

A Concept for Crystal Structure Determination without FOURIER Inversion: Some Steps towards Application*

Karl F. Fischer,^{a,**} Armin Kirfel,^b and Helmuth Zimmermann^c

^aTechnical Physics, University of Saarbruecken, D-66123 Saarbruecken, Germany

^bCrystallography/Mineralogy, University of Bonn, Poppelsdorfer Schloss, D-53115 Bonn, Germany

^cCrystallography/Condensed Matter Physics, University of Erlangen, Staudtstr. 3, D-91058 Erlangen, Germany

RECEIVED NOVEMBER 23, 2007; REVISED JANUARY 23, 2008; ACCEPTED JANUARY 29, 2008

Determination of a crystal structure without Fourier calculation of the scattering density (thus also avoiding the phase problem) is achieved in a fractional coordinate parameter space of dimension $3m$ where m is the number of independent atoms, reduced to equal point scatterers at rest. For demonstration of the basic ideas, two-dimensional parameter spaces (representing, *e. g.*, one-dimensional two-atom structures) are used. "Central reciprocal lattice row" reflections allow for solving one-dimensional projections of the structure, each requiring less reflections and simultaneously providing better resolution than does a corresponding Fourier summation. The projection solution can be obtained either from the common intersection of the hyper-faces in the m -dimensional parameter space defined by the chosen scattering amplitudes or by exploring the permitted "solution region(s)" that follow from the mere ranking of these amplitudes. All possible solutions satisfying the data are found, including "false minima". The reconstruction of a hypothetical three-dimensional 11 atom structure from the solutions of one-dimensional projections is illustrated in an example based on "theoretical", *i. e.* error-free data. Since most of the theoretical background is laid down in two former, refereed publications, emphasis is put on different options to cope with the computing demands in practical applications. Advantages and shortcomings of the concept are discussed.

Keywords
crystal structure analysis
unique solution
homometric solutions
phase problem eliminated
intensity inequalities

INTRODUCTION AND GENERAL CONCEPT

In the past decades, the determination of crystal structures has matured to an almost routine task driven by the successes of "Direct Methods". After allocating approximate phase values to a sufficient number of structure amplitudes, a scattering density approximation of the actual structure is obtained by Fourier summation over those structure factors and a structure model follows from the

three-dimensional maxima of the generated scattering density distribution, since these maxima are interpreted as atomic positions. Thus, the ever dominant "phase problem" is overcome by the derivation of appropriate phases whose mutual differences are structure invariants hidden in the total intensity distribution of the selected observed data. Their number is generally much larger than that of the unknown structure parameters, namely the atomic fractional coordinates x , y , z .

* Dedicated to Professor Dr. Klaus Knorr (Technical Physics, Saarbruecken) on the occasion of his 65th birthday.

** Author to whom correspondence should be addressed. (E-mail: karl.fischer@mx.uni-saarland.de)

During the same span of time, only a few "new developments" were reported most of which, however, still use the Fourier technique or are associated with it, see *e. g.* the recent "charge flipping" method by Oszlány and Sütő.¹ More strictly algebraic, geometrical or analytical approaches are those developed by Rothbauer,^{2–6} Navaza and Silva,⁷ Cervellino and Ciccariello^{8–10} and¹¹ following.¹² All those concepts were briefly addressed in two papers.^{13,14} Particularly, Navaza and Navaza¹⁵ stressed the role of atomicity for any "direct space" approach and Ciccariello and Cervellino¹⁶ proved that, given observations from a point structure, only a limited number of low order reflection intensities suffices for obtaining all other ones.

To date, in most cases, a refinable structure model results from a kind of "black box" procedure which does not require operator's intervention unless the method fails or must be sharpened or applied in some sort of non-routine way. If, for example, the data set is limited or the spatial resolution of the scattering density distribution is insufficient or if the structure itself possesses some tricky "pseudo-symmetry" (as frequently observed in "new materials" with interesting physical properties), the routine structure determination may be hampered, and even after a refinement, unsatisfactory model ambiguities may remain. Also, cases of exact "homometry" (identical sets of inter-atomic distances for two or more different structures, of which the wrong one(s) can, however, usually be identified by chemical considerations) may be approximated due to experimental errors. Then, a set of two or more "quasi-homometric" structures can be obtained, with their corresponding diffraction data merely differing within the uncertainties. In such cases, usually only one solution is found and reported without even knowing that other ones or another exist(s), let alone how these alternative structures look like. In other words, one may well end up in a so-called "false minimum".

