ greater than a specified fraction of $|F_o|$. The output cards from this run are then used for re-entry into part IV. These programs were written by C.P. Huber and F. Brisse, based on an earlier version witten in Australia by S.R. Hall, and are applicable only to the triclinic, monoclinic and orthorhombic systems.

Part IV of this set has been modified recently by Ahmed, in order to make it more automatic and to add some desirable features. In this version, the initial set of phases is to be selected by the user, and is assumed to contain the origin- and enantiomorph-defining reflexions, any reflexions with known phases from sigma-1 or the symbolic addition procedure, and a few reflexions with unknown phases. Each of the latter reflexions can be assigned 1-8 trial numeric phase values. The program reads the starting phase-set from cards, automatically permutes the trial phase values, refines and extends the phase-set, terminates unpromising permutations at any desired stage according to specified criteria, and produces a summary of the statistics for the different permutations. An option is also provided to refine or fix the phases of certain categories of reflexions, so that some or all of the starting phases and the well-determined phases can be fixed during specific cycles, while others can be allowed to refine. The parameters for accepting a newly determined phase, and the indicators for fixing or refining the phase-sets of the different categories, may be varied from one cycle to the next. These built-in features are intended to allow the user a wide range of flexibility when attempting to solve difficult structures.

During each sweep through the sigma-2 triples, the program accumulates the contributions from the known phases to $S_{\underset{\sim}{h}}$ and $C_{\underset{\sim}{h}}$, and at the end it calculates

$$|E_{\underset{\sim}{h}}|_c = K[(S_{\underset{\sim}{h}}^2 + C_{\underset{\sim}{h}}^2)^{\frac{1}{2}} / \Sigma_{\underset{\sim}{k}} w_{\underset{\sim}{k}} w_{\underset{\sim}{h}-\underset{\sim}{k}}]$$

where K is a scale factor to make $\Sigma|E_{\underset{\sim}{h}}|_c^2 = \Sigma|E_{\underset{\sim}{h}}|_o^2$ for all reflexions above the specified |E| threshold. A consistency measure $t_{\underset{\sim}{h}}$ is also calculated according to the expression

$$t_{\underset{\sim}{h}} = (S_{\underset{\sim}{h}}{}^{2} + C_{\underset{\sim}{h}}{}^{2})^{\frac{1}{2}} / \Sigma_{\underset{\sim}{k}}(w_{\underset{\sim}{k}} w_{\underset{\sim}{h}-\underset{\sim}{k}} E_{\underset{\sim}{k}} E_{\underset{\sim}{h}-\underset{\sim}{k}}) \; .$$

The newly determined phases, and those which are permitted to refine during a given cycle, are not accepted for contribution to the following iteration unless they meet four criteria, namely, that $|E_{\underset{\sim}{h}}|_c$, $\alpha_{\underset{\sim}{h}}$ and $t_{\underset{\sim}{h}}$ are all above the specified values for that cycle, and that the phase deviation (or oscillation) is not excessive. Alternatively, if the user so chooses, these conditions may be bypassed so that all calculated phases can contribute according to their assigned weights.

To assist the user in determining the overall reliability of the phase-sets, the following quantities are calculated and printed at the end of every cycle. They are listed here in descending order of their significance, and an estimate of their values for promising phase-sets is given in the right-hand column.

R-Karle = $\Sigma\lvert E_o - E_c\rvert \;/\; \Sigma\lvert E_o\rvert$	0.19-0.29
RT = $\Sigma\lvert tE_o\rvert \;/\; \Sigma\lvert E_o\rvert$	0.5-0.7
K = $[\Sigma\lvert E_o\rvert^2 / \Sigma\lvert E_c\rvert^2]^{\frac{1}{2}}$	0.9-1.1
Number of accepted phases	(8-10)N
$\Sigma\alpha_{\underset{\sim}{h}}$ and $<\alpha_{\underset{\sim}{h}}>$	?

where the sums are over all reflexions above the specified |E|-threshold, and N is the number of non-hydrogen atoms in the asymmetric unit.

The program can be instructed to compute for all the possible permutations of the input phases, or for only a few specified ones. Also, by proper setting of the acceptable R-Karle and RT limits, the program can be made to terminate the calculation of any permutations that do not seem to be progressing satisfactorily.

This revised version of part IV requires 64K words of core memory, but the user may adjust the array sizes to fit it into a 32K core by following the instructions in the write-up.

3.7. NRC-6: Editing of the standard data file

It provides the user with the following editing options:

(a) to examine, list, and/or copy the contents of the file;
(b) to replace the symmetry list (list 2);
(c) to edit the planes list (list 6), and re-order the reflexions if necessary.

3.8. NRC-7: Data manipulation for systems higher than orthorhombic

The planes list in the standard data file normally holds only the non-equivalent reflexions needed for the least-squares refinement. For systems higher than orthorhombic, this data-set must be extended and re-ordered before it is presented to the symbolic addition program (NRC-4 III) and the Fourier program (NRC-8). The programs (NRC-7) which perform this data manipulation were written by W.H. De Camp in 1973, and are included in our package.

3.9. NRC-8: Fourier maps

It evaluates 3-dimensional Patterson, electron-density, and difference maps, printing the results in sections for constant X, Y, or Z. The prominent feature in this program is its ability to print the sections with undistorted geometry, which is most convenient for direct plotting of the contour lines and for visual interpretation. This is accomplished by performing the Fourier summations at the grid points of a square net superposed on each section. The grid interval in this case may be 1/3 or $\frac{1}{4}$ Å. Alternatively, the grid lines may be chosen parallel to the unit-cell axes, which normally produces sections with distorted geometry. The user is also given these options: (a) a choice of different grid intervals in the three directions; (b) input of a sharpening curve for the Patterson amplitudes; (c) exclusion from the sums of any number of specified reflexions, or those with poorly determined phases such as when $|F_c|$ is less than $0.25|F_o|$, or those above a given $\sin^2\theta$ limit. This program utilizes the triple product form for the Fourier expressions and calculates the sums in three separate stages of single summations, making full use of the order in which the reflexions are stored in the data file, as has been described in reference (10).

3.10. NRC-9: Differential syntheses

It calculates the electron density and its first and second partial derivatives at each of the assumed atomic positions, then derives the positions of electron density maxima, as described by Booth (11). This operation is performed separately for the observed and the calculated structure amplitudes, then the observed atomic positions are corrected by the back-shift method for series-termination errors. Estimated standard deviations of the atomic positions are also calculated. Like NRC-8, this program utilizes the triple product form of the Fourier expressions, and makes full use of the order in which the reflexions are stored. The user is allowed to modify any of the structure amplitudes, or to exclude certain reflexions from the summations.

3.11. NRC-10: Structure factor least squares (SFLS)

This program calculates structure factors, refines atomic parameters and overall scale by the block-diagonal least-squares approximation, and evaluates the estimated standard deviations of the refined parameters. The mathematical procedure is based on that described by Cruickshank (12). The block sizes for the L.S. refinement are dependent on the number of parameters for each atom, as follows:

(1) 4x4 for x, y, z and B-isotropic;
(2) 5x5 for x, y, z, B, and occupancy n;
(3) (3x3 and 6x6) or 9x9 for x, y, z, and B_{ij} anisotropic
(4) 10x10 for x, y, z, B_{ij} and n
(5) 2x2 for the overall scale and overall B.

If any of the parameters has a fixed value, as for example $y = 0.5$ or $B_{12} = 0$, the block size is reduced by the program and the fixed parameters are excluded. However, the program does not take into account the dependence between related parameters such as when $y=x$. Any or all of the atoms may be included in the SF calculations but excluded from the LS refinement, as might be the case for H atoms. The imaginary component, $\Delta f''$, of the anomalous dispersion may be included for any of the atoms. Six weighting schemes with adjustable parameters are available. All shifts in the thermal parameters are corrected for their estimated interaction with the scale factor, as proposed by Schomaker (13). Calculated shifts in all parameters may be multiplied by a fudge factor to speed up conversion and to reduce oscillation.

The atomic parameters may be read from cards or from the standard data file. Poorly measured reflexions and those affected by extinction or multiple diffraction may be included in the SF calculation and excluded from the LS sums. Their structure amplitudes and obs/unobs indicators may also be modified during input. Unobserved reflexions may be excluded from all calculations or only from the LS sums, or only those with $|F_c| \geq x|F_o|$ where $x \geq 1.0$ may be included in the LS sums. At the end of the run, the program calculates the weighted and unweighted R indices and $\Sigma w(\Delta F)^2$, and produces a revised data file in the standard format containing the refined parameters and the revised structure amplitudes. The parameters are also output on cards.

3.12. NRC-12: Scan of interatomic distances and angles

Given the positional parameters for the atoms of a single molecule, the program calculates bond lengths and intermolecular distances within a specified upper limit for each, and prints them in separate lists. It then combines the distances from these two lists into a co-ordination table (which may be printed if needed for inorganic compounds), giving the nearest neighbours to each of the original atoms. Finally, the program computes the valency angles whose adjacent sides are within a specified upper limit, using the information in the co-ordination table. It also produces the standard deviations for the distances and angles.

3.13. NRC-14: Error analysis and agreement summary

This program analyzes the observed structure amplitudes for systematic and gross accidental errors, and examines the suitability of the weights used in the least-squares refinement. It calculates $\Sigma|F_o|$, $\Sigma|F_c|$, R, $\Sigma w(\Delta F)^2$, $\langle w(\Delta F)^2 \rangle$ for specified ranges of $|F_o|$, and/or for ranges of $\sin^2\theta$, and/or for levels of h, k, or l. It also lists reflexions with individual poor agreement, i.e., the observed reflexions with $|\Delta F|/|F_o| > 3R$ and $|\Delta F| > 2\ |F_{th}|$, and the unobserved reflexions with $|F_c| > 1.5\ |F_{th}|$, where F_{th} is the corresponding threshold amplitude.

3.14. NRC-21: The d-spacings

It generates hkl and calculates d(hkl) for all the non-equivalent reflexions within a selected sphere or part of a sphere in reciprocal space, excluding those prohibited by the space group symmetry. Reflexions are sorted into descending order of d(hkl) before printing, for use in the indexing of powder lines.

3.15. NRC-22: Mean plane and torsion angles

This program calculates the best mean plane through a group of atoms, examines the planarity of the group by the χ^2-test, and calculates the distances of any other atoms from the plane. It can also produce the dihedral angle between any pair of planes calculated, and has been modified recently to calculate the torsion angles, with signs according to the IUPAP convention.

3.16. NRC-23: Structure factor table

It lists the structure factors in a table of the proper proportions for photographic reduction and publication in Acta Crystallographica.

3.17. NRC-24: Projection of atoms onto a plane

The program projects the atoms of a molecule onto a given plane (or planes), for viewing the molecule from different angles or for preparation of a suitable clinographic drawing of the structure.

3.18. NRC-26: Structure factor table for deposit

It produces a structure factor table which meets the specifications for deposition of tables, as an alternative to publication in Acta Crystallographica.

4. Other programs

ORTEP by C.K. Johnson (14), MGTLS by Trueblood and Schomaker (15), and MULTAN by Germain, Main and Woolfson (16) have been obtained and implemented by C.P. Huber and M.L. Post for use with the NRC system.

5. References

1. Ahmed, F.R., Hall, S.R., Pippy, M.E. and Huber, C. P. (1966). World List of Cryst. Computer Programs, 2nd ed., p. 52. Utrecht: Oosthoek
2. Ahmed, F.R., Hall, S.R., Pippy, M.E. and Huber, C. P. (1973). J. Appl. Cryst. 6, 309. Accession numbers 133-147
3. Ahmed, F.R. (1970). Crystallographic Computing, p. 309. Copenhagen: Munksgaard
4. FORTRAN IV (1965). IBM Manual, File no. S360-25, Form C28-6504-2
5. Cruickshank, D.W.J., Sime, J.G., Smith, J.G.F., Truter, M.R., Wells, M., Rollett, J.S. and Freeman, H.C. (1964). KDF9 Programs for Crystal Structure Analysis. Private communication
6. Busing, W.R. and Levy, H.A. (1957). Acta Cryst. 10, 180
7. Karle, I.L. and Karle, J. (1963). Acta Cryst. 16, 969
8. Ahmed, F.R. (1970). Cryst. Computing, p. 55
9. Maslen, E.N. (1968). Acta Cryst. 22, 945
10. Ahmed, F.R. (1970). Cryst. Computing, p. 355
11. Booth, A.D. (1946). Trans. Faraday Soc. 42, 444
12. Cruickshank, D.W.J. (1970). Cryst. Computing, p. 187
13. Schomaker, V. unpublished
14. Johnson, C.K. (1965). ORTEP. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee
15. Schomaker, V. and Trueblood, K.N. (1968). Acta Cryst. B24, 63
16. Germain, G., Main, P. and Woolfson, M.M. (1971). Acta Cryst. A27, 368

XRAY76

James M. Stewart*

Summary

XRAY76 is a system of programs for large computers. It has been written in PIDGIN FORTRAN, a neutral subset of FORTRAN IV. The overall structure of the system, the constituting programs and the steps involved in the implementation are described.

Introduction

XRAY76 is a collection of programs based on XRAY63 and XRAY72. These programs have been collected from many sources and have been integrated into a single large overlay structure. The whole system of programs has been compiled and tested on a variety of "large" and "maxi" computers, among them the UNIVAC 1100 series, the CDC 6000, 7000 and CYBER series, the IBM 360 and 370 series, DEC PDP-10, ICL 1906A, BURROUGHS 6700 and Honeywell 68/80. The programs are documented in Technical Report TR-446 (March 1976) of the Computer Science Center of the University of Marylanc Figure I shows the table of contents of that report. From this table the variety of crystallographic computations performed through XRAY76 may be seen. The programs have been written and/or edited into the system throug the efforts of many authors. In addition, many others have helped in the testing and implementation of the resulting programs on the various machines. The credits to these authors are given in appendix 1 of TR-446

Fig. 1

GENERL	GENERAL DESCRIPTION OF THE "XRAY" SYSTEM, ITS TERMS AND USE
ABSCOR	ABSORPTION CORRECTION (TOMPA METHOD)
ABSORB	DIFFRACTOMETER ABSORPTION CORRECTION (GAUSSIAN QUADRATURE)
APHASE	SEARCH OF SIGMA-TWO RELATIONSHIPS FOR A SET OF POSSIBLE PHASES
BONDAT	GENERATION OF THE COORDINATES OF BONDED ATOMS
BONDLA	DETERMINATION OF CONTACT AND BOND DISTANCES AND ANGLES WITH ESTIMATED ERRORS
BUILDM	FIND A CONNECTED SET OF ATOMIC COORDINATES FROM INPUT SEARCH OF FOURIER PEAKS
CELRDN	CELL REDUCTION PROGRAM OF MIGHELL AND SANTORO
CONTRS	DRAW FOURIER MAPS ON A LINE PLOTTER
CRYLSQ	GENERAL CRYSTALLOGRAPHIC LEAST SQUARES PROGRAM
CRYSET	GENERATE GENERAL EQUIVALENT POSITIONS FROM SPACE GROUP SYMBOL BY LARSON ALGORITHM
DATCOS	TREATMENT OF INITIAL DIFFRACTOMETER INTENSITY DATA
DATRDN	PREPARATION OF BINARY DATA FILE AND PRELIMINARY TREATMENT OF DATA AND SYMMETRY
DELSIG	DISPAY DELTA F AS A FUNCTION OF FO OR SIN(THETA)

---

* University of Maryland, U.S.A.

Fig. 1 (continued)

DRAW	FOURIER CONTOURING PROGRAM
DUMCOP	DUMP OR COPY THE BINARY DATA FILE
EC	CALCULATE NORMALIZED STRUCTURE FACTORS,E, AND FORM TRANSLATION FUNCTION COEFFICIENTS
EGEN	ESTIMATION OF NORMALIZED STRUCTURE FACTORS, THE OVERALL TEMPERATURE AND SCALE FACTORS
FC	STRUCTURE FACTOR CALCULATION
FOUREF	AUTOMATIC BLOCK FOURIER REFINEMENT OF ATOMIC PARAMETERS
FOURR	FOURIER TRANSFORMATIONS TO GIVE PATTERSON, VECTOR, ELECTRON DENSITY, DIFFERENCE, OR E MAPS
LATCON	LEAST SQUARES REFINEMENT OF LATTICE PARAMETERS
LISTFC	LISTS STRUCTURE FACTORS FOR PUBLICATION
LOADAT	LOAD ATOMIC PARAMETERS INTO THE BINARY DATA FILE
LSQPL	CALCULATION OF EQUATIONS OF 3-DIMENSIONAL PLANES AND LINES
MODIFY	BINARY DATA FILE EDITING AND GENERATION OF PSEUDO DATA
MULDMP	DUMP "XRAY" BINARY DATA FILE TO GIVE INPUT FOR NON-SYSTEM "MULTAN"
MULTAN	XRAY ADAPTATION OF WOOLFSON-MAIN-GERMAIN MULTIPLE SOLUTION PROGRAM
NORMSF	PRELIMINARY DATA SCALING, CALCULATION OF NORMALIZED STRUCTURE FACTORS (E), AND ESTIMATION OF TEMPERATURE FACTORS
ORTEP	XRAY ADAPTATION OF OAK RIDGE PLOTTING PROGRAM OF C. JOHNSON
PARAM	LEAST SQUARES REFINEMENT OF CELL PARAMETERS FROM TWO THETA DAT
PEKPIK	SEARCH FOURIER MAPS FOR PEAKS
PHASE	SEARCH OF SIGMA-TWO RELATIONSHIPS FOR A SET OF POSSIBLE PHASES(CENTRIC CASE)
PLOT	PRODUCE PLOTS FOR "ORTEP", "PROJCT","CONTRS", CONNECT TO SHOP PLOTTER OR SIMULATE PLOTTER ON LINE PRINTER
POWGEN	POWDER PATTERNS FROM SINGLE CRYSTAL DATA
PROJCT	PRODUCE A PLOT OF THE PROJECTION OF A MOLECULE ON A PLANE-OUTPUT TO LINE PRINTER AND LINE PLOTTER
RIGBOD	GENERATION OF RIGID GROUPS FOR USE IN LEAST SQUARES REFINEMENT
RLIST	LISTS R VALUES FOR VARIOUS ZONES AND OTHER REFLECTION CLASSES
SINGEN	FORMS TRIPLES FOR PHASE DETERMINATION (CENTRIC OR ACENTRIC)
STEPRF	STEP REFINEMENT OF ATOMIC PARAMETERS AT LOW RESOLUTION
TANGEN	USE OF TANGENT FORMULA TO CALCULATE PHASES
WRITEU	GENERATION OF WRITE-UP
WTANAL	ANALYSIS OF LEAST SQUARES WEIGHTING SCHEMES
WTLSSQ	CALCULATION OF LEAST SQUARES WEIGHTING SCHEMES
APENDX-1	CONTRIBUTORS TO THE SYSTEM
APENDX-2	EXAMPLE DATA DECKS
APENDX-3	DISTRIBUTION OF THE SYSTEM
APENDX-4	PIDGIN FØRTRAN
APENDX-5	THE STRUCTURE OF THE SYSTEM
APENDX-6	THE FORMAT OF THE BINARY DATA FILE
APENDX-7	INFORMATION ON REPORT OF SUSPECTED ERRORS

Structure of XRAY76

The programs which constitute XRAY76 are written in a neutral subset of FORTRAN IV, which is referred to as "PIDGIN" FORTRAN. The details of "PIDGIN" FORTRAN are given in appendix 2 of TR-446. All of the subroutines which carry out crystallographic calculations are programmed to be compatible with and to depend upon the "control" part of the system. This control part centralizes input-output, word-packing, "overlaying" and the management of large data arrays in immediate access storage. The control part of the XRAY system is called the NUCLEUS. Practically all of the machine specific programming is concentrated in the NUCLEUS.

Even though the total size of the XRAY76 system is quite large, the structure of the system is very straight forward. Figure II shows, by means of an immediate access storage, "map" or "link edit", the overall structure of the system. If one understands how to implement the NUCLEUS, one overlay and the large data array, all the rest would be "easy". Appendix 5 of TR-446 gives details of the overlay structure of XRAY76.

Core Memory Daigram for XRAY76

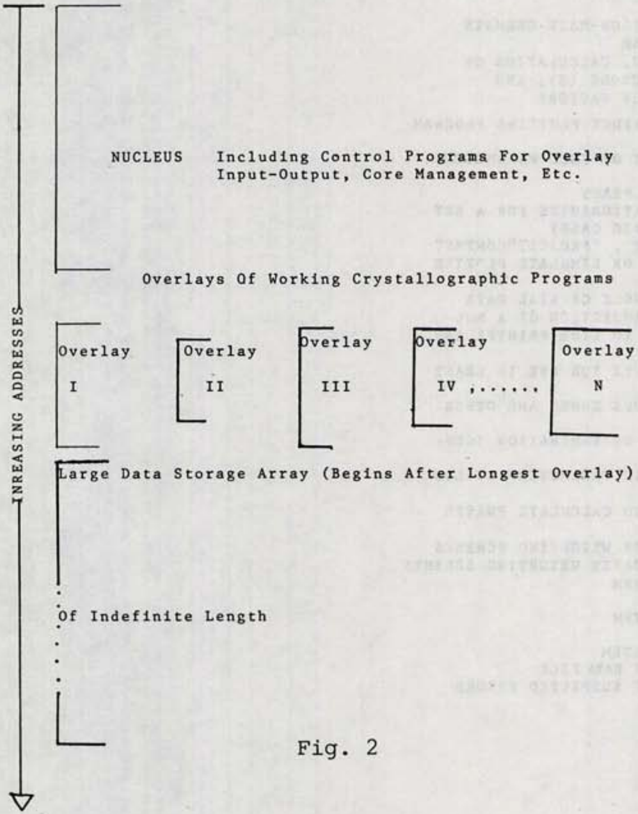


Fig. 2

Steps involved in implementing XRAY76

The implementation may be thought of in terms of seven major steps:

I. Moving symbolic PIDGIN FORTRAN, write up and test decks from the supplied tape file to local files in mass storage.
II. Editing the symbolic decks to observe local conventions, especially with respect to I/O and core management. The table of contents in the MAIN program shows which subroutines contain shop specific code.
III. Compiling the corrected FORTRAN decks and saving the "relocatable" programs which result in another local file.
IV. Adapting the overlay structure to local core availability and management practice.
V. Setting up local run requirements including establishment of data files.
VI. Running test decks and checking results against Acta Cryst. (1).
VII. Documenting the procedure for others at the installation.

It is not possible to give details of all these steps for all machines. Every machine has its own peculiarities and, moreover, each computer center is free to embellish a given machines operating system in any way they see fit.

In the case of XRAY76, the process has been carried out for a number of different machines and the results documented to help in the implementation at other laboratories using these machines. This documentation is supplied with the magnetic tape containing the symbolic decks.

References

(1) Acta Cryst. (1972), A28, 365-393

SHELX

George M. Sheldrick*

Summary

SHELX is an integrated program for crystallographic data reduction, structure solution and refinement. It is written in a simple subset of FORTRAN which is compatible with most large computers. The program is relatively easy to instal and use, and makes modest demands on computer time, store, file-space and the amount of output created. The program is valid for all space groups. All instructions are input in machine-independent free format. SHELX consists of about 5000 FORTRAN statements and includes absorption corrections, reflexion sort/merge with determination of inter-batch scale factors, multisolution direct methods, geometric placing of hydrogen atoms, full-matrix and blocked least-squares refinements with various rigid group, riding atom, and bond length constraints, analysis of variance, various Fourier syntheses, peak-search, unique molecule assembly and lineprinter 'plots'. A minicomputer system incorporating SHELX is available from SYNTEX.

1. Introduction to the program system

SHELX is entirely written in a very simple subset of FORTRAN, which is acceptable to most large computers and also compiles into efficient code. With the exception that 'logical IF' comparisons are allowed, the language appears to be a subset of 'PIGIN FORTRAN' defined by J.M. Stewart. Many problems affecting the transfer of the program between computers have been avoided by NEVER using the statements: EQUIVALENCE, BACKSPACE, ENTRY, BLOCKDATA and EXTERNAL. Both formatted and unformatted READ and WRITE are employed, but REWIND is only allowed for binary files (which are blocked internally by the program), and the only items which appear in the brackets after READ or WRITE are unit numbers and FORMAT statement numbers. Subroutines are used sparingly (ASSIGN is faster at run-time) and the only arguments allowed are simple variables.

*Anorganisch-Chemisches Institut der Universität, Tammannstraße 4, D-3400 Göttingen, Germany

Logical, double precision or complex variables are NEVER used in SHELX; all arrays are one-dimensional, with subscripts of the type 'I' or 'I±constant'. Extensive use is made of dynamic storage of several types of information in the same array, and to make the source program shorter variable names are restricted to one or (if necessary !) two letters. 'Blank COMMON' is set up in the main program as follows ('labelled COMMON' is not used):

```
COMMON LM,LR,LH,LI,LP,LA,LB,LC,LF,LU,LV,HA,HD,A(14026)
LM=14026
:
:
```

A similar 'blank COMMON' statement appears at the start of all major subroutines, the only difference being that 'A(14026)' is replaced by several arrays, with a sum of dimensions less than or equal to 14026. The last array in each blank COMMON statement is used to store the list(s) which may expand substantially for larger problems, e.g. the matrix for least-squares refinement. To create a version of the program for large problems (e.g. to refine more parameters simultaneously by full-matrix least-squares) only the main program (which is very short) needs to be recompiled, with LM set equal to the increased dimension of A. The value of LM tells each subroutine how much core store is available. The sort/merge and Fourier routines automatically make more use of scratch disc if the core store is limited; for direct methods the number of reflexions or phase relations is automatically reduced to fit the available store.

All installation-dependent parameters are set in the main program. LR ... LF define the unit numbers of all files and input/output devices. LU is a flag to take advantage of 'COMPRESS INTEGER' mode on certain computers in assigning blank COMMON allocations; it may only be used if the integer length is at least 24 bits and is not more than half the real length (in practice it is useful only for ICL, Telefunken and ODRA computers). LV sets the buffer size used by the program for binary files; if it is altered it is also necessary to change some dimensions elsewhere in the program (in practice it is never necessary to change it). HA and HD are the number of lineprinter characters per inch across and down the paper; they are needed to scale the lineprinter plots.

The overlay structure of SHELX is particularly simple: each segment is called once, in sequence, from the main program. All information is transmitted via blank COMMON or binary files. On modern virtual-memory machines the program is small enough to run without overlay. On very small computers without overlay it can simply be split up into up to six separate programs, each with its own main program, all information being transferred via the binary files.

2. Installation problems

After modifying the main program as described above, SHELX-76 has been found to compile and run correctly first time on the following computers: IBM 360, 370; DEC 10, 20; CDC 6600, 7600, cyber-72 to -76; UNIVAC 1106, 1108; TELEFUNKEN TR 440; ICL 1901 to 1906, system-4; SIEMENS 4004; AMDAHL 470; HONEYWELL-BULL 60 and CII SIRIS-7. With obvious minor action to avoid compiler bugs or known idiosyncrasies it also runs correctly on the CDC 3300, 6400; RIAD 20 and Burroughs 6700, 6718. The only known serious problems concern the Elliot 4130 (now probably extinct) and the ODRA 1305 (where SHELX-72 ran correctly but one subroutine in SHELX-76 fails, probably because there is no disc). The simple programming style usually enables the highest level of compiler optimisation to be used, though there are problems with opt=2 for the CDC 7600 and certain releases of the ICL XFEW compiler. The minimum core store required is in the range 100-140 k-bytes depending on the computer and compiler. SHELX-76 is totally incompatible with computers which have an integer word-length of less than 24 bits.

In practice most of the installation problems are concerned with reading a magnetic tape (or card deck) on one make of computer which was written on another.

3. Crystal data and instructions

In SHELX an attempt has been made to make the necessary input so simple that an experienced user rarely needs to refer to the instruction manual. However there is a price to pay in terms of the options available to the user and constraints on the programming. SHELX employs three scratch binary files (units LA, LB and LC defined in the main program) plus the following files or input/output devices:

LR: Card reader or input file for unit-cell and symmetry information, atom coordinates and all instructions.

LH: Input file or magnetic tape for raw intensity data (only required in a data reduction job; it is also possible to input the raw data via LR instead).

LF: Permanent or scratch file for h, k, l, F, σ(F), E etc. after data reduction. Can be modified by subsequent jobs, e.g. to renormalise E-values or suppress individual reflexions. To save file-space (and time !) each reflexion is packed into 6 real words and buffered in binary records; scattering factors are not stored in this file.

LI: Lineprinter output.

LP: Card punch or file output of refined atomic coordinates and (optionally) 'condensed data'.

'Condensed data' consists of h, k, l, F and $\sigma(F)$ (after data reduction) packed efficiently in coded form (about 9 reflexions per card) followed at the end by a checksum. It may be reinput via LR or LH, in which case no raw intensity data are required and LF may be a scratch file. It is particularly useful when permanent file-space is limited or not available, and enables all instructions and data for one job to be contained in a single compact file or card deck. Thus the same deck or file may be input either to a (slow) local computer or a (faster and larger) regional computer.

It should be noted that there is no permanent file containing unit-cell and symmetry information and atom coordinates. The system was originally intended to run from cards, which are still useful, especially for users who are not computer experts; if a good interactive editing system is available, it seems unnecessary to duplicate it. However a facility is provided in SHELX for concatenating input files, e.g. to read back coordinates output by a previous job.

All instruction and atom cards are read in FORMAT(A4,76A1). If the first four characters are not a recognised instruction the card is assumed to be an atom card with these four characters as name. A FORTRAN routine then decodes the 76A1 character string into an array of numbers; there is no distinction between reals and integers. Numbers are input in free format, separated by one or more spaces (or other characters which could not themselves be part of a number). There are facilities for continuation lines and comments. Extensive use is made of sensible default values for numbers which are not given. Special action is taken to decode instructions such as:

SYMM -X, 0.5+Y, 0.5-Z and SFAC C H N CL

Except for checking for input errors and some immediate processing of intensity data, no calculations are performed until the END card is read; thus instructions may be input in almost any order, and the program decides whether the instructions make crystallographic sense before proceeding with the calculations, which are performed in a fixed order. For example if a difference electron density synthesis is requested, a structure factor calculation is automatically performed first even if the user has forgotten to ask for it. This approach means that a user error is less likely to waste computer time, at a cost of some loss in flexibility. The dynamic packing of items into one-dimensional arrays imposes certain restrictions on card order however; e.g. SFAC cards must be given before atoms.

4. Examples

Details of the types of crystallographic calculations which may be performed with SHELX, and some of the theory on which they are based and the computing algorithms employed, will be presented in the afternoon classes of the summer school. It is useful first to obtain an overall picture by looking at some typical instruction files. The first example illustrates a routine data reduction and centrosymmetric direct methods job. It is assumed that data have been collected on a 4-circle diffractometer and have been converted to h, k, l, F and σ(F) in FORMAT(3I4,2F8.2) on file or tape LH, that L.p. corrections have already been applied, and that absorption corrections will not be necessary. At the end of LH there should be a blank record (or a reflexion with all five items set to zero).

```
TITL DIAMANTANE-4-OL
CELL 1.5418 16.704 16.704 7.922 90 90 90
SYMM .5-X, .5-Y, Z
SYMM -Y, .5+X, .5+Z
SYMM .5+Y, -X, .5+Z
SFAC C O H
UNIT 112 8 160
HKLF 3
MERG -2
OMIT 4
EEES 1.2
FMAP 7
GRID 23 23 -2 2 2 2
PLAN 20 3
END
```

The space group is $P4_2/n$, with unit-cell contents $C_{112}O_8H_{160}$. Default settings of 'centrosymmetric' and 'primitive lattice' have been assumed. Symmetry information is always specified be means of the coordinates of the general position as given in International Tables, omitting 'X, Y, Z'. The instructions TITL to UNIT inclusive would be the same for all calculations on this structure. Between HKLF and END, instructions could have been given in any order. The data are read from LH after 'HKLF 3'; MERG -2 will average equivalent reflexions, find E-values and put the F and σ(F) values on an approximately absolute scale. The minus sign is conventionally used to mean 'create more output', in this case to print a multicolumn list of reflexions after data reduction (note that all default settings save trees). Unit LF may be saved after the job as a permanent reflexion data file; all reflexions in it with F less than 4σ(F) will be flagged as unobserved (OMIT 4).

Alternatively LF could have been set up as a scratch file, and LIST 2 used to generate condensed data cards. EEES 1.2 controls the multisolution sigma-2 sign expansion. In this example the 2 origin and 12 (default setting) multisolution reflexions are selected automatically, and up to 4 (default) 'best' E-maps are calculated (FMAP 7) but not printed (which would require FMAP -7). The highest 20 unique peaks in each are then sorted into molecules, which are 'plotted' on the lineprinter. As it happens the program decides that 2 E-maps are sufficient, and the job required about 20 seconds CPU time on an IBM 370/165 or 5 seconds on a CDC 7600.

The second example concerns absorption corrections for Weissenberg data, followed by determination of layer scale factors using data collected from the same crystal about two different axes:

```
TITL ... UNIT as usual
ABSC 1.0
FACE 1 0 0 .037
FACE -1 0 0 .028
FACE -1 2 1 .148
FACE -1 -2 -1 .204
FACE 1 2 -1 .178
AXIS 1 0 0
HKLF 4  (program now reads unit LH up to first blank record)
AXIS 0 1 0
HKLF 4  (now reads rest of unit LH)
MERG -3
etc.
```

The absorption corrections are calculated by Gaussian integration to a precision of 1% (set on the ABSC card). The linear absorption coefficient and optimum integration grid are computed by the program. In the next example a regular hexagon (default C-C 1.395 Å) is fitted to six Fourier peaks (carbon atoms) and then refined as a rigid group:

```
.
.
L.S. 4
AFIX 66
C1  1 .068 .496 .239
C2  1 .145 .566 .215
C3  1 .240 .570 .117
C4  1 .257 .503 .049
C5  1 .169 .431 .072
C6  1 .075 .422 .167
AFIX 0
etc.
```

An effective rigid group refinement can also be achieved with bond length constraints. e.g. to refine a BF_4^- ion as a regular tetrahedron, but with the B-F distance free to refine:

```
:
L.S. 4
LIST 1 -1       (analysis of variance and F(o)/F(c) tables)
BOND .5 2       (bond lengths, angles, and non-bonded distances)
WGHT -0.001     (refine weighting parameter g; w=[σ^2(F)+gF^2]^-1)
FVAR 1.08 1.37  (overall scale and B-F bond length)
B1  3 .2566 .3649 .6487 11 .05        (sfac type, x, y, z, sof, U)
F11 4 .1552 .3188 .7129 11 .08
F12 4 .3436 .4216 .7253 11 .08
F13 4 .3677 .3134 .6035 11 .08
F14 4 .1582 .4063 .5580 11 .08
DFIX 21 0.001, 1 2, 1 3, 1 4, 1 5
DFIX 21.6330 0.001, 2 3, 2 4, 2 5, 3 4, 3 5, 4 5
etc.
```

This constrains B-F to 1.0 multiplied by 'free variable' 2 (which starts refining from 1.37) and F...F to be 1.6330 times 'free variable' 2. Note that there is no free variable 1, because the first number on the FVAR card is the overall scale factor; to fix a parameter (such as the site occupation factors in the above example) 10 is added. Free variables are also used to apply special position constraints, as in the following example from space group $P6_422$ with U_{12}(Sn) = 0.5*U_{22}(Sn) and U_{23}(Sn) = 0:

```
.
.
FVAR 0.96 0.24 0.38 0.03
SN  1 21 22 10.5 10.5 0.11 42 0.08 10 0.02 41
C2  3 31 32 10.5 10.5 .10
N   2 10 10.5 .42 10.5 .07
etc.
```

Sn and C2 are in special positions of the type x, 2x, 0.5 and N is at 0, 0.5, z. Although this example looks complicated compared with program systems which apply the constraints automatically, most users appear to have little trouble with the use of free variables in practice. A further application of free variables is to tie one group of site occupation factors to (say) free variable 3 (by setting them to 31) and another group to 1 minus free variable 3 (by setting them to -31) to refine a disordered structure. DFIX is also very useful for disordered structures. The final example illustrates the geometric placing and riding refinement for hydrogen atoms:

```
C8  1 -.2655 .3665 .2632 11 .04 .05 .04 .01 -.01 .00
AFIX 23
H81 2 0 0 0 11 21
H82 2 0 0 0 11 21
AFIX 0
etc.
```

5. Future developments

Although SHELX was written in about 1970 and has changed relatively little since 1974, it has only become epidemic in the last two years. It is now installed on about 200 computers, mainly in Europe. It appeals mainly to chemists (who appreciate the ease of installation and the simplicity of use) rather than crystallographers developing new methods (who find the restrictions irksome and the program very difficult to modify). The program is highly integrated and optimised, and although it is written in Fortran the style is closer to assembler. The lack of a normal array structure, extensive use of word packing, and brevity of the comments make minor extensions and modifications a major undertaking (even for the author): such is the price of efficiency. In fact, many of the features which make SHELX so fast and machine independent owe more to my ignorance of the more sophisticated features of FORTRAN than to deliberate planning! [I have subsequently found out what EQUIVALENCE means, but am still very hazy about EXTERNAL and ENTRY]. I prefer potential users to receive a current version of the program and documentation directly from me rather than mutilated versions from each other. This has enabled me to send corrections and other useful information to all (?) users. The feedback from users has been extremely valuable and accounts for many improvements to the program. The current version, used in almost all installations, is SHELX-76. Full compatibility is very desirable so that jobs may be transferred between local and regional computers etc. I have the impression that many SHELX users are happy to be using a well tested and debugged version, and that it is a mistake to issue new versions too frequently. However some extensions are desirable, so the next version to be released will be SHELX-80. I am particularly ashamed of distributing direct methods programs which are about a decade out of date (even if they sometimes solve structures), and improvements are necessary for the expansion of partial structures and the refinement of large structures. We have recently developed an empirical absorption correction procedure based on measurements of (equivalent) reflexions at different azimuthal angles, which should certainly be added.

At the time SHELX was written, I did not know about 16-bit minicomputers. Although it is a relatively small program, SHELX was optimised for machines with at least 24-bit integers (the 24-bit limit was chosen with ICL and Telefunken in mind) and about 120 k-bytes of core store. E. Oeser and I have recently written a system for the Data General Eclipse and Nova mini-computers which retains much of the philosophy of SHELX, but to a large extent was rewritten from scratch because SHELX is so difficult to adapt. We also took the opportunity to introduce various new facilities, some of which require an interactive system. The major changes are:

1. Empirical absorption corrections are included.

2. The direct methods have been considerably improved and brought up to date. The mini-computer program has now solved at least 20 structures which defeated MULTAN and SHELX in expert hands on large computers.

3. Full-matrix least-squares has drawbacks on a mini-computer because core store is limited and disc transfers very slow. We have therefore written a 'blocked cascade' refinement program which has large enough blocks to accommodate free variables, rigid groups etc., but small enough to fit into core. Structure factors are only recalculated for those atoms which are refining in the current cycle or the cycle immediately preceding it, so the algorithm is much faster than the blocked refinement in SHELX if there are more than two blocks. All the facilities of the SHELX refinement are included except for DFIX, which is incompatible with small blocks.

4. An iterative procedure for calculating difference Fouriers and then refining atom site occupation factors is included for the automatic extension of partial structures.

5. There is a facility for swapping out a job which is running, then restarting it later. Also a job which is running can be given an interrupt to tell it to complete the present refinement cycle, then stop refining but go on to bond length/angle calculations, Fourier maps, etc. This is useful if a refinement has obviously converged (R indices and times may be printed on the operators console). A similar facility applies to direct methods (stop doing permutations and start E-maps) and the partial structure refinement.

6. An interactive program is provided for molecular geometry calculations including least-squares planes, torsion angles (which are also printed out with e.s.d.s by the least-squares refinement), librational corrections (Schomaker-Trueblood), generation of symmetry equivalent atoms, molecule plots on lineprinter and/or tektronix VDU, and setting up of input files for ORTEP.

The system may be obtained from Syntex as an integrated hardware/software combination or as an extension to their XTL package (which has similar hardware requirements).

CHAPTER 2

Program systems on mini-computers

RÖNTGEN 75 - A SYSTEM OF PROGRAMS FOR AUTHOMATIC SEARCH FOR A STRUCTURAL MODEL

V.I. Andrianov*

Summary

RÖNTGEN-75 is a complete structure determination package succesfully used in the USSR. Automatic determination of structures containing 20 to 30 asymmetric atoms is easily achieved. A survey of the subprograms in RÖNTGEN-75 is given. The operations "Multisolution method", "Fourier synthesis", "Search for a structural model" and "Least-squares" are discussed in more detail.

1. Introduction

RÖNTGEN-75 is a system of programs which is to be used on the BESM-6 computer (1 μsec cycle-time). This system has been developed by Dr. Tarnopolsky and Mrs. Safina, Institute of Chemical Physics, Academy of Sciences of the USSR and by the author, Institute of Crystallography. It was our goal to make the process of search as much automatic as possible utilizing the experience available. In the USSR, nowadays, most of the structure determinations are carried out by means of the system of programs RÖNTGEN and RÖNTGEN-75; structures containing up to 40 asymmetric atoms, being deciphered by direct methods, are solved automatically. We call "automatic": the search for approximate structural model or its main fragments starting with the experimental data only without human intervention in the intermediate stages of calculation. The automatic search for a structural model takes more computer time than when the worker is optimizing the parameters of forthcoming calculations after the analysis of previous results. But, the extra computer time is fully compensated by the convenience of the automatic structure determination.

Determination of not very complicated structures consisting of 20 - 30 asymmetric atoms is easily accessible to the worker not familiar with details of the calculation procedure. The programs do not impose any limitations: neither on the symmetry of the crystal nor on the amount of experimental structure amplitudes. RÖNTGEN-75 is a system of FORTRAN programs for a computer with a 32 K core of 48 bits words. To optimize the employment of computer resources and to decrease the computer time-consumption some subroutines were written in the BESM-6 code. That is why RÖNTGEN-75 can not be demonstrated on other computers. It is planned to adjust RÖNTGEN-75 to widely accepted computers.

2. System organisation

About 30 various "operations" are designed to be carried out. "Operation" here means a system of programs for calculation of a definite

---

* Institute of Crystallography, Moscow, USSR

stage of structure determination. RÖNTGEN-75 performs the following "operations":

(i) Input and check of data from punchtape, punchcards, or from teletype;
(ii) Normalization of the structure amplitudes and collection of Σ_2-relationships;
(iii) Multisolution method for the phase determination;
(iv) Search for a structural model;
(v) Fourier synthesis;
(vi) Least-squares refinement;
(vii) Calculation of the geometry of the molecule;
(viii) Performance of service programs to add or exclude some data into tape/disk, their check, printing, transformation into various formats, etc.

Every of these operations, excluding "input the data", is managed by a few parameters which will be discussed further in more detail. Moreover, every operation can be carried out automatically: the parameters needed are not taken from the punchcards but have default values, set by the authors. These values were chosen on the basis of author's experience in the field of structure determination. The operation "Search for a structural model" seems to be of the utmost interest here. As a matter of fact, this operation represents continuous successive application of the two other operations, "multisolution method" and "Fourier "Fourier synthesis", without any predetermined parameter. The multisolution method determinates sets of the most suitable phases and then the program "Fourier synthesis" provides their analysis during electron density calculations; the most probable model of a structure is stored.

"Fourier synthesis" is adjusted in such a way that its cyclic use with small changes in parameters makes it possible to carry out automatic search for a structural model by heavy atom technique, and to refine the phases of structure amplitudes. For these reasons we consider here the operations "Multisolution method", "Fourier synthesis" and "Search for a structural model" in much more detail.

To give a complete description of our system of programs, we shall outline the facilities of the "Least squares" operation.

3. Multisolution method and search for a structural model

Multisolution method for the phase determination$^{1-3}$ is realized in RÖNTGEN-75. It includes the stages:

(i) Normalization of the structure amplitudes and collection of the Σ_2-relationships. The knowledge of the overall atomic temperature parameters, B, and the scale parameter, K, is needed for the calculation of normalized structure amplitudes E_{hkl}. The calculation procedure has been published elsewhere[4]. The B value needed, minimizes the function $\Sigma(\tilde{E}^2_{hkl}-1)^2$, where

$$\tilde{E}^2_{hkl} = F^2_{hkl}\exp(Bs^2)/\varepsilon_{hkl}\Sigma f^2(s)$$

This minimization is achieved by scanning the value of B within the range of $1 \le B \le 6$ with steps of $\Delta B = 0.5$. The K value is found from the condition $<|\tilde{E}^2|> = <|K^2\tilde{E}^2|> = 1$. The 10n largest values of $|E_{hkl}|$ are then selected, where n is the number of asymmetric atoms, and afterwards 50n of Σ_2-relationships are composed.

(ii) Fixing the origin of coordinates, enantiomorph and the choice of starting reflections.

The origin-fixing reflextion selection rules, number of this reflections and possible phase values are defined by the crystal symmetry[5,6]. It is desirable to choose starting reflections in such a way that the procedure for the determination of the other phases would proceed most successfully provided the phases of starting reflections are known. The procedure of this choice has been described previously[7]. In the RÖNTGEN-75 the 1024 possible sets of starting reflection phases are scanned, and hence not less than 5 starting reflections are usually chosen (every phase may have either four, or two values).

(iii) Multisolution method for the phase determination.

Various phase values, varied by $\pi/2$, are ascribed to starting reflections chosen, and the phases of the other reflections are calculated for every set of starting reflection phases. The process of determination of the other phases is dependent upon the presence or the absence of the centre of symmetry in the structure[8].

For centrosymmetric structures the signs are determined by the relation $S_h \simeq S_{h'} \cdot S_{h-h'}$, where $S_h = \text{sign } E_h$.

In our system:

(a) if the new sign is defined by the single relationship, then its reliability may be calculated as

$$P = \tfrac{1}{2} - \tfrac{1}{2} \tanh(\sigma_3 \sigma_2^{-3/2} |E_h E_{h'} E_{h-h'}|)$$

and the sign is accepted to be correct if $P \ge 0.97$.

(b) if the new sign is defined by several relationships (m^+ times as positive, and m^- times as negative), then its reliability is calculated according to the formula

$$\delta = \max(m^+, m^-)/(m^+ - m^-)$$

and the sign is considered to be correct if $\delta \ge 0.75$.

In non-centrosymmetric space groups the phase determination is based on the tangent equation

$$\tan \phi_h = \frac{\sum_{h'} w_{h'} w_{h-h'} |E_{h'} E_{h-h'}| \sin(\phi_{h'} + \phi_{h-h'})}{\sum_{h'} w_{h'} w_{h-h'} |E_{h'} E_{h-h'}| \cos(\phi_{h'} + \phi_{h-h'})} = \frac{P_h}{Q_h}$$

The reliability of this determination can be evaluated as

$$\alpha_h = 2\sigma_3 \sigma_2^{-3/2} |E_h| (P_h^2 + Q_h^2)^{1/2}$$

and the weights are given by $w_h = \tanh(\frac{1}{2}\alpha_h)$.
The figure of merit, $S_i = \sum_h w_h |E_h| / \sum |E_h|$, is calculated for every phase set, and the 50 best sets $\{\phi_h^1\}$, $\{\phi_h^2\}$, $\{\phi_h^3\}$ ... $\{\phi_h^{50}\}$ are stored. After all possible starting phases have been scanned, the best phase sets are mutually compared. Close or enantiomorphic sets can be found among them which may happen when some starting reflections provide weak contributions into the other reflection phases in the first steps of calculation. The simple comparison between the phase sets

$$\sum_h (\phi_h^i - \phi_h^j)_{\bmod 2\pi}/m < 0.2 \text{ or } \sum_h (\phi_h^i + \phi_h^j)_{\bmod 2\pi}/m < 0.2$$

where m is the number of the determined phases, enables one to exclude the close phase sets. 15, or less, of the phase sets left are then investigated by calculating their electron density functions.

(iv) The investigation of these E-syntheses.
1.2n highest peaks in the electron density maps are chosen during the calculation. The peaks are ordered according to their electron density and they get definite sorts in accordance with the structure to be determined. After the atoms have been places into the sites with maximal values of electron density, the factor

$$R = \sum_h ||F_{obs}| - |F_{calc}|| / \sum_h |F_{obs}|$$

is calculated. As it has already been mentioned 15 E-syntheses are studied in this way; the atomic coordinates corresponding to minimal R-factors are stored, and in most cases this procedure ensures the adequate determination of the structural model. As a result of the E-synthesis study, for every atom the 10 shortest distances to the neighbour atoms are printed each time. Therefore, the procedure for the search of structural model of moderate complexity appears to become completely automatic. Non-automatic search for a structural model based on the direct method is managed by a low number of parameters, i.e. by the number of E_{hkl}, number of Σ_2-relationships, and of starting reflections. These parameters can be input from punchcards. Rather complicated structures can be determined by the variation of these parameters. Some examples of structures determined are given in the Table 1.

Table 1

Compounds	n	Symmetry	K_E	$K\Sigma_2$	M	N	I(S)	P
$C_{50}H_{76}O_{16}$	66	$P2_1$	300	6000	1	512	4	41
$C_{33}H_{57}N_3O_9$	45	$P6_1$	230	4500	–	512	1	32
$C_{54}H_{70}O_3$	57	$P2_12_12_1$	400	5000	2	256	4	32
$C_{34}H_{60}N_4O_8$	46	$P2_1$	270	4500	–	64	1	39

Table 1 (continued)

Explanation of symbols:

n - number of asymmetric atoms.
K_E - number of E_{hkl}.
$K\Sigma_2$ - number of Σ_2-relationships.
M - number of starting phases determined from Σ_1.
N - number of phase sets.
I(S) - number of the phase set in the list of descending S estimates corresponding to the correct structure.
P - number of correct peaks of E-synthesis among n highest peaks.

In the course of a non-automatic search procedure it is possible to determine the phase for some reflections from Σ_1-relationships and the choice of the best set may be done visually from the evaluation of NQEST, which takes into account the statistical relationships between phases of four reflections[9,10]:
$\phi_h + \phi_k + \phi_l + \phi_m = \pi$, $(h + k + l + m = 0)$, provided $|E_h|$, $|E_k|$, $|E_l|$, $|E_m|$ are of sufficiently high values, and $|E_{h+k}|$, $|E_{h+l}|$, $|E_{h+m}|$ are remarkably low.
In this case, $NQEST = \Sigma B_i \cos(\phi_h + \phi_k + \phi_l + \phi_m)/\Sigma B_i$ where $B = (2/N)|E_h E_k E_l E_m|$ with N = the number of atoms in the unit cell.

4. Fourier synthesis

By means of this operation Patterson function, E-synthesis, F-synthesis or difference synthesis may be calculated after entering a corresponding code. In the two last cases the phases are calculated from the atomic coordinates which are summarized before the Fourier synthesis. In all cases Fourier synthesis can be printed. However, this possibility is rather seldom used, since, independently of the kind of calculation, coordinates of the largest peaks are printed, as well as their interdistances with all elements of symmetry taking unit translations into account.
When the Patterson function is calculated, the centre of symmetry is automatically added and translational elements are excluded.
After E-synthesis calculation, as has already been mentioned, ordering of atoms in accordance with their sorts and R-factor calculation is carried out after the peak choice. The so called procedure of peak "recognition" is utilized in the calculation of F-synthesis with the aim of keeping the previous atomic numeration: when the peak is found, close to the initial atom with number m of the sort s, then atom number m of the sort s is placed in the position of the new found peak. Hence, all "recognized" atoms will have definite numbers and sorts, non-"recognized" peaks will automatically get new numbers. This procedure enables one to applicate the heavy atom technique for the search of the structure as a cyclic repetition of the F-synthesis cal-

culation by choosing each time an increasing number of peaks.
The procedure of the phase correction[11] is also carried out by repetition of F-synthesis calculation with that particular feature that normalized electron density functions are stored together with atomic coordinates: $\mu_j = \rho_j/\rho_1$, where $\rho_1, \rho_2, \ldots, \rho_n$ - electron densities in points $x_1, x_2, \ldots, x_n$. Then the phases are calculated according to:

$$F_{calc} = \sum_{j=1} f_j(s)\exp(h, x_j)\mu_j$$

It is worthy to point out one very important moment in the procedure of calculation of a difference synthesis. Either the electron density function or its peak coordinates are usually printed in this procedure, the workers being encountered with difficulties in identification of peaks obtained, since eliminated peaks are not present in the synthesis. In our procedure peak coordinates gained from the difference synthesis and interdistances are provided; not only for "new" peaks but for eliminated atoms as well. The peaks so obtained are readily identified without the necessity of a detailed analysis of printed electron density data.
In the automatic mode of operation, E-synthesis and F-synthesis are calculated with the steps of 0.3 Å in all three directions, and 1.2 n of the largest symmetry independent peaks are chosen. When parameters are read from punchcards, the number of peaks chosen and the step may be varied.

5. Least-squares refinement

The least squares approach has two modifications in RÖNTGEN-75: the block-diagonal one and the full-matrix one. Automatic calculation, including two stages of refinement, of all atomic coordinates and of coordinates with isotropic temperature parameters, is also designed. During the full-matrix least-squares refinement it is possible to impose some restrictions on interatomic distances[12]. Some distances can be equivalenced and some of them can be set to predetermined values. The use of restrictions on interatomic distances provides a way to accelerate least-squares convergence, thus decreasing the consumption of computer time.

References

1. Hauptman H & Karle J. (1953). Solution of the Phase Problem. ACA Monograph, No. 3. Ann Arbor: Edwards Brothers, Inc.
2. Woolfson M.M. & Germain G. Acta Cryst. B24, 91 (1968).
3. Andrianov V.I., Safina Z.Sh. & Tarnopolsky B.L. Zh. strukt. khim. 15, 911 (1974).
4. Stewart J.M. Crystallographic Computing. pp. 71-74. Copenhagen, Munksgaard, 1970.
5. Borisov S.V., Golovachev V.P. & Belov N.V. DAN SSSR 3, 270 (1958).
6. Karle J. & Hauptman H. Acta Cryst. 14, 217 (1961).

7. Motherwell W.D. & Isaacs N.W. Acta Cryst. A27, 68 (1971).
8. Andrianov V.I. & Tarnopolsky B.L. Crystallographic Computing, pp. 115-119. Copenhagen, Munksgaard, 1976.
9. Hauptman H. Crystallographic Computing, pp. 120-130. Copenhagen, Munksgaard, 1976.
10. De Titta G.T., Edmonds J.W., Langs D.A. & Hauptman H. Acta Cryst. A31, 427 (1975).
11. Simonov V.I. Crystallographic Computing, pp. 138-143. Copenhagen, Munksgaard, 1976.
12. Waser J. Acta Cryst. 16, 1091 (1963).

THE SYNTEX XTL STRUCTURE DETERMINATION SYSTEM

Robert A. Sparks, Ph.D.*

Summary

A complete program system for crystal structure determination is described. The greater part is written in FORTRAN IV for Nova 1200 and 800 and Data General Eclipse computers. For reasons of economics in some versions the longer running programs like least squares refinement are written in FORTRAN V or assembler language . It is apparent that the most cost-effective method for structure determination calculation is to use a dedicated mini computer rather than a large computer.

Introduction

In 1972 (1) Syntex announced that they were developing a Structure Determination System based on the Data General Nova 1200 computer which was the same computer used to control the Syntex $P2_1$ Diffractometer. The first Syntex XTL Structure Determination System was delivered in early 1974. The minimum computer configuration for that system was a Nova 1200 or a Nova 800 with 24K - 16 bit words, a 2.5 Mbyte disk, a 12.5 ips magnetic tape drive and a Floating Point Processor built especially for the XTL.

A more powerful version is the E-XTL Structure Determination System. The minimum computer configuration for this system is a Data General Eclipse computer with 32K - 16 bit words, two 2.5 Mbyte disks, a 12.5 ips magnetic tape drive and a standard Floating Point Processor.

Many of the XTL systems actually use the same Nova 1200 computer for controlling the FORTRAN $P2_1$ diffractometer and for solving crystal structures. In March of 1978 this combination was replaced with the more powerful R3 Crystallographic System.

The minimum computer configuration for the R3 is a Data General Nova 3 computer with 32K - 16 bit words, a 10 Mbyte disk, a 12.5 ips magnetic tape drive and a standard Floating Point Processor.

*Syntex Analytical Instruments, Cupertino, California, U.S.A.

A very useful option selected by most customers is a Versatec 11" printer-plotter.

The programs for the XTL are written mostly in FORTRAN IV. The special Floating Point Processor was designed to execute these FORTRAN IV programs as efficiently as possible. It was also designed to execute as fast as possible the inner loop of the Full Matrix Least Squares which was written in assembly language.

Data General's FORTRAN V compiles code which is usually more efficient than FORTRAN IV. For this reason, the longer running programs (BLOCK, FLMS, and FOUR) are written in FORTRAN V for the E-XTL and R3 system. In addition, MULTAN has been converted to FORTRAN V for the R3 system.

Except for the Full Matrix Least Squares the algorithms for all of the programs are the same for the XTL, E-XTL and R3 systems.

The Syntex Structure Determination Systems process single crystal data to determine molecular structures. Calculations are handled by a comprehensive set of programs: from data reduction and structure solution through least squares refinement and table and diagram preparation. These programs are written in the widely used FORTRAN language, so that modifications and additions can be made easily.

The programs are conversational and utilize unformatted (or free-form) input. The user is prompted on each input by a print-out of possible responses. An automatic updating procedure minimizes user intervention and eliminates redundant user inputs and calculations by individual programs. Information pertaining to a particular crystal is stored in a central data file set. Each program in the system uses whatever data is required from this set of files, updating and appending the files as calculations are made. Thus, frequently used input parameters (e.g. cell dimensions, space group information, empirical formula) need only be entered once. And, since files are rewritten automatically as new information is available, calculations made by one person can be utilized by the others. The user rarely needs to intervene in going from one program to the next. In most cases, the specification of the program name, the data file identifier (common to the entire data file set), and sometimes a small number of input parameters, are all that is required to run any program in the system.