The aim of this paper is to present a concept for structure determination which circumvents reciprocal space in order to derive a suitable model (or at least valuable structure information) from generally less than the usual reflections, and this with a higher direct space resolution as compared to conventional structure determination methods. While some principles of the underlying ideas have been published before (see Refs. 13 and 14 as well as some conference abstracts), we aim here at showing that and how a structure determination can be conducted without Fourier inversion. Even though some parts of the process are presented only in a preliminary state and in need of further more detailed exploration, particularly with respect to data error allowances, we believe that a "first lap completed" report may interest chemists and bio-crystallographers as a potentially helpful tool in solving problem structures. This implies that we neither wish nor can offer a "competition" of any kind with existing program

systems. We merely illuminate a point of view that is different from, but complementary to the established one(s) and as such to be considered.

THEORETICAL BASIS

The basic concept is rather simple, if not naive and can be outlined in four statements:

- 1) As described in more detail in Ref. 13, we use a structure reduced to equal point atoms at rest so that their scattering power may be $f = 1$ throughout reciprocal space. For this model, the so-called "geometrical structure amplitudes" of a centric structure are given by

$$g(hkl) = \left| \sum_j \cos 2\pi(hx_j + ky_j + lz_j) \right|$$

with $j = 1, \dots, m$ for m independent atoms. Correspondingly, measured structure amplitudes must be reduced to "experimental $g(hkl)$ " which inevitably enhances the experimental errors by some errors due to approximations. Details see below ("Data reduction and its errors").

- 2) Each of the $3m$ unknown positional parameters of the point structure plots on one axis of a $3m$ -dimensional ortho-normal ("hyper-cubic") "parameter space", thus defining the $3m$ components $x_j, y_j, z_j [0; 1]$ of a "structure vector" \mathbf{X} that represents the atomic arrangement within the asymmetric part of the unit cell. (For simplicity, we use space group $P\bar{1}$ and $P1$ throughout this paper.) Due to the equal scatterers, the parameter space \mathbf{P}^{3m} possesses permutation symmetry of order $m!$ with the consequence that only one asymmetric part $\mathbf{A}^{3m} (1/m!)$ of the $3m$ -dimensional volume of the hypercube-shaped parameter space) must be considered for finding \mathbf{X} . (A possible asymmetric part can be simply defined by a given sequence of these atoms, see below: "One-dimensional projections")
- 3) Mapping a geometrical structure amplitude into the parameter space yields a $(3m-1)$ -dimensional hyperface called "iso-surface" and denoted as $G[h; g]$. This imposes additional symmetry to that of the parameter space. – To most of us, the inability of imaging multi-dimensional spaces leaves us with computations. A two-dimensional parameter space (representing a one-dimensional structure of two atoms, here with inversion symmetry) can, however, be nicely illustrated and thus further the understanding of the basic ideas that apply also to higher dimensions. Figure 1(a) and 1(b) show that for reflection orders $h > 1$, any set of iso-surfaces for $g(h)$ is periodic, repeating the pattern of $g(1)$ or its mirror image on linear scale $1/h$. (For more details see Figure 3(a) and 3(c) in Ref. 13.) – Compare also the "graphical abstract" of this paper representing the parameter space \mathbf{P}^3 for