Another feature that adds to the user's convenience is that most of the programs can be executed in batch mode, where input is read from a previously prepared list, usually stored on the disk. In this way, a number of different users may stack or queue many tasks in a job stream to be executed sequentially, unattended. This feature is useful where, for example, a number of long jobs are to be run overnight.

A large disk provides the necessary storage for extensive data manipulation, program storage, swapping storage, and as an extension of core storage for the full matrix least squares program. For convenience, infrequently used files such as source programs, source file for scattering factors, or data sets not in current use, can be stored on magnetic tape to be loaded onto the disk as necessary. The magnetic tape is also useful for intermediate, off-line storage during transfer of programs or data from one disk pack to another. Optionally, a second disk drive can be incorporated to increase speed.

Data Files

All program files, data files and scratch files reside on disk. A list of the data files and their contents are shown in Table I. Because more than one data set can reside on disk at any one time, the data files contain as part of their names a four letter code identifying the structure to which they belong. For a given structure, this four letter code is substituted for the letters "USER" in the file names shown in Table I.

The program file interaction is shown in Figure I. The names of the major programs are shown in rectangles, the files in ovals.

XTL PROGRAM-FILE INTERACTION

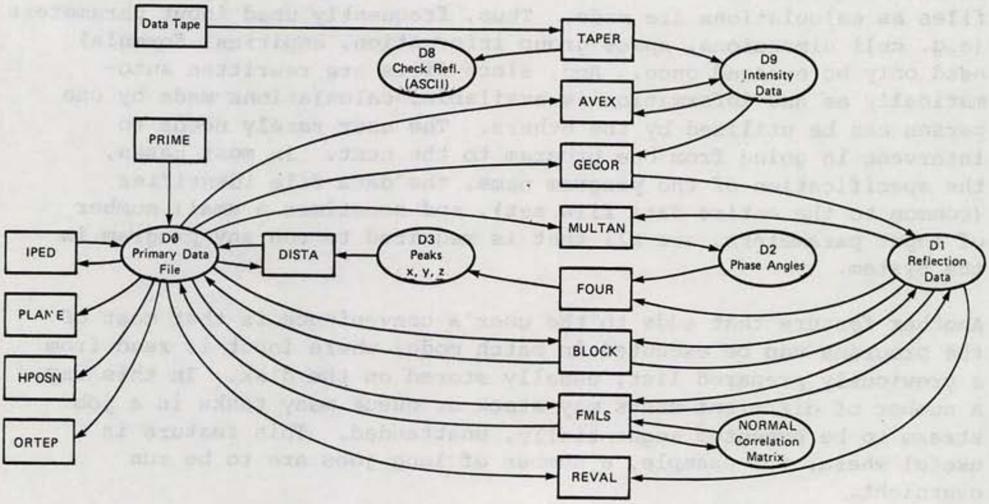


TABLE I

List of the data files and their contents

File	Name	File Contents
Primary Data File	D0USER.DA	Title, cell dimensions, space group equivalences, scattering factor types, overall scale and temerature factors, weighting scheme, atomic coordinates, etc.
Reflection Data File	D1USER.DA	h,k,l,E,F,σ,A_c,B_c, scattering factors, etc.
E phase Data File	D2USER.DA	E phases for each MULTAN solution
Peak File	D3USER.DA	Peak information generated by Fourier, E-map, or Patterson
H,K,L File	D6USER.DA	File of h,k,l values for $P2_1$ FORTRAN programs
$P2_1$ Parameter File	D7USER.DA	Orientation matrix, wave lengths, centering, indexing, collection, etc. parameters for $P2_1$ FORTRAN programs
Check Reflection File	D8USER.DA	h,k,l, intensity, time for check reflections
Raw Intensity Data File	D9USER.DA	h,k,l, intensity, σ
Scattering Factor File	SAIASF.TB	Cromer-Waber scattering factor tables
Normal Equation File	NORMAL.TM	Full-matrix normal equations

Structure Determination Programs

Large programs like MULTAN and the Full Matrix Least Squares (FMLS) were divided into many overlays so that each part would fit into 24K or 32K of core memory. MULTAN required 15 segments and FMLS required 5 on the Nova 1200 and 7 on the Eclipse and Nova 3. To minimize the number of disk transfers (which are time consuming) a very thorough understanding of a given program was necessary in order to make an efficient division into segments.

What follows is a brief description of the major programs.

PRIME

One of the central features of the XTL is use of a primary data file to store crystal information such as cell dimensions, equivalent positions, atom coordinates, etc. All information necessary for creating this file should be available to the user following data collection and prior to the initiation of the structure solution.

Another feature is that PRSYM, the second part of PRIME, is written so that it is not necessary to enter lattice type, lattice centering, equivalent positions, and other information subject to input errors. All information about a space group, including the general equivalent position set, is derived from the space group symbol using a routine originated by A.C. Larson.

An example of the input and output for PRIME is shown in Figure 2. The user inputs are underlined.

FIGURE 2

```
PRIME

S.A.I. XTL PROGRAM                    PRIMARY DATA FILE SETUP (2)

ENTER DATA FILE ID (4 CHAR.): TEST
ENTER COMPOUND NAME (30 CHAR. MAX): NEW TEST CRYSTAL
WHAT KIND OF X-RADIATION? (CU,MO,AG,FE,CR): MO

ENTER CELL CONSTANTS TO AS MUCH PRECISION AS POSSIBLE
WITH LENGTHS IN ANGSTROMS AND ANGLES IN DEGREES
ORDER - A,B,C,ALPHA,BETA,GAMMA - SEPARATE WITH COMMA'S
:13.712,15.241,10.334,90,92.17, 90
ENTER STD. DEV. FOR CELL DIMENSIONS
:.001,.002,.001,0,.02,0
IS THIS SECTION CORRECT? (YES OR NO): YES
```

FIGURE 2 (CONT.)

```
ENTER SPACE GROUP SYMBOL: P 21/C
   CRYSTAL CLASS: MONOCLINIC  LAUE GROUP: P 2/M CENTROSYMMETRIC
   NUMBER OF GENERAL POSITIONS: 4  POLAR AXIS: NONE
   UNIQUE 2-FOLD AXIS: Y

   EQUIVALENT POSITIONS:
        X,       Y,      Z,      -X,1/2+Y,   1/2-Z

ENTER ATOMIC NAMES IN STANDARD FORMAT. GIVE THE NUMBER OF EACH
IN THE ASYMMETRIC UNIT WHEN REQUESTED. FINISH WITH "END".

ATOM TYPE: FE+2
HOW MANY? 2
ATOM TYPE: C
HOW MANY? 18
ATOM TYPE: O
HOW MANY? 3
ATOM TYPE: H
HOW MANY? 8
ATOM TYPE: END
IS THIS SECTION CORRECT? (YES OR NO): YES

                         ATOM LIST

COMPOUND: NEW TEST CRYSTAL                      IDENT: TEST

                      A      B      C     ALPHA BETA  GAMMA VOLUME
CELL DIMENSIONS: 13.712 15.241 10.334 90.00 92.17 90.00 2158.1

ATOM   NUMBER/   ATOMIC    ATOMIC     WEIGHT      ABSORPTION
        CELL     NUMBER    WEIGHT      (%)          COEFF.

FE+2     8.       26.       55.85      29.1          13.2
C       72.        6.       12.01      56.3           0.4
O       12.        8.       16.00      12.5           0.2
H       32.        1.        1.01       2.1           0.0

NUMBER OF ATOMS IN THE UNIT CELL = 124.
UNIT CELL SCATTERING F (000) =       768. ELECTRONS
UNIT CELL VOLUME =                 2185.1 A3
UNIT CELL MASS =                   1535.8 AMU
CALCULATED DENSITY =                  1.18 G/CM3
ABSORPTION COEFFICIENT =              13.9 CM(-1)

PRIMARY DATA FILE SETUP COMPLETE            FILENAME: D0TEST.DA

R
```

TAPER

TAPER is an XTL utility program which reads Syntex $P\bar{1}$, $P2_1$ or P3 data collection tapes and stores the information on two data files: D9USER.DA and D8USER.DA.

D9USER.DA is a binary file and contains the raw intensity data. Information for each reflection is stored in 9 words as follows: sequence number, h, k, l, intensity (2 words), standard deviation (2 words), and integer code word (ICW) which identifies certain characteristics of the reflection.

D8USER.DA is an ASCII file and contains the check reflection information and the elapsed data collection time. For each reflection, the format is (I7,3X,3I5,2F10.0) for the following information taken directly from magnetic tape: sequence number, h, k, l, intensity, elapsed data collection time.

Raw intensity data is checked for bad tape records and for very strong and very weak reflections. Each reflection is also scanned for background imbalance, and the peak profile is examined for peak centering and asymmetry. Each error condition is reported in the printout and the appropriate flag set in the data file (D9USER.DA).

An empirical absorption correction can be made to the data, using a reasonable number of reflections (maximum = 20) distributed over the entire range of 2θ values used in data collection, each reflection with 36 points taken by a full rotation around the diffraction vector in steps of 10° (beginning $\psi = 0°$; ending $\psi = 350°$; ψ increment = 10°). Each reflection so collected will define an absorption curve of intensity vs. ϕ corrected for ω and χ. The curve nearest in 2θ to a regular reflection will be interpolated to correct that reflection for absorption. These reflections should be on the first file of the data tape. The second file contains the normal data collection records.

A subroutine HKLSG is called, once for each reflection, so that systematically absent reflections can be eliminated from further calculations by setting the ICW to 20. The user is asked to type in conditions defining these absences. The user is also asked for the sequence numbers of too strong reflections in order to re-enter intensity, standard deviation, and scale factor for those reflections if collected at some later time. This subroutine can also be modified by the user to manipulate the data as desired, for example, to transform reflection indices, or to delete specific reflections by setting ICW to some negative number.

AVEX

AVEX uses the equivalent positions in the D0USER.DA file to flag space group extinct reflections in the D9 file. It will also average the data to give a D9 file containing unique data only. For centrosymmetric space groups, all equivalent reflections will be averaged. For non-centrosymmetric space groups, the user may choose whether or not to average Friedel pairs (anomalous scattering). This program should always be run before using MULTAN to ensure a unique data set.

GECOR

GECOR performs geometrical corrections on intensity data for Lorentz-polarization factors, calculates a Wilson plot, and outputs the scale and temperature factor. It also accumulates intensity statistics and calculates normalized structure factors (E's). A reflection data file (D1USER.DA) containing F's, E's and atomic scattering factors is created with this program.

MULTAN

MULTAN is an adaptation for a small computer of Main, Woolfson, and Germain's program for direct method solution of crystal structures. It automatically selects origin and starting phases to perform multiple solution tangent refinement.

MULTAN consists of three parts: Part I (SIGMA1) sets up phase relatinships; Part II (CONVERGE) finds starting reflections for tangent formula, and Part III (FASTAN) determines phases for all reflections from each starting set produced by Part II using the tangent formula.

Both Part I and Part II output all of their results to disk files. The program can be started at the beginning of any section provided the previous section has already been run. As many sections as desired may be executed in a single run and any section may be repeated on a subsequent run. The previous results on the disk files will merely be replaced by the latest results. Only those data pertaining to the sections being run will be requested.

FOUR

This program calculates Fourier summations and Patterson functions. Either a normal Fourier synthesis or a difference Fourier synthesis can be calculated. Symmetry generation and Fourier summations are based on the algorithms developed by Bernt Klewe at the University of Oslo. E-maps can be calculated from the results output by MULTAN.

Origin removal and/or sharpening are options when calculating Patterson functions. The Patterson function is scaled so that the origin peak is ≃ 10,000. The peak picking routine does a 19 point interpolation according to J.S. Rollett's algorithm (2) and sorts the peaks according to height. Peak coordinates can be transformed to "one molecule." Distance and angles and peak plots are printed out as an option.

IPED

IPED allows the inputting and editing of overall parameters (e.g. overall scale factors) and atom parameters on the primary data file. It assumes that the first six records of the primary data file have been generated by PRIME.

IPED uses a conversational approach, and each record is written onto a temporary file as it is edited. When all necessary editing is completed, the edited records can be written onto a new file, retaining the old one, or written onto the old file and replacing it completely. Any or all of the primary data file can be edited, and portions of the file that do not require editing can be easily skipped using the WRITE option (which writes an entire record, as is, onto the new file), the COPY option (which, during the editing of atom parameters, copies the remaining records onto the new file), and the SEARCH option (which proceeds directly to the specified atom, copying the ones preceding it).

BLOCK

A Block-Diagonal Gauss-Seidel Least Squares program has been written for the Syntex XTL by Gilbert (3). The method was first proposed by V. Schomaker and J. Waser. This method is to obtain derivatives for one independent atom while calculating structure factors; solve the resulting nine equations in nine unknowns; correct the calculated structure factors forming a new set of nine equations for the next independent atom and so forth. After the last atom parameters have been adjusted in this way, the scale factor is corrected by the same procedure. This is followed by a new adjustment of the first atom parameters, etc. This method is very much like including all of the off-diagonal elements below the diagonal of the full matrix without the need to actually calculate them.

FMLS

The Full Matrix Least Squares program is an extensive modification of the Trueblood, Gantzel, Sparks UCLALS4 program. Formation of the normal equations is written in assembly language to achieve fast execution speed. All other parts are written in FORTRAN. For the Nova 1200 and Nova 800 XTL Systems the following algorithm for the formulation of normal equations was developed. Derivative vectors for three reflections are generated and stored in core. Then one half-track (768 elements) of the normal equation matrix are read into core from disk. As soon as the first element arrives in core the three corresponding products from the three derivative vectors are added to it. The next element is then processed in the same way. The arithmetic operations (three floating point multiplies and adds) are slower than the disk transfer rate. As soon as the disk has transfered the first half-track it reads in the second half-track into a second array in core. At the end of reading the second half-track the processing for the first half-track will be finished and can be written back onto the disk. Finally at the end of processing the second half-track it also is written back onto the disk. The floating point processor operates independently of the central processing unit and the program is written so that time for the address arithmetic calculations (done by the CPU) are almost completely overlapped by the floating point processor time. Thus, the floating point processor is busy about 80% of the time.

The Eclipse floating point processor is faster than the special processor designed for the Nova 1200 and Nova 800; however, the disk transfer rate is about the same for all of these computers. Therefore, for the E-XTL it is necessary to process more than three derivative vectors at a time. The algorithm used is the following. Derivative vectors for many reflections (several hundred) are written onto a disk file. When this file is filled a block of the normal equation matrix (3072 elements) is transferred to core. This block is then updated with contributions from all of the vectors in the derivative file. Then that block of the normal equation matrix is written back on disk. The derivative vector must be read for each 3072 element block of the normal equation matrix. Depending on the total number of reflections the derivative vector file may have to be written more than once and each time the normal equation matrix must be updated. To minimize disk seek times, it is necessary to have the derivative vector file and the normal equation file on different disks.

On the Nova 1200 version of the Syntex XTL two cycles of a 217 paramter, acentric problem with four equivalent positions and 1152 reflections took 117 minutes. It took 102 minutes on the R3, 96 minutes on the Nova 800 XTL and 44 minutes on the Eclipse E-XTL.

Up to 500 parameters can be handled by the XTL, E-XTL and R3 Full Matrix Least Squares programs.

DISTA

DISTA calculates all distances and angles within specified ranges, printing the results in a connectivity table format. A second version, DISTAS, also calculates standard deviations for the distances and angles.

PLANE

The plane program calculates mean planes through a set of atoms specified by their identifying symbols in the data files, calculates distances of atoms from the plane, and performs a χ^2 text on atoms in the plane.

The program was originally written by M.E. Pippy for Farid Ahmed's NRC (Canada) system, from the procedure of B.M. Blow (4).

HPOSN

HPOSN predicts the positions of hydrogen atoms based on various types of geometry, given that the atoms which define the geometry are already in the data file.

The program is adapted from that of Hakon Hope, University of California at Davis.

REVAL

REVAL calculates R factors, weighted sums-of-squares and related statistics as functions of reflection sequence number, F, sin θ/λ, and the various classes of Miller indices. It is very useful for evaluating the quality of data in different ranges and examining the correctness of weighting functions. Optionally, REVAL will generate a full listing of observed and calculated structure factors and their differences.

STRUCTURE PLOT

The Syntex Structureplot is a software package based on C.K. Johnson's ORNL ORTEP-II program and is designed for compatibility with all $P2_1$ FORTRAN, XTL, E-XTL and R3 systems equipped with the high speed line printer/plotter. It includes a series of programs for generating ORTEP instructions from DØUSER.DA and D3USER.DA files, for plotting, as well as for manipulating the binary raster data file.

User Questionnaire

A questionnaire concerning changes and additions to the Syntex Structure Determination System has been sent to all users. A summary of their responses will be presented.

Comparison of Minicomputer and Large Computer

A comparison of the minicomputer and large computer for crystallographic calculations has been made (5). Although the cost of minicomputers has remained low, their computational capability has increased enormously. It is now apparent that the most cost-effective method for structure determination calculations is to use a dedicated minicomputer.

References

1. Sparks, R.A., "Computational Needs and Resources in Crystallography", pp. 66-75, Washington, D.C.: National Academy of Sciences, 1973.

2. Rollett, J.S., Computing Methods in Crystallography, p. 35, Pergamon Press, 1965.

3. Gilbert, M., "Operations Manual Syntex XTL Structure Determination System", Syntex Analytical Instruments, Inc.

4. Blow, B.M., Acta Cryst. 13, p. 168, 1960.

5. Sparks, R.A., ACS Symposium Series, No. 57, pp. 94-105, 1977.

THE ENRAF-NONIUS CAD 4 SDP - A REAL-TIME SYSTEM FOR CONCURRENT X-RAY DATA COLLECTION AND CRYSTAL STRUCTURE SOLUTION

Bertram A. Frenz*

Summary

The Enraf-Nonius Structure Determination Package (SDP) is a coherent collection of crystallographic programs for the solution and refinement of small molecule structures. The SDP complements automatic x-ray data collection on the CAD 4 diffractometer. Together, in a real-time operating system, they provide a complete structure determination facility. This paper highlights the principal advantages of the CAD 4 SDP with particular emphasis on those features which are not present in other crystallographic computer program packages.

1. Introduction

At Enraf-Nonius the development of crystallographic programs on the mini-computer was an outgrowth of the manufacture and sale of the computer-automated CAD 4 diffractometer. This factor has been responsible for the direction taken in the design of the SDP. From the outset the intention has been to develop a time-sharing facility that would provide a continous flow from data collection through structure solution and refinement with a minimal need for operator intervention. At the same time the real-time operating system would allow simultaneous operations thereby maximizing the usage of the diffractometer and computer hardware.

The SDP was first released in mid-1974 in a single-user mode and was introduced for multi-user real-time operation in the autumn of 1976. Throughout this period, and subsequently, the SDP has been in continuous development to take advantage of new software and hardware advances. At the present time the SDP can be used on any of nine different models of Digital Equipment Corporation's PDP-11 computers and a large variety of disk and tape drives and terminals.

2. The operating system

The operation of the real-time system is imperceptible to the user. Executable programs, called tasks, are transferred in and out

*Molecular Structure Corporation, P. O. Box DF, College Station, Texas 77840

of memory according to a system of priorities and hardware interrupts. Although the central processing unit (cpu) and the peripherals can each only service one operation at a time, they can all operate independently. Thus at any given moment of time a disk transfer, a floating-point calculation, a terminal message and a line printer output can occur simultaneously. How this affects the operation of the CAD 4 SDP can be seen in the following examples:

Example 1. Concurrent execution of the CAD 4 task and a least-squares task.

The CAD 4 task is initiated when the system is booted and it remains active, although not necessarily memory-resident, at all times. Operational commands are issued from a terminal interfaced to a PDP-8 computer which, in turn is interfaced to the PDP-11 computer. (The PDP-8 contains 4K of memory, but no peripherals for data storage.) The commands automatically cause the CAD 4 program to be rolled into PDP-11 memory, the required calculation is performed, and a request in sent to the PDP-8 to cause the CAD 4 to drive motors, set the aperture wheel, open the shutter, measure intensity, etc. Meanwhile, the PDP-11 is free to perform other functions.

SDP tasks are initiated from a PDP-11 terminal and remain active until completion. If a least-squares task is active at the same time as the CAD 4 task the latter, having the higher priority, will occupy the cpu for the few seconds necessary for its calculation. Then the CAD 4 task will be rolled out to the disk and the least-squares task will be rolled into memory. The least-squares will execute while the CAD 4 is driving motors and measuring intensities. After completion of the intensity measurement the CAD 4 task is again called into memory and the next set of instructions is calculated. The CAD 4 task is required twice per reflection; the remainder of the time is available for the least-squares task. Operation of the CAD 4 during a least-squares calculation increases execution time of the least-squares by <u>ca</u>. 9% relative to a stand-alone situation. For example, a least-squares calculation that takes 7.0 minutes if run independently will take 7.6 minutes if run simultaneously with CAD 4 data collection. If sufficient memory (38K) is available for both tasks to be memory resident then the increase in execution time is only 7%.

Example 2. Concurrent execution of an interactive SDP task and a computational SDP task.

Although it is possible to operate the SDP in batch mode from a card reader in a manner similar to that of large computer facilities it is not the most practical method. Instead it is much simpler to set up and execute jobs through interactive dialogue on a terminal. Short jobs such as data reduction or table production combine conversational interaction with execution; longer jobs such as Fourier syntheses or least-squares refinements are divided into a setup program which is entirely interactive and a calculation program which contains no dialogue. One of the advantages of the two-part job is that one user can set up a series of calculations and another user, unacquainted with the structure, can initiate their execution at a later time. These jobs can also be stacked in a queue for overnight runs.

The disadvantages of interactive programs in a single-user operating system is that a significant amount of wall clock time is spent waiting for the user to decide on the answer to a question and to type in the response. During this time the computer is idle. However in a real-time operating system a computational task can execute parallel to an interactive task. The computational task proceeds almost as if it had exclusive use of the cpu. It is only interrupted briefly when the interactive task receives an 80 character buffer of information for interpretation and sends a new question to the terminal. The interactive task also proceeds as if it had exclusive use of the cpu. Because it is assigned a higher priority the response time for the interactive task is seemingly immediate. The net effect is that the real-time operating system greatly enhances the throughput of the computer.

Example 3. Concurrent execution of two computational SDP tasks.

The minimum supportable central processor for the CAD 4 SDP has 32K words of memory, of which 12-13K words are assigned to the operating system and 19-20K words are available for CAD 4 and SDP tasks. Most SDP tasks require from 12-18K words; hence only one task may reside in memory at one time. For this reason there is no timesaving advantage to running two computational SDP tasks simultaneously in the minimum 32K system. However, in configurations with 48K, 64K, or more memory there is approximately a 13% increase in throughput as a result of two computational jobs residing in memory simultaneously and sharing the cpu and peripherals.

Example 4. Multi-user and multi-terminal tasks.

Another advantage of a real-time system is that it permits the use of more than one terminal and more than one job per terminal. With single-user operating systems larger crystallographic groups have difficulty arranging schedules for terminal usage. With real-time systems and multiple terminals more users can enter jobs during normal working hours and the slack hours can be used by the computer to catch up on the backlog.

3. The design of the SDP

During the development of the SDP a number of basic principles were established in order to ensure that the finished product would combine versatility, speed, and ease of operation. Although many programs readily can be transferred from a large computer facility to a minicomputer the true benefits of the minicomputer are not realized without carefully integrating the entire program package and taking advantage of the special hardware architecture. Some of the more important concepts are outlined here.

Disk files

Much of the power of the minicomputer resides in the fast-access, mass-storage disk drives that serve as reservoirs for program libraries and data files. The SDP programs operate on a variety of disk files containing intensity data, atomic coordinates, symmetry information, least-squares parameters, etc. Initial data, such as the empirical formula and space group number, are entered during interactive dialogue. From this information other parameters such as scattering factor tables and symmetry codes are determined. This information, as well as results determined throughout the study, is stored on a disk file. Since this file serves as a permanent record there is no need to reenter information to repeat a task. For example, setup parameters for a difference Fourier calculation are entered only once and need not be reentered for successive calculations at the various stages of structure solution. Parameters stored in the disk files can be readily recalled. For instance, a single command will reveal the R factors and maximum parameter shift on the most recent least-squares cycle or the percent composition of the compound under investigation or the setup values for a bond distance and angle calculation. In this manner disk files greatly aid in the use of the minicomputer to solve and refine structures interactively.

Program input

The input to SDP programs is kept to an absolute minimum. Thus scattering factor tables, point group, equivalent positions, absorption coefficient, etc. are not required as input since these can be derived from the empirical formula and space group number. Also, data determined in one program is made available via disk files for subsequent programs. For example, distances and angles between peaks in a difference Fourier map can be calculated without entering peak coordinates since the Fourier program writes these on a disk file at the end of its calculation.

Parameters entered through the keyboard terminal are in response to questions posed by conversational tasks. Most of these programs offer a choice between detailed dialogue for the beginning user and brief dialogue for the experienced user. On subsequent runs of an interactive program only values which the user desires to change need be entered; other parameters will retain their previous values.

Program output

In addition to outputting disk files with intermediate results that can be used in subsequent programs, the SDP programs produce optional line printer and terminal hard copy output. Through the use of virtual switches the user, before or during runtime, can a) direct the output to either the input terminal device or a line printer, b) adjust the output to 80 or 132 column format, c) suppress or reinstate listing (e.g., structure factor listing), and d) spool the output file to the disk for listing at a later time. Wherever possible the output is printed in a format suitable for direct submission to journals, complete with table titles, column headings, and appropriate footnotes.

Assembly-language subroutines

There are a number of simple operations that are relatively time-consuming at the FORTRAN level but are very fast when programmed at the assembly-language level. For example, setting an array of variables to zero takes 53.4 μsec per variable in FORTRAN compared to 6.4 μsec per variable in assembly language. To a large extent, the speed of the SDP programs is a result of using assembly language subroutines in the number-crunching sections of the programs.

FORTRAN program sources

With the exception of a few assembly language subroutines described above, all SDP programs are written in FORTRAN thereby simplifying program modifications. The program sources include numerous in-line comments to aid the user in following the program logic. Also, the 275-page SDP User's Guide heavily documents the program algorithms and the use of disk files in addition to providing a guide to running the programs. Although the intention is that the SDP be complete in inself, it is also readily acknowledged that each user's group has its own particular method of performing certain calculations (e.g., least-squares weighting schemes) and every effort is made to accommodate these modifications with a minimum of programming effort.

4. The SDP programs

In this section each of the principal SDP programs is described in the order in which they are usually executed. References refer to sources of numerical values or to descriptions of similar, but not identical, programs. The first dozen programs beginning with ALPHA, are for initiation of disk information files, data reduction, and data corrections.

ALPHA - accepts input of empirical formula, cell constants, and space group number.

SCATER - calculates scattering factor tables for elements included in the empirical formula;[1] obtains anomalous dispersion terms;[2] calculates absorption coefficient[3] and similar values.

SYMTRY - adds symmetry information to the disk file for the specified space group number.

DATARED - reduces intensity data to observed structure factors corrected for background, attenuator factor, and Lorentz and polarization factors.

STDPLT - processes intensity-control reflections and draws plot of intensity versus time for the standards.

CHORTA - corrects intensity data for anisotropic decay of data crystal.

SAC - applies spherical absorption correction to intensity data.

CRYSTAL - calculates parameters related to crystal faces and dimensions (including a drawing of the crystal) in preparation for absorption correction.

ABSCOR - applies numerical absorption correction to intensity data.[4]

FAME - adds normalized structure factors to data file; calculates and draws Wilson plot; performs centrosymmetric/non-centrosymmetric statistical tests and parity analyses.

REJECT - rejects user-specified reflections or classes of reflections from the data set, e.g., non-unique reflections, systematically absent reflections, weak reflections.

ATOMS - accepts input of atom parameters from terminal or from previous E-map or difference Fourier map.

MULTAN - determines phases of reflections by direct methods;[5] handles all space groups; is currently dimensioned for up to 500 input reflections, 2000 relationships and 128 phase sets.

FOURA - accepts directions for type of Fourier calculation to be performed and portion of unit cell to be examined.

FOURB - calculates Patterson function, Fourier, difference Fourier, E-map; performs peak search and writes peaks on disk file for subsequent use in other programs.

SEARCH - performs connectivity search on Fourier peaks and/or refined atom positions; plots projection drawing of peaks and atoms; prints distances and angles.

LSA - accepts choices for least-squares options, such as atoms to be included, number of cycles, scale factor, dampening factor, extinction coefficient.

LSB - refines structure by full-matrix least-squares; handles up to 1000 variable parameters; refinement can include scale factor, coordinates, isotropic and anisotropic thermal parameters, occupancy factors, and extinction coefficient; constraints on occupancy factors can be applied; parameter shifts can be dampened; parameters can be held constant and/or reset to other parameters.

LSF - calculates structure factors and adds them to data file for subsequent Fourier synthesis.

LSG - inverts least-squares matrix.

CORCOE - calculates correlation coefficients.

ISOAN - converts thermal parameters from isotropic to anisotropic.

WEIGHT - calculates weights for reflections to be included in least-squares (three weighting schemes are available); allows elimination of weak and unobserved reflections or reflections within specified ranges of $\lambda^{-1}\sin\theta$.

ESD - calculates bond distances and angles with e.s.d.'s derived from elements in the inverted least-squares matrix.[6]

RMS - calculates root-mean-square amplitudes of thermal vibration.

PLANE - calculates weighted or unweighted least-squares planes, distances of atoms from planes and dihedral angles between planes.

ORTEPA - accepts plotting instructions through interactive dialogue.

ORTEPB - produces plot file for drawings of molecules and unit cells.[7]

ORTEPC - interfaces ORTEP plot instructions to Versatec or Calcomp software.

LIST - sorts indices in specified order and lists observed and calculated structure factors in a form suitable for publication.

PARLST - lists atomic positional and thermal parameters with their respective standard deviations.

TEMPER - lists general temperature factor expressions for anisotropically-refined atoms.

DRILL - calculates setting angles for Supper ball drilling machine for making molecular models.

NZTEST - performs N(Z) statistical test.

TRACER - performs Delaunay reduction.

Acknowledgement

Much of the credit for the initial design of the SDP goes to Professor Yoshi Okaya of the State University of New York at Stony Brook. The creation of the CAD 4 program under the real-time operating system was principally the work of Dr. Hans Bruggeman of Enraf-Nonius Delft. These contributions and their continuing interest in the development of the CAD 4 SDP are gratefully acknowledged.

References

1. D. T. Cromer and J. T. Waber, "International Tables for X-Ray Crystallography," Vol. IV, The Kynoch Press, Birmingham, England, 1974, Table 2.2B.

2. D. T. Cromer, "International Tables for X-Ray Crystallography," Vol. IV, The Kynoch Press, Birmingham, England, 1974, Table 2.3.1.

3. a) "International Tables for X-Ray Crystallography," Vol. III, The Kynoch Press, Birmingham, England, 1962, Table 3.2.2B.
   b) D. T. Cromer and D. Liberman, J. Chem. Phys., 53, 1891 (1970).

4. Ph. Coppens in "Crystallographic Computing," F. R. Ahmed, Ed., Munksgaard, Copenhagen, 1970, p. 255.

5. G. Germain, P. Main, and M. M. Woolfson, Acta Crystallogr., Sect. A, 27, 368 (1971).

6. W. R. Busing, K. O. Martin, and H. A. Levy, Report ORNL-TM-306, Oak Ridge National Laboratory, Oak Ridge, Tenn., 1964.

7. C. K. Johnson, Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tenn., 1965.

THE CRYSTAN SYSTEM

H. Burzlaff, R. Böhme, M. Gomm*

Summary

CRYSTAN is a crystallographic program system for minicomputers written in FORTRAN IV language. A survey of the structure of the system is given. Some special features in handling symmetry and the structure determination procedure with direct methods are discussed in detail.

1. Introduction

The present version of CRYSTAN is developed for a PDP 11/45 system with 24 K core, 262 K fixed head disc, two magtapes TU 10, medium speed line printer and teletype key board. The structure of the system, however, is such that changes to other configurations with other peripheral equipment are rather simple.

In the present version, CRYSTAN I, the system may be divided into five blocks (see fig. 1, structure of CRYSTAN). Block I provides for data handling with the programs INPUT, INCAL, ABSOR and DARED, block II takes care of structure determination with direct methods using the programs ENORM, SIGMA2, START, SYMBAD, FASTAN, EQUAL and SAYFO, block III contains the least squares procedures, block IV contains the programs FOUPAR, PREFOU, FT1, FT2 and FT3 for Fourier procedures and finally block V comprises the programs DISTAN, PARAM, MOLEC, OPLA and THERM for tabulation and interpretation of structure data. Facilities for operating the system are shown in fig. 2. It is assumed that the complete system with unchangeable library data files like scattering factor tables is on magtape or disc. The system always starts with the program INPUT which collects information on the structure from the user and library data files. Measurement data is worked up by INCAL. Absorption corrections for crystals shaped cylindrically or spherically (ABSOR) are possible. DARED averages multiple measurements and those of symmetrically equivalent reflexions, it takes care of Lorentz-polarization corrections.

ENORM produces E-values for direct methods, tests symmetry conditions and computes the absolute scale factor and overall B. If direct methods are wanted to be used, all triplet relations are computed by

---

*Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Angewandte Physik, Loewenichstr. 22, 8520 Erlangen, BRD

SIGMA2. START selects a good start set of triplet invariants and reflexions. Symbolic additions can be started with SYMBAD (only in centrosymmetric cases) or with FASTAN. The resulting solutions

	CRYSTAN I	CRYSTAN II
DATA HANDLING	1.INPUT 2.INCAL 3.ABSOR 4.DARED	-analytical and experimental absorption correction -profile analysis
STRUCTURE DETERMINATION	1.ENORM 2.SIGMA 2 3.START 4.SYMBAD 5.FASTAN 6.EQUAL 7.SAYFO	-additional criteria for structure invariants -simultanious test of different structure proposals - Ψ -test
LEAST SQUARES PROCEDURES	1.DFM 2.DFTO 3.DFIS 4.DFAN 5.DFSTR	-extinction corrections -multiplier refinement -full matrix least squares
FOURIER PROCEDURES	1.FOUPAR 2.PREFOU 3.FT1 4.FT2 5.FT3	-corrections of anomalous dispersion
INTERPRETATION PART	1.DISTAN 2.MOLEC 3.THERM 4.OPLA 5.PARAM	-plotter software

Fig. 1. Structure of CRYSTAN

from FASTAN are tested for equivalence by EQUAL. The initial set of phases is extended and refined by SAYFO for centric and noncentric structures. An E-map is calculated via FOUPAR. This program provides for Fourier parameters. For Patterson interpretation FOUPAR is called directly, followed by PREFOU. All Fourier calculations are done in FT1 and are interpreted in FT2 and FT3. It is possible to compute Patterson functions, E-maps, electron density maps, difference maps and electron density maps for valence electrons only. Data output is suitable for least squares or another Fourier.

Least squares procedures are managed by DFM, it chooses the optimal program to save computing time.

DISTAN produces distances and angles, MOLEC connects the atoms of the asymmetric unit to a molecule, THERM computes thermal ellipsoids from

anisotropic temperature factors and prepares information for a plot, OPLA determines optimal planes for an input set of atoms and computes angles between pairs of different planes. PARAM sets up a table of coordinates and temperature factors for publication.
The last column of fig. 1 indicates the improvements being in development.
In the meantime about 30 structures have been solved by CRYSTAN. A report on the experiences with the system will be given at the Warsaw congress.
A complete description of CRYSTAN with full output list of the source programs, with specifications how to handle it and with examples and their results is available [3].

2. Philosophy of operation

The most important disadvantage of a minicomputer is the lack of core. There are two possibilities to overcome this problem: extended use of overlays and subroutines or independent standalone programs connected by a chain job.
In the first case program maintenance is more difficult and it is not possible to make optimal use of the core. In the second case data transfer from one program to the next one must be organized using data files. In CRYSTAN the last procedure is prefered.
The system avoids program parts written in assembler language except in three cases, this is enforced by the file structure. These three exceptions are:

(i) Create a file:
CALL SETFIL (N, XY, O, P, I)
This is special DEC FORTRAN; N means the channel number, XY the file name, P the device name (MT or DF), I the unit number. Other operating systems have similar statements.

(ii) Delete the file that was used on channel number N:CALL DEL(N)

(iii) Set magtape no.N to zero:
CALL ZERO(N)
These routines are available for the PDP11.

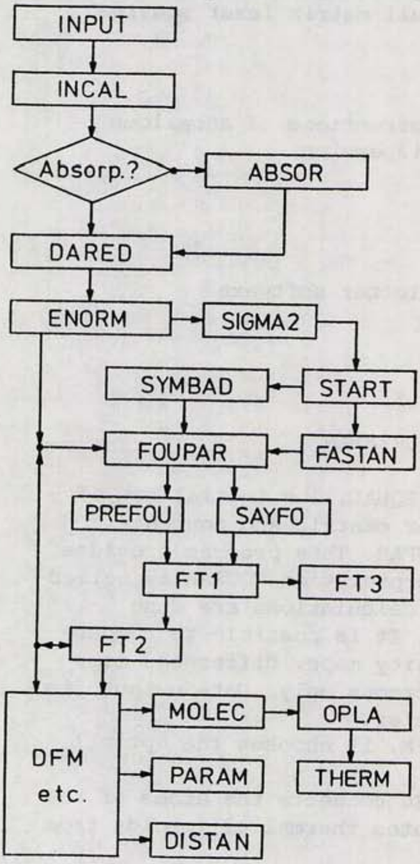


Fig. 2 Flow chart of CRYSTAN

Driving a minicomputer the crystallographer often will be the only user. So is possible to make optimal use of dialogue technique. This implies a very high comfort in running the system.

3. Special properties of the system

The efficiency of a crystallographic program does not only depend on the physical speed of the computer and the ingenious programming but also on the number of failures the crystallographer has to endure before he gets a job running. Some efforts of CRYSTAN to reduce this trouble will be discussed on the following pages.

(a) Derivation of symmetry operations

In the CRYSTAN system the symmetry operations are derived directly from the space group symbol. So it is allowed to use settings of space groups that are not tabulated in the International Tables, for example P 2_1/a instead of P 2_1/c; the only information that must be given is the shortened or full space group symbol. A space group symbol contains full information on a set of generators of the group [1]. Symmetry operations may be represented by pairs of matrices (R,T). R is called the rotational part, it is identical with the corresponding R of the related point group. T is called the translational part and can be divided into a glide component T_g and a site component T_s and $T=T_g+T_s$.

T_g determines the glide of a point parallel to the symmetry element related to the operation. It distinguishes between rotations and screws or mirror planes and different kinds of glide planes. T_g is derived by projecting T into the set of points that is invariant under the motion described by R. For each symmetry element of the space group symbol the glide component T_g is indicated, so T_g may be derived easily.

The component T_s depends on the site of the corresponding symmetry element with respect to the origin. As the origin is not specified by the space group symbol it is only possible to derive the relations among the translational parts T resulting from the product relations of the symmetry operations:

$$(R_1,T_1)\cdot(R_2,T_2) = (R_1\cdot R_2,\ R_1\cdot T_2+T_1) = (R_3,T_3)$$

This relation connects site and glide components of different symmetry operations, so that restrictions enter. The choice of origin, however, is still arbitrary. If a suitable origin is chosen the values for the site components become simple fractions of unity. For all centrosymmetric space groups it is most convenient to use a centre of symmetry for the origin. In other cases suitable origins are determined by points with high point symmetry, these rules, however, do not cover the choice of origins used in the International Tables and so in some cases special investigations are necessary to get a tabulated origin.

To demonstrate the procedure an example will be given. Space groups belonging to the point group 2/m may be generated by the twofold rotation or screw axis and the mirror or glide plane. The centre of symmetry is always the product of the two:

$$(R_{\bar{1}},T_{\bar{1}})=(R_2,T_2)\cdot(R_m,T_m)=(R_2\cdot R_m,R_2T_m+T_2)$$

$$R_{\bar{1}}=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{pmatrix},R_2=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&-1\end{pmatrix},R_m=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix},T_{\bar{1}}=\begin{pmatrix}t_{\bar{1}}^1\\t_{\bar{1}}^2\\t_{\bar{1}}^3\end{pmatrix}\quad T_2=\begin{pmatrix}t_2^1\\t_2^2\\t_2^3\end{pmatrix}\quad T_m=\begin{pmatrix}t_m^1\\t_m^2\\t_m^3\end{pmatrix}$$

The set of points invariant under the operation $\bar{1}$ consists of the origin only. So the projection of the translational part comprises the origin (000), that means a centre of symmetry has always only site components, $T = T_S$. If the centre of symmetry lies in the origin, T_S equals (0,0,0).
The set of points invariant under 2 is given by the line $\{(0,y,0)\}$, so the projection of T_2 into the y-axis is $(0,t_2^2,0)$. The set of points invariant under m is given by the plane $\{(x,0,z)\}$, so the projection of T_m into this plane is given by $(t_m^1,0,t_m^3)$. The product relation among the operations demands

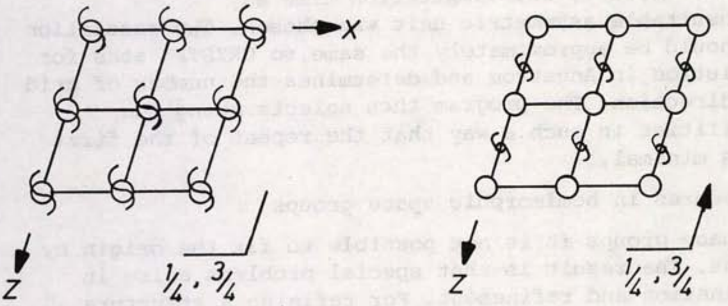
$T_{\bar{1}}=R_2T_m+T_2$ and we get $t_2^1 = t_m^1, t_m^2 = -t_2^2,\ t_2^3 = t_m^3$.

The independent component t_2^2 equals 0 for a rotation axis 2 or 1/2 for a screw axis 2_1. In the same way we get from the space group symbol the values 0 or 1/2 for t_m^1 and t_m^3 depending on the letters m, n, a or c in the space group symbol. Space groups with a nonprimitive Bravais lattice may be handled like such with a P-lattice. In these cases it may occur that a shift of origin is necessary. If S is the vector for the shift of origin, the new translational part T is determined by $T = T + (E-R)\cdot S$.

(b) Choice of asymmetric units for Fourier calculation

CRYSTAN includes the facility that an asymmetric unit for Fourier calculations is chosen without the user's intervention. The Fourier expansion employed favours orthorhombic or lower symmetry, higher symmetry is handled in the structure factors. In these cases the choice of the asymmetric unit is restricted to the orthorhombic or monoclinic subgroup. Thus in high symmetrical space groups the part of the unit cell that will be computed is too large. This part is rather a suitable piece of the cell that can be computed easily than an asymmetric unit. The handling of hexagonal space groups is only preliminary in the present version. A better solution is in preparation.
There are three conditions that give restrictions to the choice of asymmetric units. (i) The volume of the asymmetric unit equals to the quotient of the unit cell and the number of equipoints in the general position. (ii) This fraction of the unit cell cannot be chosen arbitrarily. The type of symmetry elements and their site relative

to each other may enforce or allow reductions of the unit cell. If there is a mirror plane normal to a given direction the volume must be halved in this direction. The possible reduction factor is determined by the sequence of the symmetry elements, it equals to the smallest fraction of a translation vector in x-, y- or z-direction, that is necessary to go from one symmetry element to the next one.
(iii) The distribution of symmetry elements may enforce different origins for the unit cell and the asymmetric unit. Centres of symmetry, rotation axes and mirror planes must not intersect the selected volume, they are only allowed to lie on the surface.
To demonstrate these considerations we regard the possible asymmetric units of the monoclinic space groups P 2_1/m and P 2_1/c. Fig. 3 shows the distribution of the symmetry elements



$\underline{x}$ y $\underline{z}$	0	0	0	1	1	1	0	0	0	1	1	1
$\underline{x}$ $\underline{y}$ $\underline{z}$	0	1/2	0	1/2	1/2	1/2	0	1/2	1/2	1/2	1/2	1/4
x $\underline{y}$ z	0	0	0	1/2	1/2	1/2	0	0	0	1/2	1/2	1/2
x $\underline{y}$ z	0	1/2	0	1	1/4	1	0	1/2	1/2	1	1/4	1/2
min.				1/2	1/4	1/2				1/2	1/4	1/4
	P 2_1/m						P 2_1/c					

Fig.3. Site components are surrounded by a line

and the general positions for both space groups. The volume of the asymmetric unit is a quarter of that of the unit cell. In P 2_1/m we get the reduction factor 1/2 for the x-and z-direction, for the y-direction the factor must be at least 1/2, it can be 1/4. Hence we find three possibilities:

$$1/4 = 1 \times 1/4 \times 1 = 1/2 \times 1/2 \times 1 = 1 \times 1/2 \times 1/2.$$

For the last two possibilities the origin must be shifted by a quarter in y-direction.
In P 2_1/c we get the possible reduction factor 1/2 for the x-direction and 1/4 for y and z; hence we find five possibilities:

$$1/4=1x1x1/4=1x1/4x1=1x1/2x1/2=1/2x1x1/2=1/2x1/2x1$$

These considerations may be comprehended to a formal procedure: Set all zero components of the translational parts to one, each site component must be halved; take the minimum of all components refering to the same direction.
The modified translational components are tabulated in fig. 3 in column 3 and 5. In most cases there are different possibilities to choose an asymmetric unit; so we have a chance to optimize the Fourier computation. The Fourier grid used allows a resolution of maximal 240 grid points per unity in each direction. The arrays for the first and second summation, however, only extend to a resolution of 1/30. Hence if a higher resolution is demanded the asymmetric unit has to be splitted up and the first two summations must be repeated for each part. The management of this splitting is done by the program without user's interference, but computation time may be wasted if a unsuitable asymmetric unit was chosen. The resolution for x, y and z should be approximately the same, so CRYSTAN asks for the minimal resolution in Angstrom and determines the number of grid points for each direction. The program then selects among the different possibilities in such a way that the repeat of the first two summations is minimal.

(c) Special procedures in hemimorphic space groups

In hemimorphic space groups it is not possible to fix the origin by symmetry relations. The result is that special problems arise in structure determination and refinement. For refining a structure in such a space group it is necessary to fix it with respect to the hemimorphic directions. The structure will be stabilized if the coordinates of one atom refering to these directions will not be varied in the least squares procedure. CRYSTAN recognizes a hemimorphic problem and asks only for the name of an atom to fix the structure. Another problem arises in structure determination by direct methods. In those space groups the origin must be fixed by choosing an arbitrary phase for reflexions properly related to the hemimorphic directions. This selection will be done by CRYSTAN automatically. It should be emphasized that in the tangent formula the selected phases are not allowed to be varied. We have evidence that a rather large set of nearly correct phases may diverge if this condition is violated.

(d) Handling of special positions

In least squares procedures it is customary to use key numbers designating free, coupled or fixed parameters. In the CRYSTAN system these key numbers are derived by the program.
The parameters of an atom are its coordinates x,y,z and the components of the anisotropic temperature factor $\beta_{11}, \beta_{22}, \beta_{33}, \beta_{23}, \beta_{13}, \beta_{12}$. If the atom lies on a special position then it is invariant under a certain set of symmetry operations of the space group:

$$X_a \equiv R\cdot X_a + T \bmod (G) \text{ and } B = R^t\cdot B\cdot R,$$

$X_a = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $B = \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{12} & \beta_{22} & \beta_{23} \\ \beta_{13} & \beta_{23} & \beta_{33} \end{pmatrix}$ and G a lattice vector.

The result of this relation is:

(i) coordinates can be fixed, they may adopt special values like 0, 1/8, 1/4, 1/2 etc.;
(ii) coordinates can be coupled by linear functions like y=x, y=2x, y=1/2x etc.;
(iii) non-diagonal β_{ik} may be zero or coupled;
(iv) diagonal β_{ii} may be coupled only.

Reading the parameters of an atom in the least squares procedures CRYSTAN examines the parameters by the aid of all symmetry operations. The result of this inspection is fixed in the key numbers, one for each parameter. If there are no restrictions the key number equals to the arrangement number, that means for x it is 1, for y it is 2 etc. If a parameter is fixed the key number is 0. If a parameter depends on another with a lower arrangement number, this lower number is used. The factors 2 and 1/2 instead of 1 in the linear functions are indicated by adding or subtracting 10 or 20 to the key number.

4. Direct methods in the CRYSTAN system

Because there are some special features in the application of direct methods in the CRYSTAN system detailed information on some programs of the structure determination part will be given.
Structure determination is initiated by the program ENORM that provides for normalized structure factors, symmetry tests, scale factor and overall B. Data files needed by the following programs are prepared. For the computation of E-maps 30 % of all reflexions are stored in E2.DAT. This set is reduced to the file E1.DAT by selecting the strongest reflexions, seven for each atom in the asymmetric unit. The method employed for normalisation is proposed by J.KARLE [2]. ENORM investigates the distribution of $\langle E^2 \rangle$ in the reciprocal space with respect to the directions (100), (010), (001), (110), ($\bar{1}$10), (011), (01$\bar{1}$), (101) and (10$\bar{1}$) by averaging all reflexions that are inside a narrow cone surrounding the direction lines. It is our experience that great differences in the averages predict difficulties during structure determination.
SIGMA2 sets up a Σ_2-list for each reflexion of E2.DAT on the base of E1.DAT. The program START first prepares a start set of reflexions. Approximately one half of the reflexions of E1.DAT is selected. They are distributed as uniformly as possible on the different parity classes and connected as strongly as possible with respect to their

triplet relations. From these a set of triplets is selected that holds the following conditions:

(i) They should be as reliable as possible.
(ii) The phases of all reflexions of the reduced E1.DAT should be determinable by at least n contributions and a sufficiently large α_{est} on the base of the reflexions involved in these relations. The minimal number of contributors n and α_{est} can be controlled by the user.
(iii) The number of triplets should be minimal.

All reflexions of these triplet relations belong to the start set. They are subdivided into dependent and independent reflexions because always three are combined to a triplet; the phase of one may be expressed by the other two and the structure invariant of the triplet. The least important is always chosen to be dependent. Different solutions are determined by the different phase values of the independent reflexions.

All possible values of phases are allocated to each independent reflexion. If it is a general reflexion in a noncentric structure the values are restricted to $\pm\pi/4$ and $\pm 3\pi/4$. Phase restrictions that follow from the choice of origin and the enantiomorphic form are taken into account by the program. Hence all different solutions are determined by different combinations of possible phase values.

This multiple solution technique allows an optimal adjustment to difficulties in structure determination. If a structure is simple it may be solved correctly by using only two contributers and a low α_{est}. The number of solutions will be small because only a small number of triplet relations and hence only few independent reflexions will be necessary. In complicated cases the user may require more contributers, say three or four, and a higher α_{est}. Consequently the number of triplets will be larger and so the number of independent reflexions and the number of solutions will increase. More security has to be paid with more solutions.

If the correct solution cannot be achieved even if a fairly high degree of security was used the reason should be that the values of the structure invariants for the selected triplets do not equal zero. In this case FASTAN provides for the possibility of changing the structure invariant in the fast tangent formula process. All possible solutions are tested by FASTAN. Equivalent solutions are detected by EQUAL. Those solutions that are suggested to be correct can be handled in SAYFO extending the phases to the reflexions of E1.DAT and then to those of E2.DAT. Proceeding in this way the user is aware of the fact that he is walking on thin ice at the beginning and he can choose the degree of security for the second part. Dangerous steps are marked and hence under control. So this procedure may be called the method of the calculated risk.

References: [1] H.Burzlaff,H.Zimmermann,Symmetrielehre,Stuttgart 1977
[2] J.Karle,International Tables IV, 350 ff.
[3] H.Burzlaff,R.Böhme,M.Gomm,CRYSTAN, Erlangen 1977

MINIMUM COMPUTER SYSTEM FOR COMPLETE CRYSTAL STRUCTURE DETERMINATION

Allen C. Larson* and E.J. Gabe**

Summary

A system of computer programs is described which enables to execute all stages of a crystal structure determination by means of a minimum computer configuration. The system includes programs for diffractometer control, absorption correction, datareduction, file editing, fourier mapping, application of direct methods, least squares refinement and preparation of tables. The system is reasonably well tested and compared with the speed of the intensity measurements pretty fast.

Introduction

The system which will be described was developed at the National Research Council of Canada (1). The initial reason for developing the system was the apparent high cost of computing at NRCC when using the central computing facility in Ottawa. It was felt that for a fairly reasonable increase in investment it should be possible to do most of the calculations needed for crystal structure determination and refinement. The computer to be upgraded was the Digital Equipment Corporation PDP-8e which was being used to control the Picker diffractometer.

The system that has been developed is what is felt to be essentially the minimum needed to carry out the entire process of crystal structure solution and refinement. The hardware requirements for this system are a PDP-8e (or 8a) with 32K core, a 1.6M word RK05 disc, a DECwriter (or equivalent) and a second file storage device (DECtape, magnetic tape, or a second RK05). A highly desirable optional extra is the floating point processor, FPP-8a, which will decrease computing times by an approximate factor of twenty. Other optional equipment includes a line printer and a visual display unit such as the Tektronix 4010-1.

There are several limitations imposed upon a crystal structure software package by the PDP-8 hardware and the OS/8 operating system (2). The FORTRAN IV language is a full ANSI set with a few special features such as random access I/O using files on file structured devices such as the disc or DECtape. The 12 bit word

* P.O. Box 5898, Santa Fe, NM 87502 USA
** Div. of Chem., National Research Council of Canada, Ottawa, Canada K1A OR9

of the PDP-8 causes some minor problems in that it limits array size to a maximum of 4096 floating point words. Incidently on the PDP-8, FORTRAN IV defines all data words with the same format. Data words require three PDP-8 words, one word for the characteristic and two words for the mantissa. The array size limitation, however created problems only in the least squares routine where it limits full matrix refinements to ninety parameters or less.

Outline of the system

The crystal structure package consists of thirteen programs which access two main data files and nine subsidiary files. A schematic of the programs and their usage of the data files is given in Figure 1.

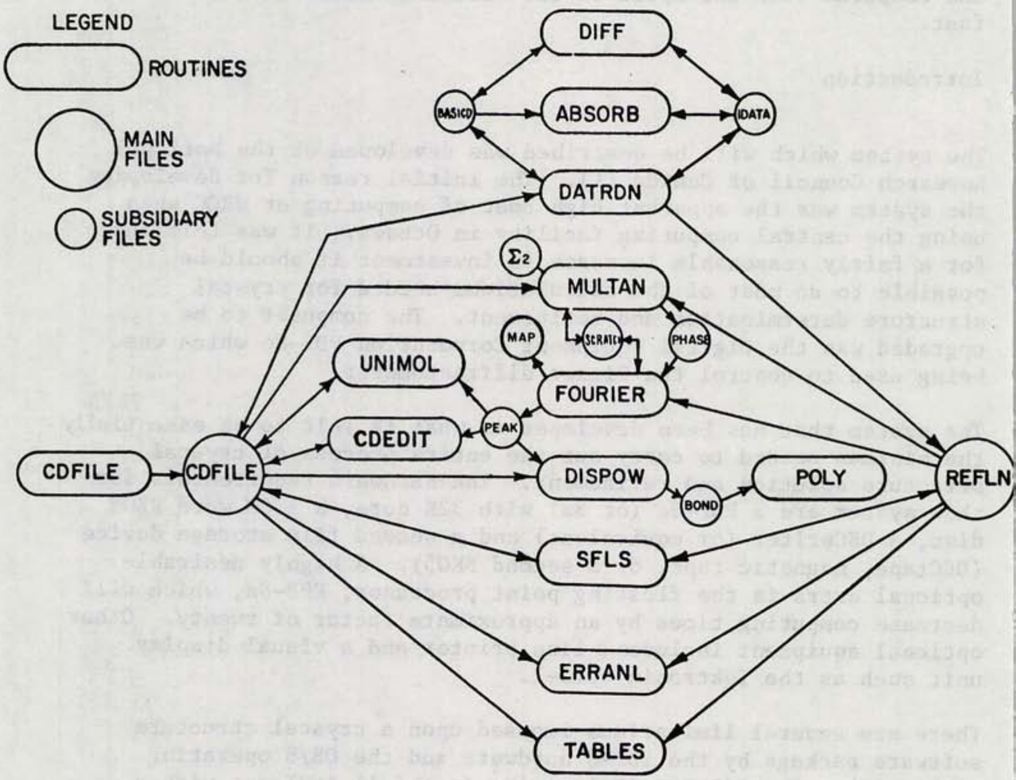


Figure 1. Overall Routine Linkage & File Usage

The CDFILE contains unit cell data including lattice constants and their sigmas, space group information including the space group symbol and the matrices which perform the operations of the group,

unit cell content information, scattering factors and the atomic parameters. The REFLN file contains all of the reflection information required by the system programs. A detailed description of each program in Figure 1 will be given later.
All data is entered into the various programs in response to prompts at the terminal. Every effort has been made to make these prompts as informative as possible without being too verbose. Most control options are requested so as to require a yes/no answer.

The system handles all space groups. This includes all non-standard setting of monoclinic, orthorhombic and tetragonal lattices. The operations of the group are determined by examination of the space group symbol and all improper and/or incomplete symbols are rejected. The constraints upon least squares parameters are determined by the programs and the user can add to them, but he cannot remove any of them.

MULTAN

First, let us examine some of the problems encountered in implementing MULTAN, a well known and widely distributed large machine program, on the PDP-8. Figure 2 shows the current memory layout for MULTAN. The program has been broken into three overlay levels, 0, 1, and 2. The system takes the first 10000_8 words, level 0 the next 53000_8 , level-1 the next 5000_8 words and level-2 the next 3400_8 words. An execution time file definition routine is used to open and close data files and it uses 400_8 words each for a device driver and for buffer space or 1000_8 words per file at the top of memory starting with UNIT 9 file space. Four data files are needed by MULTAN at one time and we are left with 400 words of storage for load time file definition buffer space. This may be used either for a batch input file or for a spooled output file. Teletype input/output and line printer output do not use high core buffer space.