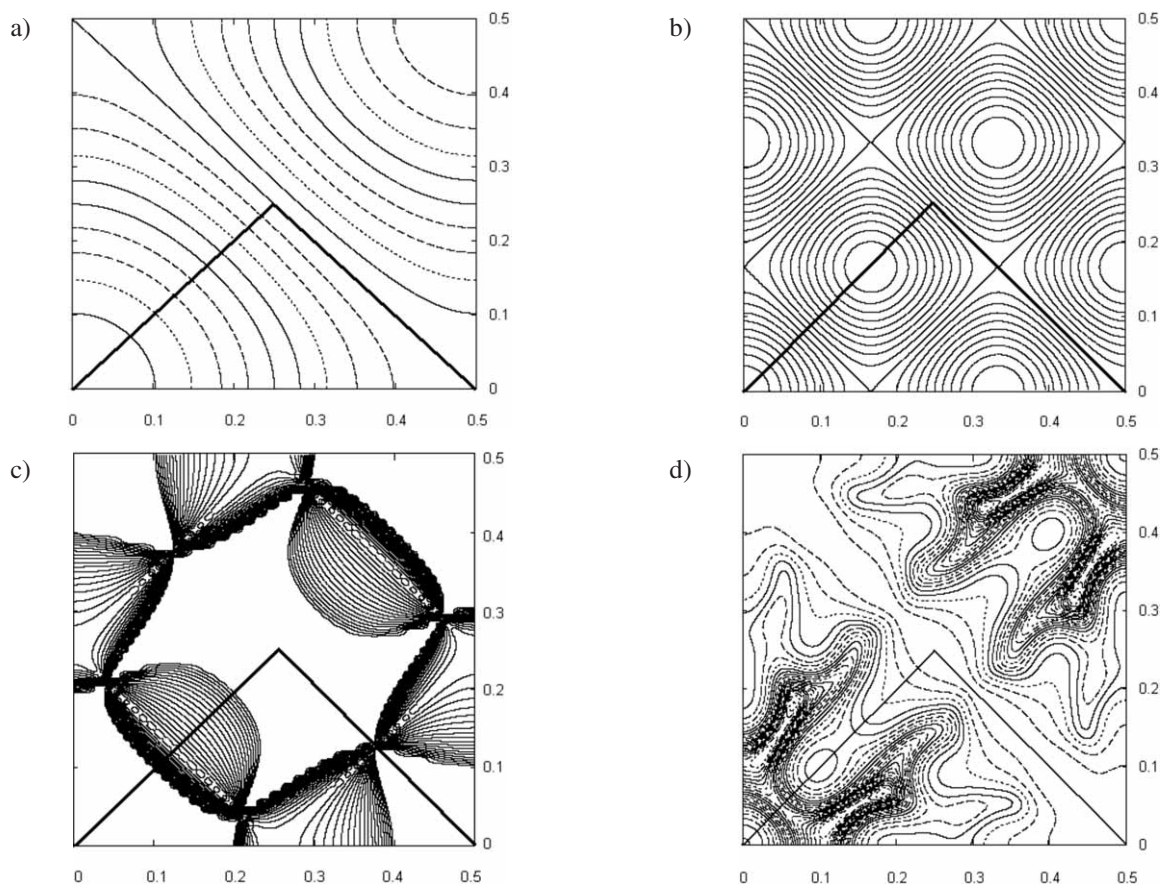


Figure 1. Parameter space $[0; \frac{1}{2}]$ for 1-dimensional, centrosymmetric 2-atom structures: each vector (x_1, x_2) represents a structure with these coordinates x_1, x_2 . Origin (depicting $x_1 = x_2 = 0$) at lower left corner; x_1 along horizontal coordinate axis, x_2 vertical coordinates. The asymmetric part \mathbf{A}^2 is one half of the square shown: triangle $(0, 0), (\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, 0)$. Half of the asymmetric part \mathbf{A}^2 is indicated by the smaller triangle (two heavy lines and horizontal coordinate axis) and suffices for description of all possible structures (the other half being symmetrically equivalent due to the inversion centre at $(x_1 = x_2 = \frac{1}{2})$ which may eventually be used as alternative "origin").

1(a) Iso-surfaces $G[1; g]$, contours at 0.2. $G[1; 0]$ is depicted as diagonal from upper left corner ($x_1 = 0, x_2 = \frac{1}{2}$) to lower right corner ($x_1 = \frac{1}{2}, x_2 = 0$); $G[1; 2]$ coincides with the lower left corner ($x_1 = x_2 = 0$) and upper right corner ($x_1 = x_2 = \frac{1}{2}$).