The level-0 overlay contains the main program, all library subroutines, all 8-mode subroutines and COMMON storage. Of the 53000_8, 22016_{10}, words in this level 44751_8, 18921_{10}, are in the COMMON storage. This amounts to only 6307 data words. Since for N E-values, MULCVG(CONVEG) requires 11*N words of data storage in COMMON, excluding the sigma-2 table, it is apparent that the sigma-2 table must reside on the disc. A direct access file is used, with blocks defined as single buffer loads (256 PDP-8 words or 85 data words) allowing up to four blocks for each reflection's sigma-2 table. Because of the need to place the sigma-2 table on disc, it appeared that using MULTAN's overall E3MIN limit would be difficult to implement and the time to generate the triply redundant sigma-2 set from the set of unique sigma-2's MULTAN normally generates would be excessive. Therefore the sigma-2 generation routine was changed to generate the full 3-fold redundant set and since each reflection was limited to no more than

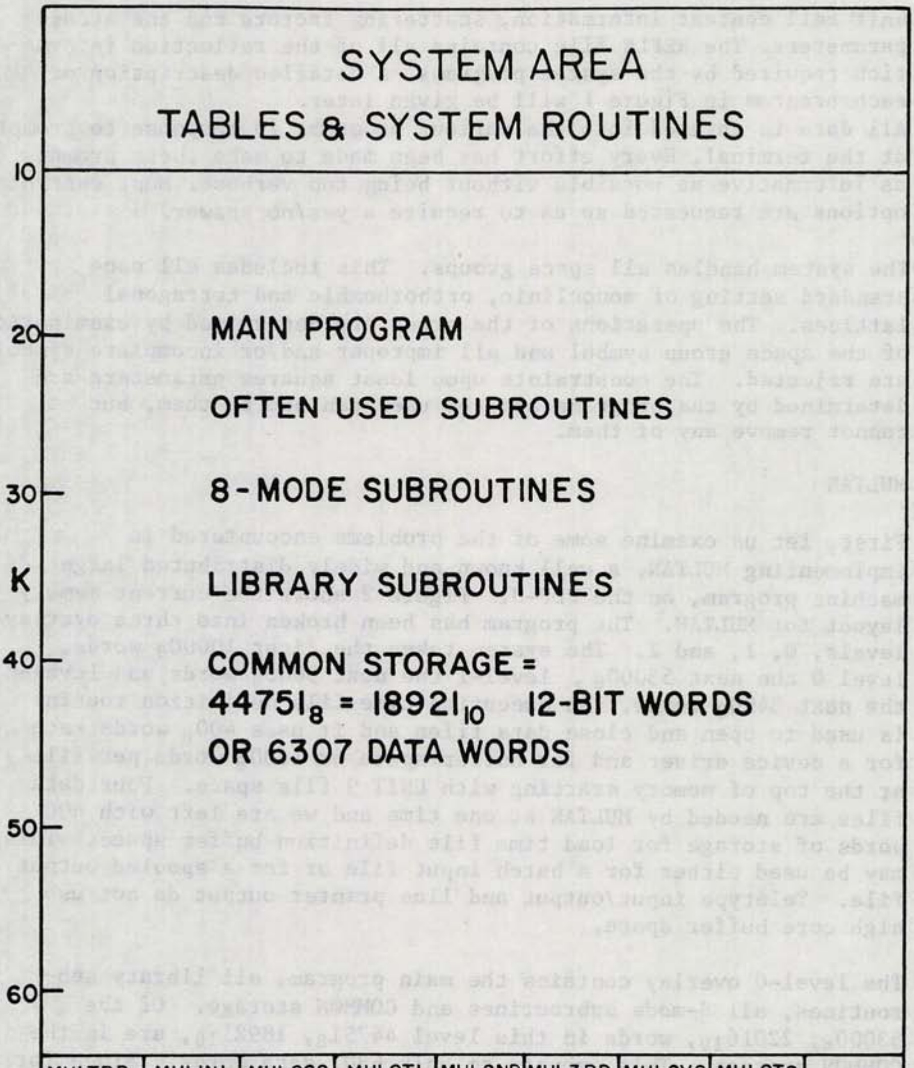


Figure 2. MULTAN Memory Layout

84 sigma-2's to determine an E3MIN for each reflection. This set of E3MIN values is required to avoid the premature removal of reflections with large numbers of sigma-2's during the converge process. It might be noted that this implements one of the improvements suggested by Lessinger (3) in his recent paper on MULTAN failures.

It was found that the routine FASTAN was too large for its proposed place in the overlay scheme and it was also observed that it could be easily split into two pieces and hence MULOT3 came into being. The input portion of FIRST was replaced with MULTRD and the control portion placed in the main program. MULTRD also took over the data reading for SECOND and THIRD since for a long running program with interactive input it is best to try to get as much of the input into the program at the start as possible. This routine has grown continuously as the prompts have been improved for interactive data input. It is now the largest routine in level-1 and calls MULFLS in level-2 which requests the names of the necessary files and opens them.

In the search for ways to reduce the total code in MULTAN and hence maximize the available storage we noted that the simple functions MOD, ABS and MIN which are normally inline functions in the FORTRAN generated code were externals and the code for each was moderately large (40 to 100 words). Since these functions are easily written explicitly into the FORTRAN, this was done and the routines were eliminated. It was a pleasant surprise to find that the explicit code for these functions was several words shorter than the calling sequence for the external functions and that the code ran faster with the explicit routines.

The program was very extensively tested during and after its implementation on the PDP-8. This was done because much of the data packing in MULTAN was changed. In the course of this testing several bugs in MULTAN were located and removed. Also it was found that for data sets ordered on <u>hkl</u>, the routine ORIGIN can at times take an excessive amount of time. To eliminate or at least reduce this it was decided to order the data on the magnitude of E with the largest E's first.

The program as it is presently set up has some limitations. Reflection indices are limited to magnitudes of less than 100, this limitation applies to the entire system. MULTAN can handle up to 500 reflections, each with up to 84 sigma-2 relationships. The maximum number of sigma-2's, 42987, is determined by the limit on the size of a file under the OS/8 system, 2047 blocks. Run time can become fairly long for large problems. On a difficult problem with 500 E-values generating about 18000 sigma-2's about 40 minutes were required for each of parts 1 and 2 and about 20 minutes per FASTAN solution. Since 64 solutions were computed this took nearly 24 hours of computer time. For a smaller problem with

340 E-values the time was about 40 minutes for parts 1 and 2 and about 14 minutes per FASTAN solution. As a result, an interrupt capability has been put into MULTAN so that these long running tasks can be run overnight for as many nights as are needed to complete the run. The 24 hour run was completed in two nights.

Now let us look at the entire system and its attributes. All data is entered into the various programs in response to prompts at the terminal. Every effort has been made to make these prompts as informative as possible without being too verbose. Most control options are requested so as to require a yes/no answer.

The other programs of the system

The program DIFF drives the NRCC variant of the Picker diffractometer to measure the lattice constants and collect diffraction data. It reads and stores the crystal orientation and unit cell information on the BASICD file and writes the measured intensity data to the IDATA file. The routine can collect data with 2θ and w scans and the results can be treated with a profile analysis routine. There is also a mode in which the measurement precision can be controlled both for scanning and peak top measurements. Using profile analysis with 2θ scanning typically 1200 reflections/day can be collected.

The program ABSORB can then be used to modify the reflection intensity data for absorption using the Gaussian integration method. Up to 3200 integration points may be used and the crystal may be described by up to 24 faces. The crystal orientation data are read from the BASICD file. The program can draw a stereo picture of the crystal on the Tektronix 4010-1 terminal to enable the user to check the representation of the crystal.

The program CDFILE is used to create the cell data file (CDFILE) for a compound. This program requests the name of the file to which you wish to write the information. It then opens this file and requests a data set identification record and the space group symbol. After determining the operations of the group and the Laue symmetry the program requests the necessary lattice constants and their sigmas. It then computes the cell volume and the reciprocal cell parameters and requests the radiation used and data concerning the contents of the unit cell (atomic symbols and number of each in the unit cell). With this information the program locates the scattering factors (5), dispersion terms (6), atomic absorption coefficients (7,8) and atomic weights (9) from internal tables and computes the linear absorption coefficient and density. It then proceeds to zero 100 atom record blocks.

The data reduction program, DATRDN, processes the intensity data file using space group information from the CDFILE to create the REFLN file. It can average equivalent reflections from multiple data sets and it processes the Friedel pairs separately placing them on the same record on the file. Statistics for the K(s)-curve are accumulated and E-values are derived from the F(<u>hkl</u>) data.

CDEDIT is the program used to modify the cell data file. Changes can be made in the settings of the overall refinement flags for such things as extinction and weighting scheme as well as adding, editing, and deleting atom records. Atomic parameters may be read from the terminal, from the PEAK file output by FOURIER, or they may be generated as hydrogen atom positions in tetrahedral or trigonal positions about carbon atoms. The program determines the site symmetry for each atom position and sets up the constraints due to site symmetry for the least squares program. The routine is highly interactive and prompts the user about the necessary commands. In addition the routine does its own decoding of numerical input and will not drop back to the OS-8 monitor if a mistake is made while typing a number.

The program FOURIER calculates all of the usual Fourier maps used in crystallography including the three usual Patterson vector maps, normal, sharpened, and origin removed, observed and calculated electron density maps, difference electron density maps and E-maps. Dispersion terms are properly accounted for in all electron density maps (1,10). The program checks the region of the unit cell given and if it is not a complete asymmetric unit prints a warning and then gives the user a chance to try again if he so desires. The peak picking section of the program eliminates both peaks off the edge of the map and duplicates of the same peak. The position of the peak is determined from a Gaussian fit of 27 points centered at the peak maximum (11).

The program DISPOW is the distance and angle computing program in the system. This routine is also the main editing program for the Metals and Alloys Structure Data file. Among its several options is the capability to compute powder patterns with d-spacings and intensities in several forms. The program can be used to analyze the nearest neighbor polyhedron to give the Frank and Kasper polyhedron (12) as redefined by Cromer, Larson and Roof (13). In the computation of the standard deviations the unit cell variances are included, as well as the correlation terms for parameters on the same atom, but those between atoms are neglected.

The program POLY uses the bond file from DISPOW to draw the coordination polyhedra on the Tektronix 4010-1. It has proven to be extremely useful in the study of alloy and inorganic structures.

The structure factor least squares program, LSTSQ, can handle up to 100 atoms. For problems with 90 parameters or less the user has the option of using either the block diagonal method or the full matrix method for the least squares. The refinement parameters are defined by flags set in the CDFILE by CDEDIT. One may refine an isotropic extinction parameter in addition to the usual set of positional and thermal (either isotropic or anisotropic) parameters. Dispersion terms are explicitly included in the structure factors calculations and sufficient data is written to the REFLN file that FOURIER can compute observed electron density maps corrected for both dispersion and extinction.

The program ERRANL is used to carry out weight and agreement analyses as functions of F(obs), F(obs)^2, sin θ/λ, $\underline{h}$, $\underline{k}$, and $\underline{l}$. These analyses can at times be of great help in understanding the cause of refinement problems. They are also a very good indicator of the correctness of a structure.

The program UNIMOL reads the CDFILE and generates a new atom position set in which all atoms belonging to a molecule are relocated to form a bonded set. This new atom position set may be written to the CDFILE if desired. When this is done the original atom order may be preserved or the atoms may be written in the order in which the bond list is printed. This ordering starts with the atom on the site of lowest multiplicity. Among atoms on sites of equal multiplicity, the atom with the largest number of bonds is chosen, and among atoms with equal numbers of bonds the one with the longest shortest bond is selected as the molecular origin. Having chosen a starting point the program follows the chain of shortest bonds to its end and then backtracks to the first unused bond each time a chain is exhausted.

The last routine in the package is TABLES. This program is used in the final stage of a crystal structure analysis, namely in preparing the report on the structure. It generates tables of crystal data and reflection data for publication.

Comments

Some indication of times, with hardware floating point processor, is shown in the following example. $P2_1/c$ structure with 67 atoms and 2000 reflections.

MULTAN with 340 E values
40 mins for parts 1 & 2; 14 mins/solution

Fourier with 23000 x,y,z points 30 mins

Least squares with 418 parameters 55 mins/cycle.

The system is thought to be reasonably well debugged, having been used for the solution and refinement of about 20 metallic, inorganic and organic structures in a wide variety of space groups. A recent 34 atom structure was solved and refined in one 8-hour working day.

References

(1) Gabe, E.J. (1976) *Crystallographic Computing Techniques*, pp.468-476. F.R. Ahmed, Editor. MUNKSGARRD, Copenhagen.
(2) *OS/8 Handbook* (1974) Digital Equipment Corp. Maynard, Mass, U.S.A.
(3) Lessinger, L. (1976) *Acta Cryst*. A32, 538-550.
(4) Larson, A.C. (1969) *Acta Cryst*. A25, S1.
(5) Cromer, D.T. and Waber, J.T. (1974) *International Tables for X-ray Diffraction* Vol. IV, Table 2.2B, pages 99-101.
(6) Cromer, D.T. (1974) *International Tables for X-ray Diffraction* Vol. IV, Table 2.3.1, pages 149-150.
(7) Koch, B. and MacGillary, C.H. (1962) *International Tables for X-ray Diffraction* Vol. III, Table 3.2.2A, pages 162-165.
(8) Hubbell, J.H., McMaster, W.H., Del Grande, N.K., and Mallett, J.H. (1974) *International Tables for X-ray Diffraction* Vol. IV, Table 2.1C, pages 61-66.
(9) Lonsdale, K. and Rieck, G.D. (1962) *International Tables for X-ray Diffraction* Vol. III, Table 2.1.4, pages 46-56.
(10) Larson, A.C. (1976) *Crystallographic Computing Techniques*, F.R. Ahmed, Editor. Pages 396-398. MUNKSGARRD, Copenhagen.
(11) Rollett, J.S. (1965) *Computing Methods in Crystallography* Oxford: Pergamon Press.
(12) Frank, F.C. and Kasper, K.S. (1958) *Acta Cryst*. 11, 184-190.
(13) Cromer, D.T., Larson, A.C., and Roof, R.B. (1964) *Acta Cryst. 17*, 272-276.

CHAPTER 3

Program systems for direct methods

RECENT DEVELOPMENTS IN MULTAN

Peter Main [*]

Summary

A number of recent improvements in Multan have both increased its power and usefulness. Magic integer phase permutation has replaced quadrant permutation to generate values for the unknown phases of the reflexions used to begin the phase determination. This allows more reflexions to be put into the starting set thus increasing the power of the multisolution method. A new tangent formula weighting scheme has been devised which takes into account the lack of independence of the separate terms in the summation. The phases therefore correspond more closely to the theoretical distribution of the Σ_2 relationships. This is particularly useful when dealing with some problems caused by pseudo symmetry and also in recycling partial structures. E-maps are interpreted automatically and the molecules found can be compared with the expected molecule. This will lead eventually to the automatic recycling of partial structures. Chemical information can be used to give improved estimates of the phase and variance of phase relationships. In particular, this allows misplaced molecular fragments found in an E-map to be used in a redetermination of the structure. The solution of structures containing heavy atoms is made easier by determining the heavy atom positions, then subtracting them statistically from the diffraction pattern. The difference (equal-atom) structure is then solved in the normal way. This avoids the problems caused by the introduction of additional symmetry by the heavy atoms alone.

1. Introduction to the multisolution method

Multan is a system of computer programmes which solves crystal structures automatically from X-ray diffraction data by direct determination of the structure factor phases. Phase determination is carried out using the tangent formula of Karle & Hauptman (1). The unknown starting phases are given a number of different values, thus producing several plausible sets of phases as first described by Germain & Woolfson (2). These phase sets are ranked in order by figures of merit and the E-map calculated using the "best" set of phases very often reveals the structure.

The reason for using explicit values rather than symbols for the unknown phases is that different indications for the same phase can always be combined. In symbolic addition, phase indications can only be combined when the symbolic part of the separate indications is the same. Different indications lead to relationships among the symbols

---

[*] Physics Department, University of York, York YO1 5DD, England

which are later used to evaluate the symbols. However, in the multisolution method such information can be used immediately, especially in the early critical stages of phase determination when few phases are known and single phase relationships may not always be reliable. The additional reliability in phase determination is paid for by extra computing time, because the multisolution method generates many more sets of phases than the symbolic addition method.

Another advantage of using explicit phase values is that because a number of different values are tried, it becomes possible to obtain a good value for a newly determined phase from inaccurate starting phases using unreliable phase relationships. This, of course, happens accidentally, but it is a common observation that the set of phases which reveal the structure do not always come from the most accurate starting point.

2. The Multan system

The operations carried out by Multan are shown below in the order in which they occur. Some of the newer or more important operations will be described in later sections of this paper. For equal-atom structures containing up to about 70 independent atoms, Multan is expected to reveal most of the structure in a single computer job. Larger structures can be solved, of course, but several attempts may have to be made.

Operation	Remarks
Calculate $\lvert E\rvert$'s from $\lvert F_{obs}\rvert$ values	Molecular geometry may be used to calculate spherically averaged molecular scattering factors (3). These lead to more accurate scale and temperature factors and improved estimates of expected intensity. Oriented atomic groups can also be used if the spherical average is inappropriate.
Set up Σ_2 phase relationships	Estimates of the phase and reliability of each relationship can be obtained from a knowledge of unit cell contents, molecular geometry, orientation of atomic groups and/or actual atomic positions (3).
Find the reflexions used to start phase determination	Origin-defining reflexions together with a small number of other reflexions are carefully chosen so that they lead to multiple indications of new phases as quickly as possible and make use of strong phase relationships. This is done by the convergence procedure of Germain, Main & Woolfson (4).
Phase determination	Values are assigned to unknown phases by magic integer phase permutation (5). A weighted form of the tangent formula is used (4) which is also modified according to Main (3) if chemical

Operation	Remarks
	knowledge has been incorporated into the phase relationships. The weighting scheme can also be modified according to Hull & Irwin (6) to obtain more accurate phases and to avoid some problems due to pseudo symmetry.
E-map	Maps are calculated in order of the figures of merit. Simple stereochemical criteria are applied to the peaks found in order to produce a chemically sensible molecule. This can be compared with the expected molecule and corresponding peaks and atoms identified (7).

3. Choosing the starting reflexions

The computer algorithm used to choose the origin-defining reflexions which form a good starting point for phase determination is the convergence procedure of Germain, Main & Woolfson (4). This algorithm is highly successful and will normally find a better set of starting reflexions than can be done by hand unless the structure has special difficulties. Since it is important, a flow diagram of the algorithm is shown in Figure 1.

The quantity α_{est} which appears in the flow diagram is given in the programme by

$$(\alpha_{est})_h = \sum_k K_{hk} \frac{I_1(K_{hk})}{I_o(K_{hk})} \qquad (1)$$

where $K_{hk} = 2\sigma_3\sigma_2^{-3/2} |E_{\bar{h}} E_{h-k} E_k|$

and $\sigma_n = \sum_j Z_j^n$

It can be calculated without knowledge of the phases and is a measure of how reliably each phase can be determined in terms of all the others remaining in the data set at any one time. The least reliably determined phase is eliminated at each stage, leaving to the end those reflexions which are strongly linked together and which will lead to reliable phase determination by giving multiple indications using strong relationships. Phases are determined later by the tangent formula in reverse order of elimination in the convergence procedure.

One of the reasons why the algorithm is successful is that it looks at the way reflexions are linked together through phase relationships and does not look at each reflexion or relationship in isolation. In terms of more recent theories, it may be said that the convergence technique takes account of the higher neighbourhoods of structure invariants (8, 9,10,11). For example, the second neighbourhood of a Σ_2 relationship

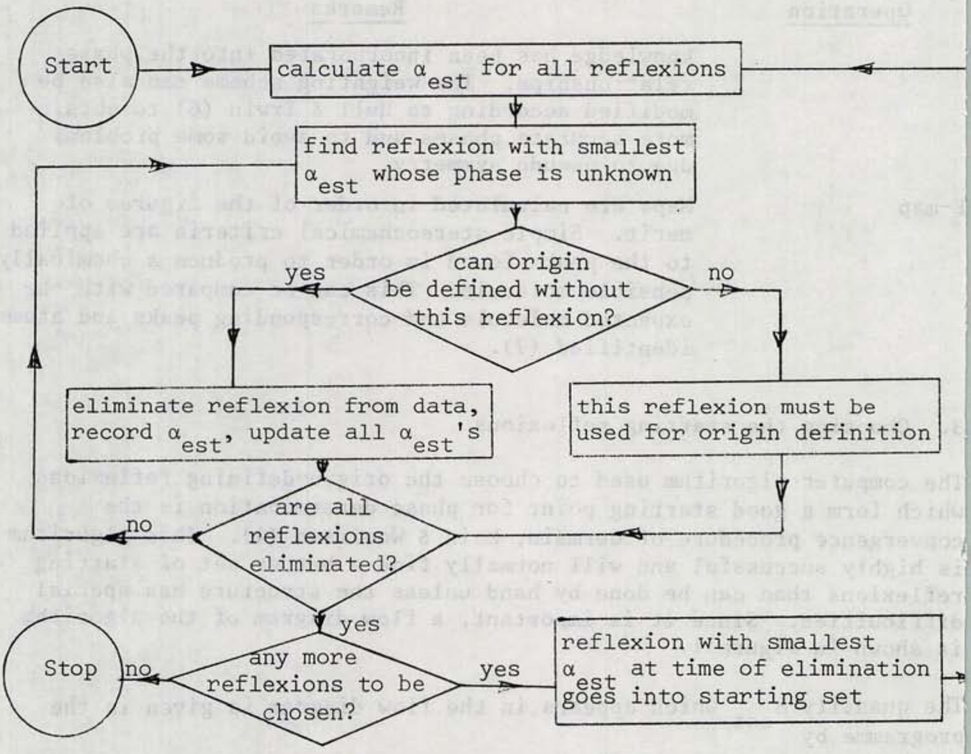


Figure 1. Flow diagram of the convergence procedure

contains the 10 reflexions with indices h, k, h+k, ℓ, h±ℓ, k±ℓ, h+k±ℓ where ℓ can have any value. Using these reflexions, the 13 relationships shown in Table 1 can be set up. Similarly, the second neighbourhood of a quartet contains the 7 reflexions h, k, ℓ, h+k+ℓ, h+k, k+ℓ and ℓ+h which are associated with the 6 Σ_2 relationships shown in Table 2. If all these reflexions are strong, both the Σ_2 relationship and the quartet will be particularly reliably estimated (Notice that they are highly interdependent.). In addition, the reflexions are clearly interlinked very strongly and will therefore tend to be eliminated only towards the end of the convergence procedure. It is expected, therefore, that not only the strongest relationships remain till the end of the convergence but also those that are particularly reliable. Although the effect is not strong, this is indeed observed practice. A deeper analysis of this effect may lead to modifications which give still better starting reflexions.

The actual number of reflexions required in the starting set can now be chosen by the programme. If any reflexion is eliminated with $\alpha_{est} = 0$, i.e. its phase cannot be determined from those reflexions remaining at that point, this represents a weak link in the phase determination process. It is removed by taking that reflexion into the starting set. The number of starting reflexions is therefore increased until all such weak links are removed or until the reflexions chosen give rise to a maximum number of phase sets.

Table 1 The 13 Σ_2 relationships from the 10 reflexions in the second neighbourhood of a single Σ_2 relationship

h	k	$\bar{h}-k$
h	$\pm\ell$	$\bar{h}\pm\ell$
h	$k\pm\ell$	$\bar{h}-k\pm\ell$
k	$\pm\ell$	$\bar{k}\pm\ell$
k	$\pm\ell+h$	$\bar{h}-k\pm\ell$
$h+k$	$\pm\ell$	$\bar{h}-k\pm\ell$
$h+k$	$k\pm\ell$	$\pm\ell-h$

Table 2 The 6 Σ_2 relationships from from the 7 reflexions in the second neighbourhood of the quartet h, k, ℓ, $\bar{h}-k-\ell$

h	k	$\bar{h}-k$
h	ℓ	$\bar{h}-\ell$
h	$k+\ell$	$\bar{h}-k-\ell$
k	ℓ	$\bar{k}-\ell$
k	$\ell+h$	$\bar{h}-k-\ell$
ℓ	$h+k$	$\bar{h}-k-\ell$

4. Magic integer phase permutation

The concept of magic integers was introduced by White & Woolfson (12). For a sequence of n integers, denoted by $m_1, m_2 \ldots m_n$, n phases are represented by the equations

$$\phi_i = m_i x \mod(1) \qquad (2)$$

where the phases, ϕ_i, are in cycles and the equations are approximately satisfied by some value of x in the range $0 \leqslant x < 1$. Main (13) gave the theory of magic integers and described a recipe for deriving a magic integer sequence $m_1 \ldots m_n$ of any length which minimises the r.m.s. error in the represented phases. Multan uses magic integers for assigning values to the unknown phases chosen by the convergence procedure to start phase determination (14). Only phases whose values are unrestricted by space group symmetry are represented by the magic integers, while restricted phases are given their two allowed values in the phase permutations.

To assign phase values, the variable x in equation (2) is given a series of values at equal intervals in the range $0 \leqslant x < 1$ cycles. A simple illustration of phase permutation is given in Figure 2. The 16 phase combinations produced by quadrant permutation on two unknown phases are plotted in Figure 2(a). If these phases are now represented by the magic integer sequence 2 3 and x takes on values at intervals of 1/12 cycles, the resulting 12 phase combinations are shown in Figure 2(b).

Clearly, the magic integer permutation results in a saving in the number of phase sets generated. There are two reasons for this. The first is that the 12 sets in Figure 2(b) cover the 2-dimensional phase space on a more closely packed lattice than the 16 sets in Figure 2(a). Phase space is therefore covered more efficiently by magic integer

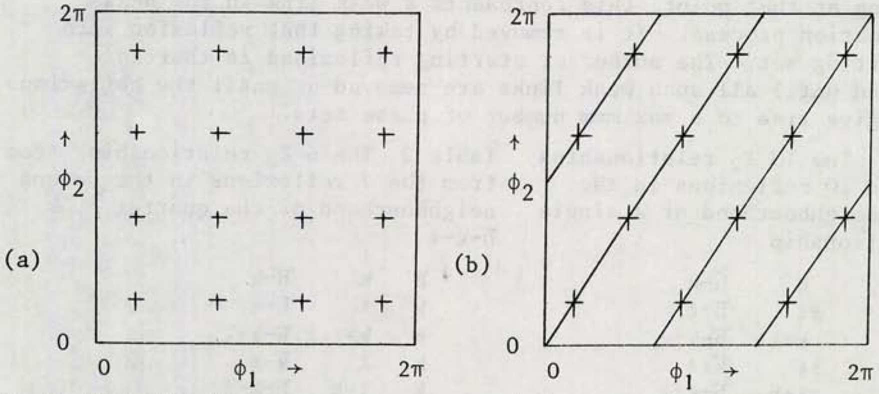


Figure 2(a) The 16 phase sets generated for two unknown phases by quadrant permutation. (b) The 12 phase sets generated for two unknown phases by magic integer permutation (magic integer sequence is 2 3)

permutation and this will be true for any magic integer sequence set up as described by Main (13). The second reason for the saving in the number of phase sets is that the r.m.s. error of the magic integer phase is slightly higher than the error resulting from quadrant permutation. For the example given, the r.m.s. error in the magic integer phases is 29.3^o which is more than the r.m.s. error of 26.0^o from quadrant permutation.

The reduction in the number of phase sets becomes very large as the number of permuted phases increases. This is shown in Table 3 where the number of phase sets produced by quadrant permutation is compared to the number currently produced by Multan using magic integers. The magic integer sequences used are based on the Fibonacci series [see (13)] 1 1 2 3 5 8 ... These numbers form the differences between adjacent numbers of the magic integer sequences which are shown in Table 4.

Table 3 Number of phase sets generated by quadrant permutation (QP) and magic integer permutation (MI) as a function of the number of unknown phases (n). Also shown is the r.m.s. error of the starting phases in degrees.

n	QP	ϕ_{rms}	MI	ϕ_{1b}
1	4	26.0	4	26.0
2	16	26.0	12	29.3
3	64	26.0	20	36.8
4	256	26.0	32	41.5
5	1024	26.0	50	44.7
6	4096	26.0	80	46.6
7	16384	26.0	128	48.0
8	65536	26.0	206	49.0
9	262144	26.0	332	49.8
10	1048576	26.0	536	50.5

Table 4 Magic integer sequences used in Multan for phase permutation

n										
1	1									
2	2	3								
3	3	4	5							
4	5	7	8	9						
5	8	11	13	14	15					
6	13	18	21	23	24	25				
7	21	29	34	37	39	40	41			
8	34	47	55	60	63	65	66	67		
9	55	76	89	97	102	105	107	108	109	
10	89	123	144	157	165	170	173	175	176	177

The interval of x used in equation (2) to produce the different phase sets is chosen to give a closely packed lattice in n-dimensional phase space as in Figure 2(b) for the 2-dimensional example. This interval is given by Main (14) as

$$\Delta x = \frac{2}{\sqrt{\pi}} \left[\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\sum_i m_i^2\right)^{n/2}} \right]^{1/n-1} \text{cycles} \qquad (3)$$

and this gives rise to the number of phase sets shown in Table 3.

For N sets of phases, the lower bound of the r.m.s. error of the starting phases is given (14) by

$$\Delta\phi_{1b} = 2\sqrt{\frac{\pi}{n+2}} \left[\frac{\Gamma\left(\frac{n+2}{2}\right)}{N} \right]^{1/n} \text{radians} \qquad (4)$$

and these quantities are also shown in Table 3. There is ample evidence that r.m.s. errors as high as these are easily tolerated by Multan. No structure has been found which was solved by quadrant permutation that could not be solved using magic integer permutation. Rather, the reverse is true. In addition, even though the starting phase sets are spaced more widely apart in phase space, duplicate phase sets still arise after tangent formula refinement. It is not yet known how high the errors of the starting phases can be before the technique fails to solve a structure that is otherwise soluble. However, this is being investigated because even a small increase in r.m.s. error results in a dramatic decrease in the number of phase sets. The main effect of the reduction in number of phase sets is to allow more reflexions in the starting set for the same computing time as before. Since these are very carefully chosen by the convergence procedure, this greatly increases the power of the multisolution method. This is probably the single most powerful improvement to Multan described in this paper.

Note that the introduction of one extra magic integer phase increases the number of phase sets by a factor of 1.6. This is cheaper than adding a restricted phase to the starting set since this will double the number of sets generated. The number of phase sets shown in Table 3 is modified in the usual way according to the requirements of origin and enantiomorph definition. The enantiomorph is defined by exploring only the appropriate half of the n-dimensional phase space (n = number of permuted phases). This means that all n phases are used simultaneously to give a strong enantiomorph definition.

5. Tangent formula weighting

Multan has always used a weighted form of the tangent formula, but recently an improved weighting scheme has been developed by Hull & Irwin (6). The weighted tangent formula may be written as

$$\phi_h = \text{phase of } \left\{ \sum_k w_k w_{h-k} E_k E_{h-k} \right\} \tag{5}$$

where w_h = weight associated with ϕ_h

The variance of ϕ_h may be calculated according to Karle & Karle (16) from

$$V_h = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \frac{I_1(\alpha_h)}{I_o(\alpha_h)} \tag{6}$$

$$\text{where } \alpha_h = 2\,\sigma_3\sigma_2^{-3/2}\,|E_h|\,\left|\sum_k w_k w_{h-k} E_k E_{h-k}\right|$$

When there are few contributors to the summation in equation (5), the variance of ϕ_h may be quite high so the phase must be used with caution when determining other phases. The correct weight to use is inversely proportional to the variance and, to a very good approximation, this is proportional to α_h. The weights used, therefore, are given by

$$w_h = \min(0.2\alpha_h,\ 1.0) \tag{7}$$

The constant of proportionality (0.2) is arbitrary, but it has been optimised by trials over a large number of structures. Multan applies no rejection criteria for doubtful phases; it depends upon such phases being given a very low weight.

Although this weighting scheme plays an important part during the first stages of phase determination, all the weights quickly become unity and so have no effect during phase refinement. The new Hull-Irwin scheme, however, is effective at all times. It attemps to match the α_h calculated from equation (6) with that expected from the probability distribution of the Σ_2 phase relationships. The expected α_h is given by Germain, Main & Woolfson (4) as

$$(\alpha_h^2)_E = \sum_k K_{hk}^{\;2} + \sum_{\substack{k \\ k\neq\ell}} \sum_{\ell} K_{hk} K_{h\ell} \frac{I_1(K_{hk})}{I_o(K_{hk})} \frac{I_1(K_{h\ell})}{I_o(K_{h\ell})} \qquad (8)$$

where K_{hk} is defined in equation (1) and the weight is calculated from

$$w_h = A \exp(-x^2) \int_o^x \exp(t^2)\, dt \qquad (9)$$

where $x = \alpha_h / (\alpha_h)_E$

and A is adjusted to make the maximum of w_h equal to unity.

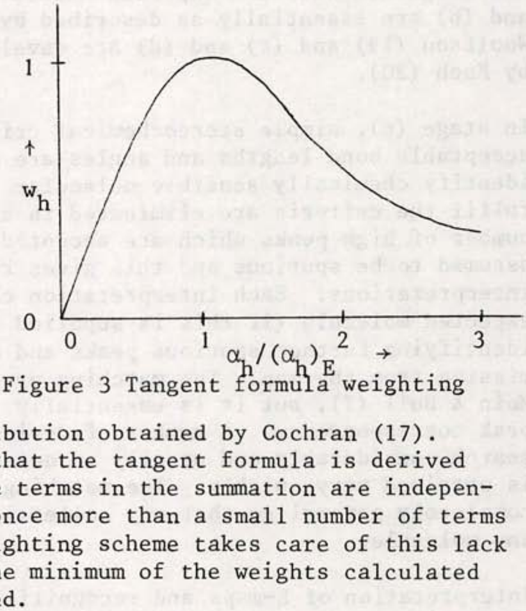


Figure 3 Tangent formula weighting

A plot of w_h as a function of $\alpha_h/(\alpha_h)_E$ is shown in Figure 3. It can be seen that if α_h is too large, $w_h < 1$. This will be the case if the phase relationships are too consistent and the weights are specifically designed to prevent this from happening. That is, rather than trying to force each individual Σ_2 relationship to have a phase of zero, the weighting scheme attempts to make the phases of the relationships conform to the theoretical distribution obtained by Cochran (17). Another way of looking at it is that the tangent formula is derived under the assumption that all the terms in the summation are independent. This is clearly not true once more than a small number of terms appear in the summation. The weighting scheme takes care of this lack of independence. In practise, the minimum of the weights calculated from equations (7) and (9) is used.

The Hull-Irwin weighting scheme is particularly useful in preventing the loss of enantiomorph in a polar space group. Unweighted tangent formula refinement of phases in $P2_1$, for example, can sometimes introduce a mirror plane through the molecule perpendicular to the 2_1 axis. This is now largely prevented without having to take special precautions. Another operation in which the new weighting scheme has proved its use is in the tangent formula recycling of partial structures (18). The weighting scheme gives more accurate phases and much more of the structure appears in the next map than is otherwise the case. At the time of writing, the use of the weighting scheme has lead to the solution of one structure that had previously resisted solution ($P2_12_12_1$, 66 independent equal atoms). However, for small structures and for those where the Σ_2 relationships are not too consistent, the new weighting scheme makes little difference to the final phases.

6. Interpretation of E-maps

Apart from the usual peak search routines, Multan now contains routines for interpreting the E-map peaks it finds in terms of a chemically sensible molecule. It can also compare the molecule it finds with that expected and corresponding atoms and peaks identified (7). The complete interpretation of the E-map is divided into four distinct stages: (a) peak search; (b) separation of peaks into potentially bonded clusters; (c) application of simple stereochemical criteria to identify possible molecular fragments; (d) comparison of the fragments with the expected molecular structure. The stages (a) and (b) are essentially as described by Declercq, Germain, Main & Woolfson (19) and (c) and (d) are developments of the work described by Koch (20).

In stage (c), simple stereochemical criteria of maximum and minimum acceptable bond lengths and angles are applied to the E-map peaks to identify chemically sensible molecular fragments. Peaks which do not fulfil the criteria are eliminated in such a way as to maximise the number of high peaks which are accepted. Different peaks may be assumed to be spurious and this gives rise to alternative interpretations. Each interpretation can be matched with the expected molecule (if this is supplied to the programme) thus identifying further spurious peaks and determining which atoms are missing from the map. The matching process is described in detail by Main & Hull (7), but it is essentially an exhaustive search for atom-peak correspondence. A number of devices are employed to speed up the search considerably and usually an acceptable atom-peak correspondence is obtained very quickly. The matching routines are efficient and completely general so that any collection of peaks can be compared to any molecule.

Interpretation of E-maps and recognition of the structures they contain is something a human being can do much better than a computer. However, the computer can now perform simple recognition tasks without human intervention. This makes it possible to use partial structural information in an automatic recycling process such as a weighted Fourier or tangent formula recycling. It is now feasible to design a programme system that will present the user with a complete structure in a single run rather than a structure with a small number of atoms still to be found by further calculations. Multan does not yet perform this automatic recycling, but it is hoped to develop it to the point where it will.

7. Use of chemical information

Chemical information about the structure can be used in Multan to help in the structure determination. Four types of information can be handled as described by Main (3): (a) unit cell contents; (b) molecular geometry; (c) orientation of atomic groups; (d) known atomic positions. Knowledge of unit cell contents is used in all direct methods and Multan is no different in this respect. However, the use made of other types of information is described in the following sub-sections.

Molecular geometry

The main use of molecular geometry is in the calculation of E's from F's. Normalised structure amplitudes (E's) may be defined by

$$|E_h|^2 = \frac{I_h}{\langle I \rangle} \tag{10}$$

where $\langle I \rangle$ = expected intensity.

The expected intensity can be obtained as the actual average intensity at the appropriate value of $\sin\theta/\lambda$ or can be estimated from a statistical distribution of the atoms. For example, if the atoms are assumed to be randomly distributed in the cell, Wilson (21) gives the expected intensity as

$$\langle I \rangle = \sum_i f_i^2 \tag{11}$$

which is modified by space group symmetry and the Debye-Waller factor in the usual way. If some molecular geometry is known, we may assume there are atomic groups in random positions and orientations in the cell. For each group separately, Debye (22) gives the group scattering factor as

$$g_i^2 = \sum_j \sum_k f_j f_k \frac{\sin t\, r_{jk}}{t\, r_{jk}} \tag{12}$$

where $t = 4\pi \sin\theta/\lambda$

and r_{jk} = distance between atoms j and k.

These g's are then combined like the atomic scattering factors in equation (11) to give $\langle I \rangle$. If a substantial part of the molecular geometry is known, molecular scattering factors can be used instead of atomic scattering factors in the Wilson plot, thus increasing the accuracy of the scale and temperature factors. The results of using molecular geometry can be seen in Figure 4 which shows the Wilson and Debye curves for pyrene-benzoquinone (23) illustrated in Figure 5. The Debye plot is a straight line within experimental error whereas the Wilson plot is not.

Care must be taken when calculating E's for structures containing such planar, highly regular molecules. If <I> is obtained from equation (11) using g's from equation (12) instead of f's, it will be found that the resulting E's are better for phase determination than if molecular scattering factors are not used. The only good alternative is to use a K-curve which follows all the "bumps" that appear in the Wilson plot. However, great care must be taken when extrapolating the K-curve at low sinθ/λ values.

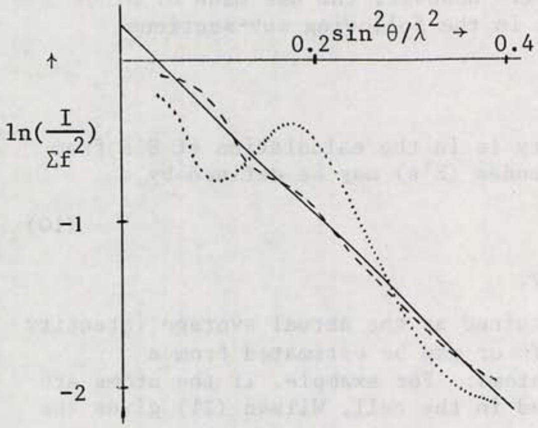


Figure 4 (Wilson (...) and Debye (---) plots for pyrene-benzoquinone. The least squares straight line through the Debye plot is also shown

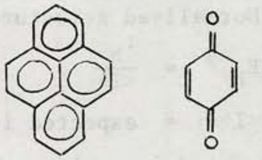


Figure 5 The two molecules in the asymmetric unit of pyrene-benzoquinone

Oriented atomic groups

If the orientation of a group of atoms is known, this information is most effectively used in calculating improved estimates for the phase and variance of the phase relationships (3). The Cochran distribution (17) of the Σ_2 relationships can be generalised to become

$$P(\phi_{hk}) = \frac{\exp\left\{2Q_{hk}|E_{\bar{h}}E_{h-k}E_k| \cos(\phi_{hk}-q_{hk})\right\}}{2\pi\, I_o\,(2Q_{hk}|E_{\bar{h}}E_{h-k}E_k|)} \qquad (13)$$

$$\text{where } Q_{hk}\exp(iq_{hk}) = \frac{\sum_i g_i(h,k)}{<|F_{\bar{h}}|^2>^{\frac{1}{2}}\,||F_{h-k}|^2>^{\frac{1}{2}}<|F_k|^2>^{\frac{1}{2}}}$$

$$\phi_{hk} = \phi_{\bar{h}} + \phi_{h-k} + \phi_k$$

and g (h,k) is a product of three atomic group scattering factors. It will be seen that the maximum probability occurs when $\phi_{hk} = q_{hk}$ so that the Σ_2 relationship becomes

$$\phi_{\bar{\underline{h}}} + \phi_{h-k} + \phi_{k} \approx q_{hk} \qquad (14)$$

In the case of randomly distributed atoms

$$g_i(\underline{h},\underline{k}) = f_i(\underline{h}-\underline{k})\, f_i(\underline{k})\, f_i(\bar{\underline{h}}) \qquad (15)$$

and the formula (13) reduces to the Cochran distribution. For an oriented group of atoms

$$g_i(\underline{h},\underline{k}) = \sum_p \sum_j \sum_k \sum_\ell f_j f_k f_\ell \exp 2\pi i[\underline{h}C_p(\underline{x}_k-\underline{x}_j) + \underline{k}C_p(\underline{x}_\ell-\underline{x}_k)] \qquad (16)$$

where the C_p are the point group operators and the summations over j,k,l are taken over all the atoms in the group. In this case, $g_i(\underline{h},\underline{k})$ is complex and q_{hk} in equation (13) can take on any value. This gives an estimate of the phase of a Σ_2 relationship other than zero. The modified phase relationships are used in a modified tangent formula for phase determination and refinement. The tangent formula becomes

$$\phi_h = \text{phase of } \{\sum_k Q_{hk} \left|E_k E_{h-k}\right| \exp i(\phi_k + \phi_{h-k} - q_{hk})\} \qquad (17)$$

which can also be weighted as in equation (5).

The first application of this was to the structure of ergocalciferol (24). The space group is $P2_12_12_1$ and there are two molecules of $C_{28}H_{44}O$ in the asymmetric unit (i.e. 58 non-hydrogen atoms). After many trials, two plausible molecular fragments were obtained from an E-map, consisting of 14 and 21 peaks. The usual methods of developing the structure failed because both molecular fragments were wrongly positioned in the cell. However, using the oriented fragments to modify the Σ_2 relationships and redetermining the phases by Multan, gave a map which showed similar fragments in new positions. These could then be developed into the complete structure using weighted Fouriers. An examination of the original fragments showed they were both incorrectly positioned by different amounts. Also, the Σ_2 relationships used to determine the first 100 phases had an average error of 52° before the introduction of the oriented fragments and 45° afterwards. Clearly, the use of chemical information improves the estimates of the phases of the Σ_2's. The latest version of Multan can solve this structure much more easily, although it still has to make use of chemical information in the way described.

Translation functions (25,26) also seek to determine the correct position of oriented molecules. Their advantage over the modification of the Σ_2's described here is that the calculations are very much smaller. However, translation functions can only deal with single fragments, whereas any number of oriented fragments can be used simultaneously with the Σ_2 method. In addition, translation functions tend to give several plausible answers, whereas the modification of the Σ_2's has always given a unique correct result in all cases tried so far, even if the fragments are small.

Known atomic positions

When a structure contains heavy atoms, the heavy atom positions are usually obtained fairly easily. However, if they possess additional symmetry to that of the space group, determination of the remainder of the structure can be difficult. This is especially true if the form of the light-atom structure is completely unknown. The determination of such structures is made easier by subtracting the heavy atom contributions statistically from the diffraction pattern and then solving the difference (equal-atom) structure in the normal way (27,28, 7). The implementation in Multan is that of Hull & Irwin (7) and will be described here with reference to the structure of potassium alborexin (29). The space group is $P2_1$ and the asymmetric unit contains one potassium atom and 62 light atoms (excluding hydrogen). This was originally solved by Busetta with great difficulty because of the mirror symmetry produced by the potassium atom. The potassium atom contribution, F_{calc}, is substracted from the F_{obs} as shown in Figure 6. The expected value of θ, the difference in phase between F_{calc} and F_{obs} is given by the Sim weighting scheme (30) as

$$\langle\cos\theta\rangle = \frac{I_1(x)}{I_o(x)} \qquad (18)$$

where $x = 2|F_{obs}||F_{calc}|/\sum_i f_i^2$

and the summation is over the atoms not included in the structure factor calculation. The magnitude of F_{diff} can now be obtained by solving the triangle in Figure 6. E's are recalculated from the F_{diff} values and used in Multan to determine what is now approximately an equal-atom structure. In the case of potassium alborexin, the first E-map gave 52 out of the 63 atoms of the structure and the problems of additional symmetry produced by the potassium were completely avoided.

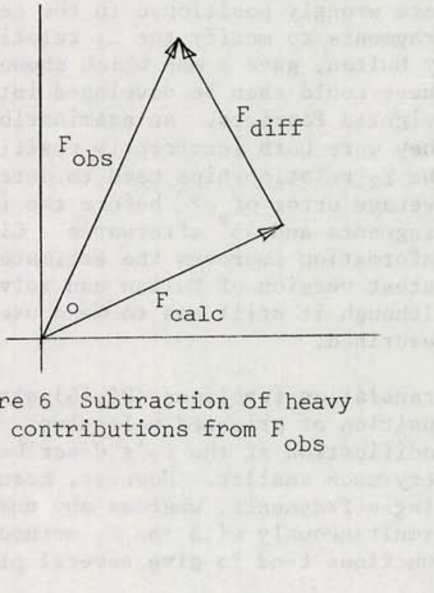


Figure 6 Subtraction of heavy atom contributions from F_{obs}

References

(1) Karle, J. & Hauptman, H. (1956) Acta Cryst., 9, 635-651.
(2) Germain, G. & Woolfson, M. M. (1968) Acta Cryst., B24, 91-96.
(3) Main, P. (1976) Crystallographic Computing Techniques, ed. F. R. Ahmed, pp. 97-105. Copenhagen: Munksgaard.
(4) Germain, G., Main, P. & Woolfson, M. M. (1970) Acta Cryst., B26, 274-285.
(5) Main, P. (1978) Acta Cryst., A34, 31-38.
(6) Hull, S. E. & Irwin, M. J. (1978) Acta Cryst., in press.
(7) Main, P. & Hull, S. E. (1978) Acta Cryst., A34, 353-361
(8) Messager, J. C. & Tsoucaris, G. (1972) Acta Cryst., A28, 482-484.
(9) Hauptman, H. (1976) Lecture, Symposium on Direct Methods, Buffalo, U.S.A.
(10) Hauptman, H. (1975) Acta Cryst., A31, 671-679.
(11) Hauptman, H. (1975) Acta Cryst., A31, 680-687.
(12) White, P. S. & Woolfson, M. M. (1975) Acta Cryst., A31, 53-56.
(13) Main, P. (1977) Acta Cryst., A33, 750-757.
(14) Main, P. (1978) Acta Cryst., A34, 31-38.
(16) Karle, J. & Karle, I. L. (1966) Acta Cryst., 21, 849-859.
(17) Cochran, W. (1955) Acta Cryst., 8, 473-478.
(18) Karle, J. (1968) Acta Cryst., B24, 182-186.
(19) Declercq, J.-P., Germain, G., Main, P. & Woolfson, M. M. (1973) Acta Cryst., A29, 231-234.
(20) Koch, M. H. J. (1974) Acta Cryst., A30, 67-70.
(21) Wilson, A. J. C. (1942) Nature, 150, 151.
(22) Debye, P. (1915) Ann. Phys. Lpz., 46, 809-823.
(23) Bernstein, J., Regev, H., Herbstein, F. H., Main, P., Rizvi, S. S. H., Sasvari, K. & Turcsanyi, B. (1976) Proc. Roy. Soc. Lond. A347, 419-434.
(24) Hull, S. E., Leban, I., Main, P., White, P. S. & Woolfson, M. M. (1976) Acta Cryst., A32, 538-550.
(25) Tollin, P. (1966) Acta Cryst., 21, 613-614.
(26) Karle, J. (1972) Acta Cryst., B28, 820-824.
(27) Beurskens, P. T. & Noordik, J. H. (1971) Acta Cryst., A27, 187-188.
(28) Van den Hark, Th. E. M., Prick, P. & Beurskens, P. T. (1976) Acta Cryst., A32, 816-821.
(29) Busetta, B. (1976) Acta Cryst., A32, 139-143.
(30) Sim, G. A. (1960) Acta Cryst., 13, 511-512.

SIMPEL

Ok Overbeek and Henk Schenk*

Summary

Simpel is an interlinked set of programs for solving centrosymmetric and non centrosymmetric crystal structures in triclinic, monoclinic and orthorhombic space groups, by means of direct methods. The starting set is generated using triplets and seven magnitude quartets, the phase extension is carried out by means of the symbolic addition method and the correct numerical phases can be found using various figures of merit, such as those based on the Σ_2-consistency, Σ_1-relation and negative quartets. The program can be run in batch mode or in interactive mode, the latter allows for easy trouble shooting and subsequent user intervention.

Introduction

The strongest and most frequently used relation in the Direct Methods is the so-called Σ_2 or triplet relation:

$$\phi_H=\phi_K+\phi_{H-K} \qquad (1)$$

(ϕ is the phase of reflection $H(=h_1 k_1 l_1)$).
The relation is most likely to be correct if the normalized product of the three E-values:

$$E_3= N^{\frac{1}{2}}\left|E_H \mathrm{x} E_K \mathrm{x} E_{H-K}\right| \qquad (2)$$

is large.
Starting with a small set of reflections with known phases ϕ, it is in principle possible to provide a large set of strong reflections with a known phase and to refine these phases, by application of a weighted tangent formula:

$$\tan\phi_H = \frac{\sum_K W_{K,H-K} \mathrm{\ x\ } \sin(\phi_K+\phi_{H-K})}{\sum_K W_{K,H-K} \mathrm{\ x\ } \cos(\phi_K+\phi_{H-K})}. \qquad (3)$$

For a successful application of (3) the choice of the weights $W_{K,H-K}$ is very inportant as will be discussed in the paper of Main in this book.

The three major problems in Direct Methods are

1. to find a good starting set of (symbolic) phases e.g. a starting set which will yield correct phases at the end of the procedure.
2. to design a reliable procedure for phase extension and/or phase

* Laboratory for Crystallography, University of Amsterdam
Nieuwe Achtergracht 166, Amsterdam, The Netherlands

refinement.
3. to use the adequate figure(s) of merit in order to select the Σ_2-solution with the correct phases.

The methods based on 1 and 3 can be separated into two conceptionally different groups:
1. multisolution and
2. single solution methods

Both methods are alike with respect to one procedure: a starting set of reflections is chosen and this set will be such, that in principle the phases of all other reflections can be calculated by means of (1). In the first group several (multi) sets of phases are calculated by inserting trial phases in this starting set, attempting to cover all possible solutions. Afterwards the best solution is selected by means of figures of merit, e.g. consistency.

Opposed to these methods are the single solution methods. A typical example is the Symbolic Addition method (Karle & Karle, 1966). While beginning with the same starting set, only one phase extension is done after assigning dummy values (symbols) to the phases in the starting set. Then, in a way similar to the multisolution methods, the best set of numerical phases for the symbols will be selected and the phase extension be executed for a second time with this set only.

A disadvantage of the latter method is, that it sometimes requires user intervention and understanding, whereas multisolution methods may be used as a black box. Opposed to this is the considerable saving of computing time, by combining all possible solutions by means of symbolic phases.

Description of SIMPEL

The program system SIMPEL is completely devoted to the Symbolic Addition method. As there are considerable differences between centro- and noncentrosymmetric space groups, regarding the application of Direct Methods, the system has been split up in two dedicated parts. The minimal requirement to get SIMPEL working, is a set of triplet relations between a number of the strongest reflections. If the structure to be solved contains N independent atoms, triplets will be calculated between the strongest 10 x N reflections; an attempt will be made to produce a set of about 100 x N non-redundant triplets. For starting set building purposes seven magnitude quartets between the ± 40 strongest reflections are usually added to the set of triplets. Apart from triplets the program can make effective use of Σ-1 relations, negative quartets and Harker-Kasper relations. These relations are all calculated once, to be used several times, if necessary.

In the end SIMPEL produces a set of h, k, l, E values with phases, which can be fed into an E-map calculating program with subsequent interpretation.
Eventually the atomic positions can be tested on correctness and be

refined by a fast diagonal least squares program. In case only part of the structure appeared in the E-map, options are available to recycle the tangent refinement with new phase information from the least-squares program.

So far the procedure is essentially the same for both centro- and non-centrosymmetric structures. The main differences will be discussed in the following two sections.

The programs have been written in FORTRAN IV for the CDC Cyber 73 Computer. The local specifications for interactive use of the computer give a limit of 24 k 60 bits words of core, and all programs are designed to meet this limit.

The centrosymmetric part

Originally the application of Direct Methods in Amsterdam was completely focussed on centrosymmetric structures. As a natural result today, most structures, containing up till about 60 independent atoms, can be solved in a routine-like fashion.

A starting set is generated by execution of a converge procedure for the triplets and quartets between the 30-50 strongest reflections. Care is taken, that the origin will be fixed; as many symbols as necessary are assigned (the maximum number of symbols is 30). Then a 'diverge' is done with the same small set of triplets and quartets, thus determining a considerable number of the (symbolic) phases of the strongest reflections.

With this extended starting set the symbolic addition routine is entered. Using only triplets for all strong reflections, it tries to symbolically phase the greater part of these reflections. New symbolic signs are accepted only, if the E_3 values of the triplets, by which they are determined, satisfy the acceptance criteria. The latter, being very strict in the beginning, are weakened every following cycle. Reflections, that would appear to get conflicting symbolic signs, are rejected for the time being. After execution of the symbolic addition these give rise to relations between the symbols and the main criterion for selecting the best solution is based on these relations. Of course these contain nothing more or less than a measure of the consistency of any possible solution. So, for space groups, where the most consistent solution is most likely to be the correct one, this solution can easily be found by trying out, which combination of signs for the symbols fits best in the relations available. In addition, in all space groups, use is made of a criterion based on all Σ-1 relations (Overbeek & Schenk, 1976).

However, in symmorphic space groups, like $P\bar{1}$, the most consistent solution does not usually lead to sensible set of signs. The application of the Σ-2 relation is completely consistent, when all signs are

+. The problem is to get hold of relations, that give rise to negative signs and use these in a figure of merit. In the SIMPEL system, one can apply three different figures of merit of this kind, based on

1. The above mentioned Σ-1 relations. Roughly one half of the Σ-1 type reflections are predicted as being negative.
2. Negative Quartets (Schenk, 1975). The sign product: s(H)xs(K)xs(L)xs(-H-K-L) is most likely to be negative, when the so-called crossterms $|E_{H+K}|$,$|E_{H+L}|$ and $|E_{K+L}|$ are small.
3. Harker-Kasper relations (Schenk & De Jong, 1973), using the probability, that the product s(H+K) x s(H-K) is negative, the larger $|E_H|$,$|E_{H+K}|$ and $|E_{H-K}|$, and the smaller $|E_K|$ and $|E_{2H}|$.
   Of course this is the result of a special negative quartet, the main terms being $E_{\bar{H}}$, $E_{\bar{H}}$, E_{H+K} and E_{H-K}.

For all centrosymmetric cases, two or more sets of criteria are calculated for all possible combinations of signs for the symbols, eventually resulting in a combined figure of merit.

The noncentrosymmetric part

The fact that, for noncentrosymmetric structures, the phases of the general reflections can have any value from 0 to 360^o, implies a much greater number of possible solutions than in a comparable centrosymmetric case. SIMPEL tries to tackle these problems in the following way.

A starting set is generated as in the centrosymmetric case (maximum number of symbols is 10). Only in this case the origin will not necessarily be fixed. In addition not even an attempt will be made to fix the enantiomorph.

Symbolic addition is carried out with this starting set for a limited number of cycles. All the necessary special attention is paid to symbols, that had been assigned to phase-restricted reflections as well as to the symbolic phases given to this kind of refections.

If the origin has not been fixed (completely) at this point, SIMPEL gathers all symbolic phases, assigned to phase-restricted reflections, and assembles them according to their parity group. By inserting one or more of these relations, from the right parity group(s), in the original starting set and thus in fact eliminating a symbol, the origin may be fixed. Symbolic addition is then done again with the new starting set.

If the origin has been fixed, one usually ends up with a set of consistency relations between the symbols, similar to the centrosymmetric case. In order to search for the most consistent solution(s), test values for the symbols are inserted and associated criteria calculated; the usual set of test values is: 0,90,180,270^o. As the consistency relations are in fact linear equations, one can easily apply a least squares refinement to the symbol values, starting from several good-looking solutions. After a typical refinement one is left with a set of solutions, that appear to be grouped two by two according to their figure of merit. These will normally lead to two essentially the same E-maps, one being the enantiomorph of the other, thus providing an easy way to discriminate them afterwards.

In all (small) structures, solved up till now, the first or the second best solution yielded the correct structure after tangent refinement. It is also possible to use a figure of merit based on the negative quartets to discriminate between the possible solutions.