1(b) Iso-surfaces $G[3; g]$: note the repetition of $G[1; g]$ and its mirror images after reduction of scale by factor $h = 3$.

1(c) Iso-surfaces $Q[2; 3]$ for an observed inequality $0 < q(2, 3) \leq 1$, i. e. $g(3) \geq g(2)$. The iso-surfaces for $q = 1$ are the borders of the white areas that represent all "forbidden" structures with $q > 1$. Contours at 0.05.

1(d) Iso-surfaces $E[1; 6]$, $e = 0$ along diagonal as for $G[1; 0]$ in Figure 1(a), contours at ≈ 0.2 .

centrosymmetric one-dimensional structures of three atoms with the iso-surface for $g(1) = 0$.

- 4) These iso-surfaces are the essential "tools" for finding \mathbf{X} , because, provided error free $g(h)$, the point of intersection of (at least) $3m$ independent iso-surfaces is \mathbf{X} . The iso-surfaces as well as those generated from combinations of experimental data – see below "Series of g, q or $e \dots$ " – can also be used for excluding from \mathbf{A}^{3m} those parts that cannot contain the solution \mathbf{X} , e. g. by statements like "a calculated $g(hkl)$ must be larger (or smaller) than a given threshold value derived from the experimental data".

Looking geometrically simple in the above case, in general, the common intersection of a number of iso-surfaces stands for solving (at least) $3m$ equations for $3m$

unknown scatterer coordinates which – *in praxi* – does not yet appear feasible. Instead, we have tried various approximations or "trial and error" tests. One of them, a "linear approximation" technique, has been described in more detail in Ref 13., including some examples. Here, we treat mainly those trial-and-error tests that appear the most successful so far.

SOME PRINCIPAL ADVANTAGES OF THE PARAMETER SPACE CONCEPT

Before entering practical considerations, some advantages of the above concept may be noted:

- 1) There is no need for phases of structure factors, i. e. the phase problem plays no role. (About origin definition, see below: "One-dimensional projections".)

2) Locating \mathbf{X} yields the complete solution, *i. e.* the coordinates of all scatterers are found simultaneously.

3) For comparable resolution, less reflections are needed than for Fourier summation. Series termination effects are avoided and in principle, even infinite resolution would be obtained if the above-mentioned $3m$ $g(hkl)$ were error free.

4) The solution is either unique or if not, then all possible atomic arrangements satisfying the experimental data, *i. e.* at the same time all so-called "quasi-homometric" structures are also determined.

Looking at both the problems and the potential for determining crystal structures from a parameter space point-of-view may thus offer some new insights and shed light on old and sometimes trivial knowledge that is buried in the actual, most effective structure solution programs.

SOME PRACTICAL PROBLEMS

There are two main questions that arise immediately:

- (i) Why should we prefer "trial and error" procedures over approximation techniques, and
- (ii) how can we obtain the basically needed data $g(hkl)$ on absolute scale given a relatively limited set of measured intensities?

For reasons discussed below, we can continue this discussion to a projection of the point structure onto one of the three normalized lattice vectors, *e. g.* a/a , implying that a structure projection has to be solved from, say, the first n $g(h00)$ amplitudes (abbreviated as $g(h)$) and their corresponding iso-surfaces $G[h; g]$ (see Figure 1(a), (b)).

- 1) Then, there are two options for absolute scaling of the point structure amplitudes g derived from un-scaled experimental data:
 - (i) The quotients $q(h_1, h_2) = g(h_1)/g(h_2)$ of two geometrical amplitudes are always on absolute scale and their iso-surfaces $Q[h_1, h_2]$ can be illustrated as shown in Figure 1(c) for general centrosymmetric cases ($m = 2$). More details are found in Ref. 14.