References

Karle,J. and Karle,I.L. (1966), Acta Cryst., 21,849
Schenk,H. and De Jong, J.G.H.(1973), Acta Cryst., A29, 31.
Schenk,H.(1975), Acta Cryst., A31, 259.
Overbeek, A.R. and Schenk, H.(1976), Proc.Kon.Ned.Ak.Wet.B79, 341.
Overbeek,A.R., Olthof,G.J., V.d.Putten, N.H. and Schenk, H., SIMPEL-MANUAL.

THE BUFFALO SYSTEM FOR DIRECT METHODS

D. A. Langs and H. A. Hauptman*

Summary

Theoretical contributions to the reliable estimation of the values of crystal structure invariants, notably towards estimates for three phase (triple) and four phase (quartet) universal structure invariants, have permitted the development of a set of computer programs based on the multisolution tangent formulation. These programs, known as QTAN, have been optimized for a small computer (DEC PDP 11/45, 32 K core, 16 bit word), are capable of inspecting large numbers (2^{14}) of phase sets in a finite amount of time (8 hours), and have been useful in a number of structure determinations which had been otherwise intractable. Two important features of the program are [1] the use of a negative quartet figure of merit (NQEST) as a sieve to discard unpromising phase sets at intermediate stages of phase extension, and [2] the formal separation of tangent extension and tangent refinement of phases, permitting global, fixed point computation techniques with attendant savings in both computing time and computer storage requirements. Incorporation of cosine estimations into the design of an optimal phase extension route has proved to be a powerful and, in some cases essential, ingredient in successful structure determination.

Introduction

A number of approaches to overcoming the phase problem are currently in use and development at the Medical Foundation of Buffalo, Inc. Given our interest in the reliable estimation of the values of structure invariants and structure seminvariants (1), research conducted here, primarily by Drs. Hauptman, Green and Fortier, and given the proven utility (2) of accurate estimations of these quantities in structure determination, seen in the work of Drs. Duax and Weeks, many techniques based on these theoretical contributions and practical applications have evolved. Because of limitations of time and space only one of these, QTAN (3), will be discussed here. A historical account of the developments leading to QTAN will serve to underscore the importance of cosine analysis to new direct methods programming.

---

*Medical Foundation of Buffalo, Inc., 73 High Street, Buffalo, New York 14203, U.S.A.

Techniques for estimating $\cos(\phi_{\underline{h}} + \phi_{\underline{k}} + \phi_{\underline{-h-k}})$, such as MDKS (1b) and TPROD (1a), have existed in the lab since 1970. It was clear that useful information was contained in these estimates and highly personalized approaches to their use were developed. For example, relatively large, self-consistent basis sets of phases could be developed with their use; these basis sets served as input to tangent procedures. Additionally, their use could lead to enantiomorph specification, and the identification of groups of phases which had the property that members were tightly bound to one another through highly reliable phase relationships. Our goal was to incorporate these estimates as actively as possible in phase determination procedures. A critical step in realizing this goal, in some measure, was to come from a theoretical breakthrough in the cosine analysis of quartets, $\cos(\phi_{\underline{h}} + \phi_{\underline{k}} + \phi_{\underline{\ell}} + \phi_{\underline{-h-k-\ell}})$.

Schenk and De Jong (4) had discovered an extremely interesting property of pairs of triples with one common phase. Using the Harker-Kasper inequalities they showed that $\phi_{\underline{h+k}} + \phi_{\underline{h-k}} - 2\phi_{\underline{h}} \approx \pi$ if $|E_{\underline{h}}|$, $|E_{\underline{h+k}}|$ and $|E_{\underline{h-k}}|$ are strong and $|E_{\underline{k}}|$ is weak. The reliable π estimate, recognized as a goal of structure invariant analysis, was in sight. Hauptman (5) showed that the relationship discussed by Schenk was a special case of the universal structure invariant $\phi_{\underline{h}} + \phi_{\underline{k}} + \phi_{\underline{\ell}} + \phi_{\underline{-h-k-\ell}}$, with $\underline{h} = \underline{k}$, and that the universal quartet was likely to sum to π if $\{|E_{\underline{h}}|, |E_{\underline{k}}|, |E_{\underline{\ell}}|, |E_{\underline{-h-k-\ell}}|\}$ were large and $\{|E_{\underline{h+k}}|, |E_{\underline{k+\ell}}|, |E_{\underline{h+\ell}}|)\}$ were small. Schenk (4,6) employed the special quartets early on as a figure of merit for symmorphic space groups; with more general quartet formulation in hand we developed the NQEST figure of merit (7). The initial tests of NQEST as a figure of merit, and the recognition of its independence with respect to phase sets generated by tangent techniques, suggested two things: [1] large numbers of phase sets, if these could be economically generated, could be tested by NQEST and the promising phase sets would be identified, and [2] it might not be necessary to carry out tangent extension for each phase set over the full range of phases; *i.e.*, there was a distinct possibility that NQEST was sufficiently powerful as a figure of merit and sufficiently independent of the phasing process that phase sets could be tested at intervals along the tangent extension pathway and hopelessly incorrect phase sets identified and rejected well before the end of the pathway. This latter property, the filtering of phase sets, could be important in keeping down the amount of computing time spent per phase set since an early cycle of extension and refinement over a number of phases is shorter than a late or final cycle over a larger number of phases. This is due to the necessity of carrying out tangent refinement over all previously extended phases after each extension cycle is performed.

With the tools in hand a program package was assembled. Some parts of the package, such as the Σ_2 generator, were dictated by the nature of the tangent formula itself. Some parts, notably the establishment of a phase route, were chosen from a variety of well established procedures in the literature. For instance, one option for routing is closely patterned on the CONVERGE procedure introduced by Germain, Main and Woolfson (8); another, which we call DIVERGE, has some of the elements of an inverse-convergence mapping. In the latter procedure a basis set of phases, usually identified by a previous convergence listing, is selected and phases are added, one at a time, so that the next phase to be determined has the highest $(\alpha_{\underline{h}})_{est}$. The strategy for the tangent formula implementation was actually dictated by the requirements of the filtering procedure.

As originally envisioned, after the selection of a basis set of phases and solution sets, tangent extension and refinement would be carried out along the established phase route until a significant (~30) number of negative quartets were phased. NQEST would then be calculated for the P plausible phase sets and extension/ refinement *continued* for the best, say P/2, phase sets as judged by NQEST. It was quickly realized that storage requirements on the PDP 11/45 would be severely taxed by such a strategy since it would require the storage of intermediate phase values (say 75-125) for P/2 phase sets. We thought that some determinations might require that P be as large as 1000, making the storage of 75,000 to 125,000 integer words necessary. This was clearly out of the question at the time. We realized that it was actually just about as quick to start the tangent extension *ab initio* for the P/2 phase sets as it would have been to read the intermediate phases from the disk and continue extension from the point where the previous extension cycle had ended.

A practical consideration in the design of the tangent strategy was whether to filter using NQEST after a cycle of extension/ refinement or after a cycle of extension alone. Some testing with a prostaglandin data set suggested that the former was preferable. This did not dictate, as such, the decision to formally separate the tangent refinement and tangent extension, but it did suggest it as being practical. Another consideration which favored this formal separation was the realization that both computing time and storage requirements would be cut significantly by adopting it. When the decisions to filter after refinement and to separate the refinement and extension were made the program began to take form.

Initial tests were encouraging as long as we focused on previously determined structures. Indications were that the program was very fast, large numbers of phase sets could indeed be examined quickly, the sifting was working without complication, and refinement results indicated that the basic algorithms were correct. However, QTAN was doing no better than dice for the tetrahymanol hemihydrate (THYM) structure which had been resisting every effort at solution (including Patterson search techniques) which had been thrown at it. Large self-consistent basis sets of phases had been selected via cosine analysis techniques (20 to 30 reflections), the tangent program was modified in such a way as to permit large basis sets to be expressed in terms of the permutable values of a few chosen symbols, and yet THYM was just not yielding to QTAN. The most encouraging QTAN run involved an examination of 2048 possible phase sets; 5 essentially equivalent phase sets had NQEST values of -0.20 to -0.26. Maps produced from these sets showed two molecules of tetrahymanol (there are two molecules per asymmetric unit). Calculation of structure factors for the h0ℓ data give a residual of 25%. Naturally this was encouraging; unfortunately it was also incorrect. The residual for the full data set was 36% and the trial structure would not refine. At this point it was realized that the two molecules determined were of opposite handedness (*i.e.*, a d,ℓ pair), an impossibility for this natural product. Efforts to elaborate a solution given the information in hand were futile.

The turning point in the determination of THYM was the decision to use the cosine analysis to throw away certain Σ_2 relationships which were likely to differ significantly from the 0° estimate. Our motivation to do so was well justified in view of what we knew about the molecular orientations and how the reliable values of the cosine invariants were partitioned on a molecular basis, but apart from this there was no *a priori* justification for rejecting triples. If anything there were real considerations which suggested that removing triples from the list of Σ_2 relationships would only weaken chances for solution. For instance, removing triples from consideration in the construction of a convergence map could introduce more 'weak links' in the map, which would require the incorporation of more ambiguities into the basis set; and there was a real chance that editing a convergence map in this way would bias the tangent process in unknown and undesirable ways. Additionally, such editing was not necessary in previous, successful applications of cosine analysis techniques. Nevertheless, and in some measure in desperation, the THYM triple file was purged of 1000 of the strongest (largest A value) 2000 triples; *i.e.*, one quarter of all of the triples were discarded from further consideration. Using this edited list of triples a convergence map was constructed; this map was compatible with the same basis set of phases employed in the QTAN run which yielded the d,ℓ THYM pair. This basis set and new route were tested over another 2048 ambiguities. Three phase sets survived the sifting; they had NQEST values of ~-0.45 and they revealed the crystal structure of THYM. The editing of triples files *via* cosine invariant analysis has led to successful structure determination in a number of subsequent situations where intractable problems were encountered.

Experience with THYM has suggested the following points for consideration [1] a blind increase in the number of ambiguities tested will not always improve the chances of securing a solution, [2] cosine invariant analysis can markedly improve the chances of securing a solution, and [3] if the information inherent in the cosine invariant analysis is used in some forceful way, even intractable problems can be solved in many instances.

QTAN has evolved continuously in response to the variety of problems it has faced in the lab. For instance, the options to use molecular fragment recycling procedures (9), magic integers (10), and A or α-weighted phase extensions have been incorporated. An important off-shoot of the recycling procedure has been the identification of phases most sensitive to the unknown parts of the structure and the active use of these phases in further defining the basis set. These options are all under user control and have proven to be helpful in various situations.

It is by examining figures of merit that optimal phasing strategies can be designed for particular situations. The figures of merit are symptomatic of various difficulties which are sometimes encountered in structure determinations and, when diagnosed, the difficulties are often either avoidable or can be minimized. Space considerations do not permit a meaningful exploration of all of the possibilities open to the user of QTAN, but some of these will be examined in the study session setting.

Acknowledgement

The foremost peculiarity of the Buffalo direct methods system is that it has never been coded with FORTRAN, but rather with people, most of whom have managed valuable contributions in addition to the other demanding research responsibilities required of them. Time and space would only permit a feeble description of the scope of direct methods activity conducted by various individuals and would not do justice to the people concerned. Their number include W. L. Duax, C. M. Weeks, E. A. Green, S. Fortier, G. D. Smith, D. L. Dorset, D. C. Rohrer, S. A. Potter and M. Duffy-Fronckowiak. Also acknowledged are Drs. G. J. Kruger (National Physical Research Laboratory, Pretoria, South Africa) and J. W. Edmonds (Dow Chemical Corporation, Midland, Michigan). The authors curiously thank Dr. G. T. DeTitta who wrote this abstract and supplied the original motivation behind the development of the QTAN computer package.

This work was supported in part by Grant No. HL-15378 from the National Heart, Lung and Blood Institute, DHEW and MPS73-04992 from the National Science Foundation.

References

1. (a) Hauptman, H., Fisher, J., Hancock, H. and Norton, D. A. (1969). Acta Cryst. B25, 811-814.
(b) Fisher, J., Hancock, H. and Hauptman, H. A. (1970). Naval Research Laboratory Report 7132; Naval Research Laboratory Report 7157.
(c) Hauptman, H. (1970). Acta Cryst. B26, 531-536.
(d) Hauptman, H. (1972). Crystal Structure Determination: The Role of the Cosine Seminvariants. Plenum Press: New York.
(e) Hauptman, H. and Duax, W. L. (1972). Acta Cryst. A28, 393-395.
(f) Duax, W. L., Weeks, C. M. and Hauptman, H. (1972). Acta Cryst. B28, 1857-1863.
(g) Hauptman, H. (1972). Acta Cryst. B28, 2337-2340.
(h) Hauptman, H. and Duax, W. L. (1972). Acta Cryst. B28, 2622-2623.
(i) Hauptman, H. and Weeks, C. M. (1972). Acta Cryst. B28, 3166-3173.
2. (a) Weeks, C. M., Cooper, A., Norton, D. A., Hauptman, H. and Fisher, J. (1971). Acta Cryst. B27, 1562-1572.
(b) Duax, W. L. (1972). Acta Cryst. B28, 1864-1871.
(c) Duax, W. L. and Hauptman, H. (1972). Acta Cryst. B28, 2912-2916.
(d) Weeks, C. M. and Hauptman, H. (1972). Acta Cryst. B28, 3331-3339.
3. Langs, D. A. and DeTitta, G. T. (1975). Acta Cryst. A31, S16, Abstract 02.2-14.
4. Schenk, H. and De Jong, J. G. H. (1973). Acta Cryst. A29, 31-34.
5. Hauptman, H. (1974). Acta Cryst. A30, 472-476.
6. Schenk, H. (1973). Acta Cryst. A29, 480-481; Schenk, H. (1974). Acta Cryst. A30, 477-481.
7. DeTitta, G. T., Edmonds, J. W., Langs, D. A. and Hauptman, H. (1975). Acta Cryst. A31, 472-479.
8. Germain, G., Main, P. and Woolfson, M. M. (1970). Acta Cryst. B26, 274-285.
9. Karle, J. (1968). Acta Cryst. B24, 182-186.
10. White, P. S. and Woolfson, M. M. (1975). Acta Cryst. A31, 53-56; Main, P. (1977). Acta Cryst. A33, 750-757.

CHAPTER 4

Automatic data collection by film and diffractometer techniques

DATA COLLECTION BY FILM METHODS AND THE PROCESSING OF FILM DATA

Oliver Lindqvist*

Summary

The measurement of single crystal data is usually performed most easily and accurately with a computer-controlled diffractometer. There are, however, important structural problems which still require collection of film data. The main such field is protein crystallography, but structure determinations from powders and of order-disorder crystals are also performed advantageously with film techniques. Common single crystal cameras are surveyed together with developments within powder diffraction methods. The on- and off-line computer evaluation of integrated single crystal intensities using optical microdensitometers is discussed.

1. Intoduction

The use of computer-controlled X-ray diffractometers now completely dominates the collection of intensity data for the determination and refinement of ordinary crystal structures. This is a result of the improved precision usually obtained with counter techniques as compared to film measurements, but, at the present state of the art, also due to economic reasons. For an average inorganic, metallo-organic or organic crystal structure it is usual to obtain diffractometer intensities with a statistical variation of 3-6%, while values of 8-10% are to be expected for film intensities. The ultimate precision for diffractometer and film techniques may be within 2% and 6%, respectively. Economically, there would seem to be important factors in favour of the use of film methods. The capacity of one or two screenless cameras and an automatic film-scanner is in principle 10 times greater than that of a conventional computercontrolled diffractometer, the costs of the two systems being comparable. Regional centres for automatic film evaluation, serving several crystallographic laboratories, would appear to be a relatively inexpensive way of collecting single crystal data. However, there is no system for automatic determination of crystal setting and registration of the intensities, available commercially, which can compete with the simplicity of the diffractometer systems. At present the adjusting of the crystal and the processing of films usually results in larger costs for film intensity evaluation. Of course,

---

*Department of Inorganic Chemistry, Chalmers University of Technology and the University of Göteborg, P.O.Box S-402 20, Göteborg 5, Sweden

the two step character of the film method in principle hinders the interaction between a computer and the crystal diffraction, but it still ought to be possible to develop more general film evaluation systems than those available at the present time.

For many applications film data collection is superior to usual diffractometry. The most important such field is protein (macromolecular) crystallography, where, due to unstable crystals and large unit cell dimensions, better total accuracy is often obtained with film data. Other cases where it may be advantageous to use film data, again due to the lack of high quality single crystals, are the elucidation of three-dimensional crystal structures from order-disorder (1) crystals or from powder diffraction diagrams.

In the following sections film intensity registration and evaluation will be commented on without any further comparison with diffractometer techniques.

2. Single crystal cameras

The X-ray cameras used for the collection of integrated intensity data utilize crystal rotation or precession motions of the crystal. The rotation axis is perpendicular to and the precession cone axis is coincident with the primary X-ray beam. One instrument involving a stationary crystal, the convergence camera, is also of interest.

Screened methods

Film scanners, supported with on- or off-line computer programs permit indexing of complicated non-screened intensity patterns, but it is still often convenient, and sometimes necessary, to collect and process ordinary equi-inclination Weissenberg (2,3) and Buerger precession (4) films. From such films it is possible to check the intensity distribution in reciprocal space and to index the reflections by inspection. The screened methods are especially valuable for the solution of structures which give classes of very weak reflections, due to pseudo-symmetries, or when a structure investigation has to be performed with twinned or order-disorder crystals.

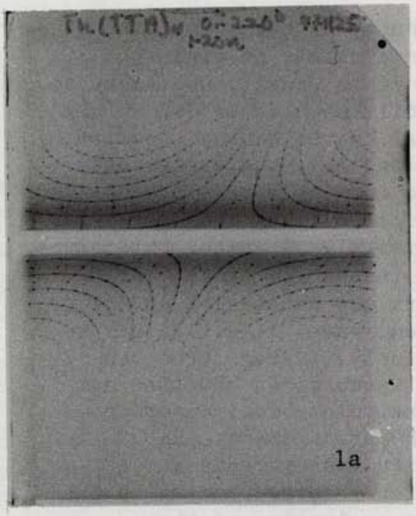
In practice it is relatively easy to adapt film-scanner software systems to precession films, while the evaluation of Weissenberg films is more complicated. At the high 2θ angles recordable with cylindrical film geometry, for well-ordered crystals appreciable $\alpha_1-\alpha_2$ splitting occurs and crystal adjustment errors may cause large shifts from the calculated film coordinates for high angle reflections. In fact Weissenberg films with high angle scattering require more elaborate centering routines than most other camera geometries.

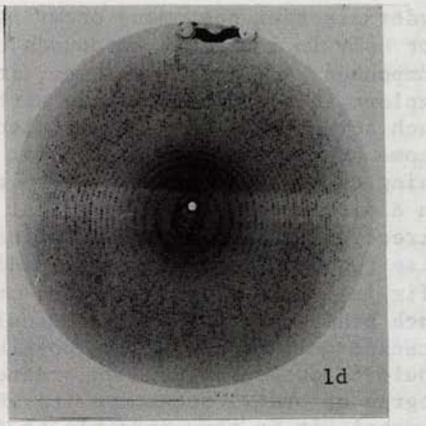
Order-disorder phenomena occur rather frequently in the solid state, not only in inorganic compounds, but also in crystals of organic compounds and, sometimes even, in proteins. It is usually necessary to explore the reciprocal space with screened film methods to determine such structures. In an order-disorder crystal the molecules, ions or atoms are ordered in only one or two dimensions (1), the latter case being the most common. The resulting ordered rods or layers are stacked in a disordered way. The crystal thus lacks periodicity in a given direction, resulting in an infinite cell axis in real space. This gives rise to a continuous intensity distribution in the reciprocal cell (fig 1a). In the simplest case ordered layers may be stacked on top of each other in one of two chemically equivalent ways. If all layers were stacked only in one of the possible ways a fully ordered structure would be obtained. Such an ordered model is known as an MDO (maximum degree of order) structure (1). For each particular order-disorder crystal, it is thus possible to define an MDO arrangement, the structure analysis of which requires intensity measurements at positions along the continuous intensity streaks corresponding to the reciprocal lattice of the MDO structure. The evaluation of the continuous intensity distribution in reciprocal space is a problem which can be handled adequately with a microdensitometer.

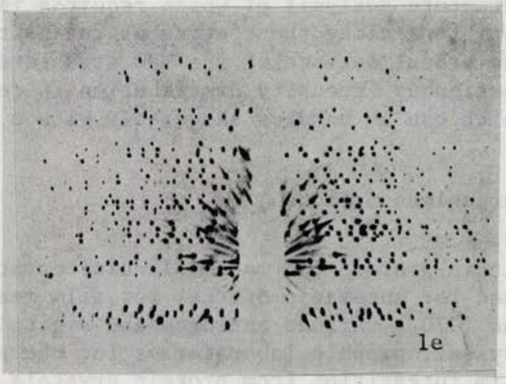
Screenless methods

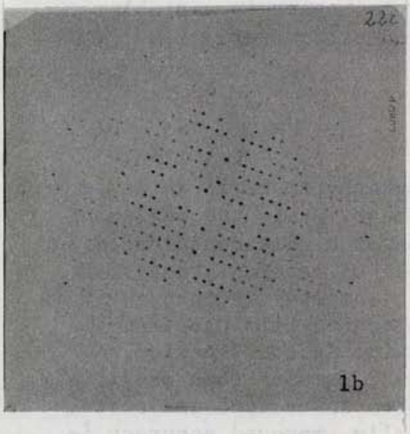
Within protein and macromolecular crystallography there is an acute need for automatic devices for film measurement. Up to the beginning of the 1970's various analogue and digital apparatus were constructed in crystallographic laboratories for the purpose of reading precession film intensities from protein crystals. With computer techniques it is possible to index and measure the more complicated film patterns obtained with screenless methods, thus allowing data collection to proceed with a speed of an order of magnitude greater than previously. Although the elevated speed in data collection saves time and costs, the most important factor is that substantially improved accuracy is obtained in the data sets, since the number of unstable crystals needed for the collection of one complete data set is reduced considerably.

The Arndt-Wonacott oscillation camera (6) is, no doubt, the most effective x-ray detector for macro-molecular crystals available commercially. Each exposure covers an oscillation of a few degrees, resulting in films such as those shown in fig 1c. Automatic collection of several subsequent exposures is possible. The other widely used screenless method in protein crystallography is the small-angle precession method (7), cf. fig 1d. The drawbacks with this method as compared with the Arndt-Wonacott oscillation procedure are: 1) the reflections are split in the precession method, the more the higher the Bragg angle and starting seriously at $2\theta = 20^\circ$; 2) rather severe overlap at high 2θ angles; 3) strongly varying Lp factor; and









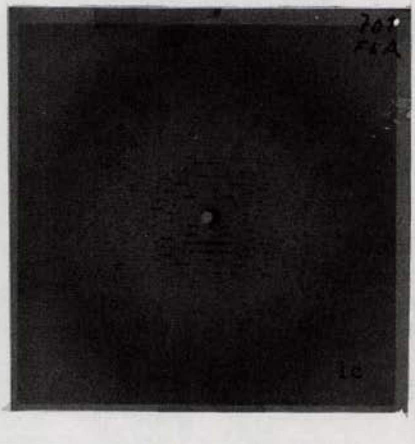


Fig. 1. Examples of film patterns

a) Weissenberg hlξ zone of $Th(C_8H_3F_3O_2S)_4$ showing order-disorder intensity distribution (5).

b) Precession hk0 film (F=6 cm, $\mu=25^o$) of $Na_2C_4H_4O_6.2H_2O$, used in the IUCr microdensitometer project.

c) Screenless precession film (F=10 cm, $\mu=0.75^o$) of lysyl tRNA synthetase from yeast to 5Å resolution (22).

d) Oscillation film (1^o, F=10 cm) of Satellite Tobacco Necrosis Virus to 2.7Å resolution (23).

e) Step-oscillation film of 2,5-dimethyl-2,5-endo-thio-1,4-dithiane (24). The oscillation angle of 45^o has been divided into four steps.

4) complicated geometry for partially recorded reflections. At high resolution studies, the overlap problem is severe, but up to a resolution of, say, 3.5 Å careful computer analyses of the above effects (8) give reliable results.

Screenless methods have also been designed for the collection of x-ray intensities from small or moderately sized unit cells. The step oscillation camera (9) makes it possible to collect one or two thousand reflections on one cylindrical film (fig 1e), which may comprise all, or a large part, of the independent diffraction from an ordinary structure investigation.

Another interesting application is the convergence method (10,11). Cameras for both ordinary (12) and protein (13) crystallography have been designed. The film geometry is similar to that of the oscillation method, but instead of the crystal being rocked, it is kept stationary and the convergent line focus of the x-ray source is used (fig 2). The advantages of the convergence method are the simple camera construction and the improved intensity yield (12). An apparent drawback is possible inhomogeneity in the line focus.

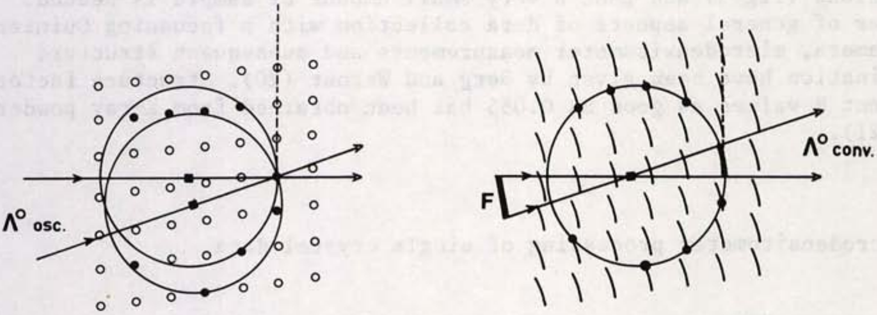


Fig. 2. Corresponding oscillation and convergence geometry of a plane in reciprocal space. After Kulpe and Schulz (12)

3. Powder methods

Many crystal structures cannot be solved in the usual way due to difficulties in preparing suitable single crystals. However, in several cases it is possible to prepare microcrystalline samples. Recent development of powder indexing (14,15) and microdensitometer measurements (16,17) in combination with profile refinement procedures (18,19) has enabled the solution of structure problems with up to 50 parameters. The physical advantages of using a focussing powder camera compared to

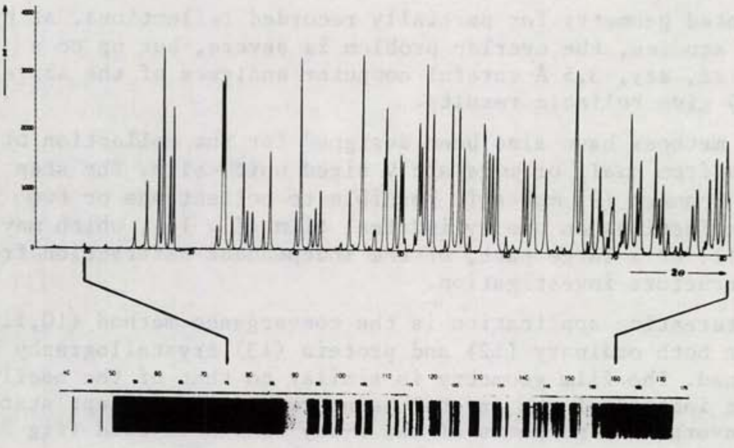


Fig. 3. Microdensitometer output from a Guinier-Hägg powder film. Reproduced from Journal of Applied Crystallography, 1977 (19)

powder diffractometer techniques are improved resolution of the reflections (fig 3) and that a very small amount of sample is needed. A number of general aspects of data collection with a focussing Guinier-Hägg camera, microdensitometer measurements and subsequent structure determination have been given by Berg and Werner (20). Structure factor agreement R values as good as 0.055 has been obtained from X-ray powder data (21).

4. Microdensitometer processing of single crystal data

The monograph "The Rotation Method in Crystallography", edited by Arndt and Wonacott (6) covers most aspects of microdensitometry. This book is very useful not only to protein crystallographers, but also to all active and potential users of film methods in crystallography. Reference is made in the book to three program systems for screenless oscillation data, the Cambridge, Harvard and Munich systems.

The considerations in the present paper are based on work with an Optronics P-1000 rotating drum microdensitometer and on-line (Hewlett-Packard 2100) and off-line (IBM 360) treatment of Weissenberg, screened and screenless precession films in Göteborg. Many of the programming ideas have been obtained from Peter Schwager, Munich.

Linearity correction

The optical density is defined as $D = -\log(I_t/I_o)$, where I_o is the intensity of the light beam incident on the film and I_t the intensity of light transmitted through the film. However, the X-ray intensity exposed on the film is not a linear function of D, although it is nearly linear up to $D = 1.2$. Morimoto and Ueda (25) have investigated the linearity of a variety of films. There are three methods which have been used in microdensitometry to correct for the non-linear dependence of the intensity with the measured optical density: 1) parabolic scaling of a film-pack (26), 2) direct biasing of the microdensitometer logarithmic amplifier circuit (27) and 3) numerical calibration of each density read by the densitometer. The third method can be used for all purposes. A separate calibration curve has to be constructed for each film quality used. An easy way of performing the calibration is to prepare an intensity scale of timed exposures on an X-ray film of the relevant quality. It is essential that the spots are homogeneous in intensity over a relatively large area and that they cover the cross-section of the microdensitometer light beam. The primary X-ray beam, attenuated in a way that does not change the wavelength characteristics, may be used. Typical curves for Ilford G film are shown in fig 4.

Choice of light beam size

The microdensitometer incident light beam is rectangular (round in some instruments) with a size usually in the range 50 × 50 μm to 200 × 200 μm. The scanning raster (beam positioning) should be selected so that the total film area of interest is transmitted. For most purposes a light beam size of about 100 × 100 μm is optimum (26), i.e. 10,000 measured values will be obtained per cm^2. For small reflections a steep intensity gradient may result in poor correspondence between the actual average optical density in a measured area and the transmittance obtained (Wooster effect). On the other hand, if a very small area is measured, the variance will increase due to sensitivity to the distribution of silver grains in the film (28). A comparison of integrated and non-integrated Weissenberg films measured with 50 × 50 μm and 100 × 100 μm, respectively, has demonstrated a clear improvement with the smaller beam size (29); cf. fig 5. The time needed to scan a given film area is four times longer with 50 × 50 μm than with 100 × 100 μm. However, the computing time and the memory requirements are approximately the same, since the 100 μm program can be used, modified with an assembler routine creating 100 × 100 μm density values as $D_{100} = (D_1 + D_2 + D_3 + D_4)_{50}/4$. $D_1, \ldots\ldots, D_4$ are the four relevant 50 × 50 μm density values.

5. Programming considerations

The evaluation of integrated intensities from single crystal X-ray films may either be performed a priori or by using precalculated film coordinates for each reflection exposed on the film. In the first case

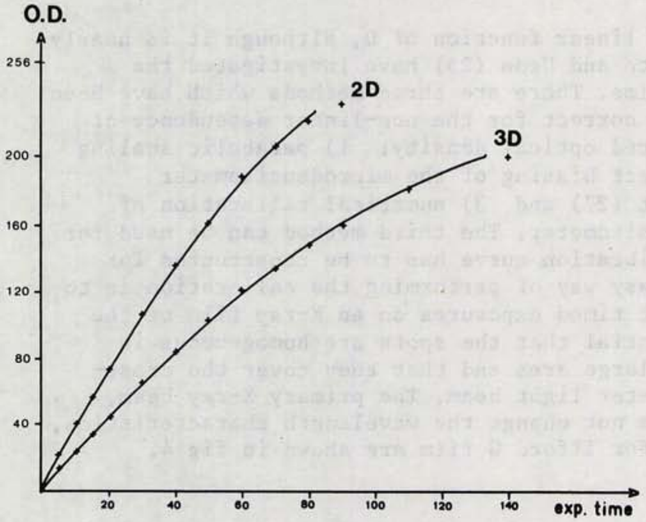
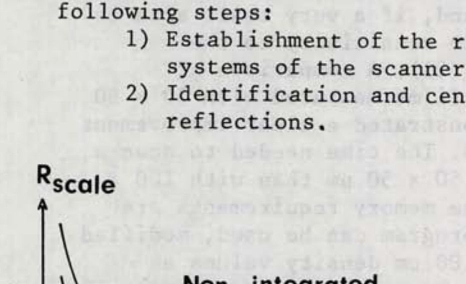


Fig. 4. Optical densities of a standard scale measured at different levels of densitometer sensitivity. The maximum output value, 256, corresponds to a filled eight bits binary register

scanner coordinates are stored for each significant reflection found, and the indexing, and the refinement of cell dimensions and the crystal setting are subsequently evaluated with computer programs. In principle this method is straightforward, but in practice artifacts, such as film damage, dust and white radiation cause problems, and, more important, weak reflections are lost. The second method, on the other hand, involves some extra measurement, but permits optimum evaluation of weak reflections, and is therefore to be recommended. A complete microdensitometer software system (on-line or off-line) can include the following steps:

1) Establishment of the relationships between the coordinate systems of the scanner, the film (= camera) and the crystal.
2) Identification and centering of a number of reference reflections.



R_{rep}

	100×100	50×50
Integrated	1.30	1.20
Non-int	1.50	1.25

Fig. 5. A comparison of measurements of sharp Weissenberg reflections, using different light beam sizes.
Reproducibility: $R_{rep} = \Sigma|I - I'|/\Sigma I$.
Film factor between adjacent films:
$R_{scale} = \Sigma|I_1 - k.I_2|/\Sigma I_1$

3) Refinement of crystal setting and cell parameters from the positions of the reference reflections.
4) Calculation of accurate film positions of the reflections relevantly exposed. Calculation of Lp factors and parameters to be used in absorption correction.
5) Evaluation of integrated intensities after individual centering of each reflection.
6) Scaling, output.

Coordinate systems

An automatic microdensitometer may operate either with step-wise positioning on the film, or by reading the densities of the whole film sequentially. It is only practical to perform on-line intensity evaluation with the first application, but both are suitable for magnetic tape recording for subsequent, off-line, computer treatment. The scanner (Y,Z) coordinate system is orthogonal and is usually defined in raster (positioning) steps. The film coordinate system, (y_f, z_f), may also be defined orthogonal and related to the camera in a conventional way, so that the x_f axis is perpendicular to the film plane and directed along the primary beam and the y_f axis is parallel to the rotation (spindle) axis, thus leaving z_f pointing upwards on most camera arrangements.

A reliable way of relating scanner and film systems is the use of fiducial marks on the film, i.e. three well-defined spots which may be exposed through holes in the backside of the film holder. The marks should be made at peripheral positions, so as not to be confused with intense low angle reflections. For each camera, the fiducial marks are more or less accurately related to the film center, $(y_f, z_f) = (0,0)$, and to the y_f (spindle axis) direction. After the film has been mounted in the scanner, the (Y,Z) positions of the fiducial marks are measured, whereupon the relationship between the scanner and film coordinate system may easily be established.

An alternative and sometimes more flexible way of orientating the film in the scanner is to drill a small hole (e.g. $\Phi = 0.4$ mm) approximately at the centre of the film and a second hole defining the y direction. Although each film needs separate consideration it is usually easy to position the holes accurately enough to be able to identify reference reflections. Another apparent advantage of this method is that it is not possible to confuse the holes with any other data on the film.

The relationship between the crystal and film (or scanner) coordinate systems involves more complicated geometrical computation. It is however, always possible to calculate film coordinates for each reflection brought to diffraction if camera constants (oscillation or spindle angles, inclination angle, film to crystal distance etc.) cell dimensions and crystal settings are known. The geometrical formulae required for the screened and screenless precession (4,7), for the Weissenberg (3) and for the oscillation (≈convergence) (6) methods are available in the publications cited.

Reference reflections

When the film is mounted in the scanner a number of reference reflections may be identified and centered, establishing a set of accurately observed film coordinates. The number of reflections should be adequate with respect to the number of parameters to be varied in the subsequent refinement, up to 15 reflections usually being sufficient. In practice it is suitable to select reflections, representatively spaced on the film with optical density significantly greater than that of neighbouring reflections in order to avoid indexing errors. Approximate film coordinates can be measured with a manual device before mounting the film in the scanner or, directly on the scanner. The reflections may be indexed either by inspection of the film, or, if necessary, by comparison with film positions calculated initially. (A fully automatic procedure in which arbitrary crystal setting and unknown cell dimensions are determined directly from the measured film coordinates, similar to single-crystal diffractometer techniques, does not seem feasible in most cases. However, for small molecule systems, such as the cylindrical step oscillation camera (9) such programming might be profitable.

Positional refinement

Camera constants such as crystal-to-film distance and orientation of rotation (spindle) axis with respect to the film plane or the film cylinder axis, may be determined from measurements on a standard crystal. Usually, the manufacturer's specifications of camera geometry are sufficiently precise, but for old and home-made cameras, it may be wisest to determine the camera constants.

In a general case of film evaluation, the film centre, film orientation, relevant crystal setting angles and the cell dimensions have to be refined from the measured positions of the reference reflections. It is important to consider to what extent a parameter will influence the positions of the reflections and the refinement can be optimized if relevant fudge factors can be assigned. The centre and orientation of the film are easy to refine, but it is not necessary to do this if fiducial marks are used. In many cases the cell dimensions are known accurately and thus need no refinement.

The relationships between the film coordinates and crystal setting angles are complicated for most camera geometries. Sometimes it may be cumbersome to establish the normal equations needed for ordinary least-squares refinement. A relevant way of solving the problem is to use an imlpicit refinement procedure, i.e. the simplex method (30). If the film coordinates for N reflections are calculated as a function of m variables, $y_{f,calc} = F_1(v_1 \ldots v_m)$ and $z_{f,calc} = F_2(v_1 \ldots v_m)$, the simplex starts from m + 1 approximate sets of coordinates $v_1 \ldots v_m$ and effectively minimizes the least-squares sum $\Sigma\{w_{y,i}(y_{f,obs} - y_{f,calc})_i^2 + w_{z,i}(z_{f,obs} - z_{f,calc})_i^2\}$ to give the best set of parameters $v_1 \ldots v_m$. The general simplex can be programmed so that any variable may be fixed.

The simplex method has been used to refine the crystal orientation for screenless precession films to 5Å resolution from unstable crystals of lysyl tRNA synthetase from yeast (22). In this case the refinement was performed after the intensities had been evaluated. It was possible to obtain only one film pack, $\mu = 0.75^{o}$, for each crystal, and a complete data set was obtained from 30 crystals. The Lp factors are extremely sensitive to setting errors in ϕ_1 (around y_f) and ϕ_2 (around z_f) for reflections near the recordable limits. The imlicit function $U\{\phi_1(a),\phi_2(a);\ \phi_1(b),\phi_2(b)\} = \Sigma w_{\varepsilon}(|F_a| - |F_b|)^2$was therefore minimized for overlapping reflections from exposures of two crystals a and b. The ε value is a measure of how close a reciprocal point will pass the recordable border (31) and $w_{\varepsilon} = 0$ for partially recorded reflections, otherwise $w_{\varepsilon} = 1/\varepsilon$. In general the intercrystal $R = \Sigma||F_a| - |F_b||/\Sigma F_{av}$ was 6-15% before the refinement and became 5-6% after the refinement (no absorption correction had been performed).

Effective explicit least-squares procedures for the refinement of setting angles for the screenless precession and the Arndt-Wonacott oscillation cameras based on partially recorded reflections are available (8,6).

A simplified way of correcting the initially calculated film coordinates is the (sometimes poor) approximation for refining the a_{ij} coefficients in the equations:

$$(y_{f,obs} = a_{11}\cdot y_{f,calc} + a_{12}\cdot z_{f,calc} + a_{13})_i$$
$$(z_{f,obs} = a_{21}\cdot y_{f,calc} + a_{22}\cdot z_{f,calc} + a_{23})_i$$

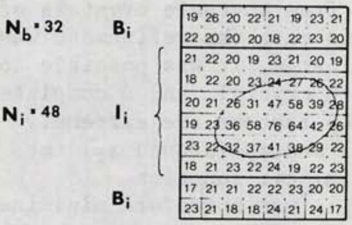
Errors in film orientation, cell dimensions and setting angles may be smeared out in the $a_{11} - a_{22}$ parameters, while errors in the film centre determination are corrected for by a_{13} and a_{23}. The approximation is, of course, better for flat films than for cylindrical films, and it can be used, at least for small cells, if the reflections are centered during subsequent intensity evaluation.

Calculation of film positions

If refined values for crystal orientation and cell dimensions are available the calculation of the film position to be used for the final intensity evaluation is straight-forward. In general the reflections which are fully recorded in the actual exposure should be included in the list. The high mechanical precision of the Arndt-Wonacott oscillation camera permits use of partially recorded intensities i.e. the two parts of a reflection measured on successive oscillation photographs from the same crystal can be added to give the correct intensity. This is of great importance for very large unit cells.

Integrated intensities

It is sufficient for most purposes to use for integration a rectangle consisting of $n_y \times n_z$ raster points well covering the reflections (fig 6). For films with variable spot size, the integration rectangle



19	26	20	22	21	19	23	21
22	20	20	20	18	22	23	20
21	22	20	19	23	21	20	19
18	22	20	23	24	27	26	22
20	21	26	31	47	58	39	28
19	23	36	58	76	64	42	26
23	22	32	37	41	38	29	22
18	21	23	22	24	19	22	23
17	21	21	22	22	23	20	20
23	21	18	18	24	21	24	17

$\Sigma B_i = 668 \qquad \Sigma I_i = 1377$

$I_{int} = \Sigma I_i - N_i/N_b \cdot \Sigma B_i = 375$

$Var\langle I\rangle = Var\langle B\rangle \cdot [\, N_i + N_i^2/N_b \,] = 576$

$\sigma\langle I\rangle = 24$

Fig. 6. Evaluation of the integrated intensity using a rectangular integration window

may be a programmed variable. The background can be estimated from raster points closely outside the integration rectangle, and may be taken from different areas according to the appearance of the film pattern. It may be a frame, on two sides only or outside the corners of the rectangle. An example of integration is shown in fig 6. It may be noticed that $\sigma(I)$ depends only on the number of raster points used and on the intensity variation between the background values. The standard deviations of the optical density values are usually constant over the range covered by the densitometer and has values of 0.02 - 0.03 O.D. units.

When white radiation streaks run through the reflections, the background should be evaluated over the white radiation to give correct I_{int} and $\sigma(I)$ values.

Centering

The centering during normal intensity evaluation can be performed around the calculated central raster point (y_C,z_C). The optical density readings are summed in a rectangle of size $(n_y - p) \times (n_z - p)$ centered around the nine neighbouring points (y_C-1,z_C-1), (y_C-1,z_C), ... (y_C+1,z_C+1). If n_y and n_z have optimum values, p should be 2. The center of gravity of the reflection corresponds to the raster point giving the highest sum of optical densities. The shift from the calculated centre can be memorized and added to the next reflection to be measured within the same film area. This centering procedure is necessary if the calculated film positions are unrefined or refined by an approximate procedure only. It is important that the centering shift is utilized only for reflections over a certain intensity limit, since otherwise the intensity of the weak reflections will be overestimated.

The centering of the reference reflections may be performed in a similar way. However, in this case it is useful to determine the centre of gravity more accurately than in terms of whole raster units, using a 9-point parabolic extrapolation procedure.

The author wishes to thank Drs Gustaf Olsson and Lennart Sjölin for thei participation in the development of the microdensitometer system, and Dr Susan Jagner for revising the English text. Grants from Knut and

Alice Wallenberg's Foundation and the Swedish Natural Science Research Council (NFR) are gratefully acknowledged.

References:

1. Dornberger-Schiff,K.(1966). Acta Cryst.A24,390.
2. Buerger,M.J.(1942). X-ray Crystallography, John Wiley & Sons, N.Y.
3. Buerger,M.J.(1960). Crystal Structure Analysis, John Wiley & Sons, p.120-125, New York.
4. Buerger,M.J.(1964). The Precession Method, John Wiley & Sons, N.Y.
5. Lenner,M. and Lindqvist,O.(1978). To be published.
6. Arndt,U.W. and Wonacott,A.J. (1977). The Rotation Method in Crystallography, North Holland Publishing Company, Amsterdam.
7. Xoung,N. and Freer,S.T.(1971). Acta Cryst.B27,2380.
8. Schwager,P., Bartels,K. and Jones,A.(1975). J.Appl.Cryst.8,275.
9. Abrahamsson,S., Aleby,S. and Innes,J.(1968). J.Sci.Instrum. Ser.2,1,869.
10. Kratky,O.(1930). Z.Kristallogr.76,261.
11. Kratky,O.(1930). Z.Kristallogr.76,517.
12. Kulpe,S. and Schulz,B.(1972). Kristall und Technik,7,463.
13. Wyckoff,H.W. and Agard,D.(1977). In the Rotation Method in Crystallography, p.187.
14. Kohlbeck,F. and Hörl,E.M.(1976). J.Appl.Cryst.9,28.
15. Werner,P.-E.(1976). J.Appl.Cryst.9,216.
16. Malmros,G. and Werner,P.-E.(1973). Acta Chem.Scand.27,493.
17. Sonnenveld,E.J. and Visser,J.W.(1975). J.Appl.Cryst.8,1.
18. Rietveld,H.M.(1969). J.Appl.Cryst.2,65.
19. Malmros,G. and Thomas,J.(1977). J.Appl.Cryst.10,7.
20. Berg,J.-E. and Werner,P.-E.(1977). Z.Kristallogr.145,310.
21. Marinder,B.-O.(1978). Chemica Scripta, In press.
22. Lagerkvist,U., Rymo,L., Lindqvist,O. and Andersson,E.(1972). J.Biol.Chem.247,3897.
23. Strandberg,B. and Waara,I.(1978). Private communication.
24. O'Connell,A.M.(1967). Acta Cryst.23,623.
25. Morimoto,H. and Ueda,R.(1963). Acta Cryst.16,1107.
26. Matthews,B.W., Klopfstein,C.E. and Colman,P.M.(1972). J.Phys.E: Sci.Instrum.5,353.
27. Xoung,N.(1969). J.Phys.E: Sci.Instrum.2,485.
28. Wooster.W.A.(1964). Acta Cryst.17,878.
29. Sjölin,L., Olsson,G. and Lindqvist,O.(1978).To be published.
30. Nelder,J.A. and Mead,R.(1965). Computer J.7,308.
31. Leijonmarck,M., Rönnquist,O. and Werener,P.-E.(1973). Acta Cryst.A29,461.

ELECTRONIC AREA DETECTORS FOR SINGLE-CRYSTAL DATA COLLECTION

U.W. Arndt[*]

Summary

Of the various types of area detectors proposed for recording X-ray diffraction patterns multi-wire proportional chambers and television detectors are the most promising. The characteristics of these detectors are compared and their mode of use for automatic data collection is discussed.

1. Introduction

The traditional single-crystal diffractometer offers an efficient method of measuring the intensities of X-ray reflexions only from crystals with relatively large reciprocal lattice dimensions : during data collection from such crystals it needs to 'look' only at those points in reciprocal space where significant information is known to be present. However, when the direct lattice is large, or when short-wavelength radiation is employed, reciprocal space is densely populated and many reciprocal lattice points lie on or near the Ewald sphere simultaneously. Table I shows the number of reciprocal lattice points which cross the Ewald sphere during a 1° rotation of the crystal as a function of unit-cell edge and desired resolution (minimum spacing to be measured), for CuKα and MoKα radiation. In single-crystal diffractometery the crystal is commonly turned through about 1°

Table 1 Number of reciprocal lattice points which cross the Ewald sphere during a 1° rotation of the crystal

	Minimum spacing	Unit Cell Edge						
		10Å	20Å	30Å	40Å	50Å	60Å	100Å
CuKα	3.0Å	0.9	6.9	23	55	110	190	860
	2.25Å	2.0	16	53	130	250	430	2000
	1.5Å	6.4	51	170	410	800	1400	64000
MoKα	3.0Å	9	70	240	560	1600		
	2.25Å	20	160	540	1300	2600		
	1.5Å	65	520	1700	4200	8000		

[*] MRC Laboratory of Molecular Biology, Cambridge, England

in order to measure the peak-intensity of a given reflexion over the full reflecting range and to record the background on both sides of the reflexion; the table, therefore, shows the factor by which the efficiency of data collection could be improved by a detector capable of recording all these reflexions.

Photographic data collection, utilising the screenless rotation or precession methods (Arndt & Wonacott, 1976) offers a more efficient procedure especially for macromolecular crystals with CuKα radiation. For MoKα X-rays, film is relatively inefficient. However, apart from a number of secondary disadvantages, film has one basic disadvantage when used to record a changing pattern such as that from an oscillating or rotating crystal: all temporal information in the pattern is lost; thus, on an oscillation or rotation photograph, for example, it is not possible to determine the precise ϕ-setting of the crystal at which a given number of X-ray quanta arrived in a given area of film. In order to keep to a minimum the total number of photographs required for complete three-dimensional data collection the ϕ-range of any one photograph is usually chosen to be several times greater than the reflecting range for most of the spots on the film. Consequently, an area of the film where a diffraction spot is recorded has been accumulating 'background' quanta during the whole of the exposure period while 'signal' quanta, belonging to the diffraction spot, have arrived only during the time corresponding to the reflecting range of the crystal. The signal-to-noise ratio, and thus the statistical precision, are, therefore lower, often by a large factor, than in an 'ideal' diffraction experiment.

In an ideal diffraction experiment one would record, for every detected quantum, in addition to the spatial co-ordinates of the point of detection, also the ϕ-position of the crystal at the moment of diffraction; one might also want to discriminate against quanta lying outside a defined energy range so as to allow the use of powerful sources of white radiation without having to employ a monochromator.

In principle electronic area detectors can be designed to perform this ideal diffraction experiment.

2. Area Detectors

A satisfactory electronic area detector must have the following properties.

a. The spatial resolution must be adequate for the pattern to be recorded.

b. The absorption efficiency must be close to unity.

c. The statistical precision of its output must be as close as possible to that set by quantum counting statistics : if the device produces an analogue output which is digitised the detector must introduce a minimum of noise so that it has a high Detective Quantum Efficiency (Jones, 1958).

d. The detector must be capable of dealing with a wide range of counting rates, i.e. it must have a large dynamic range.

The following are among the area, or position-sensitive, X-ray detectors which have been proposed:

a. Multi-wire proportional chambers.
b. Slow-scan television detectors.
c. Fast-scan television detectors.
d. Charge-coupled, self-scanned arrays.
e. Micro-channel plates.

Only the first three of these have been tested in practice for diffraction studies and only the first has been used successfully for actual single-crystal studies, by Xuong and his co-workers (Cork et al, 1975). We shall first discuss the less likely detectors.

Charge-coupled devices can serve as solid state X-ray-sensitive television cameras (Lowrance, 1978), but in spite of their promise for other applications they are unlikely to see use for single-crystal studies because of their small size (∿10 mm x 10 mm) which will probably not be increased in the future.

Micro-channel plates have been developed as imaging X-ray detectors, especially for use on board space-craft (Schnopper, 1978) but they have an unacceptably low efficiency (<5%) for CuKα and harder radiations.

The only real contenders, at present, therefore are either multi-wire proportional chambers (MWPCs), or else phosphors which convert the X-ray diffraction pattern into a light image which, in turn, is read out by slow- or fast-scan TV cameras. We shall discuss these devices further below, after considering area detector characteristics in general terms.

3. Resolution

The resolution of an area detector is determined by the physics of the detection process, by the practical limitations of what can be constructed and by the complexities of the read-out circuitry and the size of the required store.

Thus in a MWPC the precision with which an electron avalanche produced by an X-ray quantum can be determined depends on the range of ions in the filling gas and on the spacing of the cathode wires. In a polycrystalline phosphor screen the size of the scintillation is determined by the distance over which the light photons produced by the X-ray quantum are scattered. If the resultant light image is viewed by a television camera the resolution may be further degraded by the limitations of the spatial resolution of this camera.

The resolving power of the detector for neighbouring diffraction spots depends on the angular resolution, that is on $\frac{r}{s}$, where r is the linear dimension of the resolution element and s the crystal to detector distance. As a very rough rule of thumb it may be assumed that a diffraction spot must occupy an area of at least 2 x 2 resolution elements and that the space between neighbouring spots must be at least equal to the spot diameter. Then, if the length of the unit cell of the specimen crystal is L and the minimum spacing which is to be measured is d_{min} the number of orders which must be resolved is $\pm \frac{L}{d_{min}}$ or $\frac{2L}{d_{min}}$ across the dtector diameter. The total number of resolution elements required is thus approximately $\frac{8L}{d_{min}} \times \frac{8L}{d_{min}}$ and the diameter of the detector must be $\frac{8Lr}{d_{min}}$.

The difficulties of constructing the detector itself and of interfacing it to read-out circuits as well as the cost of the electronic circuitry increase rapidly as the number of resolution elements increases. At the time of writing it would be rash to contemplate a detector with more than 512 x 512 = 2^{18} resolution elements.

Xuong's MWPC (Cork et al, 1975) and the television detector described by Arndt & Gilmore (1976) have, respectively 256 x 256 and 300 x 300 resolution elements. They can, therefore, resolve about ±32 orders and ±37 orders.

While the detectors have very similar angular resolution they have quite different linear dimensions. In a gas counter r ≃ 1 mm and thus the whole counter must be about 250 mm in diameter. To collect data out to 3 Å resolution such a counter would be used at a distance of about 200 mm from the crystal. A higher resolution is possible only by increasing the physical size of the detector and a 1 m square detector, built up from four smaller chambers is near completion at La Jolla (Xuong, private communication).

In a polycrystalline phosphor screen, on the other hand r is about 150 μm. and adequate angular resolution is possible with a screen less than 40 mm in diameter. A detector of this size has been successful for the recording of muscle and fibre patterns in the hands of the Princeton group (Reynolds, 1978; Milch, 1978; Gruner, 1977). For single-crystal measurements the signal-to-background ratio improves as the distance between crystal and detection plane is increased; in Arndt & Gilmore's television detector the phosphor is deposited on an 80 mm diameter demagnifying image intensifier which is coupled to the television camera tube. In these television detectors the resolution is limited by that of the camera tubes and of the electronic circuits.

4. X-ray Absorption

Usually MWPCs are filled with xenon to atmospheric pressure. In view of the large window areas of these detectors high-pressure chambers are extremely difficult to engineer (Sobotka, 1978). The La Jolla detectors have an effective depth of 0.8 cm which leads to an absorption of 60% in the gas for CuKα X-rays and 11% for MoKα radiation. Reasonably efficient gas detectors can, therefore, be constructed quite readily for the former, but not for the latter radiation. It should be noted that a long gas column leads to a broadening of the diffraction spots at oblique incidence : gas chambers, therefore, lend themselves best to studies on relatively large crystals with relatively large primary X-ray beams.

Solid state phosphors in quite thin layers (~ 12 mg cm^{-2} for ZnS) can readily absorb $\sim$ 60% of either CuKα or MoKα radiation. Note that the high absorption for hard radiation leads to a poor discrimination between CuKα radiation and radiation of half the wavelength and requires the use of monochromators which suppress the $^{\lambda}/_2$ component.

5. Counting Rate Considerations

The biggest problem in the application of area detectors in single crystal studies is set by the very high counting rates which occur even with conventional X-ray tubes, let alone with synchrotron radiation sources which can be expected to be utilized on an increasing scale in the next few years (Rosenbaum, Holmes & Witz, 1971 ; Doniach, Lindau, Spicer & Winick, 1975; Phillips, Wlodawer, Yevitz & Hodgson, 1976).

Somewhat surprising is the fact that the diffraction pattern even from a well-ordered single crystal contributes a relatively small proportion to the overall counting rate seen by the detector: in protein crystal diffraction patterns, for example, about 90% of the counts are in the background and are due to incoherent scattering, and scattering from the disordered part of the specimen crystal, from the mother liquor and from the specimen tube. In a typical diffraction pattern from such a crystal, obtained with a rotating anode tube the total background counting rate amounts to $\sim 10^5$ quanta sec^{-1}, summed over the area of the detector (Arndt & Wonacott, 1976).

In estimating the counting losses of an area detector it must be realized that there are two different types of dead-time which lead to quite different losses. Most detectors, including all multi-wire proportional chambers, and their electronic circuits, are inactive for a time τ after every detected event no matter where it occurs in the detector. If the overall counting rate for the whole of the detector is N counts per second and the rate of arrival of quanta in one particular counting element is n counts per second, the fraction of each second during which the system is inactive will be $N\tau$ and the number of counts lost in the element under consideration will be $nN\tau$. Consequently the proportion of lost counts in the element will be $N\tau$ which is independent of n and constant, as long as N does not vary. As we have seen above, in large-molecule crystal patterns N is largely made up of background counts and so can be regarded as virtually constant. (This is not necessarily true in small-molecule patterns). Up to moderate counting rates, therefore, the counting losses in this type of counter do not affect the relative values of the spot intensities in some patterns.

However, to prevent saturation of the counter, $N\tau$ must not exceed 0.2 or 0.3: with typical MWPC dead-times of 1 to 3 μsec the upper limit of the counting rate is about 10^5 sec^{-1}.

Much higher overall counting rates are possible if the individual elements of the detector are independent of one another so that the dead-time, and thus the counting losses in a given element depend only on the counting rate in that particular element. This situation arises in a MWPC with an individual read-out amplifier on each wire which, in turn is connected to its own buffer counter (Charpak, 1977; Baru <u>et al</u> 1977). Counting rates of 10^6 can then be achieved (Bolon, 1978). Even faster counting is possible with television detectors (Arndt & Gilmore, 1978).

6. Operation of MWPC

A MWPC is a gas-filled enclosure which contains a central anode plane between two parallel cathode planes one of which may be the flat entrance window of the chamber. The anode and at least one of the cathodes consist of a series of parallel wires which are horizontal for one plane and vertical for the other. An X-ray quantum entering the chamber ionises the counting gas; when the resultant electrons have drifted into the high-field region near the anode they produce further ionisation by collision with the gas molecule. The drift of the positive ions away from the anode produces a negative pulse on the anode : simultaneously positive pulses are induced on a few of the cathode wires which are nearest to the original absorption event. The amplitudes of these pulses falls off with distance from the event so that the centre of gravity of the pulses localises the event. Detection is from two sets of orthogonal wires one of which is a cathode plane; the other may be the anode plane as in the La Jolla chamber or it may be, as is more usual, a second orthogonal group of cathode wires (figure 1).

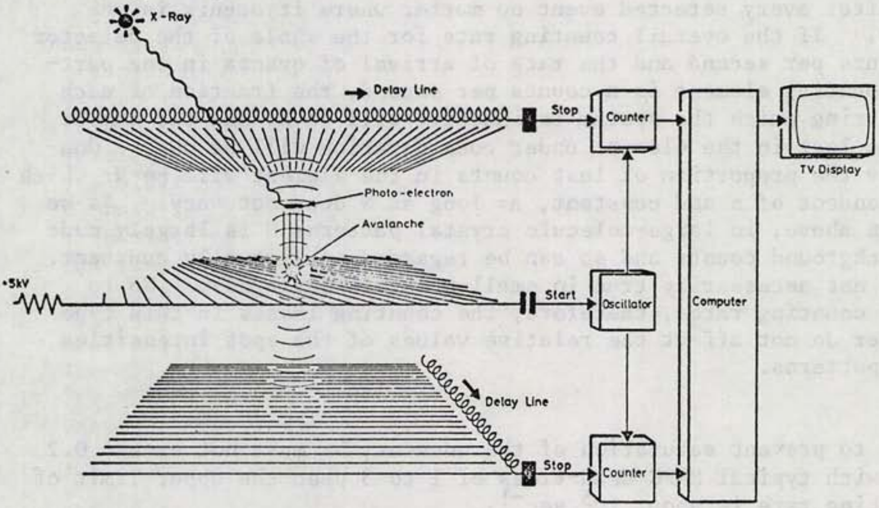


Figure 1 MWPC with Delay-Line Read-out

The pulses may be read out via amplifier-discriminators connected to individual wires or groups of wires into separate buffer counters (Charpak, 1977). Alternatively, the wires may be connected to a delay-line so that the time of arrival of a pulse at the end of the delay line is a measure of its position. The 'prompt' pulse, which starts the timing, may come from the anode as in figure 1 or it may be the undelayed pulse from the core of the delay line (Cork et al, 1974).

In a direction orthogonal to the anode wires the spatial resolution of the chamber is determined by the wire spacing; in a parallel direction the centre of gravity of the pulse group can be determined with an uncertainty which is less than the wire separation.

The temporal resolution of MWPCs has already been discussed. It should be noted that in the delay-line read-out method it depends on the oscillator clock frequency and the read-modify-write cycle of the store. In the other method it also depends on the speed of the ultimate store and on how many wires are grouped together. In either method the complexity and cost of the circuitry rises steeply with counting rate capability.

7. Television Detectors

In television detectors the diffraction pattern is formed on a phosphor which is coupled, either directly or via an intermediate image intensifier, to the photocathode of a television camera tube. The number of light photons which emerge from one side of a polycrystalline ZnS phosphor when an 8KeV X-ray quantum is absorbed is between 100 and 200; photocathodes produce about 0.15 photoelectrons per light quantum. At least 10 photoelectrons must be produced per X-ray quantum to ensure that the statistical fluctuations in this number has a small effect on the variance of the output signal from the TV camera. A high coupling efficiency between phosphor and first photocathode is, therefore, necessary. This can be achieved only with fibre-optics coupling and not with conventional lens objectives and certainly not with a de-magnifying optical system. The size of the input plane of the detector is, therefore, determined by the largest available diameter fibre-optics-input image intensifiers or TV camera tubes; at present this diameter is 80 mm.

It can be shown (Arndt & Gilmore 1978) that the fluctuations in the output of the television camera are produced essentially by two effects only, namely the statistical fluctuations in the number of primary X-ray quanta and the largely white electronic noise introduced by the camera tube and its associated amplifiers.

In all practical diffraction experiments the X-ray flux is such that a statistically significant number of X-ray quanta is recorded in any given picture element only in a time period long compared with a normal television frame period (20 msec in Europe, 16 msec in America). The diffraction pattern must, therefore, be integrated over many such normal frame periods. Two methods are available of carrying out this time-integration.

a. The scanning electron beam may be switched off for a prolonged 'exposure' period which is followed by a single read-out during which the output of the TV camera is digitised (Arndt, Gilmore & Boutle 1972; Reynolds, Milch & Gruner, 1977). The read-out, and the digitisation rates are preferably slow enough to be compatible with a direct transfer to a computer store (Milch, 1978).

b. The camera may be operated at normal TV rates and the image digitised at a sampling rate which then must be at least 6 MHz in order to be summed digitally into a large frame store (Arndt & Gilmore, 1976). This store must cycle faster than is possible with present-day computer stores and is most conveniently constructed in the form of a large shift register. (Figure 2).

The first of these methods is somewhat cheaper and simpler but suffers from the disadvantage, for single crystal methods, that the exposure is not continuous; with high-intensity patterns the duty cycle (ratio of exposure to read-out time) may be quite low.

The second method requires relatively expensive fast electronics and, in particular, a special-purpose frame store. Because of the large bandwidth required it is comparatively more difficult to achieve a good

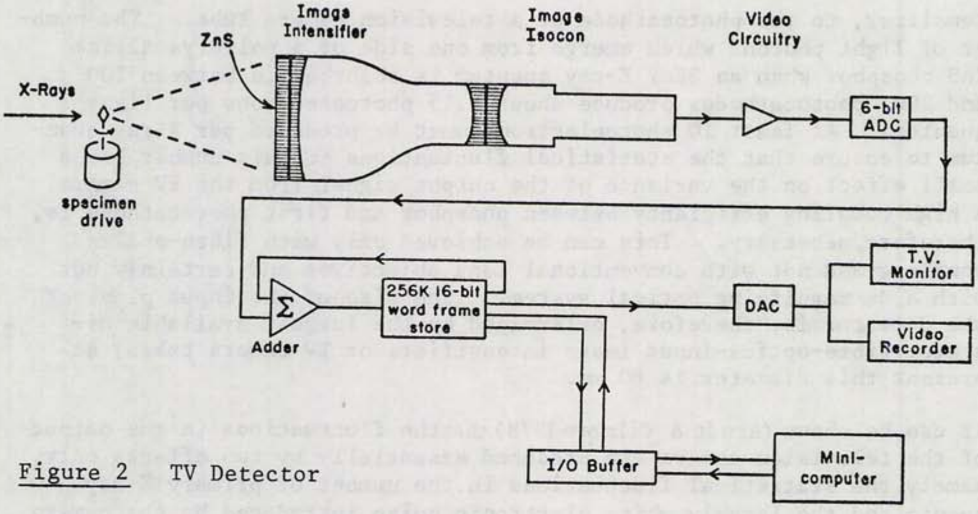


Figure 2 TV Detector

signal-to-noise ratio. However, the television camera is operated under the conditions for which it was designed and one can utilise any camera tube, selected only for the desired spatial resolution and signal-to-noise ratio, so long as an image intensifier of adequate gain is provided (Arndt & Gilmore, 1978).

8. Strategy of Data Collection

The use of an electronic area detector permits a rather simpler method of data collection than is possible with photographic film. Once the exact orientation and unit-cell parameters of the crystal are known it is possible to compute the co-ordinates of every reflexion which will be recorded during a complete rotation of the ϕ-axis. At the same time one can compute the three angles ϕ_1, ϕ_2, ϕ_3 and ϕ_4 for every reflexion, where ϕ_1 is the angle at which the reciprocal lattice 'point' is one diameter away from the surface of the Ewald sphere, ϕ_2 and ϕ_3 are the angles at which it starts and finishes its passage through the sphere and ϕ_4 is the angle at which it is once more one diameter removed from the surface (see Figure 3). The integrated intensity of the spot is the count recorded between ϕ_2 and ϕ_3 less the mean of the two background counts between ϕ_1 and ϕ_2 and between ϕ_3 and ϕ_4, summed over those resolution elements covered by the spot.

A list is computed in which each reflexion is entered four times at the four values of ϕ and the entries are listed in order of increasing ϕ-angles. The crystal is set into rotation at a constant rate and the accumulating counts are stored in a large store; whenever ϕ reaches one of the values in the pre-computed table the appropriate accumulated count is extracted from the store.

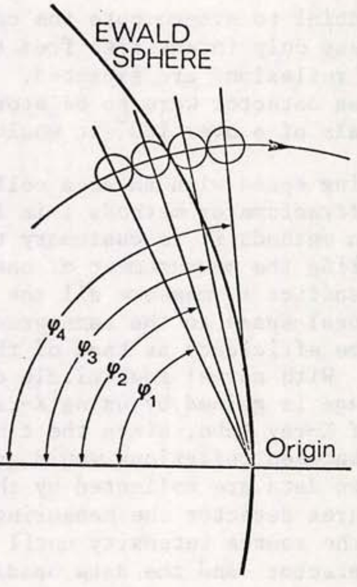


Figure 3 Recording of reflexion and background (see text)

It is quite possible to locate the centres of gravity of selected reflexions at suitable ϕ intervals and to use these to check or refine crystal orientation.

The original determination of the crystal orientation is also carried out by finding the exact co-ordinates of the centres of a number of reflexions; the refinement is very much easier than with photographic recording where the exact values of ϕ can be determined only for partially recorded reflexions.

9. Additional Applications

An electronic area detector allows the digital recording of a diffraction pattern in real time and thus the addition or subtraction of two patterns obtained at two slightly different times. This possibility offers a number of very interesting prospects. Applications to the study of changing structures would be relatively easy, especially if the structure can be made to go through regularly repeated cycles.

10. Computing Considerations

An electronic area detector must be interfaced to a mass store of some kind: the most economical as well as the most efficient way of controlling the store is by means of a computer.

It is virtually essential to pre-compute the co-ordinates of all reflexions and to process only intensities from those regions in reciprocal space where reflexions are expected. If the data from a 512 x 512 element area detector were to be stored as 8-bit intensity data at 0.01° intervals of ϕ over 180° it would occupy 5000 megabytes of storage!
The potential measuring speed with an area collector can be estimated from the speed of diffractometer methods : in large-molecule diffractometery by ω-scan methods it is customary to rotate the crystal at about 1° min^{-1} during the measurement of one reflexion. At this rate 3 hours should suffice to measure all the reflexions within one hemisphere of reciprocal space to the same precision using an area detector with the same efficiency as that of the single counter of the diffractometer. With normal four-circle diffractometers relatively little advantage is gained by using X-ray sources more powerful than a sealed-off X-ray tube, since the time taken in moving crystal and counter between reflexions would quickly exceed the actual measuring time. When data are collected by the screenless oscillation method with an area detector the measuring speed can be increased in proportion to the source intensity until the counting rate limitations of the detector and the data handling capability of the computer are reached.

The computer, then, must handle the following tasks:

a. Determination and refinement of unit cell parameters and crystal orientation.

b. Computation of the positions of all reflexions.

c. Assembly of the measured intensity ordinates into integrated reflexions.

d. Correction for spatial distortions, non-uniform response and non-linearity of the detector by means of look-up tables.

e. Correction of the measured integrated intensities for Lorentz, polarisation and absorption errors.

f. Checking for crystal movement and drift in the response of the system and, possibly applying corrections.

g. Controlling the stepping motors and input/output registers of the detection system.

h. Checking the correct functioning of the equipment, e.g. by exercising the store.

i. Providing interaction between operator and instrument.

A fairly powerful computer is needed to satisfy all these requirements.

11. Conclusions

Both MWPCs and TV area detectors have reached a state of development where their routine use for data collection from medium- and large-molecular single crystals is imminent.

References

Arndt, U.W. (1976) Ch. 17 in The Rotation Method in Crystallography Eds. U.W. Arndt & A.J. Wonacott. Amsterdam : North Holland.

Arndt, U.W. & Gilmore, D.J. (1978) Paper submitted to J. Appl. Cryst.

Arndt, U.W. & Gilmore, D.J. (1976) Adv. Electron. Electron Phys. 40B, 913.

Arndt, U.W., Gilmore, D.J. & Boutle, S.H. (1972) Adv. Electron. Electron Phys. 33B, 1069.

Arndt, U.W. & Wonacott, A.J. (1976) Eds. The Rotation Method in Crystallography, Amsterdam : North-Holland pp. 218, 222.

Baru, S.E., Proriz, G.I., Savinov, G.A., Sidorov, B.N. Shuvalov, B.N., Kabakhpashev, A.G. & Yakovlev, V.A. (1977) Internat. Conf. on Synchrotron Instrumentation & Developments. LURE, Orsay. To be published in Nucl. Instrum and Meth.

Bolon, C. (1978) Paper presented at Workshop on X-ray Instrumentation for Synchrotron Radiation Research, SSRL, Stanford University.

Charpak, G. (1977) Internat. Conf. on Synchrotron Instrumentation & Developments, LURE, Orsay. To be published in Nucl. Instrum & Meth.

Cork, C., Fehr, D., Hamlin, R., Vernon, W., Xuong, Ng. H. & Perez-Mendez, V. (1974) J. Appl. Cryst. 7, 319.

Cork, C., Hamlin, R., Vernon, W., Xuong, Ng. H. & Perez-Mendez, V. (1975) Acta Cryst. A31, 702.

Doniach, S., Lindau, W., Spicer, W. & Winick, H. (1975) Acta Cryst. A31, S232.

Gruner, S. (1977) Ph.D. Thesis, Princeton University

Jones, R.C. (1958) Photogr. Sci. Eng. 2, 57.

Lowrance, J.L. (1978) Paper presented at Workshop on X-ray Instrumentation for Synchrotron Radiation Research, SSRL, Stanford University

Milch, J.R. (1978) ibid.

Phillips, J.C., Wlodawer, A., Yevitz, M.M. & Hodgson, K.O. (1976) Proc. Nat. Acad. Sci. 73, 128.

Reynolds, G.T. (1978) Paper presented at Workshop on X-ray Instrumentation for Synchrotron Radiation Research, SSRL, Stanford University.

Reynolds, G.T., Milch, J.R. & Gruner, S, (1977). IEEE. Trans. on Nucl. Sci. NS-24, 501.

Rosenbaum, G., Holmes, K.C. & Witz, J. (1971) Nature 230, 434.

Schnopper, H. (1978) Paper presented at Workshop on X-ray Instrumentation for Synchrotron Radiation Research, SSRL, Stanford University.

Sobotka, S. (1978) ibid.

SPACE DETECTORS IN CRYSTALLOGRAPHY

G. Bassi*

Summary

A brief survey of multidetector use in neutron experiments is given, together with the use of the data processing facilities needed in these particular cases.

Introduction

Most commonly available detection devices in crystallographic experiments are based on single X-ray or neutron detectors. During the past ten years, the need for more sophisticated system has increased and some space detectors are now available. This paper will present some neutron experiments based on space multidetectors, operated at the Laue-Langevin Institute and the Nuclear Research Center, both in Grenoble, France.

Space detectors are used when simultaneous measurements may lead to a decrease in overall time spent to achieve an experimental phase. The gain in time may occur either for dynamical reasons, (metastable phases, transition studies, etc...) or for economical reasons, often important with neutron experiments (1).

More than ten years ago, some neutron diffractometers have been designed with two BF_3 detectors instead of one. The angle between the two detectors was typically 15 to 30°, thus permitting to measure simultaneously two parts of a powder diffraction diagram. The gain in time was about a factor of two.

The idea of multiplying the number of separate detector could not lead to a much larger benefit, because, as the number of counters increases, mechanical problems could not be solved easily.

Instead of multiple single detectors, the solution to the problem was found in the design of single multidetectors, where one single device is able to measure the place where the event has occured (2). It is clear that the need for mechanical motions controlled and operated during the experiment disappears in most cases, thus leading to a gain in simplicity and reliability of the experiments.

---

*
Centre d'Etudes Nucléaires de Grenoble - DRF/CPN
85 X - 38041 Grenoble Cédex (France)

The multidetector used are filled with BF_3 gas and operated on the proportional system. The geometry and filling pressure of each detector are adapted to each particular problem.

Linear multidetectors

The first use of multidetectors has been to replace the "two theta" motion by a fixed linear multidectector. This was done at the Nuclear Center of Grenoble (Siloe reactor) with a matrix of 400 elementary detectors fitted in a single curved wontainer filled with BF_3 gas (3). The electrodes are interconnected in a matrix of 20 anode channels x 20 cathode channels. This arrangement leads to 20 + 20 = 40 analog detections so that two signals appearing in due time on one anode channel Y and one cathode channel X indicate that an event has occured in the cell number 4X + Y. Some problems arise when a event occurs at the limit between two cathodes because the detector delivers more than one characteristic signal. A complementary logic is then used to eliminate parasitic signals and to make a choice between the best signals.
The general characteristics of this instrument are :
number of cells : 400
total length : 2.5 m
angular theta range : 40°
diameter of detection wires : 40 µm
distance between wires : 5.3 ± 0.15 mm
When compared to a "normal" powder diffractometer, this linear detector operates roughly with a gain of 30 in time.
It appears clearly that such a system may be used only when the 400 cells give about the same answer (number of counts) when placed under the same radiation conditions (number of incident thermal neutrons). The performance of this device has been carefully controlled with a thin collimated beam of neutrons. But such an experiment cannot been commonly performed. To determine the coefficients of each cells, and control the stability of the system, a diffraction experiment is performed using a sample, such as water or vanadium, giving a white diffraction spectrum (4).
The associated on line data processing system is a device called Plurimat, based on a french minicomputer, the MULTI 8, made by Intertechnique. This computer allows the use of 3 Direct Memory Accesses (DMA). One may be used for disc access, and the other two for particular devices such as those needed in this case. One DMA performs the +1 in Random Access Memory (RAM), using the signals delivered by the associated electronics, and the other DMA allows for continuous reading of the RAM space devoted to the current diagram to provide for live visualization. The RAM capacity is 12 Ko, permitting for four diagrams, the associated programs, and the array containing the normalization coefficients.
Programs may not be run during data acquisition phases, but only bet-

ween experiments. The experimental phase is itself under the control of a particular program, working in stand alone mode. The most important programs used are as follows :

- Data acquisition : the +1 occurs in three separate arrays containing respectively the heavy, medium and light weight of the 400 channels. This phase may be controlled either by the internal clock (time controlled experiment) or by an external signal delivered by a scaler connected to a radiation monitor (constant precision experiment).
- Print : point-out the specified diagram on the teletype.
- Punch : punch-out the specified diagram, either in ASC11 or in binary mode. If the binary mode is used, the tape will contain the image of the memory : three blocks of information containing the separate weights of the diagram.
- Determination of the normalization coefficients : these coefficients must be present when performing an experiment. They can be read from a punched tape and punched from RAM. They must be determined by comparison between a standard "white" diagram, obtained on a single detector diffractometer, and an experiment conducted with the same sample. Some parts of this diagram may be eliminated to avoid errors occuring from region of the diagram not absolutely "white".
- Diagram algebra : two diagrams may be added or substracted, each beeing multiplied by a coefficient supplied on the keyboard. This may be used to determine magnetic reflexions by diffrence between a diagram obtained at ordinary temperature and another one at low temperature, after background corrections.

This system has worked for several years and will be shortly replaced by a new linear multidetector device of 800 cells instead of 400. The distance between wires will be about 2.5 mm, thus leading to an angular spacing of 5/100° in theta space. The associated on line system will be an "IN90" supplied by Intertechnique and based on a "Multi 20" computer, providing for simultaneous calculations during experimental phases, and some more sophisticated experimental programs. The matrix detections will not be used any more.

Another linear multidetector is used at the Laue-Langevin Institute (I.L.L.), in the experiment called D1B. This experiment is very similar to the proceeding one, and provides for 400 cells at 1.5 m from the center of the sample that may be placed in a cryostat or in a high temperature device. The diagram is directly registered in a multichannel analyzer (DIDAC). This analyzer is connected to a general data acquisition system called CARINE, based on a remote T2000 computer. The system allows for the same possibilities as in the case of the Plurimat system, with more sophisticated possibilities because programming may be written in FORTRAN. Diagram outputs are made on magnetic tapes.

Two dimensional multidetectors

At the I.L.L., two experiments called D11A and D17, have been designed to use an X-Y detector device. These experiments are both intended to supply small angle scattering facilities but D17 may also be used for diffraction experiments.

The D11 small angle camera (5) is a large instrument with a total length of 80 m. The distance from the sample to the detector may be 2, 5, 10, 20 or 40 m, depending on the physical place of the multidetector vessel that can be transfered from one place to another thanks to a movable crane. The radiation used is from a cold neutron source. The neutrons are transmitted from the source through a curved neutron guide.

The multidetector is a two-dimensional matrix of 64 x 64 elementary detectors fitted in a single welded container filled with BF_3 gas. It operates now by gas amplification. The four corners of the square are not radiation sensitive : thus only 3808 individual cells are available. The surface of each cell is 10 mm x 10 mm. The sensitive volume is determined by the inner geometry : two cathodes and one anodes glass planes, each bearing 64 band-like electrodes. The detection space is thus doubled (2 x 10 mm).

The high incident data rate (up to 300 KHz) and the corresponding large spectrum width require a preliminary data reduction system to avoid overloads in the data storage system. The multidetector is connected to a coding unit transforming the 64 possible addresses in X and Y into a 12 bit address (6). This coding unit feeds a multichannel analyzer built in the CAMAC standard. An event is determined by the adress of the event in the multidetector array (12 bits) and some extra informations such as time-of-flight (7 bits) and other parameters (3 bits). Using either the symmetry of the spectrum or any other structural considerations, the 12 bit-word from the detector can often be reduced. The channels can therefore be grouped so that a +1 is performed in a corresponding word of a memory module of 4 Kwords of 16 bits capacity. The list of these addresses is written in another memory module directly fed by the associated on-line computer. Data reduction is then performed on line, in the CAMAC system, independently from the associated computer. The associated on-line computer is a DEC PdP11/40, itself linked to a PdP10. The size of the memory of the 11/40 is 80 Ko and the mass memory is made of two RK05 cartridge discs. A graphic unit is linked to the system through a 9600 bauds line.

All CAMAC functions may be realized by any program, located in the PdP11/40 (7,8), by calling the corresponding subroutines from the LIBCAM library. Most programs are available under the general name of CRESUS system. Unfortunately the PdP11/40 is rather too slow and the FORTRAN language cannot be always used except when measurements are not controlled in real-time. The data-reduction are performed only in the PdP10. All the necessary functions are available in the local computer to control the CAMAC system, dialog with the display unit and use the file management system. The disc capacity allows for not more than 100 spectra of 4 Kwords (9).

The D17 experiment is very similar to the previous one (10). The main difference comes from the multidetector that is an array of 128 x 128 cells, thus giving an array of 16K cells of 5 mm x 5 mm. The total surface of the detector is the same as in D11. A measurement phase lasts typically 1 hour. Reading the external memory to write on the disc through the file management lasts about 15 seconds. Often the resolution is reduced to 1 cm^2 by grouping four neighbouring cells into one. Each measurement is made twice : with and without the sample ; this leads to determine accurately the center of the of small angle diffraction rings.

The normalization of the cells can be performed thanks to a measurement of a diagram obtained with an isotropically diffusing sample, such as water or plexiglass. The spread on 16K cells leads to a poor accuracy of each count rate. Therefore the cells are gathered four by four, in a 4K array where a typical number of counts is 5000 in 8 hours. A correction is then applied if necessary, in the hardware (voltage, gains, etc...) but not by program. This operation must be done much more often than in the case of linear detectors.

This experiment will also be used for diffraction (8). This needs of a special program in the PdP11 to perform measurements knowing data on existing reflexions. The mode is Omega scan only. The program will run the following steps : go to a given Omega value ; one or more reflexions occurs (max. 10) and the measurement is performed ; keep only a square region around the computed X-Y position of each reflesion ; integrate over the smallest possible zones. It is hoped to keep the sample under radiation not more than 30 minutes for a diffracting volume of 1 mm^3. The background will be determined by measurement of one point left and one right out of the reflexions.

The data from D11A and D17 are stored normally on magnetic tapes. These data can be treated in the PdP10 for reduction or further computing (11). Some standard programs are available. At some particular stages the data may be plotted, listed, or punched on cards or written on magnetic tapes for further use on any other computer. The programs are divided into three classes :

- Raw data programs : listing of the raw detector data cell by cell, and filtering of cell contents out of given thresholds.
- Radial distribution programs : the purpose is to determine a radial distribution function I(R) from an experimental map. The Central cell is given by the user. The programs determine some other quantities such as the average intensity of the 25 central cells, that is usefull in the case of transmission measurement. Some cells may be ignored by using different filters. A background correction can be applied.
- Analysis of anisotropic scattering data : the system handles 120 spectra each of 4096 values of data. Arithmetic operation can be performed on spectra (addition, division). Contours and printings are possible. Cells can be eliminated, integrations over a chosen segment at a selected angle to the detector axes are possible. Are also available programs for fitting to a given function, and for partially integrating the spectra in different ways.

Neutron multidectectors can be used only in an environement permitting large data processing facilities, both on-line and off-line. X-rays multidetectors are now available too, and the problems are not fundamentally different, except for background corrections that requires a much more sophisticated process. Replacing mechanical motions by multidetector devices increases always speed and reliability of the experiments provided these detectors can remain stable for long periods of measurements.

References

(1) R. Allemand, J. Bourdel, E. Roudaut, P. Convert, K. Ibel, J. Jacobe, J.P. Cotton, B. Farnoux, Nucl. Instrum. Meth. 126 (1975) 29-42.

(2) J. Jacobe, Report on a discussion meeting on applications of position sensitive detectors to thermal neutron scattering, Institute Laue-Langevin, Janvier 1975.

(3) P. Lecuyer, Nouveau diffractomètre pour neutrons. Multidétecteur LMT LC 400CP, Revue technique Thomson CSF (1977) vol. 9 n° 4.

(4) E. Roudaut, Private communication.

(5) K. Ibel, J. Appl. Cryst. (1976) 9, 296.

(6) R. Klesse, G. Kostorz, Crystallographic Computing Techniques, Munksgaard, Copenhagen, 1975.

(7) M. Grevaz, Automatisation des expériences D11 et D17, I.L.L., 76G28T.

(8) M. Grevaz, Utilisation de Dinosor, 76G18J

(9) M. Goeltz, Private communication.

(10) M. Roth, Private communication.

(11) R.E. Ghosh, G. Göltz, J.S. Higgins, G. Kostorz, I.L.L., 76G124T.

THE IN-HOUSE CRYSTALLOGRAPHIC LABORATORY USING A RESOURCE SHARING COMPUTER SYSTEM

Y. Okaya*

Summary

The use of a mini-computer as the center of an in-house crystallographic laboratory was discussed with special reference to recommended hardware and executive software features. A system based on the Stony Brook in-house system was presented as an example. The structure of an ideal diffractometer control program was discussed in view of the ease with which modifications be made on the control strategy. Additional benefits the Stony Brook system provides to the Department were pointed out.

I.Introduction

The cost-performance ratios of mini-and micro-computers have improved drastically in the recent years. It is now possible to make use of relatively small local computers, either by themselves or in hierarchical structures, to carry out tasks which were, a decade ago, only able to be executed on centralized large systems. This development and efforts made by instrument manufacturers as OEM (original equipment manufacturer) users have enhanced the further exploitations of the potential of small computer systems.

In view of these developments, it seems worthwile to make critical analysis of the roles and utility values of such computer systems, especially with respect to the crystallographic tasks.

II The crystallographic task

II-I The data collection

Before going further into discussions of crystallographic applications, it is worthwhile to discuss a method of classifying various experiments with respect to their interactions with the computer.

The first consideration to be made is whether an experiment is time-dependent or not. In general, one can classify an experiment into one of the two categories. The first, the demand-response type, works in such a way that the subsequent step of the experiment must await operation complete of the present step. In this case, the performance of the experiment depends on the operational speeds of various components. Many efforts are made for such experiments to preserve

---

* Chemistry Department, State University of New York,
Stony Brook,11794, U.S.A.

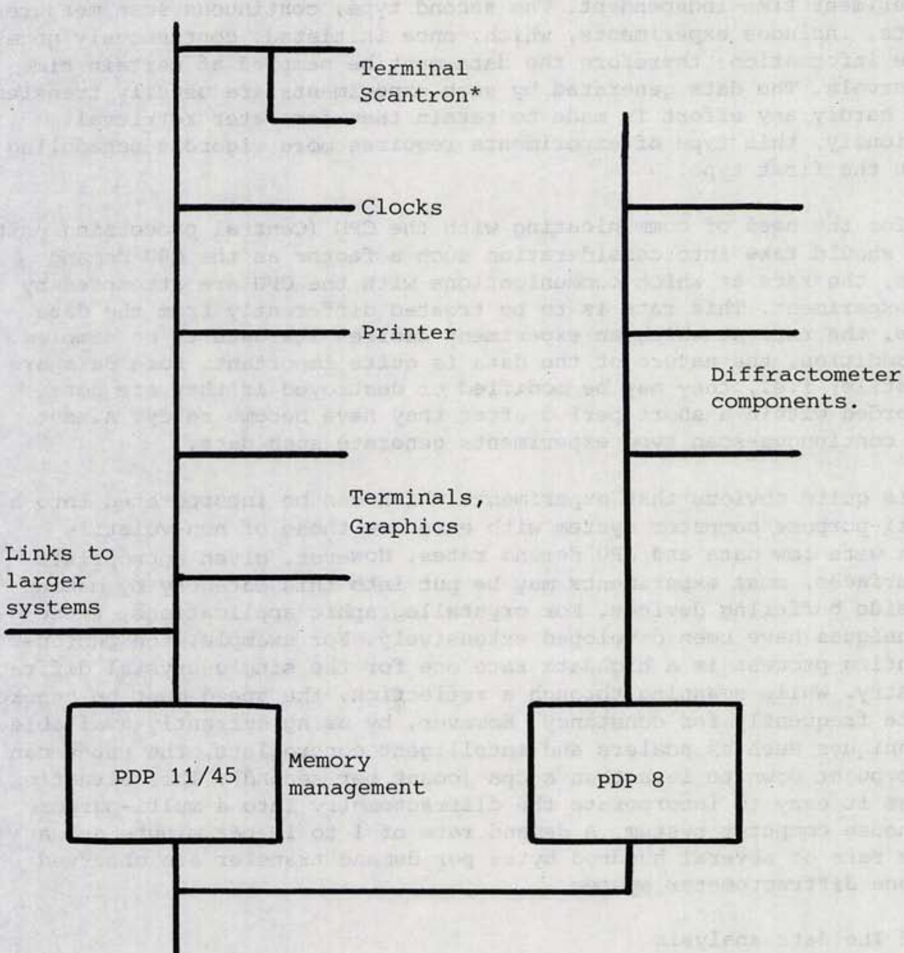
data so that any transient condition may be saved, thereby making the experiment time-independent. The second type, continuous scan measurements, includes experiments, which, once initiated, continuously generate information; therefore the data must be sampled at certain time intervals. The data generated by such experiments are usually transient and hardly any effort is made to retain them for later retrieval. Obviously, this type of experiments requires more rigorous scheduling than the first type.

As for the need of communicating with the CPU (Central processing unit), one should take into consideration such a factor as the CPU demand rate, the rate at which communications with the CPU are attempted by an experiment. This rate is to be treated differently from the data rate, the rate at which an experiment desires its data to be sampled. In addition, the nature of the data is quite important. Some data are volatile, i.e., they may be modified or destroyed if they are not recorded within a short period after they have become ready. Almost all continuous-scan type experiments generate such data.

It is quite obvious that experiments which can be incorporated into a multi-purpose computer system with ease are those of non-volatile data with low data and CPU demand rates. However, given appropriate interfaces, most experiments may be put into this category by making outside buffering devices. For crystallographic applications, these techniques have been developed extensively. For example, the photon-counting process is a highdata rate one for the single-crystal diffractometry. While scanning through a reflection, the speed must be checked quite frequently for constancy. However, by using currently-available techniques such as scalers and intelligent controllers, the rates can be brought down to less than a cps (count per second). This situation makes it easy to incorporate the diffractometry into a multi-purpose in-house computer system. A demand rate of 1 to 10 per minute and a data rate of several hundred bytes per demand transfer are observed in one diffractometer system.

II-2 The data analysis

The nature of crystallographic data analysis tasks is well defined. The tasks include frequent executions of large and long programs (mainly least-squares analyses of atomic parameters), interleaved by other short programs including house-keeping utility tasks such as file manipulations. Almost all computational programs make extensive use of floating hardware. Except for certain portions of least-squares programs, the tasks are heavily I/O (input-output) oriented. In order to expedite the performance of an in-house computing system, it is quite important for the executive system to provide certain services as will be discussed below. For example, a long-lasting job (such as least-squares) can be temporarily interrupted when it is desirable to run a short program which requires terminal operations, and therefore human presence. The interruption of a long-lasting job might actually involve preempting it to an auxilliary storage medium. A pro-



*OMR (optical mark reader)

FIG.1.

gram which resides in the core memory but is inactive can be activated when a running program (which coexists in the core) is forced into a dormant stage mainly due to I/O delays. With enough memory, the system can execute more than one long-lasting program by making each program into active state using the I/O delay or time-slicing round-robbin schemes.

III The in-house crystallographic system : an example

At this stage, it seems desirable to introduce one in-house crystallographic laboratory system based on the Stony Brook system. Although the Stony Brook system does not possess all the hardware as shown in Fig.1, it seems worthwhile to add planned future additions to the system of the current configuration. Those anticipated additions are placed in parentheses.

The disposition of the core memory is shown in Fig.2. As one can see from Fig.1, the central PDP-11/45 possesses a memory-management hardware system and most of the programs are dynamically loaded in a large area and given opportunities for their executions under various priority schemes.

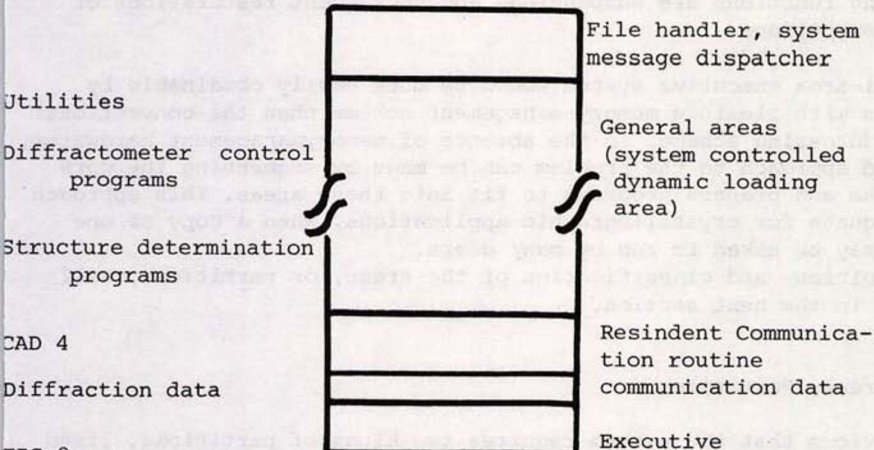


FIG.2

IV Useful in-house crystallographic systems: suggestions

IV-1 The executive system

The requirement for useful crystallographic computer system can be summarized as follows.

1. More than one active program can be resident in core memory at one time.

2. Existence of either interrupt-handling routines or resident routines which can be used as communication directors to outside instruments.
3. Programs of higher priority can be introduced into core when those with lower priority are residing. In case no appropriate space is available for the high-priority programs, the low-priority program must be preemptied to auxiliary storages.
4. Programs resident in core at the same time must be given chances for becoming to active states according to priority levels which might be modified by other programs, statuses of I/O operations (or flag-waiting conditions) of other programs, and some priority-shifting schemes.

The above requirement may be satisfied by executive systems which contain all or some of the features described below.

IV-1-a The multi-area executive system

The important requirement for multi-purpose in-house systems is that the executive system must be a multi-area, multi-programming system so that many active programs are able to reside in core at the same time. These active programs, using various executive options, carry out many independent and unrelated tasks. Particulary important program sequencing functions are suspensions and subsequent restorations of program executions.

The multi-area executive system would be more easily obtainable by computers with flexible memory-management scheme than the conventional direct addressing scheme. In the absence of memorymanagement hardwares, a limited approach to the problem can be made by segmenting the core into areas and prepare programs to fit into these areas. This approach is inadequate for crystallographic applications, when a copy of one program may be asked to run by many users.
The definition and classification of the areas, or partitions, will be given in the next section.

IV-1-b Areas, Partitions

It is obvious that the system requires two kinds of partitions, fixed and dynamic partitions. The main attribute of fixed partitions is that only one task can be executing in the partition at one time regardless of memory size and requirement. This kind of partitions is used mainly for resident routines which communicate with outside random events and associated data areas. As is evident from the nature of the routines their presences must be guaranteed when the outside event occurs. In some systems, one can gather frequently-used reentrant subroutines in a fixed partition so that they can be shared by many programs.

Dynamic partitions are contiguous portions of memory that are managed and dynamically allocated to programs by the executive system. As many tasks as possible are allowed to run simultaneously in a dynamic partition. Programs are allocated regions in the partitions and relocated and shuffled by the memory-management hardware. If a program with higher execution priority cannot enter the partition, lower priority residing program may be preempted.

IV-I -c Checkpointing

Checkpointing is a technique by which programs not residing in memory can gain access to the CPU (central processing unit) by given opportunities to become core residing. If a partition contains a program of lower priority and this program is checkpointable, the executive system can move the task out of memory to an auxilliary storage medium to make room for higher priority programs. (One must remember that not all the low priority programs are checkpointable. The checkpointability of a program must be assigned with great care.)When the checkpointing high priority program exits out of core, the checkpointed program is allowed to enter the partition (may not be the same region as before the checkpointing) and the processing is reinitiated from the point at which it was interrupted. This roll-out, roll-in process is called checkpointing or preempting.

With a memory-management hardware, the system can dynamically load and shuffle programs in a partition,thereby making best use of available processor resources. There is no need to emphasize that for crystallographic applications, the checkpointing service of the executive is probably the most important one in improving the turn-round condition of many short computing tasks which usually require presence at terminals of operators. These short tasks can safely interleaved into long-lasting jobs without waiting for their completions.

IV-1-d Round robbin scheduling and swapping

When more than one competing program are resident in a partition, the executive usually tends to give more time to those programs which entered earlier into the partition. This may not be too desirable in certain cases. For example, when more than one least-squares refinement job are to be run, it would be advantageous that they all finish nearly at the same time. For this purpose, the executive may provide a scheme called round-robbin scheduling. Under this option, the executive system employs a scheduling algorithm which periodically rotates programs of equal or similar priority.

A similar algorithm may also made to make programs core resident by checkpointing programs of the same (or similar) priority so that other programs of the same (or similar) priority may enter into core.

IV-1-e Priority alterations

The desirable executive features described in IV-1-a to -d are based on the proper assignment of the program and execution priority. Although the priority of a program is usually assigned at the time of its building, the actual run time priority may be dynamically changeable based on the importance, readiness and resource availability status of the program. The following short list indicates some of the frequent use of priority alteration capability.

1. Increase in priority of an experiment (diffraction) control program when the experiment demands the executions of the next stage of the operation.
2. Automatic lowering of execution priority while waiting for significant events such as I/O completions or ready signals from experiments.
3. Changein priority by programs (either by itself or by some other directive programs).
4. Periodic lowering of the execution priority of a running program so that other competing dormant programs can be given opportunities.

IV-2 Suggested hardware features

IV-2-a Central computers

The recent developments in the computer hardware are such that most of the commercially available mini-computers can serve as a central resource-sharing computer around which a successful in-house crystallographic system can be built. The following list indicates some of important desirable features which must be put into consideration in selecting computers for our purposes. It is also noted that not all the mini-(or midi-) computers are adequately equipped with such features.

1. A large number of machine registers and a LIFO (last-in-first-out) stack feature using the registers.
   In general, instructions dealing with machine registers are faster than those involving individual locations. In addition, the registers and the LIFO feature are of great help in making reentrant and position-independent programs. Ideally, a duplicated set of registers can be assigned to each priority hardware level, but, in practice one should avoid computer systems which have less than eight registers.
2. Powerful floating point hardware.
   There is no need to stress the importance of fast floating point hardware with an extensive library of instructions.
   The computer system can be designed so that such floating point operations can overlap in time with CPU operations. It is also possible to construct function generators (e.g., sine and cosine) and attach them to the bus and take advantage of parallel processings. Look-ahead programming techniques are required to make full

use of the feature.

3. Memory-management hardware systems.
   As discussed extensively in the previous sections, in order to attain the efficiency and versatility of an in-house system, one must employ a multi-area, multi-programming executive system. The operation of such system is enhanced by an incorporation of memory-management hardware systems which translate virtual addresses of a program into physical addresses by properly selecting the base address for the program. The hardware also sets a window while the program is being executed so that any possible memory access violation can be detected. All the programs can be constructed by making its base address at a constant location (e.g., 0) and they are relocated dynamically at run time with ease.
4. Interrupt structure.
   The interrupt feature, upon a receipt of a signal representing an interrupting condition, makes an automatic branch from the normal sequence to another program sequence as requested under such a condition. After the completion of the required processing, the interrupted program is allowed to resume its execution. It is quite important to have flexible interrupt-handling capability for the computer system one selects for the in-house system. One of the important features is forced saving and subsequent restoration of machine registers without programming need.

IV-2-b Small instrument control computers

After the analyses of multi-purpose in-house systems, the role of small instrument control computers (or microprocessors) has become clear. It is well known that the cost of these small computers have been reduced drastically in the recent years. Some of them, being mass-produced, are becoming less expensive than some of custom-made instrument interface hardwares. Although the cost of logic and memory portions of the microprocessor have been reduced, the cost of reliable I/0 units is still high and there is no trend of future reduction. Under these circumstances the following usage has been developped. Connect microprocessors on-line to a central computer as its satellites and treat them as intelligent areas of multiarea executive system. Then, the intelligent areas may use their own instruction repertoire and the interrupt-handling capability. When an experiment on a satellite has arrived at such a stage that communication with other areas (usually those in the central computer) or I/O units of the central computer is desired, it can do so via the connection between the central and the satellite computers. The communication may involve data transfers of program acquisition from the central computer which acts now as an algorithm distributor. Programs for the satellites may be assembled and build in the central computer and stored in its storages. The programs may be passed on to the satellites when requests are made.

The available choice of the microprocessors is wide and in deciding which one of them is to be used for a particular application, one must weight the relative ease of interfacing with the central computer. Obviously it is advantageous to select microprocessors from the same manufacturer as that of the central computer if the manufacturer can supply well tested (hardware and software wise) off-the-shelf communication systems.

The relationship between the larger and smaller computer systems is not limited to one hierarchy relation. Any number of levels can be constructed to make best use of the computer resources.

V. Experiment control programs

V-1. Organization

Since it is not possible for experiment control programs to occupy a large portion of core permanently, one must devise an experiment sequence control function which connects segments of experiment control schemes in such a way as to maintain orderly flow of logic necessary to carry out the experiment. The extent of the need for the logic depends on the nature of the experiment. For example, those of high-data rate with short duration are relatively simple in nature and may not require sequencing except for calling a data storing routine at the end of measurements. In contrast, single-crystal diffractometry requires a rather elaborate sequencing scheme. The flexibility in sequence may become an important factor if the next step to be taken depends not only on the current step but also on the performance of various parts of the apparatus as well as the data generated by them. This is important for experiments such as diffractometry that requires many components to be successfully controlled before meaningful data can be accumulated. The flexibility is also of importance when modifications to control programs become necessary to attain different methods of data collection.

In sequencing the experimental steps, two questions must be answered. The first one concerns the possible next step. The second one is the readiness for the experiment to the next stage. The experiment may physically connected to the computer at all times but becomes logically active only when certain function ordered by the control program have been completed; all other times, it is dormant and does not (and should not) interact with the computer. In essence, an operation complete condition makes the experiment active and ready for the next stage.

Several methods have been developped and used for solving the sequencing problem. Probably, the best solution so far advanced is to make use of the checkpointing service of the executive system. For example, when a diffractometer control program has reached such a stage that it issues a command to the diffractometer and then waits for the completion of the command, the program lowers its program priority so that other programs competing for core residency, if any, may preempt the control program. A later response to the command from the diffractometer is transmitted to a resident communication routine (or interrupt handler) which raises the priority of the pre-empted control program so that its residency in core and eventual execution is ascertained. Another useful technique is to maintain a set of experimental status words in an inter-program common data area. The experiment control program passes proper information to those status words before exiting from the core. The words, the contents of which may be modified by the condition of the experiment, are used by the system in selecting the next program to be run for the experiment control.

V-2 Modularity

Contrary to process-control applications or very simple data-acquisition instrumentation, where the control methods remain unchanged after their successful establishments, ways of conducting usual scientific experiments, therefore their control program, change quite frequently. Although for the single-crystal diffractometry, various data-acquisition methods have been studied extensively and certain set of reflection scanning techniques established, it is quite desirable that one can easily modify control programs in such way that a new or improved measurement method can be introduced and tried. For example, for the determinations of the precise positions of reciprocal lattice points from bent or distorted specimens, one must use different algorithm from that used for unstrained or normal crystals. Another example of program modifications is to dynamically correlate results of data analysis with data acquisition programs so that optimum utilization of diffraction equipments as well as available crystals can be made. In case one wishes to carry out other types of diffraction studies such as measurements of critical (or diffuse) scattering on the computer-controlled diffractometer, one must generate his own program, preferrably by making relatively minor modifications to the available single-crystal diffractometry programs. The shape of the volume which must be scanned for critical scattering measurements varies according to the type of structure defects involved and the segment of the control program which generates these non-integral reciprocal lattice points to be sampled may have to be written for each example.

In order to facilitate these changes and modifications involved in further studies, it is quite important to make the basic diffractometer control programs as modular as possible. Every functionally different logical segment must be confined to its own block. Also it is strongly recommended (or requested, to be more precise) that the structure of the program is well documented and available to the user. There is no need to mention the importance of the use of high-level languages such as FORTRAN for those program segments pertaining to the logical flow of the control program. All those program segments must be readily availavle to the user in the source language form.

Further discussions on active control programs and their modifications will be made at an afternoon session.

VI. Use of the in-house system

At present there exist a number of in-house systems which are successfully and efficiently utilized for crystallographic applications. The use of these systems, especially in view of various commercially available structure determination program packages will be discussed extensively at working session by many lecturers; therefore in this contribution, discussion will be limited to computing or other applications different from normal crystallographic tasks.

VI-1 Interactive graphics

One can incorporate into the in-house crystallographic system interactive graphic capability by choosing a graphic terminal which contains some intelligent controllers such as hardware rotation capability and regeneration circuitry. One existing problem is a relatively high cost of such systems and there seems to show no trend of future reduction in the cost. The usefulness of interactive graphic capability is undebatable, specially for structural studies of large biological molecules and of longrange orderings in critical phases. An excellent system has been established around a PDP-11/40 using a Vector General display terminal part of CRYSNET.

VI-b On-line connections of other experiments

A multi-channel analyser attached to an NMR(nuclear magnetic resonance) spectrometer is connected on-line to the PDP-11/45 of the Stony Brook in-house system through a parallel interface card (DR11C). Occasional continuous-scan type data acquisition tasks are attempted with the analyzer. It is also planned to connect an LSI-11 (a small or microprocessor version of PDP-11) which is driving an optical spectrometer. The main rationale for the connection is for the 11/45 to become an algorithm distributor to the LSI-11 whose programs have to be modified quite frequently to achieve flexibility in the spectroscopic experiment system. It is anticipated that the demand and data rates generated by the above additions are so small that there will be no overall degradation of crystallographic throughput of the in-house system.

VI-c Incorporation of small but long-lasting programs

There are a number of programs used by theoretical chemists here which are relatively small in size but require enormously long execution time. One such example is a Monte-Carlo method of optimizing a set of parameters describing surface reaction phenomena. Executions of this type of programs as background, low-priority, checkpointable tasks are efficient methods of mopping up remaining unused CPU time with little or no effect on overall efficiency of the whole system other than dedication of a small portion of system disk space to store programs and intermediate data. Our previous experience using another in-house system (an IBM 1800, with MPX (multi-experiment programming executive system) was quite encouraging. By proper segmentation of programs and periodic storing of key parameters and results mainly for automatic restarting after inadvertent interruptionsof processing, we were able to run a set of programs (which requires several-hundred hours) simply by granting opportunities under a priority scheme.

VI-d Use of an on-line OMR (optical mark reader) for student test scoring

This application of the Stony Brook in-house computer system is actually a simpel computational task using an on-line instrument; however, it might be a good example of possible extra services which a relatively small computer system can render to the host educational institution and might be used as a good selling point in acquiring such a computer system.

The OMR (optical mark reader) used by Stony Brook system is connected to the PDP-11/45 through an interface sharable with an ONTEL graphic terminal. Marked test sheets can be read by using a normal terminal I/O driver command. The images of the test sheets thus read are then processed by a series of scoring, listing and correlation programs developped by the Stony Brook group. There is no question about the advantage of on-line scoring of test sheets. We have tried many types of questions which are different from usual multiple-choice ones. For example, we can ask questions whose answers are logical (AND, OR or XOR) combinations of choices. Each choice may have its own merit (or demerit, in case of wrong choice) point. These new features have made substantial improvements in the way questions are composed and presented. The answer to one question may be correlated to that of another question so that the logic used by the student may be tested. Another advantage is an easy scrambling of test forms to avoid unwanted inter-student communications (euphenism for cheating).

A number of batch-stream command files have been developped so that scoring processes can be carried out by unskilled personnel. In any event, this is an interesting business application of the in-house system.

CHAPTER 5

Molecular interactions

THE DETERMINATION OF THE INTERMOLECULAR INTERACTION ENERGY BY EMPIRICAL METHODS

K. Mirsky*

Summary

A survey of empirical methods for the determination of the intermolecular interaction energy is given with special emphasis on the description of the atom-atom potential method and its uses. Methods of deriving potential parameters from crystal data are described and criteria for selecting the best potential functions are given. A table of recommended potential parameters for several atoms is provided.

1. Introduction

The deep and widespread interest in the elucidation of the nature and origin of cohesive forces in solids, liquids and gases results from the fact that physical, chemical and even biological properties of substances largely depend on these forces. During the last twenty or thirty years the main emphasis of theoretical and experimental investigations of these forces has gradually shifted from metallic and ionic substances to more complicated systems made up of large organic molecules. This has been a consequence of the more extensive practical use of such compounds as well as the influence of rapidly developing molecular biology.

Only a limited number of experimental methods are available for measuring intermolecular forces directly. Also there are considerable computational difficulties in applying quantum-mechanical theory to systems with a large number of electrons. Fortunately, however, because of the special nature of the weak van der Waals forces between molecules (which were interpreted in a consistent way about forty years ago by London[1]) rather simple models of intermolecular interactions can be constructed. This has led to the development of so-called semi-empirical methods, which are based on the conclusions of quantum-mechanical theory, for calculating the energy of van der Waals interactions. These, at present, are the most useful for understanding and even predicting the structure and properties of organic and biological molecules and crystals.

At the beginning of the era of semiempirical calculations of intermolecular forces, only small molecules were considered such as

---

*The Weizmann Institute of Science, Israel

N_2, O_2, CO_2, CH_4, each of them being treated as a point center of attraction. However, London[1] recognized that for extended molecules (e.g. polymers) the assumption of central forces cannot be valid, and Müller[2] tried to break up a paraffin molecule into force centers, considering atomic force centers as one of three possible models.

Nowadays one can still find in the literature three different approaches to the problem of intermolecular interactions ([3,4] and references therein). In the first, each of two interacting molecules is considered as one point force center; in the second, a molecule is divided into groups of atoms (CH_2, CH_3, etc.) or valence bonds, and in the third, every atom in the molecule is considered as a force center. Among those, the atom-atom approximation has proved to be the most convenient and has been most widely used.

The idea of the atom-atom approximation is supported by organic chemical crystallography[5,6]. The method of calculation of the intermolecular energy based on this approximation became known as a method of atom-atom potentials (AAP)[6]. Systematic investigation of the validity of the method for calculation of the energy of interaction between large organic molecules was initiated by Kitaigorodsky and Mirskaya[7,8], and continued by Williams[9], Momany, Vanderkooi and Scheraga[10], Warshel and Lifson[11], and others (see[12] and references therein).

It is the purpose of this paper to give an idea of the AAP method and its current development, and to provide the reader with some practical hints on its uses.

2. Model of atom-atom potentials

Underlying the AAP method are the following assumptions:

a. The energy of interaction between molecules is equal to the sum of the interactions of the constituent atoms (additivity).

b. Central forces operate between atoms which means that the interaction energy depends only on the interatomic distance, r. (A somewhat more complicated model may be introduced if necessary).

c. Generally, the interaction potential of two atoms, u(r), can be broken up into two terms, van der Waals and electrostatic

$$u(r) = \phi(r) + u^{elect}(r) \qquad (1)$$

In case of hydrogen bonding, a special hydrogen-bond term may be included in eq. (1).

d. The potentials $\phi(r)$ depend only on the type of atom, no matter in what molecule they occur and what their valence state may be (transferability). Some authors have tried to distinguish among identical atoms in different chemical groups (examples are given in Tables 1 and 2), but the results are not conclusive.

e. The $\phi(r)$ term takes account of repulsive and attractive dispersion forces. Various analytical expressions may be adopted to represent this potential. The Buckingham (6-exp) and Lennard-Jones (6-n) potentials having the forms $\phi = -Ar^{-6}+Be^{-\alpha r}$ and $\phi = -Ar^{-6}+Br^{-n}$, respectively, are most widely used. In recent publications they are often written in the generalized Buckingham-Lennard-Jones form,

$$\phi = -Ar^{-6} + Br^{-n} \exp(-\alpha r) \qquad (2)$$

(where setting either $\alpha=0$ or $n=0$, leads to the 6-n or to the 6-exp potential). Eq. (2) can be expressed in terms of alternative parameters, ε, r_o, n, λ:

$$\phi = \frac{\varepsilon(n+\lambda)}{(n+\lambda)-6}\left\{-\left(\frac{r_o}{r}\right)^6 + \frac{6}{n+\lambda}\left(\frac{r_o}{r}\right)^n \exp\left[\lambda\left(1-\frac{r}{r_o}\right)\right]\right\} \qquad (3)$$

In this equation, ε and r_o have the physical meaning of the depth of the potential well and the equilibrium interatomic distance, the other two, n and λ ($=\alpha r_o$) govern the steepness of the repulsion in the 6-n or 6-exp potential curves, respectively.

f. Adjustable parameters, A, B, α, n, or ε, r_o, λ, n are to be determined on the basis of experimental data.

g. The electrostatic term has usually been calculated as a sum of the interactions either of molecular multipoles or of point charges placed at atomic nuclei (Coulomb law). It is well known, however, that the choice of the charges assigned to each atom of the molecule raises many problems and is ambiguous even if ab-initio quantum mechanical methods are used for their evaluation. Recent progress in the experimental determination of electron density distributions from X-ray and neutron diffraction and their comparison with theoretical molecular charge densities[13] has led to the following more sophisticated, and more relaisic, approach to the calculation of intermolecular electrostatic interactions. According to the stockholder recipe of Hirshfeld[14], the molecular charge density can be decomposed into atomic fragments. Each atomic fragment is then represented by its calculated net charge, dipole and quadrupole moments localized at the atomic nucleus. The intermolecular electrostatic energy can be calculated as a sum of the interactions of such atomic charges and electrical moments. Such an approach takes account of the asphericity of the atomic electron cloud caused by intramolecular bonding. The challenging question of transferability of numerical parameters characterizing the charge distributions of atoms in various chemical groups can only be answered by detailed comparison of related molecules. Some comparisons of this sort, using both theoretical and experimental charge distributions, have already been undertaken.[14,15] Intermolecular energies calculated with three different models of electrostatic interactions (molecular moments, atomic point charges, and atomic moments) have been shown to differ appreciably for intermolecular separations shorter than the molecular dimensions.[16]

h. For hydrogen bonds, Lippincott-Schroeder[17] potentials are often used[18]. There are attempts to use a simple charge-charge Coulomb law in this case, too.[19]

3. Parameters and combining rules for $\phi(r)$ potential functions

After the validity of the AAP model had been confirmed by extensive calculations ([6,12] and references therein), the problem of deriving numerical parameters for the potential curves for various atoms of the periodic table became central for the successful applications of the method.

The atom-atom approach to the calculation of non-bonded van der Waals interactions was first used almost exclusively in conformational analysis of molecular structures.[20]

However, a more straighforward way to derive non-bonded potential curves is to fit them to the crystalline structures and properties. For crystals, the intermolecular-energy expression contains, as unknowns, non-bonded potential parameters alone (if there are no hydrogen bonds), whereas in the molecular strain-energy equation, such terms as the angular deformation energy, torsion energy, etc., containing additional adjustable parameters, must be included.

Empirical fitting of the A, B, α or n (or, alternatively, ε, r_o, λ or n) constants (eqs. 2 and 3) usually involves considerable difficulties owing to the limited number of suitable experimental data. The most readily available data are the structures and heats of sublimation of molecular crystals. Comparing calculated equilibrium crystal structures with the experimental data and lattice energies with the heats of sublimation one can get reliable coordinates for the minima of interatomic potentials (ε and r_o), but the shape of the curves (λ or n) can only be derived approximately. More detailed information on the "steepness" parameters can be obtained from lattice-dynamics data or from elasticity moduli of crystals, which depend on the second derivatives of the energy with respect to molecular positions and orientations.[21]

Strictly speaking, all the above data must be known as a function of temperature (for molecular crystals thermal-expansion anisotropy and other thermal effects may be quite significant[6]) so that they may be extrapolated to absolute zero. At the latter temperature the effects of molecular and lattice vibrations are minimized and crystal properties are mainly determined by the lattice potential energy. (Even then one should be cautious with zero-point vibrational corrections, especially for crystals with strong hydrogen bonds[6]). Unfortunately, low-temperature data for molecular crystals are still scarce. When experimental data at high temperatures are taken for comparison, the free enery of a crystal should be calculated and minimized.[6,11]

The number of numerical parameters to be fitted to experimental data depends on the form of the potential used and on the number of atom-types involved. If the 6-n potential is used, the n value is usually taken equal to 12, leaving only two parameters, ε and r_o, for each curve to adjust. The 6-12 potentials are more convenient for computer calculations. However, the 6-exp potentials, with three parameters (ε, r_o, λ) are more flexible. The exponential form of the repulsion is also more satisfactory from the experimental and quantum-mechanical points of view[22]. Currently, this latter form of the interatomic potential function is in wider use.