- (ii) For a batch of (un-scaled) $g(h)$, one can define "quasi-normalized amplitudes" $e(h, n)$, for example as:

$$e(h, n) = g(h) / \left\{ (1/n) \cdot \left[\sum_i g^2(h_i) \right]^{1/2} \right\}, h, h_i = 1, \dots, n$$

where the $g(h)$ are from a "central reciprocal lattice row". These e values are also on absolute scale and likewise their associated iso-surfaces $E[h, n]$, see Figure 1(d). (See also Ref. 17 for a similar scaling. More details will be given in a separate paper.)

Both q and e data were mostly used in test calculations, simply because they are on absolute scale and thus preserving the relative intensity distribution, therefore re-

flecting – though not unambiguously – the structure information.

- 2) Comparing the three different types of iso-surfaces (Figure 1(a), (c), (d)) one may conclude from the different shapes of their corresponding iso-surfaces that $G[h; g]$ can be approximated (*e. g.* by linear or spherical hyper-faces) within wider ranges of the parameter space than $Q[h_1, h_2]$ and $E[h, n]$, which exhibit less smooth curvatures within \mathbf{A} . Therefore, using Q - and/or E -isosurfaces, a combination of systematic and random assessment appears the best choice which directly leads to a third question.
- 3) How many systematic tries are necessary? – In principle, a huge number N_p of "test structure vectors" \mathbf{X}_t has to be tested by comparing calculated q or e values with "experimental" ones: Using a grid of p equidistant points along each principal direction of the parameter space, the number of hyper-cube-shaped test volumes N_p in the asymmetric part \mathbf{A} of \mathbf{P} is roughly $p^{3m}/(m!)$ for $p > m$ (for a more accurate, but here unnecessary and more complicated estimate, the borders of \mathbf{A} must be accounted for properly). For practical purposes, N_p should be small enough to be handled with a standard PC when a selection of different measures described in the next section is applied.

(About random assessment see below: sections "Reducing N_p by grid coarsening" and "Practical example ...".)

A centrosymmetric, one-dimensional 2-atom example ($x_1 = 0.06$; $x_2 = 0.04$) demanding substantial spatial resolution has been successfully treated applying different techniques of the parameter space concept: (i) linear approximation based on $g(h)$, see Ref. 13, example (2) and Figure 6(a), (b); (ii) graphic solution as well as (iii) grid computing based on $q(h, k)$, see Ref. 14, Figure 6 and Table II (the latter including various data errors). In all cases, between 3 and 9 reflection amplitudes were employed, whereas a Fourier summation with 10 error-free reflection amplitudes and correct signs provided but hardly useful results (see Ref. 13, Figure 6(c)).

SELECTION PROCEDURES

One-dimensional Projections

A first step for a substantial reduction of N_p is the use of data from "central reciprocal lattice rows" from which, of course, only one-dimensional projections of the structure can be obtained. This measure is connected with experimental advantages, but also with two shortcomings:

- (i) strict "atomicity" may be lost in cases where projected scatterers more or less coincide, and

- (ii) homometric structures are much more probable in one-dimensional projections than in three dimensional space so that also quasi-homometric cases caused by the unavoidable errors in the observations will more frequently occur.

In a one-dimensional projection, *e. g.* onto the x direction, the equal point scatterers can be sorted in descending order of their coordinates

$$x_1 \geq x_2 \geq \dots \geq x_m \geq 0$$

thus defining an asymmetric part \mathbf{A}^m of the parameter space (details see Ref. 13). For a centrosymmetric structure projection, $x_1 \leq 0.5$ is sufficient. Also, only one half of \mathbf{A}^m needs to be searched for the solution vector \mathbf{X} corresponding to one of the two possible origins (see Fig. 1). How the solution of a structure projection is actually obtained, is described below (section "Solution of each projection").

Re-constructing the three-dimensional direct space arrangement of the atoms from projections is a task similar to, *e. g.*, generating an X-ray tomographic absorption density distribution from one-dimensional projection data. The task is, however, much less complicated because the structure model to be developed contains only a known number of discrete point scatterers. Hence, one can denote it as "point tomography"¹¹ (see also Refs. 18 and 19) which, so far, we have successfully applied to centrosymmetric structures (for acentric structures, the problem of origin conservation has still to be solved *in praxi*). How this can be worked out is demonstrated in the "Practical example ..." below. More details will appear in a separate paper.