Generally, molecules consist of two or more types of atoms. To evaluate intermolecular interaction by the atom-atom approximation one must know the potential parameters both for the interaction of pairs of like atoms and for mixed interactions. In order to decrease the number of parameters to be fitted from the scarce experimental data, those for mixed interactions are usually determined by means of so-called combining rules which connect the parameters for pairs of unlike atoms with those for like atoms. For the two-parameter Lennard-Jones potential the "empirical" arithmetic-mean combining rule for r_o and the geometric mean rule for ε are commonly used. For the three-parameter 6-exp potential curve, several mathematical expressions have been proposed as combining rules. Often the geometric mean rule of Mason and Rice[23] is used for both attractive and repulsive terms in eq. (2), leading to $A_{12} = (A_{11}A_{22})^{1/2}$, $B_{12}=(B_{11}B_{22})^{1/2}$ and $\alpha_{12}=1/2(\alpha_{11}+\alpha_{22})$. However, various experimental and theoretical results suggest that the geometric mean rule is valid only for the total interaction energy of atoms and not for its separate terms. New combining rules for the modified Buckingham-type 6-exp potential, eq. (3), have been used,[24] which apply the geometric mean rule to the total interaction energy within the framework of the AAP model. Following from this approach are

$$\varepsilon_{12} = (\varepsilon_{11}\varepsilon_{22})^{1/2} \text{ and } \lambda_{12} \simeq 1/2(\lambda_{11}+\lambda_{22}) \qquad (4)$$

(the exact value for λ_{12} may also be calculated). The arithmetic mean is retained for r_o:

$$r_{0,12} = 1/2(r_{0,11} + r_{0,22}) \qquad (5)$$

Calculations have shown that the second derivatives of the lattice energy are sensitive to the choice of the combining rules and that the new combining rules give better agreement[24] with the elasticity data than those of Mason and Rice.

Recently, Taddei, Righini and Manzelli[25] suggested the arithmetic-mean rule for both r_o and σ, the latter being the short distance at which the potential function becomes zero (σ also characterizes the steepness of the repulsion). However, σ is believed to be comparable to r_{vdW}, the van der Waals interatomic spacing in crystals (see Section 5), and for the latter, the statistical analysis of experimental data shows preference for the geometric mean rule for mixed intermolecular contacts rather than the arithmetic mean[26].

Empirical potential parameters are generally being derived from a system of observational equations for chosen model compounds, the final parameters being selected through the least-squares-optimization procedure or its modifications.[8-11,27-31] While the same in principle, the procedures used by different authors vary both in selecting experimental data for comparison and in the initial assumptions of the theory such as the shape of the force field, the number of parameters to be fitted, the combining rules, etc. They differ also in computational methods utilized. Owing to these differences and to the strong correlation between adjustable parameters, discrepancies in the values, and even in the meaning, of the same parameters arise (see [32] for the more detailed discussion of this problem and for some examples). A survey of the literature results of parameter-fitting procedures for several atoms of periodic table is given in the next section.

4. Literature survey on potential functions for C, H, O, N and Cl atoms

A diversity of absolute values have been ascribed to the same non-bonded potential parameters by different authors. The wide spectrum of parameter values, displayed in the following tables, makes it necessary to work out some criteria for the evaluation of their relative merits. Such an attempt will be made in Section 5.

In Tables 1-5, σ, r_o, ε, λ, and n values are listed for each atom type. The r_o, ε and λ (or n) parameters are independent and characterize the potential curve completely. They have definite physical meaning (see above) and, when compared, show the difference in potential curves immediately. They are also advantageous in the parameter-fitting procedure.[25,28] The corresponding A, B, and α values can be calculated using eq. (3) or found in the original papers. In all the tables the units are: Å for σ and r_o, kcal/mole for ε (1 cal = 4.184J, 1Å=10^{-8}cm).

Van der Waals potential parameters for hydrocarbons (Tables 1 & 2)

Hydrocarbons have been most extensively investigated by the AAP method. A great variety of non-bonded potential functions for C

TABLE 1. C...C potential parameters (r_{vdW}^{CC}=3.4 ÷ 3.6Å)

No.	σ	r_D	ε	λ	n	Remarks*	Refs.
1	3.36	3.80	0.067	13.6	0	AR+AL	8
1a	3.38	3.80	0.079	14.0	0	"	28
2	3.45	3.88	0.095	14.0	0	"	9
2a	3.52	3.96	0.087	14.3	0	"	30
3	3.53	3.96	0.077	0	12	AR	33
3a	3.77	4.23	0.040	0	12	AL	33
4	3.30	3.70	0.072	0	12	AR	34
4a	3.67	4.12	0.038	0	12	AL	34
4b	3.34	3.74	0.140	0	12	CR	34
5	3.16	3.62	0.184	0	9	AL	11
5a	3.88	4.36	0.039	0	12	AL	19
5b	3.28	3.75	0.042	0	9	AM	19
5c	3.62	4.06	0.149	0	12	AM	19
6	3.42	3.85	0.092	14.3	0	AR	25

* AR-aromatic carbon, AL-aliphatic carbon, CR-carboxyl, carboxylic acid, or peptide carbon (C=O), AM-amide carbon (C=O)

TABLE 2. H...H potential parameters (r_{vdW}^{HH}=2.3÷2.4Å)

No.**	σ	r_o	ε	λ	n	$\phi_{r=2Å}$	$-(\partial\phi/\partial r)_{r=2Å}$
1	2.48	2.80	0.067	13.6	0	1.63	8.6
1a	2.45	2.80	0.030	12.0	0	0.47	2.4
2	2.97	3.37	0.010	12.6	0	1.07	4.0
2a	2.84	3.24	0.013	12.1	0	0.85	3.3
3,3a	2.46	2.76	0.026	0	12	0.90	5.6
4,4a	2.61	2.92	0.036	0	12	2.80	16.6
4b	2.40	2.68	0.061	0	12	1.35	8.9
5	3.10	3.55	0.003	0	9	0.77	3.5
5a	2.45	2.75	0.037	0	12	1.23	7.8
6	2.52	2.84	0.035	13.4	0	0.98	5.0

** The numbers in this column correspond to those in Table 1; references are the same as for Table 1

and H atoms can be found in the literature derived both from conformational analysis and from crystal data ([32] and references therein). Listed below are several of those obtained from various crystal data. Even those listed display a great diversity and show how cautious one should be in selecting parameters for practical calculations.

Van der Waals potential parameters for atoms other than hydrocarbons

Much less effort than for hydrocarbons has been made to derive potential parameters for N, O, Cl, F, S and a few other atomic species. The literature values for the first three are listed in Tables 3-5. These too are mostly obtained from crystal data ("cryst."), although some conformational ("conf.") and second-virial-coefficient ("gas") results are also included.

TABLE 3. N...N potential parameters (r_{vdW}^{NN}=3.0÷3.2Å)

No.	σ	r_o	ε	λ	n	Remarks	Refs.
1	3.64	4.04	0.226	16.2	0	cryst.	35
2	2.80	3.14	0.076	14.5	0	gas	36
3	2.76	3.10	0.205	0	12	conf.	37
4	3.06	3.44	0.079	13.0	0	cryst.	38
5	2.94	3.30	0.212	0	12	conf.	39
6	3.20	3.60	0.081	0	12	gas	40
7	3.13	3.52	0.165	0	12	cryst.	33
8	3.21	3.60	0.194	0	12	conf.	41
9	3.42	3.85	0.070	14.0	0	cryst.	42
10	3.18	3.60	0.067	13.6	0	cryst.	43
11	3.38	3.80	0.076	0	12	gas	44
12	3.41	3.82	0.066	0	12	cryst.	45
13	3.50	3.93	0.055	14.3	0	cryst.	46
14	3.50	3.94	0.167	0	12	cryst.	19
14a	3.50	4.02	0.162	0	9	cryst.	19
15	3.56	3.99	0.045	0	12	cryst.	34
15a	3.12	3.51	0.107	0	12	cryst.	34
16	3.42	3.84	0.133	13.8	0	cryst.	29
17	3.30	3.71	0.077	0	12	cryst.	47

TABLE 4. O...O potential parameters (r_{vdW}^{OO}=2.6÷3.0)

No.	σ	r_o	ε	λ	n	Remarks	Refs.
1	2.50	2.82	0.135	13.7	0	conf.	48
2	2.85	3.20	0.172	0	12	"	39
3	2.70	3.04	0.232	0	12	"	37
4	3.00	3.36	0.124	0	12	gas	40
5	2.94	3.33	0.079	13.0	0	cryst.	38
6	2.88	3.25	0.124	13.6	0	"	49
7	2.68	3.00	0.230	0	12	conf.	41
8	2.76	3.10	0.325	0	12	cryst.	33
9	2.96	3.32	0.106	0	12	gas	44
10	2.88	3.24	0.094	0	12	cryst.	34
11	2.86	3.21	0.230	0	12	"	19
11a	3.18	3.65	0.198	0	9	"	19

TABLE 5. Cl...Cl potential parameters (r_{vdW}^{ClCl}=3.5÷3.6Å, as derived from Cl...Cl contacts*)

No.	σ	r_o	ε	λ	n	Remarks	Refs.
1	3.20	3.58	0.510	0	12	gas	50
2	3.30	3.70	0.314	0	12	gas	51
3	3.44	3.86	0.244	14.0	0	conf.	52
4	3.38	3.79	0.425	0	12	cryst.	53
5	3.10	3.50	0.745	13.1	0	conf.	54
6	3.26	3.70	0.612	13.0	0	conf.	55
7	3.34	3.82	0.200	11.6	0	cryst.	56
8	3.22	3.64	0.826	12.8	0	"	57
9	3.48	3.90	0.374	14.2	0	"	58
10	3.52	3.96	0.257	13.9	0	"	27
11	3.56	4.20	0.200	9.5	0	"	31

* See [31] for a discussion of anisotropy of the chlorine atom

5. Some criteria for selecting among van der Waals potentials

When considering which of many sets of potential functions found in the literature is better for solving a particular intermolecular-interaction problem, one can be guided by the following criteria:

1) The set of potentials is more reliable and versatile if a wide experimental basis was used in the parameter-fitting procedure. Each parameter of the potential curve is particularly sensitive to certain properties.[6] Thus, the equilibrium distances, r_o, are mainly determined by measured lattice constants; the potential well depth, ε, by heats of sublimation; and the steepness parameter, λ or n, by vibrational frequencies and crystal elasticity tensors.[28]

2) If in the parameter-fitting procedure, the van der Waals parameters represented only part of the total force field, the latter including other adjustable parameters like electrostatic-charge separation[30], hydrogen-bond parameters[19], intramolecular angular or torsion parameters[11], they cannot, generally, be used separately afterwards because of the strong correlation between various terms in the total energy expression.[11,34]

3) If experimental data for different temperatures were used as observables and only potential energy was calculated, the averaged thermal effects would be incorporated into the derived potential parameters. Such potential parameters cannot be used for evaluation of various temperature effects.

4) It is evident that a reliable set of potential parameters must lead to a reasonable agreement of calculated molecular and crystal properties with experimental data. These parameters must also be

transferable from one compound to another, to the extent of the validity of the approximations outlined in Section 3. However, these are not the only requirements on the values of the empirical parameters; these values must fall in the ranges suggested by a variety of experimental data. For example, it is well known that in crystals, cohesive forces lead to a contraction of intermolecular distances relative to their values in the gas phase.[6] Therefore, one can expect that the r_o values in the potentials for "free" atoms should exceed the mean distances between the nearest atoms of adjacent molecules in crystals, r_{vdW}. Our calculations show that this effect is of the order of magnitude of 5-15%.[8,31,43,49] Further, it is hardly possible that great positive work will be done in packing the atoms into crystals; thus at the shortest observed approach between atoms their interaction energy must be close to zero.[56] This means that σ should be comparable to r_{vdW}. One can easily see that many of the potentials listed in the above Tables do not satisfy these simple criteria (typical ranges of suggested van der Waals radii are given in the headings of the Tables).

5) If the least-square procedure was used formally, great and unpredictable differences may arise in the resulting sets of parameters (cf. the 5b and 5c sets in Table 1). General physical criteria, like those just described, should be used to select the best set of potentials among those well satisfying the least-squares test.[28,31]

6) Recent analysis by White and Bovill[59] of non-bonded functions for hydrocarbons employed in molecular mechanics calculations revealed that an acceptable choice of parameters for the H...H function (and hence C...C as the two are highly correlated) is restricted by the upper limits for the energy per H...H interaction at 2.0Å (1.5 kcal/mole) and for the modulus of its first derivative, $-(\partial\phi/\partial r)_{r=2\text{Å}}$ (8 kcal/mole Å). These values are given in Table 2 for the several H...H functions.

As an example, let us use the above criteria to select from Tables 1 and 2 a satisfactory set of potential parameters for hydrocarbons. First, to be sure that the selected set of parameters is consistent with the experimental data, at least for model compounds, we should take both C...C and H...H functions from one source. Having this in mind, we compare, successively, all the σ and r_o values in Tables 1 and 2 with r^{CC}_{vdW} and r^{HH}_{vdW}. Satisfying both $r_o > r_{vdW}$ by 5-15% and $\sigma \simeq r_{vdW}$ conditions, complemented by the test of White and Bovill for H...H curves, are three sets of hydrocarbon potentials: 1a, 3 and 6. Among these, for the set 3, the steepness parameters for each of two potential curves were fixed (the functions are of the 6-12 type). The other two sets are of the 6-exp type with three parameters, ε, r_o, and λ, for each curve fitted to experimental data. According to the authors' reports, all these sets of parameters give reasonable agreement of the calculated unit cell parameters and lattice energies of model compounds with corresponding experimental quantities. However, set 1a satisfies, additionally, the elasticity moduli of model compounds extrapolated to absolute zero (to obtain this set, all the data used as observables were extrapolated to 0 K). We therefore select for hydrocarbons the 1a set as satisfying all the above criteria.

6. Table of recommended potentials

In Table 6, A, B and α values are given for the above 1a set for hydrocarbons, for several other like atoms, and for some "mixed" interactions.

The N...N, O...O, and Cl...Cl parameters in this table correspond to the curves 10 of Table 3, 6 of Table 4 and 11 of Table 5. Combining rules (4) and (5) and eq. (3) were used to obtain A, B, and α values for unlike-atom pairs.

These potentials can be recommended for intermolecular energy calculations for the following reasons. All of them were fitted to crystal-structure and heat-of-sublimation data (plus elasticity data for hydrocarbons), extrapolated, when possible, to absolute zero. The zero-point vibrational corrections to the energy were introduced when necessary. For nitrogen- and oxygen-containing model compounds (CO_2 and N_2O) the electrostatic part of the lattice energy was calculated separately using experimental molecular multipole moments, so that there were no adjustable electrostatic parameters in the force field. In other cases, the experimental data were selected for crystals of "large" (polyatomic) electrically neutral molecules without permanent dipole moments. For such crystals, the electrostatic part of the lattice energy was estimated to be a much slower function of lattice parameters than the van der Waals potential[6] and its contribution to the total lattice energy to be small.[6,31] In fact, the potentials fitted to the structure and energy data for such crystals incorporate some unknown but presumably small part due to electrostatic effects. The smallness of this effect was confirmed by the transferability of the potentials from model compounds to those not involved in the parameter-fitting procedure. Such transferability was tested for all the potential functions listed in Table 6 (see original papers and [32]). Thus, in all the above cases the van der Waals parameters were independent, being the only empirical parameters of the theory. Finally, these potential functions have been already successfully used for various structure and energy calculations.[6,31,18,16,62-65]

Conclusions

Significant progress has been made recently in the determination of the interaction energy between large and complicated molecules

TABLE 6. Atom-atom potentials, $\phi = -A/r^6 + B \exp(-\alpha r)$

Atom pair	A, kcal/mole. Å^6	B x 10^{-4}, kcal/mole	α, Å^{-1}
C...C	421	7.16	3.68
H...H	29	0.49	4.29
O...O	259.4	7.77	4.18
N...N	259	4.20	3.78
Cl...Cl	2980	0.458	2.262
F...F[60]	148	4.20	4.15
S...S[61]	2346	23.5	3.49
C...H	118	1.86	3.04
O...C	339.4	7.57	3.91
O...H	88	1.95	4.23
N...C	331.4	5.53	3.73
N...H	91	1.44	4.00
Cl...C	1055	1.67	2.94
Cl...H	322	0.456	3.07
O...N	262.9	5.79	3.97
S...C	847	9.94	3.54

by empirical methods. The AAP method has proved to be especially helpful here. Many successful applications of the method have been made. It has been used not only to calculate crystal and molecular structure and properties, but also to gain insight into adsorption phenomena occurring on the surface of solids[66], to obtain second virial coefficients for gases[65], and to calculate properties of liquids.[67]

As well as other models of this kind, the atom-atom approximation has its limits of applicability. These limits and the accuracy of the AAP method are the subjects for future studies. In any case, further development of the method and the refinement of the above potential-energy functions require improved low temperature crystal data, reliable molecular electron-density distributions and well tested minimization methods.

Examples of the use of the AAP method for calculation of lattice energies (including their electrostatic parts), equilibrium crystal structures (with emphasis on the minimization problems), various potential surfaces, effects of crystal field on molecular conformations, extended defects in crystals, etc. will be given at the working sessions.

The author is grateful to F.L. Hirshfeld and J. Sussman for reading the manuscript and for valuable corrections.

References

1. F. London, Zeits.Phys.Chem., B11, 222, 236 (1930); Trans.Far. Soc., 33, 8 (1937).
2. A. Müller, Proc.Roy.Soc. (London) A154, 624 (1936).
3. R. Mason, Molec.Cryst., Liquid Cryst. 9, 3 (1969).
4. F. Mulder, C. Huiszoon, Mol. Phys. 34, 1215 (1977).
5. A.I. Kitaigorodsky, Organic Chemical Crystallography. Consultant Bureau N.Y. (1961).
6. A.I. Kitaigorodsky, Molecular Crystals and Molecules, Academic Press, N.Y.-L. (1973).
7. A.I. Kitaigorodsky and K.V. Mirskaya, Soviet Physics-Crystallography 6, 408 (1962).
8. A.I. Kitaigorodsky and K.V. Mirskaya, ibid. 9, 137 (1964).
9. D.E. Williams, J.Chem. Phys., 45, 3770 (1966); 47, 4680 (1967).
10. F.A. Momany, G. Vanderkooi and H.A. Scheraga, Proc.Nat.Acad. Sci., USA 61, 429 (1968).
11. A. Warshel and S. Lifson, J.Chem.Phys., 53, 582 (1970).
12. A.I. Kitaigorodsky and K.V. Mirskaya, Materials Research Bulletin 7, 1271 (1972).
13. Proceedings of the Bat-Sheva Seminar on Electron Density Mapping in Molecules and Crystals, edited by F.L. Hirshfeld, Israel J. of Chem. 16, Nos. 2-3 (1977).
14. F.L. Hirshfeld, Theor. Chim. Acta (Berl.) 44, 129 (1977).
15. M. Eisenstein, to be published.
16. F.L. Hirshfeld and K. Mirsky, to be published.
17. E.R. Lippincott and R. Schroeder, J. Chem.Phys. 23, 1099 (1955).
18. J.L. Derissen and P.H. Smit, Acta Cryst. A33, 230 (1977).
19. A.T. Hagler, E. Huler, and S. Lifson, JACS 96, 5319 (1974).
20. T.L. Hill, J.Chem.Phys., 14, 465 (1946); F.H. Westheimer and J.E. Mayer, J.Chem.Phys., 14, 733 (1946); F.H. Westheimer, in "Steric Effects in Organic Chemistry", M.S. Neuman, ed., Ch. 12, Wiley, N.Y. and Chapman & Hall, N.Y. (1956).
21. K.V. Mirskaya, Soviet Physics-Crystallography 17, 51 (1972).
22. R.A. Buckingham, Trans. Far.Soc. 54, 453 (1958).
23. E.A. Mason and W.E. Rice, J.Chem.Phys., 22, 522 (1954).
24. K.V. Mirskaya, Tetrahedron 29, 679 (1973).
25. G. Taddei, R. Righini, and P. Manzelli, Acta Cryst. A33, 626 (1977).
26. Yu.V. Zefirov and P.M. Zorkii, J. Struct.Chem. 15, 102 (1974).
27. J.B. Bates and W.R. Busing, J.Chem.Phys. 60, 2414 (1974).
28. K.V. Mirskaya, I.E. Kozlova, V.F. Bereznitskaya, Phys.Stat.Sol. (b) 62, 291 (1974).
29. H.A.J. Govers, Acta Cryst. A31, 380 (1975).
30. D.E. Williams and T.L. Starr, Computers & Chemistry 1, 173 (1977).
31. K. Mirsky and M.D. Cohen, Chem.Phys. 28, 193 (1978).
32. K. Mirsky, Acta Cryst. A32, 199 (1976).
33. D.R. Ferro and J. Hermans, Jr., Liquid Crystals and Ordered Fluids, p. 259 (1970). New York, Plenum Press.
34. F.A. Momany, L.M. Carruthers, R.F. McGuire, and H.A. Scheraga, J.Phys.Chem. 78, 1595 (1974).

35. E.A. Mason and W.E. Rice, J. Chem.Phys. 22, 843 (1954).
36. P. De Santis, E. Giglio, A.M. Liquori, and A. Ripamonti, Nature, London 206, 456 (1965).
37. E.A. Scott and H.A. Scheraga, J.Chem.Phys. 45, 2091 (1966).
38. C.M. Venkatachalam and G.N. Ramachandran, Conformation of Biopolymers, v.1, Academic Press, N.Y., 1967, p. 83.
39. D.A. Brant, W.G. Miller and P.J. Flory, J. Mol. Biol. 23, 47 (1967).
40. N.G. Parsonage and R.S. Pemberton, Trans. Far.Soc. 63, 311 (1967).
41. A. Warshel, M. Levitt and S. Lifson, J.Mol.Spectroscopy. 33 84 (1970).
42. T.S. Kuan, A. Warshel and O. Schnepp, J. Chem.Phys. 52, 3012 (1970).
43. K.V. Mirskaya and V.V. Nauchitel, Soviet Physics-Crystallography 17, 56 (1972).
44. M. Oobatake and T. Ooi, Progr.Theor.Phys. 48, 2132 (1972).
45. N. Jacobi and O. Schnepp, J.Chem.Phys. 58, 3647 (1973).
46. P.A. Reynolds, J.Chem.Phys. 59, 2777 (1973).
47. A. Zunger and E. Huler, J.Chem.Phys. 62, 3010 (1975).
48. N.L. Allinger, J. Allinger, L.A. Freiberg, R.F. Czaja, N.A. Le Bel, J.Am.Chem.Soc. 82, 5876 (1960).
49. A.I. Kitaigorodsky, K.V. Mirskaya and V.V. Nauchitel, Soviet Physics-Crystallography 14, 769 (1970).
50. Y. Morino and I. Miagawa, J.Chem.Soc.Japan 68, 62 (1947).
51. T.L. Hill, J. Chem.Phys. 16, 399 (1948).
52. E.A. Mason and M.M. Kreevoy, J.Am.Chem.Soc. 77, 5808 (1955).
53. S.C. Nyburg, J.Chem.Phys. 40, 2493 (1964).
54. R.A. Scott and H.A. Scheraga, J.Chem.Phys. 42, 2209 (1965).
55. A.I. Kitaigorodsky and V.G. Dashevsky, Tetrahedron 24, 5917 (1968).
56. M.A. Neuman, Trans.Am.Cryst.Assoc. 6, 111 (1970).
57. H. Bonadeo and E.D'Alessio, Chem.Phys.Lett. 19, 117 (1973).
58. P.A. Reynolds, J.K. Kjems, and J.W. White, J.Chem.Phys. 60, 824 (1974).
59. D.N.J. White and M.J. Bovill, J.Mol.Struct. 33, 273 (1976).
60. L.F. Konshina, K.V. Mirskaya, V.M. Kozhin, and I.E. Kozlova, Prikl. i. Teor. Fizika 1974 (6), 226.
61. V.V. Nauchitel and K.V. Mirskaya, Soviet Physics-Crystallography 16, 891 (1972).
62. A.I. Kitaigorodsky and N.A. Ahmed, Acta Cryst. A28, 207 (1972).
63. K. Mirsky and M.D. Cohen, J.Chem.Soc. Faraday II 72, 2155 (1976); Chem. Phys.Lett. 54, 40 (1978).
64. Hyo-gun Kim and H. Markovitz, J.Chem.Phys. 56, 3541 (1972).
65. V.G. Dashevsky and G.N. Sarkisov, Doklady Physical Chemistry 202, 152 (1972).
66. D.P. Poshkus and A.Ya. Afreimovich, Russ.J. Phys.Chem. 42, 626 (1968).
67. M.J. Huron and P. Claverie, J.Phys.Chem., 76, 2123 (1972).

AB INITIO CALCULATIONS OF VAN DER WAALS INTERACTIONS BETWEEN MOLECULES. AN INTRODUCTION

A. van der Avoird*

Summary

The problems arising in the practical calculation of Van der Waals interactions between (small) molecules directly from their quantumtheoretical formulation are discussed and the methods for performing such "ab initio" calculations are briefly described. Quantitative results for some examples, to which the references are given in this introduction, will be presented at the summer school as an illustration. The application of the interaction potentials to the calculation of the packing and lattice dynamics of molecular crystals is demonstrated.

1. Introduction

As it was first demonstrated by Eisenschitz and London in their historical paper [1], the Van der Waals interactions which determine the packing in molecular crystals, but also many other microscopic and macroscopic properties of molecular systems [2-4], can be derived from the time-independent (non-relativistic) Schrödinger equation. They showed this by making a perturbation expansion of this equation in the operator V, where V is that part of the hamiltonian that describes the (electron-electron, electron-nuclear and nuclear-nuclear) interactions between the molecules. Their formalism is valid for a wide range of intermolecular distances since it includes the (Heitler-London) exchange interactions which yield chemical bonding, as well as the long range Van der Waals attractions [5]. Also the repulsive exchange forces between closed shell molecules which lead, together with the attractive long range forces, to the occurrence of a Van der Waals minimum are described by this formalism [5]. In most of the subsequent work the short range forces related to the overlap between the molecular charge clouds and the long range (electrostatic, induction and dispersion) forces are distinguished, mainly because *in practice* they had to be obtained from separate calculations. This work and the underlying theory is covered up to 1975 in several reviews [6-10].

2. Long range interactions

The long range forces are obtained from the conventional Rayleigh-Schrödinger perturbation theory, which differs from the original formulation by Eisenschitz and London by the assumption of non-overlapping

---

* Institute of Theoretical Chemistry, University of Nijmegen,
Toernooiveld, Nijmegen, The Netherlands

molecular wave functions. The unperturbed wave functions for a system of two molecules, A and B, can then be written as products of molecular wave functions, $\phi_a^A \phi_b^B$, and the first and second order interaction energies are given by:

$$\Delta E^{(1)} = \langle \phi_0^A \phi_0^B | V^{AB} | \phi_0^A \phi_0^B \rangle \tag{1}$$

and

$$\Delta E^{(2)} = \Sigma'_{a,b} \frac{|\langle \phi_0^A \phi_0^B | V^{AB} | \phi_a^A \phi_b^B \rangle|^2}{E_0^A + E_0^B - E_a^A - E_b^B} \tag{2}$$

when E_a^A and E_b^B denote the molecular energies (for the ground states, a=0 and b=0, and the excited states) and V^{AB} is the interaction operator:

$$V^{AB} = \sum_{i \in A} \sum_{j \in B} \frac{q_i q_j}{r_{ij}} \tag{3}$$

between all particles (electrons and nucleii) belonging to A and B with charges q_i and q_j , respectively.
Usually these formulas are further simplified by writing the multipole expansion for V^{AB}:

$$V^{AB} = \sum_{\ell_A, \ell_B = 0}^{\infty} R^{-\ell_A - \ell_B - 1} M_\alpha^{(\ell_A)} T_{\alpha\beta}^{(\ell_A + \ell_B)} (\underline{\Omega}_A, \underline{\Omega}_B) M_\beta^{(\ell_B)} \tag{4}$$

where the ℓ_Ath and ℓ_Bth rank tensors $\underline{M}^{(\ell_A)}$ and $\underline{M}^{(\ell_B)}$ stand for the 2^{ℓ_A}-pole and 2^{ℓ_B}-pole on A and B, respectively, R is the distance between the molecular centres and the $(\ell_A + \ell_B)$th rank tensor $\underline{T}^{(\ell_A + \ell_B)}$ describes the dependence of the interactions between the multipoles on the molecular orientations, $\underline{\Omega}_A$ and $\underline{\Omega}_B$. The convention of summation over repeated indices, α and β, has been used in this formula. The multipole moments can be defined as spherical harmonics [11] or as cartesian tensors:

$$M_\alpha^{(\ell)} = \sum_i q_i \, r_{\alpha_1}(i) \, r_{\alpha_2}(i) \, ... \, r_{\alpha_\ell}(i) \tag{5}$$

with r_{α_k} = x, y or z in the local molecular systems of axes

The simplification consists of a separation of the matrix elements in (1) and (2) into products of monomer matrix elements:

$$\Delta E^{(1)} = \sum_{\ell_A, \ell_B} R^{-\ell_A - \ell_B - 1} T_{\alpha\beta}^{(\ell_A + \ell_B)} (\underline{\Omega}_A, \underline{\Omega}_B)$$

$$\times \langle \phi_0^A | M_\alpha^{(\ell_A)} | \phi_0^A \rangle \langle \phi_0^B | M_\beta^{(\ell_B)} | \phi_0^B \rangle \tag{6}$$

and:

$$\Delta E^{(2)} = \sum_{\substack{\ell_A,\ell_B \\ \ell'_A,\ell'_B}} R^{-\ell_A-\ell'_A-\ell_B-\ell'_B-2} \, T^{(\ell_A+\ell_B)}_{\alpha\beta}(\underline{\Omega}_A,\underline{\Omega}_B) \, T^{(\ell'_A+\ell'_B)}_{\alpha'\beta'}(\underline{\Omega}_A,\underline{\Omega}_B)$$

$$\times \sum_{a,b}{}' \frac{\langle \phi^A_0 | M^{(\ell_A)}_{\alpha} | \phi^A_a \rangle \langle \phi^A_a | M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle \langle \phi^B_0 | M^{(\ell_B)}_{\beta} | \phi^B_b \rangle \langle \phi^B_b | M^{(\ell'_B)}_{\beta'} | \phi^B_0 \rangle}{E^A_0 - E^A_a + E^B_0 - E^B_b} \qquad (7)$$

The first order energy (6) corresponds with the classical electrostatic interaction between permanent multipole moments (dipole-dipole $\sim R^{-3}$ dipole-quadrupole $\sim R^{-4}$, etc.), which are now replaced by ground state expectation values. The second order energy (7) can be split into three terms (the prime indicates the exclusion of the ground state from the summation):

$$\sum_{a,b}{}' = \sum_{\substack{b \\ (a=0)}}{}' + \sum_{\substack{a \\ (b=0)}}{}' + \sum_{a}{}' \sum_{b}{}' \; ; \qquad (7a)$$

which represent: the interaction between the permanent multipoles on A and the induced moments on B and vice-versa, and the dispersion interaction. The induction terms correspond with their classical equivalents when the quantummechanical formula for the molecular polarizabilities:

$$\alpha^{(\ell_A,\ell'_A)}_{\alpha\alpha'} = 2 \sum_{a}{}' \frac{\langle \phi^A_0 | M^{(\ell_A)}_{\alpha} | \phi^A_a \rangle \langle \phi^A_a | M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle}{E^A_a - E^A_0} \qquad (8)$$

(and the same for B) are substituted, the dispersion has no classical counterpart.

The calculation of $\Delta E^{(1)}$ and $\Delta E^{(2)}$, although it can be carried out nowadays for not too large molecules A and B, is still not an easy job. The first problem is the calculation of the molecular wave functions, ϕ^A_a and ϕ^B_b. One can take Hartree-Fock-LCAO approximations to these (i.e. Slater determinants composed of Molecular Orbitals which are Linear Combinations of Atomic Orbitals), or even better ones going beyond the independent particle (Hartree-Fock) model, for very small molecules. The next problem, if one uses the formulas (1) and (2) for $\Delta E^{(1)}$ and $\Delta E^{(2)}$, is the accurate evaluation and handling of all the one - and two - electron integrals occurring in the matrix elements over V^{AB} (3). This problem is avoided by using the multipole expansion formulas (6) and (7), where only the molecular multiple moments and transition moments have to be evaluated. One must not forget, however, that these expansions (6) and (7) are in fact asymptotically divergent and that they can only be used in practice for rather large distances. Especially the lower moments (dipole, quadrupole) are sometimes available from experiment and the transition dipole moments can be evaluated from experimental oscillator strengths or refractive index data [9]. In most cases the experimental information is not sufficient to calculate the interaction energies to the desired accuracy, however, and the data which are required for the computation of (6) and (7) have to be obtained from ab initio calculations as well. For a typical case such as the C_2H_4-C_2H_4 interaction the calculation of the monomer (Hartree-Fock) wave functions in a sufficiently good basis of A.O.'s including atomic polarization functions (or without the latter) takes 30 (6) min. of CPU time, the computation of the interaction integrals takes 200 (45) min., while the calculation of the molecular (transition) moments takes 5 (2) min. and their interactions up to all R^{-10} terms inclusive 5 (1.5) min. These timings (on IBM 370-158) are given for the interactions in one geometry of the dimer.

Another problem we have to deal with is the infinite summation in the expressions (2), (7) and (8) for $\Delta E^{(2)}$ and α. The simplest solution to this problem is to assume that the excitation energies occurring in the denominators of (7) and (8) can be replaced by a single average excitation energy, ΔE^A_{av} in expression (8) and $-(\Delta E^A_{av}+\Delta E^B_{av})$ in expression (7). Then one can use the closure relation for the complete set of molecular eigenstates and write, for (7) in the so-called Unsöld [12] approximation:

$$\Delta E^{(2)} \simeq - \sum_{\substack{\ell_A,\ell_B \\ \ell'_A,\ell'_B}} R^{-\ell_A-\ell'_A-\ell_B-\ell'_B-2}\, T^{(\ell_A+\ell_B)}_{\alpha\beta}\, T^{(\ell'_A+\ell'_B)}_{\alpha'\beta'}\, (\Delta E^A_{av} + \Delta E^B_{av})^{-1}$$

$$\times \left[\langle \phi^A_0 | M^{(\ell_A)}_{\alpha}\, M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle - \langle \phi^A_0 | M^{(\ell_A)}_{\alpha} | \phi^A_0 \rangle \langle \phi^A_0 | M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle\right]$$

$$\times \left[\langle \phi^B_0 | M^{(\ell_B)}_{\beta}\, M^{(\ell'_B)}_{\beta'} | \phi^B_0 \rangle - \langle \phi^B_0 | M^{(\ell_B)}_{\beta} | \phi^B_0 \rangle \langle \phi^B_0 | M^{(\ell'_B)}_{\beta'} | \phi^B_0 \rangle\right], \tag{9}$$

and for (8):

$$\alpha^{(\ell_A,\ell'_A)}_{\alpha\alpha'} \simeq 2(\Delta E^A_{av})^{-1}\left[\langle \phi^A_0 | M^{(\ell_A)}_{\alpha}\, M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle - \langle \phi^A_0 | M^{(\ell_A)}_{\alpha} | \phi^A_0 \rangle \langle \phi^A_0 | M^{(\ell'_A)}_{\alpha'} | \phi^A_0 \rangle\right] \tag{10}$$

The only matrix elements which have to be calculated now are the expectation values of the moment operators over the molecular ground state wave functions. If all the components of the polarizabilities (10) are available from experiment (which is mostly not the case), one can even avoid this calculation by substituting (10) into (9). For the average excitation energy one usually takes some empirical value, for instance the molecular ionization energy, I, which leads to the well-known approximate London formula for the (orientationally averaged) induced dipole-induced dipole dispersion energy:

$$\Delta E^{(2)} \simeq -\frac{3}{2} R^{-6} (I_A + I_B)^{-1} I_A I_B\, \alpha^{(1,1)}_A \alpha^{(1,1)}_B \tag{11}$$

Although this procedure yields a rather good R^{-6} dispersion energy for vo hydrogen atoms (92% of the exact result), the results are much more rong (typically by a factor of 2 to 3) for larger systems. This is ainly due to the empirical approximation $\Delta E_{av} \approx I$, however, and the se of theoretical procedures [13,14] to estimate ΔE_{av} values which are ependent on the multipole operators, ℓ_A and ℓ'_A, and their components, and α', gives good results for $\Delta E^{(2)}$. A great practical advantage f this non-empirical Unsöld procedure [13,14] is that it can be used n ab initio calculations with much smaller A.O. basis sets, since only he ground state wave functions, ϕ_0^A and ϕ_0^B, need to be described by the ases.

he Unsöld approximation which replaces the infinite sums over transi- ion moments in (7) and (8) by a single term, is a simple example of pplying the so-called sum rules [15]. If ΔE_{av} would be taken as the owest molecular excitation energy, the expression (9) can be proved be smaller (more negative) than the exact expression (7). More sophisticated application of the sum rules can yield finite expressions or $\Delta E^{(2)}$ and α which are upper and lower bounds [16,17] to the infi- te summations (7) and (8).A similar way of approximating these in- nite sums is to use Kirkwood type [18] approximations to the first rder perturbed wave function, which in combination with the Hylleraas riation principle [15,19], yield approximate second order energies nd polarizabilities [20]. In its simplest form this procedure is quivalent to the Unsöld approximation, formulas (9) and (10), with xplicit variational expressions for the parameter ΔE_{av}.

he most direct way to calculate (7) and (8) is to replace the infinite ms by finite summations over "effective excited states" [13,21,22]. ese have no relation with the spectroscopical excited states of the lecules; they are only meant to satisfy the mathematical complete- ss relations. The Hylleraas variation principle shows that these nite sums form upper bounds to the exact $\Delta E^{(2)}$ (7) and lower bounds α (8), which can be used to optimize the wave functions of these ffective excited states". These "sum over states" methods have also en applied [13,23,24] to calculate $\Delta E^{(2)}$ without invoking the multi- le expansion, using formula (2), which has the advantage that the sults are valid for somewhat shorter distances.

ing such procedures the long range interactions between small mole- les, up to the size of pyrazine [14], have been calculated. Parti- larly interesting is the orientational dependence of this interac- on energy which can be probed sensitively by beam experiments 5] and by spectroscopic studies of Van der Waals molecules [26,27]. molecular crystals this orientational dependence leads to the oc- rrence of librational phonon modes and to the possibility of orien- tional order-disorder phase transitions [28]. Theoretically it can be eated most elegantly [22,29,30] by using spherical tensor operators r the multipole moments, instead of the cartesian tensors defined in), and applying irreducible tensor formalisms [31].

Short range interactions

en the molecular charge clouds overlap new interaction mechanisms arise which are essentially of two types. The first mechanism is caused by the penetration of the electronic charge distributions which leads to classical electrostatic forces even between neutral non-polar systems such as rare gas atoms. In general these electrostatic forces, which are not accounted for by the multipole expansion are attractive, since the electron-nuclear attraction, due to the incomplete screening of the nucleii with respect to the electrons, dominates. These electrostatic forces also polarize the charge clouds, which leads to extra induction forces in addition to those caused by the molecular multipole moments, and they modify the dispersion energy, too.

The second mechanism which is effective in the overlap region is the intermolecular exchange. The electrons being indistinguishable fermions the total electronic wave function must be antisymmetric with respect to all electron permutations, including those which interchange electrons from different molecules. For closed shell molecules this leads to a decrease in electron density in the overlap region which causes a repulsive intermolecular force (Pauli repulsion). This exchange repulsion in turn leads to polarization of the molecular charge distributions, giving rise to exchange-induction and exchange-dispersion forces.

The two short range mechanisms, penetration and exchange, both depend on the overlap between the molecular wave functions and the resulting forces depend exponentially on the intermolecular distance. The exchange interactions are larger than the penetration effects, typically by a factor of 5 to 10 [21,32], so that the total short range force between closed shell molecules is repulsive.

In the short range, where these effects dominate the long range interactions, the interaction energy is usually calculated from a "supermolecule" treatment. This implies that the methods which are normally used for molecules, are now applied to the calculation of the total energy of the interacting system. The interaction energy is obtained by subtraction of the separate monomer energies and, possibly, since the monomers are described by limited basis sets, correction for the so-called basis set superposition error [24,33,34]. This supermolecule treatment can be performed at different levels of accuracy, using the (Hartree-Fock) independent particle model or extensions of this model (e.g. Configuration Interaction). Although the accuracy of the results will differ, the short range interactions will be given by all these treatments, since the usual molecular (ab initio) methods take into account the overlap between the basis orbitals and use correctly antisymmetrized electronic wave functions.

4. Intermediate region; the Van der Waals minimum

This region is defined as the range of distances where the long range and short range interactions have comparable magnitudes. For closed shell molecules where the short range (mainly exchange) forces are repulsive and the long range interactions are mostly attractive, a Van der Waals minimum occurs. Unfortunately, since most physical properties of molecular systems are determined by the interactions in this region, the calculations are the most difficult here. The long range theory is not valid since it neglects exchange,and the multipole expansion for the interaction operator V, even if it seems to converge, will not agree with the results for the exact V, since it neglects penetration effects. The supermolecule treatment of the short range cannot be applied in a straightforward manner because the interaction energy is only a very small fraction of the total energy (typically 10^{-5}, while the chemical binding energy which is a few hundred times larger is already very hard to obtain with some accuracy). Moreover, if we carry out the supermolecule calculation at the Hartree-Fock level, the dispersion interaction will not be found at all. This interaction originates namely from the correlation between the electronic motions on the different molecules (intermolecular correlation) which is not accounted for by the independent particle model. The usual way of correcting the Hartree-Fock results in molecules by Configuration Interaction will not give a dispersion attraction either, if one takes no special precautions, since all the effort of improving the wave functions will go into the energy lowering related with the intramolecular electronic correlation (which is about 100 times larger than the intermolecular correlation energy at the Van der Waals minimum).

Considering these problems two ways of treating the intermediate region suggest themselves and they have actually been worked out in methods which are applicable to (small) molecules. The first group of methods, following the original formalism by Eisenschitz and London [1] or similar derivations, adapt the long range perturbation theory to the overlap region by antisymmetrizing the unperturbed product wave functions $\phi_a^A\phi_b^B$ and/or the perturbation corrections to these. They have to cope with the problem that these antisymmetrized products are no longer eigenfunctions of the unperturbed hamiltonian, $H^A + H^B$. Moreover, the originally complete set of products of molecular wave functions becomes linearly dependent (overcomplete) by the antisymmetrization and, therefore, the perturbation expansion becomes non-unique. These methods which are reviewed in refs. [35,36], are called exchange perturbation theories or, more generally, symmetry adapted perturbation theories.

The second group of methods are essentially (variational) supermolecule treatments which are adapted to the problem of intermolecular forces by separating the intermolecular effects from the intramolecular ones. One can achieve this by using basis functions which are localized as much as possible on the monomers and ensuring that the orbital bases and the configuration wave functions are chosen such that all the relevant interactions are correctly described [37,38]. An example of the latter type is the Multistructure Valence Bond method employed by us [39-41], which is nothing but a variational expansion of the total wave function in terms of antisymmetrized products, $\mathcal{A}\, \phi_a^A\phi_b^B$, with optimized wave functions ϕ_a^A and ϕ_b^B for the molecular "effective excited states".

As one can write general perturbation expansions of variational schemes when certain interaction terms in the total hamiltonian are small [41], there is in fact a great similarity between the two types of methods. Both are using the exact interaction operator V, avoiding the multi-

Table

Results for the C_2H_4 crystal, symmetry $P2_1/n$ (Z=2).

		experiment a)	theoretical b)		
			ab initio c)	empirical d)	
unit cell	a [Å]	4.626	4.730	4.726	
	b [Å]	6.620	6.205	6.435	
	c [Å]	4.067	4.005	4.135	
	β [degrees]	94.4	88.5	94.0	
	ξ [degrees]	-27.0	-31.7	-27.0	
	η [degrees]	-14.6	-9.3	-11.7	
	ζ [degrees]	-34.3	-31.3	-31.8	
cohesion energy [kcal/ mole]	short range	-	- 4.84	- 3.08	
	dispersion	-	8.89	6.58	
	electrostatic	-	1.41	0.42	
	total	4.7	5.46	3.92	
IR	$\omega[\text{cm}^{-1}]$ e	-	$75^{(1.07)}$	$63^{(1.07)}$	Au
		$73^{(1.05)}$	$94^{(1.07)}$	$75^{(1.07)}$	Au
		$110^{(1.06)}$	$126^{(1.07)}$	$113^{(1.07)}$	Bu
Raman	$\omega[\text{cm}^{-1}]$ e	$73^{(1.22)}$	$50^{(1.21)}$	$38^{(1.22)}$	Bg
		$90^{(1.20)}$	$80^{(1.19)}$	$79^{(1.21)}$	Ag
		$97^{(1.24)}$	$105^{(1.20)}$	$90^{(1.18)}$	Ag
		$114^{(1.20)}$	$157^{(1.17)}$	$131^{(1.18)}$	Bg
		$167^{(1.36)}$	$196^{(1.41)}$	$163^{(1.40)}$	Ag
		$177^{(1.31)}$	$249^{(1.41)}$	$170^{(1.41)}$	Bg

a) Structure (at 85 °K) from [62], cohesion energy from [63] corrected for the zero-point vibrational energy, IR frequencies (at 20 °K) from [64], Raman frequencies (at 30 °K) from [65].

b) Harmonic calculation [61] at the theoretical equilibrium geometry with an atom-atom potential of the type.

$$V_{ij} = -A\, r_{ij}^{-6} + B \exp(-C r_{ij}) + q_i q_j r_{ij}^{-1}$$

c) Parameters from best fit to the ab initio potential [61]:

$A_{CC} = 876$ $A_{CH} = 132$ $A_{HH} = 20$ kcal Å^6 mole^{-1}

$B_{CC} = 27116$ $B_{CH} = 6378$ $B_{HH} = 1500$ kcal mole^{-1}

$C_{CC} = 3.16$ $C_{CH} = 3.43$ $C_{HH} = 3.70$ Å^{-1}

$q_C = -0.5274$, $q_H = +0.2637$ (charges shifted from the nucleii)

d) Empirical parameters from Williams [66]

$A_{CC} = 449.3$ $A_{CH} = 134.3$ $A_{HH} = 40.15$ kcal $Å^{6}$ $mole^{-1}$

$B_{CC} = 71461$ $B_{CH} = 14316$ $B_{HH} = 2868$ kcal $mole^{-1}$

$C_{CC} = 3.60$ $C_{CH} = 3.67$ $C_{HH} = 3.74$ $Å^{-1}$

$q_C = -0.24$, $q_H = +0.12$

e) In parentheses, the isotopic ratio: $\omega_{C_2H_4}/\omega_{C_2D_4}$.

pole expansion; in the long range limit the results are found to converge again to the R^{-n} expansion results.

5. Results; application to molecular crystals

The problems we have discussed and the methods developed to solve them are illustrated by actual quantitative calculations of the interactions between (small) molecules [20,24,42-60]. In our group we have looked at the convergence of the multipole expansions of the first and second order interaction energies [13,22,30,41], at the importance of exchange and penetration effects in the intermediate region [13,21,41], the calculation of short range exchange forces [32,39-41], the orientational dependence of the interaction energy [22,29,30] and the accuracy of theoretical (and the usual semi-empirical) Unsöld approximations [13,14]. The ab initio calculated long range interactions have been used, for ethylene and pyrazine, to study the importance of the different contributions to the cohesion energy of molecular crystals [14]; the ab initio results for the intermediate and long range Van der Waals interaction potential have been applied to perform a completely non-empirical calculation of the packing and the lattice dynamics of the C_2H_4 crystal [61]. The results of the latter calculation are summarized in the following table (next page).

We observe that the structure and the phonon frequencies are in reasonable agreement with experiment [62-65], although the results with the best empirical atom-atom potentials [66] are still somewhat better. We must realize, however, that the same type of data as calculated here, have been used in parametrizing the empirical atom-atom potentials, while the ab initio atom-atom potential contains no empirically adjusted parameters. Therefore, the crystal results are satisfactory and they form a confirmation of the accuracy reached in the ab initio calculations.

Comparing the distance and orientational dependence of our C_2H_4-C_2H_4 potential with the (empirical) atom-atom potential model, we have concluded [61] that the ab initio results for the main part justify this model. The most striking difference with the empirical atom-atom potentials [66,67] has been found in the electrostatic (point charge) interactions, which should be larger, according to our calculations. Looking more closely at the ab initio potentials, one observes clearly some deviations from the (central) atom-atom potential which are caused by the effects of the chemical bonding in the C_2H_4 molecules on the intermolecular interactions. For instance, the short range repulsion between carbon atoms has a relatively long range (small exponent) in the π-direction, indicating the relative diffuseness of the π-clouds. The latter observation is in agreement with the experimental data for π-stacked graphite layers [68] in comparison with other hydrocarbons [67] which are packed in different manners in the crystals.

Acknowledgement

The work which forms the basis of this paper has been performed in close cooperation with drs. P.E.S. Wormer, F. Mulder, P.J.M. Geurts, R.M. Berns and T. Wasiutynski at the Nijmegen Institute of Theoretical Chemistry and with drs. M.C. van Hemert (Leiden) and C. Huiszoon (Twente). Special thanks are due to Fred Mulder for his critical reading of the manuscript.

References

1. R. Eisenschitz and F. London, Z. Physik 60 (1930) 491.
2. J.O. Hirschfelder, C.F. Curtiss and R.B. Bird, Molecular Theory of Gases and Liquids, Wiley, New York (1964).
3. B. Pullman, ed., Molecular Associations in Biology,Academic Press, New York (1968), especially Pullman's introduction.
4. S. Nir, Prog. Surface Scie. 8 (1977) 1
5. This is demonstrated by quantitative calculations on different states of H_2^+ (A. van der Avoird, Chem. Phys. Letters 1 (1967) 429) and H_2 (P.R. Certain, J.O. Hirschfelder, W. Kolos and L. Wolniewicz J. Chem. Phys. 49 (1968) 24).
6. Intermolecular Forces, Discuss. Faraday Soc. 40 (1965).
7. J.O. Hirschfelder, ed., Advan. Chem. Phys. 12 (1967).
8. H. Margenau and N.R. Kestner, Theory of Intermolecular Forces, 2nd ed., Pergamon, New York (1971).
9. P.R. Certain and L.W. Bruch, in: Physical Chemistry Series 1, Vol. 1 Theoretical Chemistry, MTP International Review of Science, Butterworths, London (1972).
10. A.T. Amos and R.J. Crispin, in: Theoretical Chemistry, Advances and Perspectives, Vol. 2., H. Eyring and D. Henderson, eds., Academic Press, New York (1976).
11. A.R. Edmonds, Angular Momentum in Quantum Mechanics,Princeton Univ. Press, Princeton (1957).
12. A. Unsöld, Z. Physik 43 (1927) 563.
13. F. Mulder, M.C. van Hemert, P.E.S. Wormer and A. van der Avoird, Theor. Chim. Acta 46 (1977) 39.
14. F. Mulder and C. Huiszoon, Mol. Phys. 34 (1977) 1215.
15. J.O. Hirschfelder, W. Byers Brown, and S.T. Epstein, Advan. Quantum Chem. 1 (1964) 256.

6. S.O. Goscinski, Intern. J. Quantum Chem. 2 (1968) 761.
7. P.W. Langhoff and M. Karplus, in: The Padé Approximant in Theoretical Physics, G.A. Baker and J.L. Gammel, eds., Academic Press, New York (1970) 41.
8. J.G. Kirkwood, Phys. Z. 33 (1931) 57.
9. E.A. Hylleraas, Z. Physik 65 (1930) 209.
0. H.N.W. Lekkerkerker, P. Coulon and R. Luyckx, J. Chem. Soc. Faraday Trans II 73 (1977) 1328.
1. F. Mulder, P.J.M. Geurts and A. van der Avoird, Chem. Phys. Letters 33 (1975) 215.
2. F. Mulder, A. van der Avoird, P.E.S. Wormer, Mol. Phys., submitted.
3. E. Kochanski, J. Chem. Phys. 58 (1973) 5823.
4. T. Groen and F.B. van Duijneveldt, Theor. Chim. Acta, submitted.
5. L. Zandee and J. Reuss, Chem. Phys. 26 (1977) 345.
6. R.J. Le Roy and J. van Kranendonk, J. Chem. Phys. 61 (1974) 4750.
7. A.M. Dunker and R.G. Gordon, J. Chem. Phys. 68 (1978) 700.
8. F.A.S. Ligthart, Thesis, Amsterdam (1975).
9. P.E.S. Wormer, Thesis, Nijmegen (1975).
0. P.E.S. Wormer, F. Mulder and A. van der Avoird, Intern. J. Quantum Chemistry, 11 (1977) 959.
1. U. Fano and G. Racah, Irreducible Tensorial Sets, Academic Press, New York, (1959).
2. A. van der Avoird and P.E.S. Wormer, Mol. Phys. 33 (1977) 1367.
3. S.F. Boys and F. Bernardi, Mol. Phys. 19 (1970) 558.
4. M. Bulski and G. Chalasinski, Theor. Chim. Acta 44 (1977) 399.
5. D.M. Chipman and J.O. Hirschfelder, J. Chem. Phys. 59 (1973) 2838.
6. B. Jeziorski, K. Szalewicz and G. Chalisinski, to be published.
7. H.F. Schaefer, D.R. Mc Laughlin, F.E. Harris and B.J. Alder, Phys. Rev. Letters 25 (1970) 988.
8. P. Bertoncini and A.C. Wahl, Phys. Rev. Letters 25 (1970) 991.
9. P.E.S. Wormer, A. van Berkel and A. van der Avoird, Mol. Phys. 29 (1975) 1181.
0. P.J.M. Geurts, P.E.S. Wormer and A. van der Avoird, Chem. Phys. Letters 35 (1975) 444.
1. P.E.S. Wormer and A. van der Avoird, J. Chem. Phys. 62 (1975) 3326.
2. V. Magnasco and G.F. Musso, J. Chem. Phys. 48 (1968) 2657.
3. O. Tapia, G. Bessis and S. Bratoz, Intern. J. Quantum Chem. S4 (1971) 289.
4. C.F. Bender and H.F. Schaefer, J. Chem. Phys. 57 (1972) 217.
5. B. Tsapline and W. Kutzelnigg, Chem. Phys. Letters 23 (1973) 173.
6. U. Urban and P. Hobza, Theoret. Chim. Acta 36 (1975) 215.
7. J.W. Birks, H.S. Johnston and H.F. Schaefer, J. Chem. Phys. 63 (1975) 1741.
8. R.R. Lucchese and H.F. Schaefer, J. Am. Chem. Soc. 97 (1975) 7205.
9. B.J. Garrison, W.A. Lester, P. Siegbahn and H.F. Schaefer, J. Chem. Phys. 63 (1975) 4167.
0. G.F.H. Diercksen, W.P. Kraemer and B.O. Roos, Theoret. Chim. Acta 36 (1975) 249.
1. B. Jeziorski and M. van Hemert, Mol. Phys. 31 (1976) 713.
2. O. Matsuoka, E. Clementi and U. Yoshimine , J. Chem. Phys. 64

(1976) 1351.
53. W. Meyer, Chem. Phys. 17 (1976) 27.
54. M. Jaszunski, E. Kochanski and P. Siegbahn, Mol. Phys. 33 (1977) 139.
55. G.A. Gallup, Mol. Phys. 33 (1977) 943.
56. E. Clementi, F. Cavallone and R. Scordamaglia, J. Am. Chem. Soc. 99 (1977) 5531, 5545.
57. M. Jaszunski and E. Kochanski, J. Am. Chem. Soc. 99 (1977) 4624.
58. M.M. Szczesniak and H. Ratajczak, J. Chem. Phys. 67 (1977) 5400.
59. J. Prisette , E. Kochanski and D.R. Flower, Chem. Phys. 27 (1978) 373.
60. P. Smit, Thesis, Utrecht (1978).
61. T. Wasiutynski, A. van der Avoird and R.M. Berns, to be published.
62. G.J.H. van Nes and A. Vos, Acta Cryst. B 33 (1977) 1653.
63. C.J. Egan and J.D. Kemp, J. Am. Chem. Soc. 59 (1937) 1264.
64. M. Brith and A. Ron, J. Chem. Phys. 50 (1969) 3053.
65. G.R. Elliot and G.E. Leroi, J. Chem. Phys. 59 (1973) 1217.
66. D. Hall and D.E. Williams, Acta Cryst. A 31 (1975) 56.
67. D.E. Williams, J. Chem. Phys. 47 (1967) 4680.
68. A.D. Crowell, J. Chem. Phys. 29 (1958) 446.

THE LATTICE DYNAMICS OF MOLECULAR CRYSTALS

G.A. Mackenzie*

Summary

The method of calculating the lattice dynamics of molecular crystals in the rigid molecule approximation is outlined. Extensions of the model to include non-rigid and polarisable molecules are discussed. Applications of the model and recent experimental results are reviewed with particular emphasis on neutron coherent inelastic scattering measurements.

Introduction

The lattice dynamics of crystals in which the basic structural units are individual atoms is a well developed field in physics. The standard text on the subject is that of Born and Huang (1) and there are several good reviews in the literature e.g. Cochran (2) and Cochran and Cowley (3). In the so-called Born-von Karman formalism the procedure is to assign three positional degrees of freedom to each of the N atoms in the primitive unit cell, set up the equations of motion and solve them to obtain the 3N sets of frequencies and eigenvectors for each value of the phonon wavevector within the first Brillouin zone. The curves which result from plotting the frequencies as functions of wavevector are the dispersion curves of which there are, in general, 3N branches.