Experimental Data of One-dimensional Projections

A single spot of a Laue-type diffraction diagram (recorded with white primary radiation and both crystal and detector in fixed positions) may contain several "harmonic" reflection orders, depending on the glancing angle 2θ , the range of the primary λ -spectrum and the indices h, k, l of the first order reflection. For large d -spacings, *i. e.* short reciprocal lattice vectors, the intensities of reflections nh, nk, nl up to high orders contribute to this single spot. The separation and quantitative intensity measurement of the harmonic reflections is achieved by using a properly aligned energy-dispersive detector whose response is transformed into a sequence of n intensity peak profiles by multi-channel analyzing (MCA). Thus, at the same time, a sequence of reflection amplitudes from a "central reciprocal lattice row" can be obtained from one crystal- and diffractometer-setting,^{20,21} provided the primary intensity spectrum and all relevant correction factors are known. A kappa-type single-crystal diffractometer equipped with two independently operating detectors²² permits simultaneous measurements on two central reciprocal lattice rows, again in a fixed-crystal, fixed-detector mode (*e. g.*

at beam-line F1 at HASYLAB/DESY, now operated by Prof. Bismayer, Hamburg).

If in addition, a crystal contains an atomic species called "a" exhibiting significant anomalous scattering (*i. e.* the x-ray absorption edge wavelength λ_a is contained in the primary radiation spectrum), some of the harmonic reflections can be shifted, one after the other, across the respective λ_a by a simple ω - 2θ -scan. The intensity changes at λ_a caused by changes in f_a' and/or f_a'' of the anomalous scatterers can be used to determine the "partial structure amplitudes" of these a-atoms, *e. g.*, by the procedure of Karle.²³ Simultaneously, the harmonics with wavelengths far from λ_a help monitoring and thus in reducing the relatively large relative statistical uncertainties as a consequence of the generally small "anomalous scattering signal". (It should be noted in this context, that both f_a' and f_a'' are point scatterer based quantities and as such well complying with the underlying point structure model. – A partial structure means, of course, a reduced m and consequently its solution is a much easier task.

Series of g, q or e Sorted for Magnitude and New Inequalities for Approximating a Structure

Sorting (the relative) $g(h_i), i = 1, \dots, n$ into a sequence of decreasing amplitudes (called "data ranking") provides

$$g_{max} = g(h_1) \geq g(h_2) \geq g(h_3) \geq \dots \geq g(h_n).$$

The sequence of these "ranked $g(h)$ " is identical with that of the equivalent series of

$$e(h_1, n) \geq e(h_2, n) \geq e(h_3, n) \dots$$

and of 'neighbour' q 's:

$$1 \geq q(h_2, h_1), 1 \geq q(h_3, h_2), \dots (n-1 \text{ different ones}).$$

In Ref. 14 (Figure 4(a)) was demonstrated, how the "solution region" in the parameter space can be more and more confined by a sequence of such new inequalities $q(h_i, h_{i-1}) \leq 1$ by using iso-surfaces of $q(h_i, h_{i-1}) = 1$ as borders between 'permitted = possible' and 'forbidden = impossible' regions of the parameter space. Since for all structures, the "neighbour" q 's of the above sequence are as close as possible to 1, these borders enclose the true solution point \mathbf{X} as close as possible, tighter than any other q 's calculated from the same batch of data. From this we may infer that – apart from the quantitative use of the data – already the index sequence of h_i defines (a) region(s) of approximate solution(s) depending on how many reflections are considered. It is obvious that any additional reflection never augments the volume of the m -dimensional solution region already defined, it can only reduce it. (For high resolution by a kind of "nonius effect" see Figure 4(b) in Ref. 14) It should be noted that this technique makes merely a qualitative use of quantitative data and does not require absolute scaling as is ne-

cessary with $g(h)$. The quality of the data is unimportant as long as their uncertainties do not affect their ranking – see Eqs. (6) and (7) in Ref. 14 where also the potential and limitation(s) of these q -series are discussed.