This type of treatment is less appropriate for crystals in which the basic structural units are tightly bound molecules with relatively weak intermolecular binding. The motion of molecules as individual units is a distinguishing feature of molecular crystals and arises from the fact that the forces between molecules are much weaker than those between atoms within the same molecule. The normal modes of the system may be divided into two types: internal modes and external modes. The internal modes are those in which the atoms within each molecule move with respect to each other while the molecular centre of mass remains fixed. The external modes are those in which the molecules vibrate as whole units about their equilibrium positions. The motion is both translational and rotational. The applicability of

*Department of Physics, University of Edinburgh, Scotland and Physics Department, Risø National Laboratory, Denmark

this separation into external and internal modes depends on the relative strengths of the two types of forces and may be judged by investigating the vibrational spectra of the system. In many cases there is a clear distinction between the two types of mode with a large frequency gap between the two regimes. The external modes lie typically between 0 and 3 THz with frequencies strongly dependent on the wave-vector q. The internal modes normally have frequencies above 5 THz which are only slightly shifted from the free state values with very weak $\underline{q}$ dependence.

In the rigid molecule approximation the separation between internal and external modes is assumed to be complete. Only the molecular translational and rotational degrees of freedom enter into the calculation allowing a great simplification of the problem. The effects of non rigidity can be investigated at a later stage by reintroducing the internal degrees of freedom. It has been found that even for quite non-rigid molecules like naphthalene the effect on the frequencies is only about 10% at most. The approximation of rigid molecules is therefore quite good.

The theoretical developments in molecular dynamics have been reviewed by Venkataraman and Sahni (4) and Schnepp and Jacobi (5). The first complete lattice dynamical treatment of a molecular crystal was done by Cochran and Pawley (6) for hexamethylenetetramine (HMT). This substance is of sufficiently high symmetry that the force constants can be treated as individual model parameters which are adjusted to fit the experimental measurements. In most cases however the symmetry of molecular crystals is so low that the number of independent parameters becomes very large. The method adopted has been to use atom-atom pair potential functions from which all the required force constants can be derived. Pawley (7) presented the first complete treatment of this kind in his model of the lattice dynamics of naphthalene and anthracene. The formalism is outlined in the next section.

The rigid molecule model

(i) Basic equations

The formalism of the rigid molecule model for lattice dynamics has been described in great detail elsewhere (4,8) and only a brief outline is given here. The treatment follows Pawley (8) and is essentially classical. In the quantum mechanical treatment the displacements of the atoms are expressed in terms of creation and annihilation operators which act on simple harmonic oscillator functions. Here it is adequate to note that the energy is quantised, the quantum of vibrational energy being called the phonon. The term libration is often used to describe a molecular lattice vibration which involves molecular rotation, with a corresponding quantum the "libron". In this work the terms librational and rotational are, in general synonymous.

Within the harmonic approximation the potential energy of the crystal can be written

$$\phi = \phi_o + \sum_{\ell k i} \frac{d\phi}{du_l(\ell k)} u_i(\ell k) + \tfrac{1}{2} \sum_{\ell k i} \sum_{\ell' k' j} \frac{d^2\phi}{du_i(\ell k)\, u_j(\ell' k')} u_i(\ell k) u_j(\ell' k') \qquad (1)$$

where $u_i(\ell k)$ is the ith component of the displacement of the kth molecule in the ℓth unit cell, and the differentials are evaluated at equilibrium. The displacement $\underline{u}$ is a six dimensional vector of which the first three components represent translations and the last three represent rotations. It should be noted that the translations form a polar vector while the rotations form an axial vector. This difference gives rise to different transformation properties. The first term of the expansion, ϕ_o, is the arbitrary zero point of potential which can be chosen to be zero while the linear term involving $\frac{d\phi}{du}$ must be zero since the equilibrium position is defined as the minimum of ϕ. The only non-zero term is therefore the quadratic term.

The equations of motion for the molecule (ℓk) are:

$$m_i \ddot{u}_i(\ell k) = - \sum_{\ell' k' j} \phi_{ij}(\ell k, \ell' k')\, u_j(\ell' k') \qquad (2)$$

$$\text{where } \phi_{ij}(\ell k, \ell' k') = \frac{d^2\phi}{du_i(\ell k)\, du_j(\ell' k')} \qquad (3)$$

and m_i is the molecular mass for i = 1,2,3 and the moment of inertia about the principal axes for i = 4,5,6. The displacements $\underline{u}(\ell k)$ are expressed in the principal inertial axes coordinate system of the molecule (ℓk) so that care must be taken in evaluating the differentials. A travelling wave

$$\underline{u}(\ell k) = \underline{U}(k\underline{q}) \exp(i(q.\underline{r}(\ell k) - \omega(q)t) \qquad (4)$$

can be substituted to obtain

$$\omega^2(\underline{q}) \; (\sqrt{m}_i U_i(k\underline{q})) = \sum_{k' j} (\sqrt{m}_j U_j(k'\underline{q}))\; M_{ij}(kk'\underline{q}) \qquad (5)$$

The matrix M is the mass reduced dynamical matrix whose elements are given by:

$$M_{ij} = \sum_{\ell'} \frac{\phi_{ij}}{\sqrt{m_i m_j}}(\ell k, \ell' k') \exp i\underline{q}.(\underline{r}(\ell' k') - \underline{r}(\ell k)) \qquad (6)$$

The eigenvalues of M are the squared frequencies of the phonons, $\omega^2(\underline{q})$, and the matrix which diagonalises M is the matrix of phonon eigenvectors. The problem of obtaining the phonon frequencies and eigenvectors then consists of constructing M and diagonalising it.

(ii) The force constant tensors

For highly symmetric crystals like HMT the number of independent elements in the force constant tensors ϕ_{ij} is sufficiently small that they can be treated as variable parameters (6). For more typical molecular crystals of low symmetry the number of independent terms is very large. This is overcome by using an expression for the intermolecular potential from which the force constant tensors may be derived. The crystal potential is assumed to be a sum of pairwise interactions between molecules,

$$\phi = \sum_{mn} V_{mn} \tag{7}$$

A second pair central force assumption on a more detailed level is then applied to V_{mn},

$$V_{mn} = \sum_{ij} V_{ij}(r_{ij}) \tag{8}$$

where r_{ij} is the separation between atom i in molecule m and atom j in molecule n. The individual V_{ij} contain a small number of parameters depending on the chemical nature of the atoms represented by i and j. A commonly used form for V_{ij} is the Buckingham or '6-exp' potential which was originally used for molecular crystals by Kitaigorodskii (10),

$$V_{ij} = -\frac{A_k}{r_{ij}^6} + B_k \exp(-\alpha_k r_{ij}) \tag{9}$$

The constants A_k, B_k and α_k depend on the chemical nature of the atoms i and j. For a system with only one type of atom such as orthorhombic sulphur (10) only three parameters are involved. For a hydrocarbon such as naphthalene, $C_{10}H_8$, we must consider C-C, H-H and C-H pairs so there is a total of nine parameters.

The force constant tensor for the interaction between the two molecules labelled O and N can then be written

$$\phi(O,N) = \sum_{ij} \phi(O_i,N_j) \tag{10}$$

where i and j range over the atoms in O and N respectively.

The problem then reduces to finding the 6×6 tensors (O_i,N_j).

Since the Van der Waals attractive potential represented by the r^{-6} term is of relatively short range it is not necessary to calculate beyond a certain limiting value of r_{ij}, Figure 1. It has been found that an upper limit of 5.5 Å gives sufficient accuracy in the calculated frequencies for the case of naphthalene. This limitation of the number of atom-atom contacts is of course of great importance in the reduction of computation time.

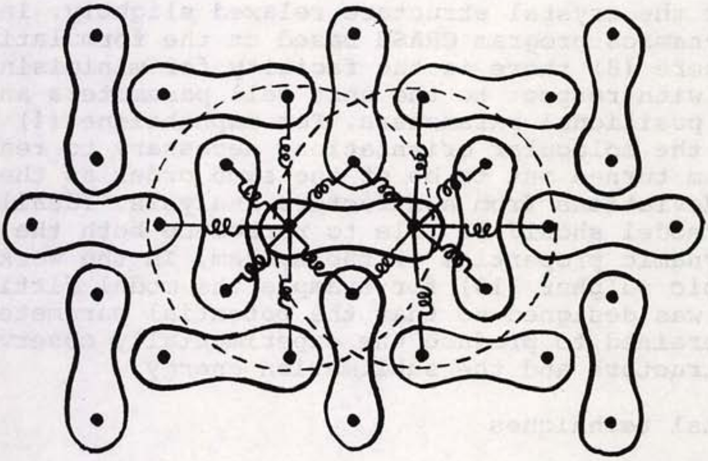


Fig. 1. The atom-atom interactions in a diatomic molecular crystal showing the limiting interaction radius

The contributions to the potential energy from interactions of greater than the limiting range are taken into account by replacing the summation with an integral, assuming a homogeneous atomic density.

It should be noted that when there are Coulombic interactions in the system it is necessary to use the lattice sum technique of Ewald (1).

The atom-atom force constant tensors are formed from the first and second derivatives of the potential V_{ij} with respect to the interatomic separation r_{ij}. These differentials are easily obtained from the analytical expression for the potential (equation 9) and can be programmed directly in analytical form. This results in a saving in computation time and a gain in accuracy over a procedure in which the differentials are calculated numerically. The analytical formulation is also advantageous if a model fitting procedure is to be incorporated in the program.

Equilibrium

The formulation of the lattice dynamics is based on the assumption that the system is in equilibrium. For an arbitrary set of force constants however there is no guarantee that this is the case. There may be forces and torques acting on the molecule. For some cases the symmetry of the system is sufficient to ensure that equilibrium exists but for naphthalene for example while there can be no net force acting on the molecule there may be a net torque. To make equilibrium certain either the force constants must be altered or the crystal structure relaxed slightly. In the lattice dynamics program CRASH based on the formulation outlined here (8) there is the facility for minimising the potential with respect to the unit cell parameters and the molecular positional parameters. For naphthalene (4) the shifts on the molecular orientations necessary to reach equilibrium turned out to be of the same order as the standard deviations from a structure analysis. Ideally the potential model should be able to reproduce both the static and the dynamic properties of the system. In the work on orthorhombic sulphur (10) for example the model fitting procedure was designed so that the potential parameters were constrained to produce the experimentally observed crystal structure and the sublimation energy.

Experimental techniques

By far the best experimental technique for investigating the lattice dynamics of crystals is the coherent inelastic scattering of thermal neutrons (3). This method requires relatively large (several cm^3) single crystals and in addition all hydrogen must be replaced by deuterium to avoid incoherent scattering which gives rise to a very high background intensity. Because of the complexity of the dispersion curve systems and the difficulty in obtaining suitable single crystal specimens, the number of molecular crystals which have been thoroughly investigated remains quite small. A review of experimental developments up to 1976 has been given by Dolling (9).

The great advantage of coherent inelastic neutron scattering is that the phonon dispersion curves can be measured at all values of the wavevector q. This contrasts with optical measurements, Raman and Brillouin scattering and infrared absorption, which are limited to the region of very small q, effectively $q = 0$, because of the great difference between the wavelengths of the probe radiation and the crystal excitations. Neutrons are unique in that both the wavelength and the energy are matched to those of the crystal excitations.

In coherent inelastic neutron scattering, normally using hydrogenous samples, can provide a measure of the density of vibrational states of the system. However quantitative interpretation of the data is difficult and the technique is of little use in the fitting of lattice dynamical models (11).

The interpretation of coherent inelastic neutron scattering measurements can be quite a considerable task, particularly for complicated molecular crystals with many dispersion branches. A great aid, both in the planning and in the interpretation of the measurements, is a calculation of the expected scattering intensity for the diffirent branches at particular points in reciprocal space. The separation between branches is frequently much less than the resolution of the measurement and it is fortunate that it is possible to find regions in reciprocal space where certain modes scatter strongly while others scatter very weakly. The scattering intensity at any point in reciprocal space depends mainly on the eigenvector of the mode, or in other words the pattern of atomic displacements. It has been found that the eigenvectors are normally relatively insensitive to the details of the crystal model used and it follows that the calculated scattering intensity is not strongly model dependent. This means that it is possible to use a rather rough model which gives poor results for the frequencies but nevertheless predicts the scattering intensity reliably. The calculations allow the particular modes to be measured. The model can then be refined to improve the correspondence between observed and calculated frequencies enabling further interpretation of the measurements. A subroutine for the calculation of the neutron scattering intensity is included in the computer program CRASH, written by Pawley and based on the formalism outlined here (8).

Extended models

(i) Deformable molecules

The validity of the rigid molecule approximation was tested by Pawley and Cyvin (12) who performed dynamical calculations for naphthalene which took into account the non-rigidity of the molecule. The forces between atoms within the molecule were independently obtained by fitting the gas-phase internal mode frequencies. This force field was then combined with the non bonded interactions and the resulting large dynamical matrix was solved to give the frequencies of all the modes in the system. The results showed that some of the external mode frequencies are shifted substantially from the rigid molecule values, in some cases by as much as 0.25 THz which is considerably larger than the typical uncertainty in a neutron scattering experiment. Consequently account must be taken of this effect if a rigid molecule model is to be used to fit measured frequencies. In most cases the effect of non-rigidity is to decrease the external mode frequencies by varying amounts depending on the mode eigenvector. Since the eigenvectors are generally quite insensitive to the details of the potential model the shifts in the modes should also be insensitive. Model fitting can therefore take the following form.

(a) All the measured frequencies are increased by the amount suggested by the calculation.

(b) These modified frequencies are used for the fitting of a rigid-molecule model.

(c) The best rigid-molecule model is then combined with the internal molecular force field and the new mode frequencies are compared with the unmodified experimental frequencies.

(d) If the result of (c) is significantly different from (a) a further iteration is done.

This procedure should then yield intermolecular potentials which can be applied to other structures.

(ii) Shell model

In the studies of ionic crystals such as the alkali halides it was first assumed that the ions move as rigid units. The work of Woods et al (13) showed that there were considerable defects in the theory based on this assumption when the results were compared with neutron scattering measurements. The theory was improved by introducing polarisable ions which could develop dipole moments as a result of the displacements involved in the lattice vibrations. The model was called the "shell model" because the ions could be considered as composed of cores and electronic shells which could move with respect to each other to produce the dipoles.

A similar model for molecular crystals has recently been developed (14) and further extended (15) to include fixed molecular multipole moments. The molecules can be considered as an array of atomic cores and shells and the displacements are assumed to occur in such a way that the core framework and the shell framework each move as rigid units. As in the ionic case an isotropic force constant is assumed between cores and shells and when there is relative movement of the two arrays dipole moments are formed at the atomic sites.

The model has been applied with some success to orthorhombic sulphur (16) and work is proceeding on applying it to the smaller molecules DCN, OCS, BrCN, ICN, I_2 and N_2.

Applications

The eventual aim of calculations and measurements of the lattice dynamics of molecular crystals is to understand the intermolecular forces. A good model of the interactions should be able to reproduce both the observed structure and the vibrational frequencies as functions of temperature and pressure. There is great interest in finding a universal potential for the aromatic hydrocarbons, benzene, naphthalene, anthracene, pyrene etc., and there have been attempts to produce potentials which can be applied over a wide range of these materials (17). The amount of data with which to compare the calculations is relatively small however and often limited to room temperature structure and optical spectroscopy measurements. As mentioned already the main factor preventing more extensive experimental study by neutron spectroscopy is the lack of suitable single crystal samples. Of the aromatics only naphthalene has been studied thoroughly by neutron spectroscopy. Measurements of the external mode dispersion curves for wavevectors along the $\underline{b}^*$ direction were made at Risø (18) using a 1 g crystal, Figure 2. The measurements were fitted using the rigid molecule model with allowance made for deformable molecules. The set of dispersion curves for the best model are shown in Figure 3.

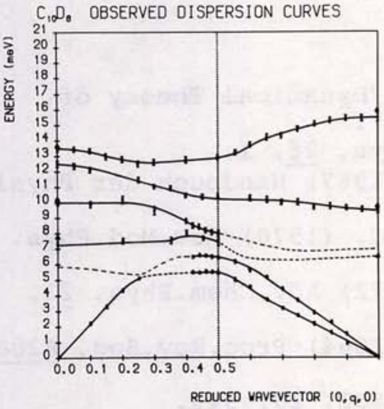


Figure 2. Observed frequencies for naphthalene-d_8. The solid lines are guides to the eye. The dotted curves were unobservable

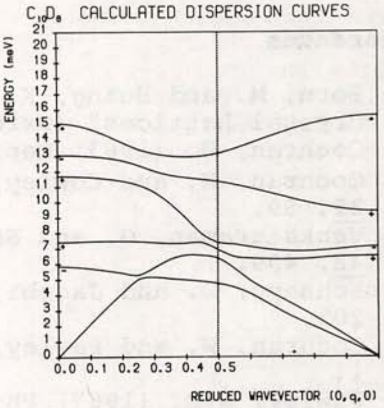


Figure 3. The dispersion curves corresponding to the best model obtained by fitting the frequencies shown in Figure 2

Recently measurements have been made at the high flux reactor at ILL, Grenoble, using a 5 g crystal. The great gain intensity has enabled the dispersion curves to be measured along all crystal directions with great accuracy at temperatures of 100 K (19) and 5 K (20). Measurements at pressures up to 3 kbar at 5 K are planned in the near future.

These very extensive measurements will allow detailed comparison between calculation and experiment and should show the inadequacies in the present simple model based on the "6-exp" potential.

Conclusions

The rigid molecule model for the lattice dynamics of molecular crystals has been found very useful in calculating statis and dynamic properties. In order to progress to more sophisticated models however it is necessary to show that the present model is inadequate and this requires extensive experimental measurements of crystal structures and phonon dispersion curves as functions of temperature and pressure. We have now reached this stage with one molecular crystal, naphthalene, and it is to be hoped that we will obtain results from other similar systems in the next few years.

Acknowledgements

The formulation of the lattice dynamics of molecular crystals outlined here is mainly due to G.S. Pawley whom the author wishes to thank for continued help and collaboration.

References

1. Born, M. and Huang, K. (1954) "Dynamical Theory of Crystal Lattices" (Oxford U.P.).
2. Cochran, W. (1963) Rep.Prog.Phys. 26, 1.
3. Cochran, W. and Cowley, R.A. (1967) Handbuch der Physik 25, 59.
4. Venkataraman, G. and Sahni, V.C. (1970) Rev.Mod.Phys. 42, 409.
5. Schnepp, O. and Jacobi, N. (1972) Adv.Chem.Phys. 22, 205.
6. Cochran, W. and Pawley, G.S. (1964) Proc.Roy.Soc. A280, 1.
7. Pawley, G.S. (1967) Phys.Stat.Sol. 20, 346.
8. Pawley, G.S. (1972) Phys.Stat.Sol. 49, 475.
9. Dolling, G. (1976) Conf. on Neutron Scattering, Gatlinburg, Tenn. 263.
10. Rinaldi, R.P. and Pawley, G.S. (1975) J.Phys.C. 8, 599.
11. Bokhenkov, E.L., Natkaniec I. and Shika, E.F. (1976) Sov.Phys. JETP 43 (3), 536.
12. Pawley, G.S. and Cyvin, S.J. (1970). J. Chem.Phys. 52, 4073.
13. Woods, A.D.B., Cochran, W. and Brockhouse, B.N. (1960) Phys.Rev. 119, 980.
14. Luty, T. and Pawley, G.S. (1974) Phys.Stat.Sol. B66, 309.
15. Pawley, G.S. and Leech, J.W. (1977) J.Phys.C. 10, 2527.

16. Luty, T. and Pawley, G.S. (1975) Phys.Stat.Sol. B69, 551.
17. Starr, T.L. and Williams, D.E. (1977) Acta Cryst. A33, 771.
18. Mackenzie, G.A. and Pawley, G.S. (1977) J.Phys.C. 10, 3723.
19. Bokhenkov, E.L., Sheka, E.F. Dorner, B. and Natkaniec,I. (1977) Sol.St.Comm. 23, 89.
20. Bokhenkov, E.L., Sheka, E.F., Kalus, J., Schmelzer, U., Natkaniec, I., Pawley, G.S., Mackenzie, G.A. and Dorner, B. (in preparation).

CHAPTER 6

Miscellaneous

MICROPROCESSORS AND MICROCOMPUTERS

G. C. Bassi [*]

Summary

A review of microprocessors and microcomputers is given, together with applications and possible applications in crystallography. The range of power of the present microprocessors does only allow calculation of smaller crystallographic problems. The most important impact of the microprocessor will be in the field of instrumentation.

Introduction

Microprocessors are not yet widely used in crystallographic computing. The application of such devices is usually considered as more convenient for automation where speed and overall throughput are needed than for scientific computing. More and more the preference is given in crystallography to specialised systems designed for large scale computers or adapted to the late generation of mega-mini computers.

Nevertheless it is possible that the influence of microprocessors will increase during the next several years, as a consequence of the performance that has almost doubled each year.

The versatility of microcomputers and associated family of chips is the starting point of more systems, not always designed for automation, but sometimes available as stand-alone general purpose computers, thus called microcomputers. Where lies the limit between mini and microcomputers? The current internal speed of microprocessors has increased by a factor of 10 or more in five years, and minis are available with much more computation power than previously.

In 1976, Parasuraman (1) gave a clear definition based on current memory cycle time:
- A minicomputer is limited by memory cycle time. The Central Processing Unit is about five times faster than its associated memory.
- A microprocessor is limited by CPU execution time. The CPU is five times slower than memory.

This latter assertion appears not so clearly true in 1978. Some manufacturers have now designed microprocessors with cycle time increased by a factor of five, and some new series of compatible chips are allowing the design of powerful microcomputers with throughputs of the same order of magnitude than some minicomputer's.

[*] Centre d'Etudes Nucléaires de Grenoble, DRF/CPN, 85 x - 38041 Grenoble Cédex, France

Another criterion could be the number of chips necessary for the CPU design, including the generation of memory and external control busses: a mini-computer central unit is often made of more than 100 chips, including the read-only memories (ROM) for microprogramming; a microprocessor and its associated chips should not extend over 20 pieces, and we may even find cases with only one chip.

These criterions appear as rather arbitrary, and in all cases closely ad apted to the technical evolution of the moment. The definition that we will use will be that, in one integrated chip, we might find at least CPU registers, arithmetic and logical unit (ALU) and signals that can be used for the generation of a data bus and of an address bus.

4 - 8 - 12 - 16 - bit: families of microprocessors

The family of microprocessors offers a wide range of possibilities, depe ding not only on the speed or the integration level, but also on the number of bits handled by the ALU and registers.

At the lower end we find the 4 - bit chip family intended to be used in low cost applications such as microwave ovens, high fidelity turntables, non-vidos games, etc.. The price of 4 - bit chips is typically less than $ 3. The internal logic is simple but does not allow for external memory nor handle many I/O signals. The most widely used units are Texas Instru ments Inc's 1000 (microprogrammable), National Semiconductor Corp's COPS family, General Instrument Corp's 1644, Intel's 4004 and 4040 and Rockwell Microelectronic Group's PPS 4 family.

Now consider the 8- bit family. The number of such devices is very large and we will distinguish two classes: control units and data processing units. Control units are mainly used for industrial applications instrumentation and machine control. They are often designed as single chips including small Random Access Memories (RAM) and ROM: this is the case of Intel's 8048 and 8085, Zilog Z-8, Rockwell's PPS 4/1 (first single-chip processor to be used in stand-alone mode), and others. The 8 - bit chips for data processing are designed as multi-chips systems allowing the assembly of self-consistent hardware systems from two to mo than twelve chips. The trend is to a larger integration so that the overall power is increasing with the decreasing number of necessary components. A typical two-chip system is Morotola's 6802/6846, which allo direct connection to a 6800 bus; another is Signetics 2650A/2656, that permits direct linking to memory and I/O packages. Among the multi-chips systems we find the popular series: Intel's 8080 and 8085 families, Zilo Z-80, Morotola 6800 and many others. These families often include software compatible systems, with increasing degrees of integration includin sometimes the clock generator. This leads to 3, 4 and even 5 MHz version of the same basic unit.

The typical applications are real-time industrial or process control, intelligent terminals, etc.. The memory capacity is 32 or 64K byte, whereas the one-chip versions do not include much more than 128 byte of RAM and 2 Kbyte of ROM.

In the 12 - bit family, we find a typical example with the Intersil's IM6100 family, designed to emulate the PDP 8/E instruction set (2). This is a particular approach to the problem of software compatibility so that users of PDP8 may find a low-cost computer with very efficient and well known software.

The 16 - bit family is a decisive step toward increased performance. National Semiconductor produced the first 16 - bit microprocessor: the one-chip PACE. Other early available 16 - bit are Texas Instrument 's TMS 9900, General Instrument 's CP1600. More recently other firms have made some announcements: Intel (8086) (4), Motorola (6809), and Zilog (Z8000). The application range tends now to cover scientific computing, real-time data processing, and other fields up to now covered by minicomputers. The memory capacity will be larger than 64 Kbyte and the throughput will lie in the 10 MHz region (5 MHz for the faster 8 - bit chip) (3).

Mention must be made of bit-sliced systems to complete this brief review of microprocessors.A bit-slice processor is not a full processor, but only a part of it. The width of the slice is typically 2 or 4 bits, and the number of functions in the ALU remains under 20. The cycle time ranges from 100 ns to 1 microsecond. These chips are designed mainly to provide fast and powerful components in process-control design and high-speed controllers for minicomputers (4).

Alternative use of microprocessors: single board computers and microcomputers

Building a computer with chips is not easy for everyone, even living within an highly technical environment. With such a wide variety of chips available, it would seem to be very simple to construct ones own computers. However despite the high level of compatibility in a family of compatible chips, it is often not interesting to spend time and efforts to build ones own intelligent hardware system. Some manufacturers are now supplying boards, including a micro-processor, together with all convenient hardware such as RAM, ROM, I/O parts, and often software such as a monitor delivered on ROM.

Many manufacturers can deliver single board computers: Intel, National Semiconductor Corp., Motorola Inc, Zilog. Corp., Mostek Corp., Texas Instruments Inc., Computer Automation Inc., etc.. The first example will be the Intel's SBC 80 family. The SBC 80/04, the cheapest of the family, (2,03 microsecond instruction cycle), is built around the one-chip 8085 microprocessor. It is delivered with two serial I/O lines (RS-232-C), 22 parallel I/O lines, one programmable 14-bit binary interval timer/event counter. The board contains 512 bytes of RAM and 4 Kbytes of EPROMS (erasable with U.V.). The power requirement is only 0.60 A at 5 V.

The SBC 80/05 contains almost the same features as the 80/04, but generates a bus, called by Intel the "MULTIBUS". This bus is a multimaster control logic allowing memory expansion and peripheral controllers. This has become an international standard.

The SBC 80/10A (1,95 instruction cycle) is built around the well known 8080A. It contains 1K RAM, 4K EPROM, 1 TTY or RS-232-C I/O port, 48 parallel I/O lines, 1 level of interrupt and the "MULTIBUS" generation. The tensions required are ± 5V and ± 12 V.

The SBC 80/20 (1,86 microsecond instruction cycle), is also built around the 8080A, contains 4K RAM instead of 1K, 8K EPROM instead of 4, 1 seria I/O port that is RS-232-C only, 8 levels of interrupt, ant 1 programmabl 16 - bit binary and bus timers/event counters.

The SBC 80/20A contains the same features as the 80/20 except that the RAM capacity is 8K.

The SBC 80 family is completed with all memory expansion boards, analog and digital couplers and peripherals, power supply, etc.. A complete microcomputer can be made in the laboratory or deleivered by Intel.

The second example of SBC will be the DEC's LSI-11: it is a complete PDP11 processor with 4K words of 16 - bit RAM, a real-time clock, an automatic dynamic memory refresh, a built-in ASCII programmer's console (optional) mounted on the same board. It generates a bus structure that provides position-dependent priority. The LSI-11 is a 16 bit minicompute that is built around four microprocessor chips custom-manufactured for DEC by Western Digital Corp (5). The microprogrammable structure of the LSI-11 allows for emulation of the whole set of PDP11 instruction in a "pipeline" mode: a microinstruction is executed while the next microaddress is computed. An optional module may contain a user written micro code. The nominal cycle time of the micromachine is 350 nanosecond and some instructions, such as register-to-register 8 - bit addition require only 1 cycle (6).

The DEC LSI-11 is no longer available: it has been replaced by a fully compatible successor called DEC LSI-11/2, that is less expensive than th former model. The size of the new board is exactly one-half the size of the LSI-11. It no longer contains a RAM, but the RAM compatible board may contain up to 32K 16 - bit words (7).

The last example of SBC will be the Micronova: it is a complete microcomputer with CPU, 2K or 4K words of dynamic RAM, an I/O controller, etc The board is built around the mN601 microprocessor: a full 16 - bit, that features NOVA 16 - bit architecture including hardware multiply/divide, multiple addressing mode, programmed priority interrupt to 16 levels... A full system of compatible chips has been developped for I/O controllers, memory address drivers, RAM and others. The Micronova architecture duplicates Nova functions, permitting full use of Nova software including the Data General's Real Time Operator System.

An application: interfacing a four-circle diffractometer

Such an interface has been realized to connect a four-circle Hilger and Watts diffractometer to a remote central computer. At that time, the only microprocessor available in an integrated form was the Intel 8008. The 8008 was slower by a factor of about 10 than the 8080. All rapid functions such as scalers or bit position controllers were wired. On the other hand, the microprocessor was itself integrated on a single board called MICRAL N; a bus, generated by the MICRAL N, could handle RAM boards, EPROM boards, mixed RAM and EPROM boards, I/O ports, external stacks, etc.. The configuration of the system was 1 CPU, 1 board with 3K EPROM and 1K RAM, suitable I/O boards, 1 TTY coupler, and an external stack providing a very fast buffer for external transmission in connection with the remote computer. The role of the microcomputer was first to decode the content of a coupler command message from the central computer, execute the corresponding instructions such as shafts positionning, attenuators and shutters commands, fixed point or step-by-step measurements. The microcomputer, at the end of the execution of these instructions, had to send the results of all these actions to the central computer that was in charge of all computations and external logics such as peak position search, matrix refinements, integration and correction of results, etc..

Microcomputer in a Data Acquisition Network

As a consequence of the experience gained with this latter use, we have developped a real time data acquisition network. It is a general-purpose network for experiments within a distance of 1 km. The basic idea is to spread the logic over the network in order to optimize the use of every computer linked within the network: thecentral processor (MITRA 125) controls the overall logic of each connected experiment and each terminal station is a microcomputer based on a 8080A, itself linked to the experiment, in order to execute all control operations, such as servomecanisms, two-state commands, D/A or A/D conversions, measurement collections... The choice of a microcomputer, rather than of a microprocessor was made because the possibility of versatility, as required by large variability of the many experiments conducted at D.R.F. We have then the possibility to use the same hardware and software systems, with adapted peculiarities, all elements being standardized at a very high level.

An intelligent oven regulator

Temperature programming and regulation is a problem often encountered in crystallographic laboratories. C. Mouget (8) designed a sophisticated regulator for a laboratory oven, permitting altogether temperature programmation and control. The system is based on Motorola 6800 microprocessor with 256 bytes of RAM and 4K bytes of EPROMS. The CMOS chips of RAM is maintained with a battery in order to save the parameters of the transfer function of the oven. The power is applied to the resistor by means of a "Triac", that is not controlled as usual by phase angle switching, but of spreading active full wave forms over 128 periods of the alternative power supply. For example, for half power heating, the triac will be switched on during one period out of two. Process control is easy when applying this technique even if the power is to be applied to a tension transformer.

The elements of the heating program are introduced in the memory (RAM) by means of a simplified keyboard. Each program step is made of an instruction linked to a numerical value. The maximum number of steps in a program is 20. The steps are interpreted by the permanent system written on EPROMS. Each control step is always logically linked to the next one thus completing the temperature evolution. For instance, a control statement defining a temperature slope together with the numerical value in degrees per hour is associated to the next instruction defining the stabilization of the temperature together with its value in degrees. This latter instruction is itself associated to the next one to keep this steady state together with its duration in hours.

The keyboard allows the modification of any instruction inside the program, even when the process control is active. The fundamental parameters of the oven can also be introduced, modified, (they are protected by a special key in order to avoid interactions from non-authorized operators) and read: they are P (proportional bandwidth), I (integration time constant) and D (derivation time constant). Other parameters can be read only: output tension and intensity, remaining duration of the steps, state of program, temperature, etc.. All these parameters appear on a "Led" board.

The system will be shortly improved with an autocorrecting program adapting permanently the fundamental parameters of the system: the tension applied will be computed then as a function of the temperature discrepancy.

Automatic diffractometer control

A two-circle diffractometer has been adapted to a particular experiment in order to measure magnetic X-rays reflections (9). Therefore a third movement was added, the axis lying along the diffraction vector. All three movements are controled by stepping motors (one step = 0.005 degree) The system is provided with all classical electromechanical devices such as attenuators, shutters, safety switches, etc., but not for datum points. A magnetic field can be applied to the sample.

The servo-control and measurement logics are supplied by a microcomputer "MICRAL S" based on the microprocessor Intel 8080A. The processor board generates a bidirectional general purpose bus, thus allowing for memory and input-output. The memory space is 20K bytes of which 7K are only used by the system itself. To perform any action on the experiments, the operator may type instruction symbols, divided into two classes.

The first class includes all instructions to be performed only once. The most significant are:

ANP - Initialization of the numerical values of the three angles (zero positions cannot be automatically determined) and angular shifts.

BUT - Definition of the angular space allowed in the software.

ATT - Stop everything and wait.

REP - After the previous command, restart from the interrupt point.

IAN - Print the setting angles.

IBU - Print the angular space and shifts.

ECR - Push into the stack a program step belonging to the second class, thus defining a full measurement program to be interpreted later.

EXE - Start the execution of the current measurement program.

The second class of instructions can be either pushed into the program stack when typed under "ECR" control, or executed at once. The main instructions are:

POS - Set the three shafts to specified angular values.

BAL - Move continuously one shaft under specified movement conditions (displacement, speed, ...).

PAP - Step by step movement of one shaft under specified conditions, value of the step, measurement parameters...).

CPL - Shaft coupling.

DCP - Shaft uncoupling.

MES - Measurement. The parameters are: number and duration of each measurement (typically 100 seconds), magnetic field inversion. A fast quasi-Fourier transform can be computed during the multi-measurement process in order to examine the first order harmonic that should appear much greater than others when magnetic diffraction occurs.

FIN - End of program.

BRV - Branch if the specified condition is true.

BRF - Branch if the specified condition is false.

The remaining memory (approximately 10 Kbyte) is used to compute an histogram of step by step measurements: when scanning many times over ar angular space, each measurement is added to the previous sum like in a multichannel process, where each channel should correspond to an angula position. This is made with two arrays. The first one containing the current results, and the other individual measurements obtained during the scan in course. When the scan is finished, then only the computer adds the second array to the first one. This allows to stop the experime at any time during a scan. The graph can be typed on the teletype.

An active data transfer mode point

A single board computer, the Z80-MCB is used to link together a PDP 10/K a PDP 11/70 and some local input/output devices such as paper tape punch and reader, an incremental plotter, a C.R.T.terminal and a PROM programmer (10).

The Z80-MCB is built around the Zilog Z80 microcomputer (158 instruction It operates at more than 2 MHz and contains 4K byte dynamic RAM, sockets for 4K bytes of PROMS, one serial I/O port (RS-232 or current interface) two 8 - bit ports (universal I/O), and a baud rate generator. A home-made associated board contains some other features such as three more serial I/O ports.

The plotter, the terminal and the paper tape devices can then be used from the two PDP remote computers. Both can perform cross assemblies for Intel 8080, Z80 and M6800 and the results can be either punched, or directly used to program the PROMS.

Impact of microprocessors in crystallography

The range of power of microprocessors does nor permit one to use them now for important calculations such as direct methods, structure refinements or Fourier series. For smaller problems, crystallographers may use some desk-top calculators that are all based on microprocessors. The mos important impact of microprocessors will be in the field of instrumentation, either as an integrated part in the internal logic of delivered apparatus, or used to control home-made experiments and special devices. The microprocessor is so powerfull that it must be prefered to any wired logic as far as its throughput remains compatible with the experiment that it should control.

References

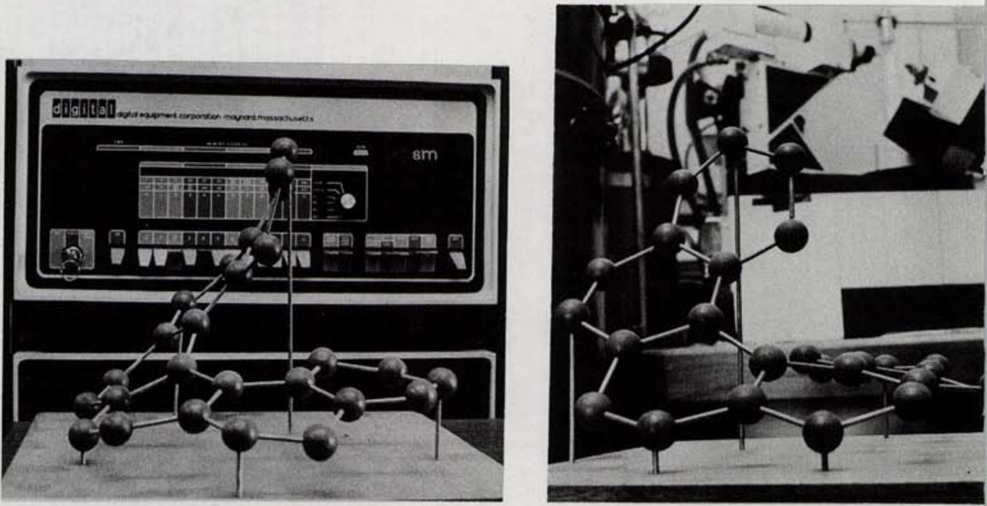
1. Bala Parajuraman: High-performance microprocessor architecture, Proc. IEEE, vol. 64, no 6 (851,859).
2. A. Thampy Thomas: Architecture and application of a 12-bit CMOS microprocessor, Proc. IEEE, vol. 64, no 6 (873,881).
3. Laurence Altman: Microcomputer families expand, part 1: the new chip

Electronics (1977), vol. 50, no 25, p. 89, 99.
4. B. Jeffrey Katz, Stephe P. Morjz, William B. Pohlman, Bruce W. Renevel: Electronics (1978), vol. 51, no 4, p. 99, 104.
5. Microprocessors-special issue: Chips and Boards, Electronics (1976), vol. 49, no 8. p. 77, 100.
6. Mark J. Sebern: A minicomputer-compatible microcomputer system; The DEC LSI-11, Proc. IEEE, vol. 64, no 6 (881,888).
7. Raymon P. Capece: Microcomputer families expand, part 2: the new boards, Electronics (1977), vol. 50, no 26, p. 65, 72.
8. C. Mouget, C.N.R.S., Laboratoire de Cristallographie, Grenoble: Private communication.
9. F. De Bergevin, M. Brunel, C, Mouget, Laboratoire de Cristallographie, Grenoble: Private communication.
0. C. Mouget, P. Wolfers, C.N.R.S., Laboratoire de Cristallographie, Grenoble: Private communication.

THE DIGITAL EQUIPMENT PDP-8 COMPUTER IN CRYSTALLOGRAPHY

Helen Goossens*

The use of modern diffractometer systems for crystal structure determinations has greatly improved the accuracy of the results, and at the same time reduced the supervision and manual intervention from the crystallographer during datacollection.
A minicomputer is a necessary and indispensable part of such a modern single crystal diffractometer system.
The CAD-4 diffractometer uses the Digital Equipment PDP-8 computer for goniometer control and initial data processing.
In one of the possible configurations this diffractometer system

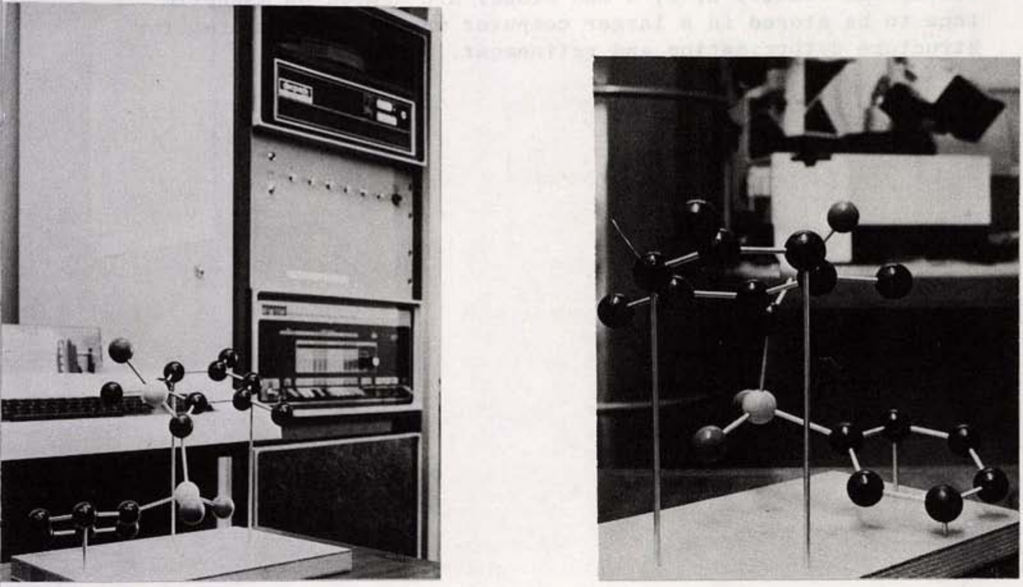


Molecular structure of a hexahelicene. Structural studies on a number of helicenes have been performed to establish the effect of substituents and the number of rings on the helicene 'core'. As a result of these studies it was found that in hexahelicenes bonds along the periphery of the helix are shortened and bonds around the inner 'core' of the molecule are lengthened. Substitution of methyl groups onto the inner carbon atoms of the end rings in hexahelicene gives rise to strong steric interactions and significantly changes the geometry of the helix. (Dr. Th.E.M. van den Hark, Dr. J.H. Noordik, Dr. W. Laarhoven; Depts. of Crystallography and Organic Chemistry, University of Nijmegen, The Netherlands.)

*DIGITAL EQUIPMENT B.V., Kaap Hoorndreef 38, Utrecht

consists of a high stability X-ray generator, radiation detection equipment, a single crystal goniometer and a PDP-8/M computer with a 16K core memory together with a RK05 disk and a magnetic tape recorder. The disk is used to store the necessary diffractometer control programs and the mathematical programs needed for the datacollection procedure and the initial data processing. The disk also serves as intermediate storage device for the temporary storage of experimental results before these are transported to magnetic tape.

During datacollection the PDP-8 controls the goniometer axes, the aperture disk, the attenuation filter, the shutter, the timer, the scaler and, if wanted, a balanced filter. Searching for reflections, optimizing of the reflection position and primitive cell calculation are done under PDP-8 control without manual intervention. The same holds for datacollection; once started, the data collection process proceeds completely



Molecular structure and geometrical configuration of E-Mesityl(phenylsulphonyl)sulphine. Structural studies on this compound and on a number of related isomeric sulphines have been performed to provide an unambiguous proof of the geometrical configuration of structurally related sulphines in the solid state. Results of these studies in the solid state were compared with the results of NMR experiments in solution, from which preferred conformations of the sulphines in solution could be deduced. (Dr. Th.W.A. Hummelink, Dr. J.H. Noordik, Dr. A. Tangerman, Prof. Dr. B. Zwanenburg; Depts. of Crystallography and Organic Chemistry, University of Nijmegen, The Netherlands.)

under PDP-8 control. Reflections to control the crystal orientation are monitored after user specified intervals and when the deviation between the measured orientation and the orientation at the beginning of datacollection exceeds a user specified maximum, automatic reorientation is done.

All the software, except a few goniometer control routines, is written in standard FORTRAN, and is executed under OS-8/FRTS. The FORTRAN programming makes user specific software modifications very easy to introduce.

Initial processing of the reflection intensity data may also be performed by the PDP-8 computer.
Although not absolutely necessary, then a hardware Floating Point Processor is a very helpfull extension of the basic PDP-8 configuration. The RK05 disk is used for temporary storage of data and scratch space for the profile analysis program.
Reflection profiles are analysed and reflections are corrected for Lp and finally h, k, l and F(obs) are dumped on magnetic tape to be stored in a larger computer where they are used for structure determination and refinement.

INDEXING POWDER DIAGRAMS

Robin Shirley*

Summary

Using computer-based methods developed in the last 5-10 years, one can now undertake with reasonable confidence to index the powder pattern of an unknown material that is a single solid phase, irrespective of its symmetry, provided that the data have been recorded properly. This means that a focussing camera or diffractometer was used, and the pattern calibrated so that zero error is less than 0.02° of 2θ.

A powder pattern does not usually yield a single solution, whose correctness is obvious by inspection. The low angle lines restrict the range of possible unit cells, but those at higher angles are relatively easy to fit to incorrect solutions. There may be a large number of plausible pseudo-solutions - as many as 500 is not unusual. Fortunately, we have figures of merit which discriminate strongly between them, and the solution with highest figure of merit is nearly always the correct one (assuming that this has been included in the list). It is thus important to examine all maxima in those parts of the solution field within which the correct solution should lie, using the semi-exhaustive methods that are now available.

The following approaches are described, embodied in the programs indicated (the name of the version used at the Summer School is also shown, if different):

1. Runge-Ito-de Wolff zone-indexing method: Visser (ITO=FZON)
2. Index-trial methods: Kohlbeck (TMO=KOHL), Werner (TREOR=TRER), Taupin (POWDER=TAUP)
3. Successive dichotomy: Louër (LOUV,LOUM)
4. Grid search: Shirley (POWDER 49); grid search by dichotomy: Louër and Shirley (LOSH)
5. Combination - zone-indexing/grid/search/dichotomy: Shirley, Louër and Visser (LZON)
6. Other methods: Smith and Kahara (QTEST)

Because of the different methods used, it is valuable to have more than one program available. Some advice is given concerning the types of problem that suit each best, and their running times, etc.

---

* Chemical Physics Dept., University of Surrey, England

Introduction

Great advances in powder indexing methods have been made in the last decade, as part of a general renaissance in quantitative powder diffraction methods. In this context, powder indexing means the reconstruction of crystal lattices from d-spacing data, without prior knowledge of the unit cell dimensions. Except in the highest symmetry cases, this can only be carried out for single solid phases, not for mixtures. Thus it is quite distinct from the more widespread uses of powder patterns for phase identification and analysis, though the results can often aid these processes. Although, as we shall see, indexing can involve substantial computations, there are many important cases where crystal form and texture prevent the use of single-crystal X-ray diffraction methods, and only powder data are available.

Powder Constants and Zones

If a lattice is described by its six direct cell constants, general d-spacing relations need implicitly to include transformations between direct and reciprocal space, and become of fearsome complexity. For indexing purposes it is much simpler to convert observed spacing data to $Q = {}^{1}/d^{2}$ and the reciprocal cell to its quadratic form, giving the six powder constants:

$$\begin{aligned} Q_A &= a^{*2} & Q_D &= 2b^{*}c^{*}\cos\alpha^{*} \\ Q_B &= b^{*2} & Q_E &= 2c^{*}a^{*}\cos\beta^{*} \\ Q_C &= c^{*2} & Q_F &= 2a^{*}b^{*}\cos\gamma^{*} \end{aligned}$$

The various $Q_{hk\ell}$ are then given by the equation:

$$Q_{hk\ell} = h^2Q_A + k^2Q_B + \ell^2Q_C + k\ell Q_D + \ell hQ_E + hkQ_F$$

Reciprocal lattice points lying in a plane through the origin constitute a central zone in reciprocal space. The corresponding powder lines are said to belong to one powder zone. Assuming a provisional assignment of axes, any such zone can be described as hk0, and needs only three powder constants:

$$Q_{hk0} = h^2Q_A + k^2Q_B + hkQ_F$$

It has become customary to multiply Q values by 10^4 to give more convenient numbers, and throughout the chapter we shall adopt these Q-units of $\text{Å}^{-2} \times 10^{-4} = 10^{16}\ \text{m}^{-2}$ (abbreviated to QU).

Figures of Merit

Experience has shown on a great many occasions that a list giving apparently good agreement between observed and calculated d-values is of little use as a guide to the correctness of a proposed cell. This is because d-spacings are so non-linear with respect to the true accuracy of measurement of 2θ, and because such a comparison can be 'improved' to any arbitrary degree simply by making a sufficient increase in the volume of the proposed cell.

These deficiencies can be corrected as follows. The discrepancy δ_i for the i-th observed line is defined as its absolute difference in QU from the nearest calculated line. This cannot exceed a 'target' ε_j which is half the distance in QU between the two calculated lines which bracket it. The ratio between mean target <ε> and mean discrepancy <δ> automatically allows for the cell volume, and forms the basis of a family of well-behaved figures of merit of order n, where n is the power to which the discrepancies are raised.

$$\text{Unweighted:}\quad Mn = \frac{\langle\varepsilon\rangle}{(\langle\delta^n\rangle)^{1/n}} = \frac{\text{Qcalcmax}}{\text{2Ncalc}}\left[\frac{\text{Nobs}}{\Sigma\delta^n}\right]^{1/n}$$

$$\text{Weighted:}\quad WMn = \langle\varepsilon\rangle\left[\frac{\langle\omega\rangle}{\langle\omega\delta^n\rangle}\right]^{1/n} = \frac{\text{Qcalcmax}}{\text{2Ncalc}}\left[\frac{\Sigma\omega}{\Sigma\omega\delta^n}\right]^{1/n}$$

Ncalc should only include potentially observable lines. Thus one should not count symmetry-equivalent lines separately, nor any space group absences, as far as they are known.

Although some of the higher Mn and WMn can be more powerful at locating the correct solution, it is the first order unweighted M1 that is usually quoted for comparison purposes, in the form of the de Wolff (1968) figure of merit M_{20}, to which it is broadly equivalent if calculated for the first 20 observed lines, all of which are counted as indexed.

$$M_{20} = \frac{\text{Qobsmax}}{\text{2Ncalc}\langle\delta\rangle}$$

It is worth checking how programs treat unindexed lines when calculating figures of merit. De Wolff defined M_{20} as being for the first 20 observed and indexed lines i.e. - the first 22 if 2 have not been included as indexed (X_{20}=2, as defined below). This guards against M_{20} apparently improving when the worst discrepancies are omitted as 'not indexed', but it is less convenient for the programmer and thus may not have been adopted.

The whole question of the treatment of unindexed lines can lead to great difficulty. One is strongly recommended to try to remove all doubtful lines from the data on a priori grounds before indexing commences, in the hope of avoiding the dubious procedure of discarding data a posteriori because they do not fit the proposed model, a practice that is as hazardous as it is tempting. De Wolff used X_{20}, the number of unexplained lines among the first 20 observed and indexed lines, as a rough measure of the extent of contamination by impurities.

As we shall see, it is not possible to define an absolute level of M_{20} which guarantees correctness, because the important thing is the relative value compared with those for other proposed cells. Nevertheless, it is rare for an incorrect cell to have $X_{20}=0$ and $M_{20}>20$.

A different figure of merit F_N has been proposed by Smith and Snyder (1977) defined (using the nomenclature here) as:

$$F_N = \frac{1}{<\delta 2\theta>} \cdot \frac{\text{Nobs}}{\text{Ncalc}}$$

where N = Nobs, the number of observed lines included (30 is recommended), and $<\delta 2\theta>$ is the mean absolute discrepancy in $2\theta^0$. F_N is intended mainly for summarising the accuracy and completeness of an observed pattern assuming the correct cell to be known, while M_{20} is for assessing the plausibility of proposed unit cells and thus reflects the fact that larger volume and lowered symmetry makes the correctness of a cell harder to establish from powder data.

In practice it is not easy to state an exact and definitive value of M_{20} (or other figures of merit) for a particular indexing solution, because of their sensitivity to small cell parameter differences. Different refinement programs (or changing the control parameters for a single program) may easily yield figures of merit differing by more than 10% for the same solution, and only differences between solutions that are clearly larger than this should be considered significant. Further work in this area leading to a firmer concensus would be welcome.

Alternative Settings and Geometrical Ambiguities

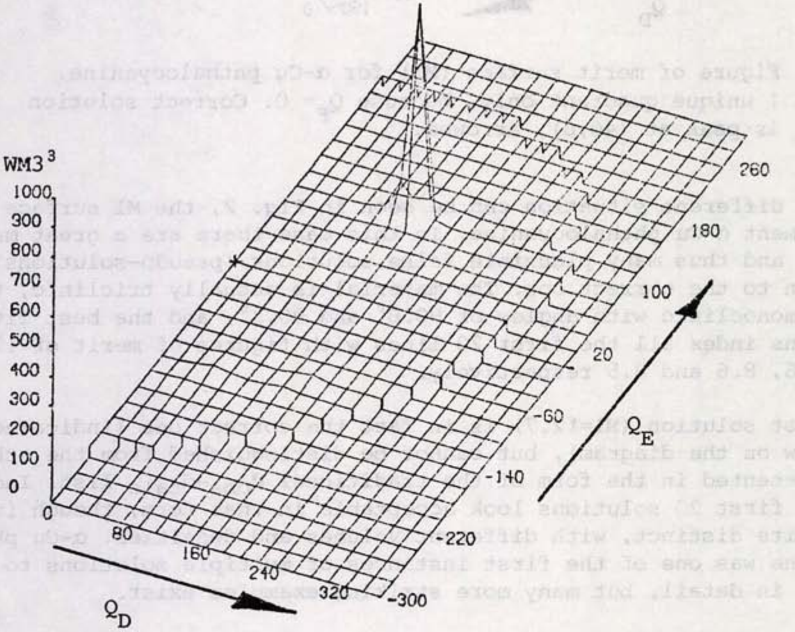
A lattice can be described by an infinite number of choices of cell within it, and every high-symmetry lattice yields numerous, less convenient cells in lower crystal systems. However, these alternative settings are easily recognised as the same solution, because they all lead to the same reduced (Niggli) cell (Mighell, 1976). Sub- or superlattices may be obtained, for example by halving or centring the true cell, but again the reduced cells will indicate that these are not independent solutions. In all these cases the cell volumes will be either equal or in a simple ratio; this is a useful recognition aid, but not as conclusive as reduced cell comparisons.

Mighell and Santoro (1975) have shown that certain distinct lattices, with different though related reduced cells, can give geometrically identical powder patterns. This geometrical ambiguity is only found when special relationships exist between the parameters of the two lattices, and cannot occur between general cells of orthorhombic or lower symmetry. However, it always occurs when the metrical symmetry of one of the lattices is higher than orthorhombic. For example, if one xamines the numerous lower symmetry cells that can account for a hexaonal pattern it will be found that some of them are genuine geometrial ambiguities.

n practice the situation is easily detected, because, as far as is nown, the volumes of such cells will always be in simple relationship. he higher symmetry solution is usually the correct one, although one annot be sure of this without studying the line multiplicities or ntensities, and cases are known where the opposite is true. For example α-$Li_4B_2O_5$ (de Wolff, 1968) can be indexed exactly as hexagonal, ut has been shown by single crystal work actually to be orthorhombic.

ultiple Solutions

f the figure of merit surface for a particular powder pattern is xamined, possible solutions show up as peaks. Fig. 1 shows the sharened figure of merit (WM^3) surface for hexamethylbenzene when Q_D and $_E$ are varied (generated by grid search using POWDER 49). It can be een that there is only a single prominent peak, and so in this case a nique solution. We are usually not so fortunate.



ig. 1. Sharpened figure of merit surface ($WM3^3$) for hexamethylbenzene. Unique area only (2 quadrants), sampling interval 20 QU.

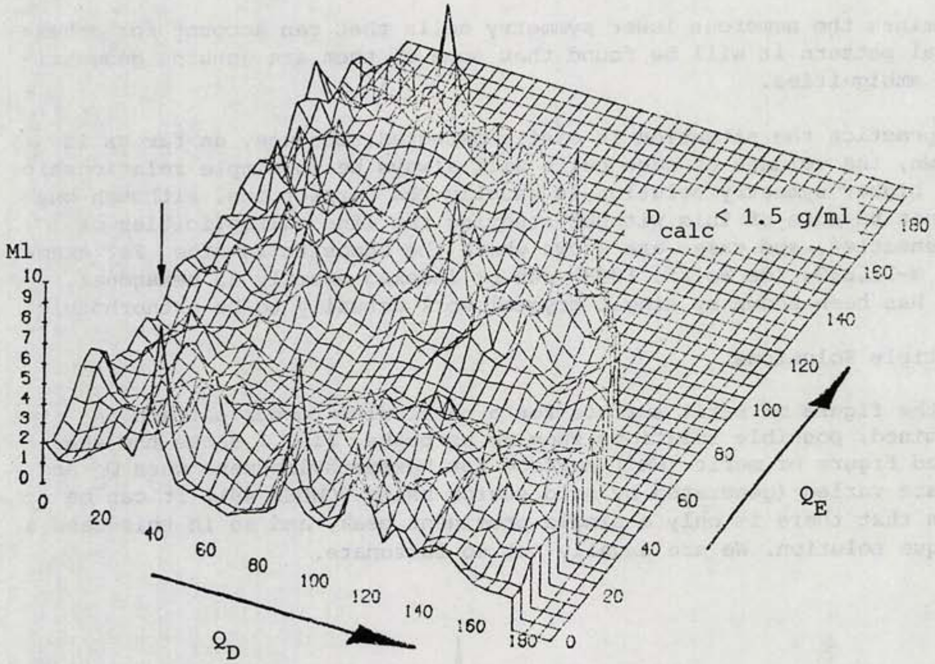


Fig. 2. Figure of merit surface (Ml) for α-Cu phthalocyanine. 1 unique quadrant only, because $Q_F = 0$. Correct solution is peak at (40,6), arrowed.

A quite different situation can be seen in Fig. 2, the Ml surface for the pigment α-Cu phthalocyanine. In this case there are a great many maxima, and thus many plausible false solutions (pseudo-solutions) in addition to the correct one. The material is actually triclinic, though pseudo-monoclinic with angles of 90.8^0 and 90.3^0, and the best five solutions index all the first 20 lines with figures of merit of 12.7, 9.7, 9.6, 8.6 and 8.5 respectively.

The first solution (Ml=12.7) is in fact the correct one (indicated by an arrow on the diagram), but cannot be distinguished from the others when presented in the form of the traditional d_{obs}-d_{calc} list. Indeed, all the first 20 solutions look acceptable in that form, though in fact quite distinct, with different volumes and densities. α-Cu phthalocyanine was one of the first instances of multiple solutions to be studied in detail, but many more striking examples exist.

The first three solutions for monoammonium urate give M_{20}=21, 15 and 13, while for monosodium urate monohydrate the best five cells give 30, 21, 20, 19 and 18, and no less than 31 cells index all the first 20 lines with a figure of merit better than 10. There are in fact some 200 tolerably plausible solutions for this pattern, although only the best

36 have been investigated in detail. More extreme examples are not uncommon: using the LOSH and LZON programs, the search for solutions hat index the first 20 lines within their 2θ error limits is often erminated at 500 cells, the present storage maximum.

n each of these cases, the solution with highest figure of merit is orrect, but can only be identified by a comparison with the entire set. here is no level of figure of merit that by itself guarantees correct-ess. In fact, false solutions can occur with any arbitrarily high igure of merit.

he powder zone whose reciprocal net cell has smallest area S will enerate most of the calculated lines near the origin, and, unless also glide plane, will index most of the low angle observed lines. When S pproaches half the geometrical mean value $(V^*)^{2/3}$, this dominant zone ill probably index about 10 out of the first 20 lines, leaving the emaining three powder constants to explain only about 10 upper-level ines. These are not very difficult to fit approximately by chance, so uch materials will tend to give numerous, plausible, multiple solu-ions, as occurs with the three examples above.

t would be nice to be able to determine all the multiple solutions by xhaustive search, but if one uses reasonable search limits and locates olutions on a 5QU grid, it turns out that more than 10^{14} different lat-ices are possible. At the rate of one per 3ms, they would take 12000 ears to search by direct exhaustive trial, and so it is not likely hat an exhaustive search will ever be feasible in the general case.

owever, it is a reasonable method for high-symmetry systems, and, iven an efficient search algorithm, for orthorhombic too. In lower ymmetry systems, semi-exhaustive methods are possible, in which the earch is restricted to 2 or 3 dimensions by carefully chosen simpli-ying assumptions.

imilar conclusions can be drawn if one approaches the problem by as-igning Miller indices to each of N observed lines where N need not xceed the number of unknown parameters, so that solutions are located n a 3N-dimensional index-space in which only integer co-ordinates are ossible.

t is clear that one should not think of tackling more than one solid hase at a time (an extra six parameters per impurity), nor inadequately alibrated ones (at least 1 or 2 more parameters).

ata Requirements

t is a waste of time to attempt to index poor data, unless the solu-ion is trivially simple. Good resolution is needed, to at least $\frac{1}{4}^{0}$ of θ between adjacent lines, which effectively rules out standard Debye-cherrer cameras.

If data errors were negligibly small, the correct solution would have an enormous figure of merit compared with the relatively modest ones of the various pseudo-solutions. Data errors degrade the true solution, but do not greatly affect the field of pseudo-solutions produced by accidental superimpositions. As one peak declines, another will improve, so that the general character of the group of false solutions is resistant to change. In the limit, the true solution is so degraded that it can no longer be distinguished from the pseudo-solutions, and the data become unsolvable.

Systematic error is much more important than random error. It is desirable to keep zero error below 0.02^0 of 2θ, while random variation with $\sigma(2\theta)$ up to 0.03^0 is tolerable (based on the findings of Rouse (1973), revised downwards in the light of later experience). The random error requirement is easily met - routine measurements ordinarily give $\sigma(2\theta)<0.02^0$. Zero error is more difficult and requires careful calibration, preferably against an internal standard, particularly in the case of diffractometers with the usual para-focussing geometry, where the effective sample position is not obvious and varies from sample to sample. The full calibration equation is usually parabolic in form. Uncalibrated patterns are unlikely to be solvable unless zero error can be estimated a posteriori from higher orders, where these can be identified unambiguously.

As has been seen, it is very desirable to eliminate spurious lines. An impurity level of <1% should be sought (and verified in a density gradient column), and lines from other X-ray wavelengths also tested for (see also the recommendations of the Commission of Crystallographic Data, 1971).

Unit Cell Volume

A good density measurement is desirable, because it can suggest a powerful constraint on cell volume, provided that one has a well-established empirical formula and a reliable value inferred for Z, and can exclude the possibility of fractional site occupancies.

Cell volume is very important for indexing because cpu time tends to rise disproportionately with it, and also because it is related to the maximum figure of merit that can be expected for data of a given accuracy (Werner, 1976):

$$M_{20}\max \simeq \frac{3md_{20}}{8\pi V\langle\delta\rangle}$$

where d_{20} is the d-spacing for the 20th observed line, m is the multiplicity factor and $\langle\delta\rangle$ is the estimated precision in $1/d^2$. This can be rearranged to give a critical volume for proposed cells beyond which the data will not on average be able to give better than a specified $M_{20}\max$:

$$V_{crit} \simeq \frac{3md_{20}}{8\pi M_{20}\max\langle\delta\rangle}$$

Estimates of the cell volume from the closeness of lines in the observed pattern are discussed by Lipson (1949) and Smith (1976). The simplest method is that suggested by Smith (1977):

$$V \simeq 13.39\, d_{20}^{3}$$

for good focussing camera or diffractometer data from a triclinic material. The estimate should be multiplied by 1.5 and 2.3 for primitive monoclinic and orthorhombic cells respectively, and be increased further if glide planes, etc., are expected to be present.

Modern Indexing Methods for Orthorhombic and Lower Symmetry

Some methods search for solutions mainly in index-space, by varying Miller indices, while others search mainly in parameterspace, by varying cell parameters in some form.

A classification more directly important to the user is whether their approach is mainly deductive or exhaustive. In broad terms, deductive methods attempt to infer the values of lattice parameters from coincidences and relations between the observed lines, achieving speed at the cost of rigour. By contrast, exhaustive methods systematically search the relevant solutionspace, gaining rigour at the expense of speed. The classification is not rigid, because programs often incorporate some aspect of both approaches. For example, deductive programs generally employ some kind of systematic combination of the deduced possibilities, while most of the newer programs are semi-exhaustive, incorporating judicious deductions to limit the solution field so as to gain speed.

The following table shows how the programs to be described can be arranged under this classification (the program name is shown in brackets, followed by that of the version to be used at the Summer School, if different).

Table

	Parameter-space	Index-space
Deductive:	Visser (ITO=FZON)	Smith & Kahara (QTEST)
Semi-exhaustive:	Shirley (POWDER 49)	Werner (TREOR=TRER)
	Louër & Shirley (LOSH)	Kohlbeck (TMO=KOHL)
	Shirley, Louër & Visser (LZON)	
Exhaustive:	Taupin (POWDER=TAUP)	Louër (LOUV, LOUM)

Some aspects are common to all the programs and may usefully be considered together; for example, data input formats. Each program requires approximately the same data concerning the observed pattern, but generally in a different format. A small utility program PDAT is available to punch data decks in the appropriate format for each program. In the longer term it is hoped that a common format such as the proposed Standard Powder Data Interchange Format (SPDIF) will be accepted.

When solutions have been found, they need to be refined, evaluated and displayed conveniently for the user. FZON has extensive facilities for doing this, which have been extracted, turned into a new subroutine FZRF and incorporated into KOHL, LOUV, LOUM nad LOSH. It is desirable that eventually these should also include reduced cell calculations. A stopping rule is another facility needed by exhaustive and semi-exhaustive programs, to terminate the run on finding a solution that is very likely to be correct. It can be a difficult decision when a dominant zone is present, and more work in this area is desirable.

Zone-Indexing Approach (Runge; Ito; de Wolff; Visser)

Formally the earliest method of powder indexing, this was discovered by Runge (1917) but little used until rediscovered by Ito (1949, 1950), when it attracted great interest because of its application to triclinic cases. Ito's method operates in index space, but de Wolff (1957, 1958, 1963) extended and generalised it so that it ceased to depend on specific trial indexings, but varied the zone angle by systematic trial. In this form it is more accurately classified as a parameter-space method. It was programmed originally in Algol 60 by Visser (1969), and has been available from 1974 in FORTRAN IV.

The program (ITO=FZON) initially seeks central zones using pairs of low-angle lines as trial Q_A and Q_B. Trial indices "hk0" with $|h|, |k| \leq 2$ are given to each observed line and a tentative "Q_F" calculated. The most frequently occurring "Q_F" values are saved and the resulting trial zones refined and sorted. Pairs of zones having a common row are then combined and renamed, leaving Q_D as the only unknown parameter. Q_D is found by varying the angle between the zones in a second one-dimensional search, analogous to that for Q_F. The resulting solutions are then refined, sorted on lines indexed and figure of merit, displayed and tested for higher symmetry.

FZON is very fast, automatic and quite powerful, but must find and combine two correct zones in order to succeed. Its success rate is as high as 75% for high quality data, and because of its speed and convenience it is usually the best choice for the initial attempt at a new problem. It is relatively insensitive to spurious lines.

Index-Trial Methods (Kohlbeck; Werner; Taupin)

Werner's (1964) program (TREOR=TRER) and Kohlbeck's (1976, 1978) program (TMO=KOHL) are semi-exhaustive programs operating in index-space. Index-space methods generally proceed by selecting sets of base lines, each of which are just sufficient in number to solve for the unknown lattice parameters in that crystal system. These are systematically assigned sets of trial indices, hence yielding trial cells, which are hecked against volume constraints, etc., then used to try to index the emaining observed lines. The relatively few surviving trial cells are hen refined and evaluated, and displayed if their figure of merit is ufficiently high.

or low-symmetry systems, this process can only be made economical if he base lines are carefully selected and their index field judiciously imited so as to retain most of the solutions while achieving the ecessary large gains in speed. Both programs are powerful and tolerably fficient, being particularly useful in those low-symmetry cases where t proves difficult to select a reliable basis zone for grid search rograms. TREOR is restricted to monoclinic systems and above, succeeding in perhaps 75% of those cases (more if the data are of very high uality). KOHL tackles any system including triclinic, seems more tolerant of data quality, and so is probably the more powerful of the two.

aupin's (1973) program (POWDER=TAUP) makes fewer assumptions limiting he index fields, etc., and in effect makes an exhaustive search by on-redundant index permutation. Although it will attempt problems of ny symmetry, cpu times tend to become prohibitive for symmetries below rthorhombic and care should also be taken to avoid excessive output. ubject to these practical limitations, it will generally find all ossible solutions, including geometrical ambiguities, etc.

uccessive Dichotomy (Louër & Louër)

he successive dichotomy method of Louër and Louër (1972) is a form of inary search, probably the optimal exhaustive strategy in parameter-space.

ach variable parameter is divided into sections of a standard size ½Å for cell sides, 2½^0 for angles), so that the whole parameter-space s divided into n-dimensional domains. Using inequalities based on the rror bounds of the observed lines, each domain is tested and discarded f it cannot contain a possible solution. Each remaining domain is then isected on every side, giving 2^n sub-domains, which are tested as before. This is repeated a total of six times, after which only a few ub-domains remain, by now small enough to fix the location of the possible solutions, and these are printed out. In the current programs, his process is carried out for successive 400 $Å^3$ shells of volume, hus finding the smaller and more probable solutions first.

he two programs LOUV and LOUM, successors to those described in the 972 paper, apply to systems down to orthorhombic and to monoclinic espectively. LOUV is exhaustive and very fast, and is probably the ost effective method available for higher-symmetry cases. LOUM is very seful for low-volume monoclinic problems, but becomes rather demanding n cpu time as it moves to shells of volume beyond 800 $Å^3$ or so. Being xhaustive, both these programs can give negative as well as positive nformation (i.e. demonstrating that solutions of certain symmetries annot exist within the stated volume and error limits).

Grid Search Programs (Shirley; Louër & Shirley)

Grid search (Shirley, 1975) is a semi-exhaustive method in parameter-space, and can be thought of as an extension of zone-indexing, requiring only one valid zone to be known - the basis zone, hk0. A reliable line not indexed by this zone can arbitrarily be given indices 00ℓ, where in the first instance ℓ is assumed to be 1, hence giving Q_C and completing the basis set - a procedure devised, independently, by Ishida and Watanabe (1967, 1971) and by Shirley. Only two unknown constants now remain: Q_D and Q_E, giving the angles α^* and β^*. They are found by a grid search of the resulting solution space, which, being only 2-dimensional, is a relatively fast process. If the data and the basis set are valid, the correct solution must be found.

The method was first implemented in the SEARCH routine of Shirley's POWDER system (currently POWDER 49) written in Algol 60, using a systematic step-scan through the figure of merit surface, and producing plots such as those of Figs. 1 and 2. The resulting peaks are then refined and evaluated.

Although such plots are helpful for understanding the nature of powder-indexing solution fields, for most applications POWDER 49 has effectively been superseded by the Louër and Shirley program LOSH, which is written in Fortran IV and uses the much more efficient dichotomy search to locate α^* and β^*, gaining greatly in speed without loss of generality. LOSH is not automatic, however, because it is left to the user to provide a suitable basis set.

An experimental program LZON (Shirley, Louër and Visser) now exists, incorporating the powerful solution-finding procedure of LOSH into the zone-finding and solution-evaluation framework of FZON, to give a fully automatic process. Results so far look distinctly promising.

Other Methods (Smith & Kahara)

The QTEST program of Smith and Kahara (1975) is relatively unusual in that it is a mainly deductive method operating in index-space. It is aimed specifically at the monoclinic system and takes advantage of special relations involving 020 that are common among monoclinic patterns. It is reported to be fast and reasonably efficient, automatically solving approximately 60% of monoclinic cases (more if supplied with high-quality data).

Which Programs to Use?

All the programs discussed usually succeed in at least 50% of cases, and many achieve better than 80%. Because they use a variety of quite distinct methods, their limitations tend to affect different problems, and so a user with three or four powerful programs available will seldom be unable to index a properly-measured pattern. In my experience, the following procedure is effective and economical, and it is offered as a suggested starting point.

For the initial run on a new pattern, I would recommend ITO(=FZON). It is automatic in operation, has no symmetry restrictions, offers a good chance of success at very little cost, and will in any case produce useful information to assist in the choice of basis zone, etc. Next if necessary, a run of LOUV will quickly determine whether there exist any solutions with orthorhombic symmetry or above. If not, the material is monoclinic or triclinic, and will usually yield to TMO(=KOHL) and/or LOSH. If both of these fail, there is probably something odd about the data, which should be checked for mispunchings, systematic errors or impurity lines, etc., before proceeding to reruns with modified settings and to TREOR, LOUM, etc.

Summary of Characteristics of Major Powder Indexing Programs

Program	Version (FORTRAN Deck)	Authors	Crystal Systems	Fully Automatic	Exhaustive	Success Rate (approx)	Execution sec (CDC 7600)
ITO	FZON	Visser	any	Yes	No	75%	1-5
LOUV	LOUV	Louër	orthorhombic+	Yes	Yes	(100%)	1-5
LOUM	LOUM	Louër	monoclinic	Yes	Yes	(80%+)	30-1200
LOSH	LOSH	Louër & Shirley	any	No	Semi	80%+	10-300
LZON	LZON	Shirley, Louër & Visser	any	Yes	Semi	?High	30-1200
POWDER 49	(Algol)	Shirley	any	(Yes)	Semi	80%	(20-300)
POWDER	TAUP	Taupin	any (triclinic impractical)	Yes	Yes	50%+	50-1200+
TMO	KOHL	Kohlbeck	any	Yes	Semi	80%+	2-200
TREOR	TRER	Werner	monoclinic+	Yes	Semi	75%	2-100
QTEST	-	Smith & Kahara	monoclinic	Yes	No	(60%)	(30)

Notes:
1) Crystal system followed by + means "and all higher systems".
2) Success rate in brackets means "within its crystal system limits".
3) Other bracketed quantities indicate an estimate, or some other qualification.
4) Execution times show the typical range for that program, but not absolute maxima.

References

Commission on Crystallographic Data (1971). J. Appl. Cryst. 4, 81-86.
Ishida, T. & Watanabe, Y. (1967). J. Phys. Soc. Japan 23, 556-565.
Ishida, T. & Watanabe, Y. (1971). J. Appl. Cryst. 4, 311-316.
Ito, T. (1949). Nature 164, 755-756.
Ito, T. (1950). "X-ray Studies on Polymorphism" (Maruzen, Tokyo), 187-228.
Kohlbeck, F. & Hörl, E.M. (1976). L. Appl. Cryst. 9, 28-33.
Kohlbeck, F. & Hörl, E.M. (1978). J. Appl. Cryst. 11, 60-61.
Lipson, H. (1949). Acta Cryst. 2, 43-45.
Louër, D. & Louër, M. (1972). J. Appl. Cryst. 5, 271-275.
Mighell, A.D. & Santoro, A. (1975). J. Appl. Cryst. 8, 372-374.
Mighell, A.D. (1976). J. Appl. Cryst. 9, 491-498.
Rouse, J.E. (1973). M.Sc. Dissertation, University of Surrey.
Runge, C. (1917). Z. Physik. 18, 509-515.
Shirley, R. (1975). Acta Cryst. A31, S197.
Smith, G.S. & Kahara, E. (1975). J. Appl. Cryst. 8, 681-683.

Smith, G.S. (1976). J. Appl. Cryst. 9, 424-428.
Smith, G.S. (1977). J. Appl. Cryst. 10, 252-255.
Smith, G.S. & Snyder, R.L. (1977). Private communication.
Taupin, D. (1968). J. Appl. Cryst. 1, 178-181.
Taupin, D. (1973). J. Appl. Cryst. 6, 380-385.
Visser, J.W. (1969). J. Appl. Cryst. 2, 89-95.
Werner, P.-E. (1964). Z. Kristallogr. 120, 375-387.
Werner, P.-E. (1976). J. Appl. Cryst. 9, 216-219.
Wolff, P.M. de (1957). Acta Cryst. 10, 590-595.
Wolff, P.M. de (1958). Acta Cryst. 11, 664-665.
Wolff, P.M. de (1963). Adv. X-ray Anal. 6, 1-17.
Wolff, P.M. de (1968). J. Appl. Cryst. 1, 108-113.
Wolff, P.M. de (1972). J. Appl. Cryst. 5, 243.

RATMAC: AN ADAPTATION OF RATFOR AND MACRO OF KERNIGHAN AND PLAUGER

James M. Stewart and R. J. Munn[*]

Summary

The purpose of this paper is to show a method of changing FORTRAN to a "structured" or "top-down" higher level language, RATFOR. This language is closely related to FORTRAN and in fact, is simply a precompiler which produces FORTRAN. RATFOR, therefore, allows an evolutionary transition from FORTRAN to PL1, PASCAL or some other structured programming language. In addition to offering this transition, the preprocessor offers a MACRO feature which provides a powerful construct for making programs transportable.

Introduction

RATFOR and MACRO are the work of B.W. Kernighan and P.J. Plauger. The details of these programs and the philosophy of their use is described in detail in their book "Software Tools" (1). The symbolic decks of RATFOR and MACRO may be obtained from Addison-Wesley. The combined modification of of these programs, RATMAC, is described in the technical report TR-675 (2). RATMAC will be used for the implementation and distribution of the new XRAY system tentatively designated XSTAL or XSTAL 80 (?).

Programming languages and transportibility of programs

The problem of writing application programs in a way that will allow their use on most digital computers is an intriguing one. The problem is particularly interesting to crystallographers because many of the algorithms they use in the course of crystal structure analysis are already clearly and efficiently programmed. It therefore seems a shame to reprogram extensively to move codes from one computer to another. The fact is that the architecture of computing machines varies widely while the computing languages vary subtly. Because of these factors, it is an interesting challenge to find programming methods which will permit ready transportibility of application programs.

At the present time, FORTRAN seems to be the most universally and commonly used "higher" compiler language. It has been used extensively within the community of scientific programmers. It is probably fair to say that the most used crystallographic programs are written in FORTRAN.

---

[*] University of Maryland, U.S.A.

RATFOR

RATFOR is an acronym for rational FORTRAN. It is a precompiler written in RATFOR which translates (into FORTRAN) a structured program written i RATFOR. Most of the syntax of RATFOR is exactly the same as FORTRAN. There are, however, certain "top-down" or "structured" control statement and the use of brackets which distinguish a RATFOR program from its FORTRAN counterpart.

For those not yet familiar with "structured" programming it may be helpf to show a few simple examples as an introduction. In a structured progra the "flow" is from start to finish. Parts of the code may be bypassed based on conditions, but in general there is no "going back". For this reason GØ TØ is not used. There are, instead, a number of control statements which serve to allow the structuring. They are:

IF
ELSE IF
ELSE
WHILE
FØR
REPEAT
REPEAT UNTIL
BREAK
NEXT

Control is maintained by the use of brackets, i.e. {}, or $($) if {} is not available on the local machine. The brackets delimit the range of th control statements.

The following trivial examples, first in FORTRAN, then in RATFOR, should serve as an introduction. The listing of the XSTAL system has more complicated examples.

The FORTRAN

```
C-----LØØP ØVER ALL ATØMS
      DØ 7 J=1, MAXAT, 3
C-----TEST FØR UNWANTED ATØMS
      IF (X(J).LT.EX) GØ TØ 7
C-----FIND HX + KY + LZ
      ARG = H1*X(J)+H2*X(J+1)+H3*X(J+2)
  7   CØNTINUE
```

Then the RATFOR to accomplish the same task:

```
FØR (J=1;J<=MAXAT; J=J+3){≠LØØP ØVER ALL ATOMS
     IF(X(J)>EX) {≠ ØNLY TREAT WANTED ATØMS
          ARG=H1*X(J) + H2*X(J+1) + H3*X(J+2)
     }≠ THAT GIVE HX + KY + LZ
}
```

The indentation is simply to help the eye in following the action. The RATFOR preprocessor takes care of all the amenities in producing the FORTRAN code.

Another example of FORTRAN: (Don't try to optimize it):

```
      IF (A.LE. 1.0) GØ TØ 5
      IF (A.LE. 2.0) GØ TØ 7
      B= 28.0
      C= 400.0
      GØ TØ 11
  5   B= 7.0
      C= 100.0
      GØ TØ 11
  7   B= 14.0
      C= 200.0
  11  CØNTINUE
```

Then the corresponding RATFOR:

```
IF(A<=1.0) {
  B=7.0
  C=100.0
}
ELSE IF (A<=2.0) {
     B=14.0
     C=200.0
}
ELSE  {
     B=28.0
     C=400.0
}
```

An example of the WHILE statement; first the FORTRAN:

```
      J=1
      JM=10
  5   CØNTINUE
C-----READ UNTIL FIRST WØRD ØF RECØRD IS ZERØ ØR LESS
      READ(9) (A(K),K=J,JM)
      IF (A(J).LE. 0.0) GØ TØ 10
      J=J+10
      JM=JM+10
      GØ TØ 5
  10  CØNTINUE
```

and then the RATFOR:

```
J=1 ≠ SET STARTING CØUNTER
JM=10  ≠ SET ENDING CØUNTER
READ (9) (A(K), K=J,JM) ≠ MAKE INITIAL READ
WHILE (A(J)>0.0) {READ UNTIL FIRST WØRD ØF RECØRD
      J=J+10 ≠ IS ZERØ
      JM=JM+10
      READ (9) (A(K), K=J,JM)
```

```
}
```

An alternative RATFOR program could be:

```
J=1
JM=10
REPEAT {
       READ (9)(A( K), K=J,JM)
       J=J+10
       JM=JM+10
       UNTIL (A(J-10)< =0.0)
}
```

Within the computer science community more and more emphasis is being given to structured programming. Languages such as PASCAL and PL1 are be urged as the way of the future. From the standpoint of a working scienti such as a crystallographer, which of these languages and their dialects will supercede FORTRAN is unclear. RATFOR solves this problem by being a preprocessor which simply "edits" the structured RATFOR into the local FORTRAN dialect. Indeed, during the transition or recording period, one may have coexisting RATFOR and FORTRAN subroutines. Once the wave of the future is clearly defined, it is possible to recode RATFOR to produce, say, PASCAL instead of FORTRAN.

MACROS

A MACRO is a "large" portion of code represented by a short descriptive string of characters. For example, if one has a group of variables used in many FORTRAN subroutines, a COMMON declaration might be made in the following way:

```
       CØMMØN/SYS/A,B,C,I,J,K(100)
       INTEGER I,J,K
       REAL A,B,C
```

The MACRO processor may be invoked as:

```
MACRØ ( SYSCØM:,[CØMMØN/SYS/A,B,C,I,J,K(100)
        INTEGER I,J,K; REAL A,B,C])
```

so that in all RATFOR subroutines, all that is required is:

```
SYSCØM:
```

and the CØMMØN and TYPE statements will be carried forward to the FORTRAN code.

For purposes of isolating the differences among all the FORTRANS that exist and among the various machines that are built, the MACRO feature of Kernighan and Plauger's work is the most beneficial. The matter of structured programming is of academic interest. The MACRO capability is the real "work-horse". It is the main reason for using RATMAC in the new XRAY system.

As a second example consider the problem of I/O arrays with the following example of the use of MACROS to isolate machine dependent quantities.

```
MACRØ(BUFSIZ:,256)
MACRØ(IØBUF:,[CØMMØN/IØ/INPUT(BUFSIZ:),OUTPUT(BUFSIZ:)])
```

In the RATFOR code the line

```
IOBUFS:
```

will result in the FORTRAN code having

```
    CØMMØN/IØ/INPUT(256),ØUTPUT(256)
```

It is possible by similar techniques using the clever built-in functions and recursive properties of the MACRO processor to isolate all machine specific, compiler specific statements. This then provides for the simplification of the editing necessary to transfer the application programs from one machine to another. Moreover, the transformation of the RATFOR programs may be made either before or after the transfer from the first machines to the second.

Quantities such as bits per word, bytes per word, bits per byte, words per output line, etc. may all be isolated in MACRO. Arrays for all major CØMMØN statements may be left indefinite in terms of word size, number of bits, etc. Input-output may be handled in general by use of MACROS. Even the problem of setting up DATA statements as 2HXX or 4HXXXX or 10HXXXXXXXXXX, may be vastly simplified and dealt with using the MACRO processor.

Application of RATMAC to the XRAY system

The power of RATFOR and MACRO is illustrated in the coding of the new XRAY system, XSTAL. In addition the RATMAC technical report gives some excellent examples of the use of Kernighan and Plauger's powerful preprocessor.

The book "Software Tools" is well worth reading. On page 319-321, the authors summarize their programming principles:

  I Keep it simple
 II Build it in stages
III Let someone else do the hard part.

These principles are disarmingly artless when written this way. But when they are applied systematically to application programs, the results they produce, are indeed artful.

Because of the way the RATMAC preprocessor has been programmed, it is readily transportable from one machine to another. For the future of the XSTAL system, the FORTRAN version of it will be distributed as the first of seven files.

FILE I	The RATMAC processor in FORTRAN tailored for the local machine
FILE II	The RATMAC processor in RATMAC. This will allow for local enhancements to be made and a new FORTRAN version prepared by use of FILE I.
FILE III	The RATMAC write-up.
FILE IV	The XSTALprogram in RATFOR. The first part of this file will b the MACROS of XSTAL. These MACROS will have to be edited to conform to local operating architecture. The rest of editing will be done by the RATMAC processor to produce the local FORT code.
FILE V	The XSTAL write-up.
FILE VI	The XSTAL test deck.
FILE VII	A collection of "machine code" or "assembler" routines which have been used to eliminate the FORTRAN I/O library on various machines. These "primitives" are to serve as models for other machines where it is desirable to enhance performance or save space. This is not a mandatory procedure. All the MACROS can b set to use standard FORTRAN I/O.

It is hoped that the application of RATFOR will be of aid to future programming efforts. The acceptance on April 3, 1978 of the new ANSI FORTRAN 77 standard means a new set of rules for FORTRAN users. New machines and operating systems are already being supplied. These standar make XRAY76 obsolete even on large scale computers. In addition the adve of "mini" and "midi" computers suggests that a restructuring and moderni tion of XRAY is needed. The use of a preprocessor promises a certain sta bility in codes not available by the use of "PIDGIN" FORTRAN. We hope th this proves to be the case for XSTAL 80.

References

(1) Software Tools, Addison-Wesley, paperback 3669 (1976)

(2) TR 675, RATMAC: Kernighan and Plauger's Structured Programming Language, R. J. Munn and J. M. Stewart, Computer Science Center, University of Maryland, College Park, MD 20742.

TDS ERRORS IN REFLECTION INTENSITIES OF MOLECULAR CRYSTALS

Peter A. Kroon and Aafje Vos*

Summary

Programs LATDYN and TDS have been written in FORTRAN to perform lattice dynamical calculations of thermal diffuse scattering for a model of harmonically vibrating, weakly interacting, rigid molecules. Calculations on naphthalene at 100 K show that errors due to the TDS from optic modes and second order TDS can be neglected. The usual long wave (LW) approximation for acoustic modes overestimates the TDS corrections. A modified LW method is described which gives good TDS corrections and requires considerably less computing time than the lattice dynamical treatment.

Introduction

When scanning through a reflection not only the Bragg intensity plus "usual" diffuse background, but also first and second order thermal diffuse scattering (TDS) is measured. Fig.1. shows that the TDS profile does not proceed linearly below the Bragg peak, but peaks at the reflection position. Due to this the TDS contribution is not eliminated by the usual background correction. The remaining TDS contribution I_{TDS} (nett, H) is given by area B and obeys the formula

$$I_{TDS}(\mathrm{nett},H) = I_{TDS}(V_p^*) - \tfrac{1}{2}\left[I_{TDS}(V_l^*)\frac{V_p^*}{V_l^*} + I_{TDS}(V_r^*)\frac{V_p^*}{V_r^*}\right] \qquad (1)$$

where $I_{TDS}(V_i^*)$ is the total TDS intensity measured by scanning a reciprocal volume V_i^*, and V_p^*, V_l^* and V_r^* are the reciprocal volumes observed by the counter during the peak and the left and right background measurements respectively.

Various computer programs are available for the calculation of TDS corrections. However, for crystals with low symmetry, only acoustical modes of vibration are taken into account in the long wave approximation. To study the influence of these approximations TDS calculations were done for a model of harmonically vibrating, weakly interacting, rigid molecules. The interactions are described as the sum of pairwise interatomic interactions. At present the programs LATDYN and TDS work with infinite resolution of the experimental set up, apart from a finite counter opening. The part of reciprocal space seen by the counter during the scan can either be calculated according to Helmholdt &

---

*Department of Chemical Physics, Rijksuniversiteit Groningen,
Nijenborgh 16, 9747 AG Groningen, The Netherlands

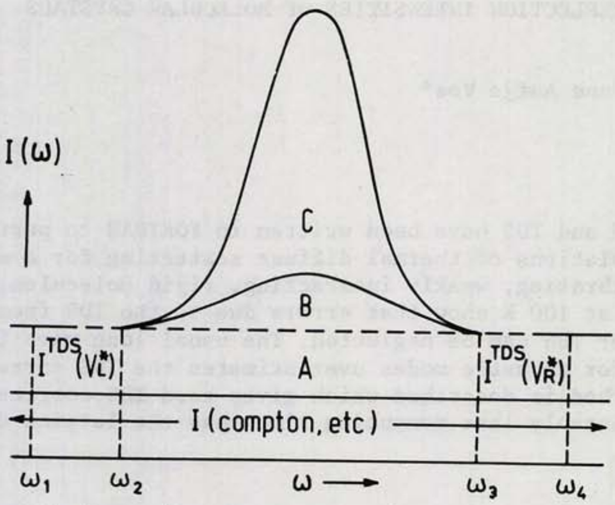


Fig.1. Scattered intensity vs. scan angle w. See text. A + B = TDS superimposed on diffuse background, C = Bragg intensity, B = TDS error

Vos[1] or can be taken the same for all reflections. Derivations of the formulae and further details of the work can be found in Kroon & Vos[2].

Models used and formulae for first order TDS

a. "Exact" lattice dynamical treatment (EX)

The lattice modes, formulated according to Cochran,[3] are characterized by an index $\sigma, \sigma = 1 - 6n$ where n is the number of molecules per cell, and by the wave vector $\underset{\sim}{k}$. All modes with wave vector $\underset{\sim}{k}$ give first order TDS at points $\underset{\sim}{S} = \underset{\sim}{H} - \underset{\sim}{k}$ in reciprocal space. For the rigid body model mentioned above, the total first order TDS intensity at $\underset{\sim}{S} = \underset{\sim}{H} - \underset{\sim}{k}$ is, in units directly comparable with the Bragg intensity $|F_B(\underset{\sim}{H})|^2$ given by

$$I_1^{EX}(\underset{\sim}{S}=\underset{\sim}{H}-\underset{\sim}{k})dv^*_{rel}(\underset{\sim}{k}) = 4\pi^2 k_B T S^2 dv^*_{rel}(\underset{\sim}{k}) \sum_{\sigma} \omega^{-2}(\sigma\underset{\sim}{k})$$
$$\Big[\sum_{ma} f(\underset{\sim}{S}ma) \times |\underset{\sim}{U}(\underset{\sim}{S}\sigma\underset{\sim}{k}ma)| \exp i[2\pi\underset{\sim}{H}.\underset{\sim}{r}(m) + 2\pi\underset{\sim}{S}.\underset{\sim}{r}(ma) + \psi(\underset{\sim}{S}\sigma\underset{\sim}{k}ma)]\Big]^2 \qquad (2)$$

$\underset{\sim}{S}$, $\underset{\sim}{H}$, k_B, $f(\underset{\sim}{S}ma)$ and T have the usual meaning, $\omega(\sigma\underset{\sim}{k})$ = angular frequency of mode $\sigma\underset{\sim}{k}$, $dv^*_{rel} = dv^*/V^*$, $\underset{\sim}{r}(m)$ = position vector for gravity center G of molecule m, $\underset{\sim}{r}(ma)$ = position vector of atom ma of m, relative to G. $|\underset{\sim}{U}(\underset{\sim}{S}\sigma\underset{\sim}{k}ma)| \exp i\psi(\underset{\sim}{S}\sigma\underset{\sim}{k}ma)$ is, apart from a scale factor and the phase factor $\exp 2\pi i\underset{\sim}{k}.\underset{\sim}{r}(m)$, the complex amplitude of the inner product $\hat{\underset{\sim}{s}}.\underset{\sim}{p}(\sigma\underset{\sim}{k}ma,t)$ where $\hat{\underset{\sim}{s}} = \underset{\sim}{S}/S$ and $\underset{\sim}{p}$ denotes the deviation of atom ma at time t due to mode $\sigma\underset{\sim}{k}$.

b. Long wave approximation for rigid molecules

(LWF or "long wave with Fourier transform" approximation)

For long wave acoustic modes ($k\rightarrow 0$) the dispersion of the acoustic velocity $v(\sigma\hat{\mathbf{k}})$ is zero. The dispersion of the angular frequency is thus given by

$$\omega(\sigma\mathbf{k}) = 2\pi k v(\sigma\hat{\mathbf{k}}) \qquad (3)$$

Lattice dynamical treatment for this situation by Born & Huang[4] shows that for mode $(\sigma\mathbf{k})$ all molecules in the unit cell move in parallel with an amplitude which only depends on the direction of $\mathbf{k}$. This implies that at $k\rightarrow 0$ the acoustic modes are purely translational. Substitution in (2) of (3) and

$|\mathbf{U}(\mathbf{S}\sigma\mathbf{k}ma)| = |\mathbf{U}^{LW}(\mathbf{S}\sigma\hat{\mathbf{k}})|$, $\psi(\mathbf{S}\sigma\mathbf{k}ma) = \psi^{LW}(\mathbf{S}\sigma\mathbf{k}) = 0$,

as equal phases may be set to zero, gives

$$I_1^{LWF}(\mathbf{S}=\mathbf{H}-\mathbf{k})dv^*_{rel}(\mathbf{k}) = [k_BTS^2/k^2]dv^*_{rel}(\mathbf{k})\times$$
$$\times \sum_{\sigma=1}^{3}[|\mathbf{U}^{LW}(\mathbf{S}\sigma\hat{\mathbf{k}})|/v(\sigma\hat{\mathbf{k}})]^2\Big|\sum_{ma} f(\mathbf{S}ma)\exp[2\pi i(\mathbf{H}\cdot\mathbf{r}(m) + \mathbf{S}\cdot\mathbf{r}(ma))]\Big|^2 \qquad (4)$$

The LWF approximation now implies that (4) is used not only for $k\rightarrow 0$, but also for larger $\mathbf{k}$ values. The F in LWF stands for the fact that (4) contains the Fourier transform $\sum_a f(\mathbf{S}ma) \times \exp[2\pi i\mathbf{S}\cdot\mathbf{r}(ma)]$ of the rigid molecules m.

c. Long wave approximation for separate atoms (LW)

By assuming that the individual atoms are the rigid bodies for the translational modes, we obtain $\mathbf{r}(m) = \mathbf{r}(a)$, $\mathbf{r}(ma) = 0$ (5)

Substitution of (5) in (4) and use of

$$F_B(\mathbf{S}) = \sum_{ma} f(\mathbf{S}ma)\exp[2\pi i\mathbf{S}\cdot\{\mathbf{r}(m) + \mathbf{r}(ma)\}]$$

gives the usual LW formula

$$I_1^{LW}(\mathbf{S}=\mathbf{H}-\mathbf{k}) = [k_BTS^2/k^2]|F_B(\mathbf{H})|^2\sum_{\sigma=1}^{3}[|\mathbf{U}^{LW}(\mathbf{S}\sigma\hat{\mathbf{k}})|/v(\sigma\hat{\mathbf{k}})]^2 \qquad (6)$$

d. Long wave approximation with velocity dispersion (LWD)

This approximation does not have a sound physical basis but it turns out to work good in practice (see below). For $I_1^{LWD}(\mathbf{S}=\mathbf{H}-\mathbf{k})$ formula (6) is applied apart from the fact that for the frequency instead of formula (3) the real (or calculated) $\omega(\sigma\mathbf{k})$ value of the phonon dispersion branches is used. In this version (6) reads

$$I_1^{LWD}(\underset{\sim}{S} = \underset{\sim}{H}-\underset{\sim}{k}) = [4\pi^2 k_B T S^2]\,|F_B(\underset{\sim}{H})|^2 \sum_{\sigma=1}^{3} [\,|\underset{\sim}{U}^{LW}(\underset{\sim}{S}\sigma\underset{\sim}{k})|/\omega(\sigma\underset{\sim}{k})]^2 \qquad (7)$$

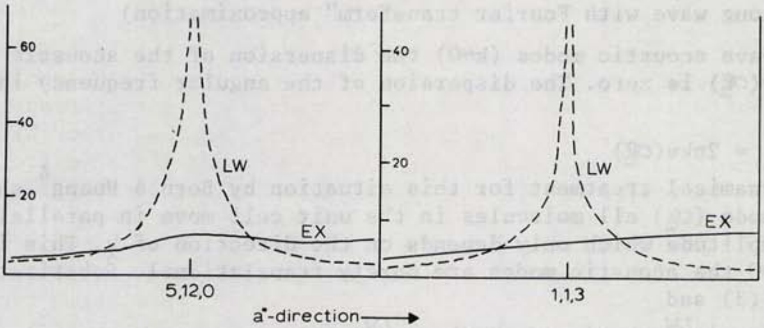


Fig.2. Second order TDS for 5,12,0 ($\sin\Theta/\lambda = 1.06\ \text{Å}^{-1}$) and 1,1,3 ($\sin\Theta/\lambda = 0.254\ \text{Å}^{-1}$) on a relative scale, in the region $\underset{\sim}{H} - \frac{1}{2}\underset{\sim}{a}^*$ to $\underset{\sim}{H} + \frac{1}{2}\underset{\sim}{a}^*$

Contribution of optical modes and second order TDS

From (2) and (3) we see that the peaking of the TDS profile at the reflection position will be caused mainly by the acoustic modes for which $\lim \omega^{-2}(\sigma\underset{\sim}{k}) = \infty$ for $\underset{\sim}{k}\rightarrow 0$ or for $\underset{\sim}{S}=\underset{\sim}{H}-\underset{\sim}{k} \rightarrow \underset{\sim}{H}$. For the optic modes where $\omega^{-2}(\sigma\underset{\sim}{k})$ is smaller and varies less drastically with $\underset{\sim}{k}$ and not always in the same direction, the peaking is expected to be small. This suggests that I_{TDS}(nett, H) for these modes may possibly be neglected. This was confirmed by EX calculations on naphthalene. For all naphthalene reflections at 100 K we calculated $|I_{TDS}(\text{nett, optic})| < 0.015\ I_{TDS}(\text{nett, acoustic})$.

Formulae analogous to (2), (4), (6) and (7) can be derived for second order TDS. If two modes with wave vectors $\underset{\sim}{k}'$ and $\underset{\sim}{k}''$ are considered, the second order TDS occurs at $\underset{\sim}{S} = \underset{\sim}{H}-\underset{\sim}{k}$ with $\underset{\sim}{k} = k' + k''$. In Fig. 2 the EX and LW second order TDS profiles are shown for two naphthalene reflections at 100 K. In contradistinction to the LW profile, the EX profile is so flat that the second order TDS contribution to I_{TDS}(nett) can be assumed to be negligibly small. This is in contrast to statements in the literature[5] where for high order reflections and low temperature (second order TDS is approximately proportional to S^4 and T^2) second order TDS corrections amounting to about 10% of the first order corrections have been calculated on the basis of the LW approximation.

From the discussion given above it was concluded that TDS from optic modes and second order TDS can be neglected for naphthalene at 100 K.

Brillouin zone summations

For the EX expression (2) the integration over $V\underset{\sim}{p}^*$ to obtain $I_{TDS}(V\underset{\sim}{p}^*)$ cannot be performed in the partly analytical way given by Helmholdt & Vos[1] for the LW method. The TDS program performs the integrations completely numerically with use of an isometric sampling grid. Without further precautions difficulties arise, however, for the acoustic modes with small $\underset{\sim}{k}$ values. From (2) and (3) we see that for these modes the summation contains terms of type

$$I_1^{ac}(\underset{\sim}{S}=\underset{\sim}{H}-\underset{\sim}{k})\Delta V^* = [D(\sigma\hat{\underset{\sim}{k}})/k^2]\,\Delta V^* \qquad (8)$$

where $D(\sigma\hat{\underset{\sim}{k}})$ is a function smoothly varying with $\hat{\underset{\sim}{k}}$. These terms become infinitely large for $k \to 0$, although their integrated value over v_o around $k=0$ is finite.
In the TDS program the contribution of v_o is estimated from the nearest 26 surrounding elements by use of the formula

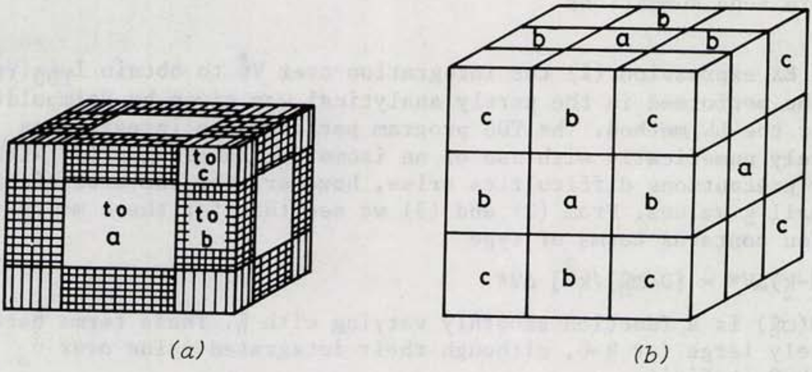
$$I(v_o) = 0.5\sum_a I(v_a) + 0.5\sum_b I(v_b) + 0.375\sum_c I(v_c) \qquad (9)$$

This formula can be derived by dividing v_o into pyramids (Fig. 3.) and assuming the functional form (8). The method was tested by numerical integration of the function k^{-2} for a spherical volume. If the number of volume elements is taken sufficiently large, surface effects can be neglected. The result of the integration is systematically too low, however, because due to the inverse quadratic form of the function especially the contribution of the volume elements with small $\underset{\sim}{k}$ is systematically too small. This effect is accounted for by multiplying $I(v_o)$ as obtained from (9) by an empirical factor ζ. Calculations with various choices of the sampling grid, showed ζ to be independent of these choices in good approximation. From the values 1.29 to 1.33 obtained for reasonable grids, the value 1.31 was adopted. The practical implementation of this method is surprisingly simple and merely consists of multiplying $I_1(\sigma, \underset{\sim}{S}=\underset{\sim}{H}-\underset{\sim}{k})$ for $\sigma = 1,2,3$ for the a and b-type volume elements by $(1 + 0.5\zeta)$ and for the c-type elements by $(1 + 0.375\zeta)$.

Choice of crystal structure and potential functions

The TDS calculations have been done on naphthalene at 100 K (space group $P2_1/a$, $n=2$). Naphthalene has the advantage that the molecules contain C and H atoms only, for which atoms various sets of potential functions are published in the literature (references are given in ref. 2). For the description of the interaction between two atoms i and j at distance r_{ij} we have chosen to write the potential functions in the Buckingham (6-exp) form.

$$V(r_{ij}) = Ar_{ij}^{-6} + B\exp(-Cr_{ij}) \qquad (10)$$



(a) (b)

Fig.3. (a) Division of the v_o element into pyramids. The figure shows the bases of the pyramids on the surface of v_o, each of them pointing to one of the surrounding volume elements of which there are three types, shown in (b)

where A, B and C are constants depending on the types of atoms considered. Interaction constants with good ability to reproduce the velocities of elastic waves for naphthalene and related compounds, were chosen. For C-H 1.09 Å was adopted. Prior to the lattice dynamical calculations the potential energy of the structure was minimized for the interactions (10) by giving small rotations to the molecules in the published naphthalene structure. This was performed with a minimization procedure analogous to that applied by Williams[6] in his program PCK 6.

Practical calculations

a. Frequencies and eigenvectors

Prior to the calculation of the TDS intensities and the $\underset{\sim}{T}$ and $\underset{\sim}{L}$ tensors, the program LATDYN was used to set up the dynamical matrix $\underset{\sim}{D}(\underset{\sim}{k})$ for 7488 $\underset{\sim}{k}$-values evenly spread in half the Brillouin zone. Each dynamical matrix has 6n eigenvalues $\omega^2(\sigma\underset{\sim}{k})$ and 6n eigenvectors $\underset{\sim}{U}(\sigma\underset{\sim}{k})$. Each eigenvector is composed of n six-dimensional vectors $\underset{\sim}{U}(\sigma\underset{\sim}{k}m)$ containing the (unscaled and mass-adjusted) translational and librational amplitudes of molecule m. In general an eigenvector has both translational and librational components which means that the translational and librational motions are coupled. Frequencies and eigenvectors for the second half of the B.Z. are obtained by the simple relations $\omega^2(\sigma\underset{\sim}{k}) = \omega^2(\sigma,-\underset{\sim}{k})$ and $\underset{\sim}{U}(\sigma\underset{\sim}{k}) = \underset{\sim}{U}^*(\sigma,-\underset{\sim}{k})$. The eigenvectors and eigenvalues were stored on a random access file.

b. $\underset{\sim}{T}$, $\underset{\sim}{L}$ and $\underset{\sim}{S}$ tensors

The naphthalene molecules lie at inversion centres. For such a case the translational and librational components of the eigenvectors have a phase difference of $\pi/2$, which makes $\underset{\sim}{S} = 0$. It should be noted that $\underset{\sim}{S} = 0$ does not mean that all modes are either purely translational or librational.
The tensors $\underset{\sim}{T}$ and $\underset{\sim}{L}$ were calculated with the formulae[7]

$$T_{ij}(m) = \mathrm{Re}[\sum_{\sigma \underset{\sim}{k}} < t_i(\sigma \underset{\sim}{k} m) t_j^*(\sigma \underset{\sim}{k} m) >]$$

$$L_{ij}(m) = \mathrm{Re}[\sum_{\sigma \underset{\sim}{k}} < l_i(\sigma \underset{\sim}{k} m) l_j^*(\sigma \underset{\sim}{k} m) >]$$

$t_i(\sigma \underset{\sim}{k} m)$ and $l_i(\sigma \underset{\sim}{k} m)$, $i = 1$-3 and $\sigma = 1$-12, are the translational and librational amplitudes for molecule m.

c. Calculation of I_{TDS}(nett)

In the range up to $\sin\Theta/\lambda = 1.2$ Å^{-1} one of each five reflections was taken giving a total of 1077 independent reflections. First integrated TDS intensities for the different approximations were calculated with use of the Vp^* volumes given by Helmholdt & Vos[1] for the ω-scan, apart from the fact that a larger counter opening of 6 x 6 mm was taken. From the results it was deduced, however, that artefacts occurred due to the course grid applied for the summation. It should be possible to avoid this problem by taking a finer grid. This was hampered, however, by an impractically large increase in computing time, especially for the EX method. We have therefore chosen to keep, for the present comparative study, the volume Vp^* the same for all reflections. Vp^* was chosen such that the maximum $\underset{\sim}{k}$ values encountered are the same as in practical experiments.
Computing times on a Cyber 74-16 for the calculation of I_{TDS}(nett) for one naphthalene reflection are for the constant Vp^* case 11 for EX, 10 for LWF and 1 c.p.u. seconds for LWD and LW.

d. Least squares refinements

To determine the changes in structural parameters due to the presence of TDS least squares refinements were done. Values $|F_o(\underset{\sim}{H})| = [|F_c(\underset{\sim}{H})|^2 + I_{TDS}(\text{nett})]^{\frac{1}{2}}$ were considered as observed structure factor amplitudes ($F_c(\underset{\sim}{H})$ is the structure factor corresponding with the naphthalene model). H atoms were constrained to C. The function $Q = \Sigma[|F_o(\underset{\sim}{H})| - |F_c(\underset{\sim}{H})|/K]^2$ was minimized with the full matrix least squares method. For all four sets of TDS calculations (EX, LW, LWF and LWD) full angle (f.a.), low order (l.o.) and high order (h.o.) refinements were performed, where l.o. was chosen as $\sin\Theta/\lambda < 0.9$ and h.o. as $\sin\Theta/\lambda > 1.0$ Å^{-1}. This results in about 400 reflections in each range. The parameter changes are given by

$$\Delta_i = p_i(\text{LS}) - p_i\ (\text{model})$$

Results

a. TDS profiles

In Fig. 4 the calculated first order TDS intensity profiles are plotted for the $\underset{\sim}{b}^*$ direction. For the reflections with $k = 2n$ the profiles clearly show that for all approximations $\lim I_1(\underset{\sim}{S}) = \infty$ for $\underset{\sim}{k} \rightarrow 0$ (or $\underset{\sim}{S} \rightarrow \underset{\sim}{H}$), as expected. Around the systematically absent reflections oko with $k = 2n + 1$ for the three long-wave cases zero TDS is calculated. For LW and LWD this is evident since $F_B(\underset{\sim}{H}) = 0$ for these reflections. For LWF it follows from (4), since the y-components $r_y(m)$ of the molecular centres differ by 0.5 b, whereas the components $r_y(ma)$ are the same for corresponding atoms in the two molecules. That in the EX case the TDS is non-zero for systematically extinct reflections is due to the librational character of the two antisymmetric modes. For these modes the y-positions of corresponding atoms move in opposite directions which disturbs the twofold screw symmetry.

b. I_{TDS}(nett) contributions and parameter changes

In Table 1 the values $\alpha(\underset{\sim}{H}) = I_{TDS}(\text{nett},\underset{\sim}{H})/|F_B(\underset{\sim}{H})|^2$ are given for a small series of reflections. From the table (and even better from the complete output) we see: (1) The TDS is strongly anisotropic, but increases on average with S. For high order reflections $I_{TDS}(\text{nett},\underset{\sim}{H})$ may become, even at 100 K, as large as $|F_B(\underset{\sim}{H})|^2$. (2) If the EX

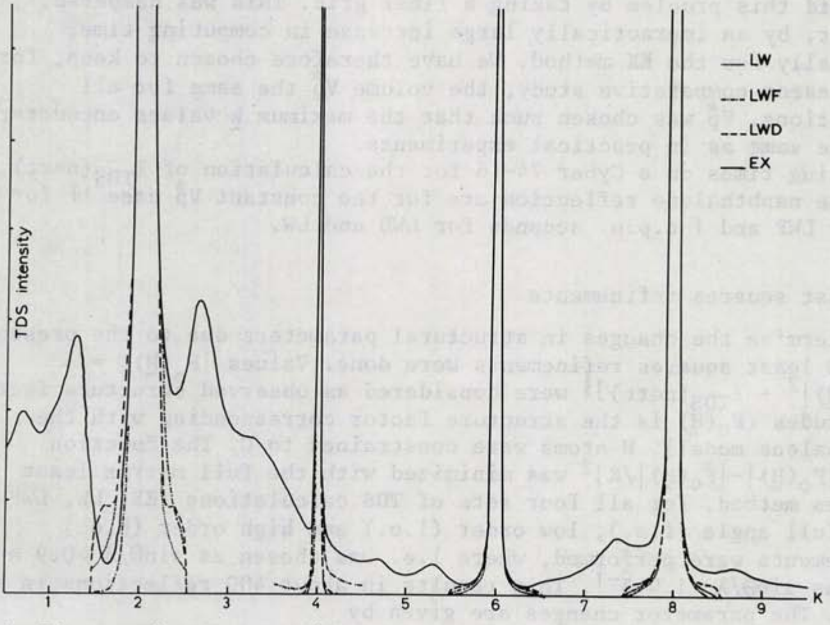


Fig. 4. First order TDS profiles along the b direction in naphthalene at 100 K. The numbers give the positions of the oko reflections*

eatment is assumed to correspond with reality the LW method erestimates the TDS corrections, while these corrections are found in od approximation by the LWD method (the LWF results are not discussed rther).

Table 2 the parameter changes are summarized. We see: (1) The rors in the coordinates, due to neglect of TDS corrections, are rprisingly small. This means that the coordinates are hardly affected the occurrence of the phases $\psi(\underset{\sim}{S}\sigma\underset{\sim}{k}ma)$ in the expression (2) for $^{X}(\underset{\sim}{S})$. This is due to the fact that in the integration over V_p^* the ctors $\underset{\sim}{k}$ and $-\underset{\sim}{k}$ are equally involved with $\psi(\underset{\sim}{S}\sigma\underset{\sim}{k}ma) = -\psi(\underset{\sim}{S}\sigma,-\underset{\sim}{k},ma)$.

) The changes in the thermal parameters and in K are considerable. glect of TDS corrections leads to underestimation of the thermal rameters. For accurate studies of the dynamical behaviour of the lecules, TDS corrections are thus necessary.

) If corrections for TDS are not made, high order refinement rameters are, for the naphthalene case, not transferable to the low der region. This should be noted by crystallographers interested in curate studies of electron density distributions, who want to find rameters not affected by bonding from high order X-ray data.

) $\Delta_i^{LWD} = \Delta_i^{EX}$ (follows from more extended table). This implies at parameter errors can be avoided if TDS corrections according to e LWD method are made.

ıble 1. Values $\alpha(\underset{\sim}{H}) = I_{TDS}(\mathrm{nett},\underset{\sim}{H})/|F_B(\underset{\sim}{H})|^2$ in per cents

k	l	$\sin\Theta/\lambda$	$\lvert F_B\rvert^2$	$\alpha^{LW}(\underset{\sim}{H})$	$\alpha^{LWF}(\underset{\sim}{H})$	$\alpha^{LWD}(\underset{\sim}{H})$	$\alpha^{EX}(\underset{\sim}{H})$
5	3	.52	406	12.6	13.2	11.1	10.7
5	$\bar{9}$	.74	4957	34.1	34.5	27.6	27.4
6	$\overline{14}$	1.00	110	24.3	23.6	23.8	23.8
4	4	1.02	186	94.9	95.8	69.8	69.0
7	6	1.16	34	100.1	103.2	76.5	73.2

able 2. Parameter changes obtained for the different TDS calculations

	$<\lvert\Delta r\rvert>$ in 10^{-4} Å			$-<\Delta U_{ii}>$ in 10^{-2} Å^2			ΔK(%)		
	h.o.	l.o.	f.a.	h.o.	l.o.	f.a.	h.o.	l.o.	f.a.
W	3	3	3	.21	.31	.29	6.1	.7	1.0
LWD	1	2	2	.19	.26	.24	4.6	.4	.7
EX	4	2	2	.19	.26	.24	4.6	.4	.7

Conclusion

As the computation of good TDS corrections is time consuming and moreover requires a correct description of the intermolecular interactions, TDS should be avoided as much as possible by doing the measurements at low, preferably He, temperatures. If TDS is reduced in that way, it may be hoped that theoretical knowledge of interaction models (or experimental knowledge of, for instance, wave velocities) is adequate to eliminate the remaining small TDS errors by making corrections in LWD or LW approximation. As the LWD approximation is much less time consuming than the EX method, for LWD a sufficiently dense integration grid can be applied to make summation over real volumes V_p^* feasible. For LW the well known partial analytical integrat method can be used.

References

1. Helmholdt, R.B. & Vos, A. (1977). Acta Cryst. **A33**, 38-45.
2. Kroon, P.A. & Vos, A. (1978a,b,c). Acta Cryst., to be published.
3. Cochran, W. (1963). In "Reports on Progress in Physics" **26**, 1-45. Physical Society, London.
4. Born, M. & Huang, K. (1968). "Dynamical Theory of Crystal Lattices". Clarendon Press, Oxford.
5. Stevens, E.D. (1974). Acta Cryst. A30, 184-189.
6. Williams, D.E. (1972). Acta Cryst. **A28**, 629-635.
7. Scheringer, C. (1973). Acta Cryst. **A29**, 554-570.