Depending on the case, more or also even less than m inequalities are needed or suffice, respectively, for obtaining one (or more) solution region(s) which, after conventional least-squares refinement, provide the solution and discard possible alternatives.

The practical application employing a grid technique requires either

- (i) computing for each of the N_p grid test vectors X_i (in \mathbf{A}^m) the associated $q(h_i, h_k)$ for $n - 1$ inequalities, comparing it with the observation in question and discarding the grid point ($= X_i$) or keeping it for the next inequality check or
- (ii) using a beforehand computed data pool of test vectors, specific to m , the desired grid density and the observed q -sequence.

Evidently, the second alternative is much faster as it substitutes the systematic grid search by a "look-up" procedure. Such data pools, however, contain $N_p \cdot N_q$ "flagged" entries (where N_q denotes the number of possible sequence permutations, see below) whose large number presently appears to preclude practical application. As a consequence, options for reducing $N_p \cdot N_q$ were explored and are discussed in the next two sections.

"Salami Tactics" for the Above Series for Reducing N_q (theoretical example)

Assume $n = 13$ reflections have been measured and ranked as above. Then, in principle, $N_q = 13! \approx 6 \cdot 10^9$ different ranking sequences are possible. (There will be somewhat less, because not all permutations are permitted due to the general restriction that $|\cos 2\pi hx_j| \leq 1$ and to known inequalities like those given by Harker and Kasper.²⁴) After cutting the sequence into 4 parts, *e. g.* $h = 1 \dots 4$; $h = 4 \dots 7$ and so on, each of the 4 batches allows at most $4! = 24$ permutations, a number that can be handled comfortably.

The data pool could thus be reduced by a factor of about $6 \cdot 10^9 / 96 \approx 6 \cdot 10^7$. Then, for each batch, one has only to select the sole permutation (of 24) that complies with the observed ranking sequence so that the region common to the once found 4 permitted solution regions defines approximate solution(s) with respect to all x_j of the structure projection.

Reducing N_p by Grid Coarsening

Choosing a smaller p (the basis of the exponential expression for N_p) provides an effective reduction of N_p and is, of course, primarily justified for low-order reflections. Combining this measure with the (qualitative) 'index sequence after ranking' provides fast separation of

'permitted' solution regions in \mathbf{A}^m from 'forbidden' ones. In a next step, only these approximations X_p to possible solutions must be subjected to a more detailed check – *e. g.* by applying a finer grid defined in a m -dimensional box with centre X_p or by randomly generating in each of the boxes a given number of X vectors. This latter approach is fast, effective and has been successfully applied in a great number of tests.

The Use of Extreme Amplitude Values

The highest possible diffraction contrast is found between $g(h_1) = m$ and $g(h_2) = 0$ (or corresponding values for q or e).

In the parameter space representation, the iso-surface of a $g(h_1) = m$ plots as one or more discrete point(s) in \mathbf{A}^m (the number increases with increasing h_1 and their positions are *a priori* known). Thus, a (scaled) $g(h)$ close to m restricts the solution regions to small 'hyper-spheres' about these points. *I. e.*, for $g(h_1) \geq m - \Delta \approx m$ (*i. e.* $\Delta \ll m$), possible X vectors are limited to be within a radius defined by Δ and h_1 .

Similarly, $g(h_2) = 0$ is a singular branch of the $g(h_2)$ -iso-surfaces providing valuable structure information (in acentric structures it is an $(m-2)$ -dimensional iso-surface). That an accidentally extinct reflection is more valuable than any other has already been discussed as early as 1927 by one of our ancestors, H. Ott,²⁵ who showed that and how a structure can even be determined from just a few zero observations. Details about using such highly contrasting observations, including error dependencies, will be dealt with in a more theoretically oriented paper.²⁶

Data Reduction and Error Sources

In order to obtain $g(hkl)$ values for a small batch of experimental data, *i. e.* observed scattering amplitudes $|\sum f_j \cdot T_j \cdot \cos 2\pi(hx_j + ky_j + lz_j)|$ where $T_j = -B_j \cdot (\sin^2\theta/\lambda^2)$ one cannot resort to Wilson statistics. Instead, one has to

- (i) estimate isotropic thermal displacement parameters B_j (or an overall one, B) and
- (ii) correct each observation by the respective average over all $f_j \cdot T_j$ for obtaining estimated $g(hkl)_{\text{obs}}$ approximating

$$g(hkl) = |\sum \cos 2\pi(hx_j + ky_j + lz_j)|.$$

B_j may be reasonably well guessed from chemical experience. Nevertheless, deviations from the "true" B_j introduce systematic errors increasing with θ . The second approximation must be expected to cause errors due to assuming that all scatterers are equal and deviations from this assumption will contribute according to the geometrical terms $\cos 2\pi(hx_j + ky_j + lz_j)$. These errors are necessarily distributed in an unknown way over all reflections so that also the in principle absolutely scaled q -

or e -values can be affected. Also, these errors increase, of course, with increasing f_j contrasts, but they may be tolerable, *e. g.* for X-ray scattering data from organic compounds neglecting hydrogen atoms. The general effect of unequal scatterers in the structure is briefly addressed in Ref. 13 (see Figure 10), namely breaking the permutation symmetry of the parameter space and distortion of the iso-surfaces.

Step (ii) involves smaller errors for neutron diffraction data than for X-ray measurements, because in the X-ray case, the f_j exhibit different dependencies on $(\sin \theta)/\lambda$ whereas in the neutron case merely the differences in the nuclear scattering amplitudes add to the errors caused by the Debye-Waller temperature factors. (For different signs of nuclear neutron scattering amplitudes see Ref. 13, Figure 11.)

PRACTICAL EXAMPLE: A SMALL THREE-DIMENSIONAL CENTROSYMMETRIC TEST STRUCTURE ($m = 11$)

One of us (A. K.) has made numerous feasibility and test computations, mainly on one-dimensional structures or one-dimensional projections of two- and three-dimensional structures. These calculations, presently still performed on the basis of error free data stemming from the point scatterer model, aimed to combine some of the tools described above and at gradually increasing the number of atoms. The result of one of them is presented in the following.

Test Point Structure and Its "Experimental" Basis

An 11-atom fragment of a molecule with 20 independent atoms (see Figure 2) was taken from a known crystal structure with symmetry $P\bar{1}$. The coordinates are indicated as "true" in Table I. The atoms were considered equal, and the geometrical structure amplitudes $g(hkl)$ of this model structure were calculated up to reflection order 7 for the i) three principal axes (*e. g.* [100]), ii) six face diagonals (*e. g.* [110]), and iii) four space diagonals (*e. g.* [111]). These data were quasi-normalized to $e(h, 7)$ for each of the 13 structure projections. Ranking each group of reflections in descending order yielded thus 13

sequences $h_1 \dots h_7$ corresponding to chapter "Series of g, q or $e \dots$ ", each serving as basis for finding a solution vector X for the respective structure projections.

Solution of Each Projection

In a first step, the asymmetric part A^{11} of a 11-dimensional grid was scanned in every direction using grid points given by coordinates $t_d = 0.02 + 0.03067 \cdot d$ ($d = 1 \dots 15$). Those vectors X_{t0} that produced a (merely qualitative!) e -ranking equal to the "observed" one were saved as survivors of a coarse, but in terms of PC-time efficient pre-selection of possible solution regions. In step two, the space within a hyper-cube of edge length 0.032 around each X_{t0} was inspected by random generation of vectors X_{t1} (*e. g.* 20,000 tries) and testing of the quantitative agreement. If, for example, for all $n \leq 7$ reflections $|e(h, n)_{\text{cal}} - e(h, n)_{\text{obs}}| \leq 0.1$, the vector was accepted for a further calculation of $R(e)$. If then, $R(e)$ was found below a given threshold (*e. g.* 0.03), the very vector was considered as promising candidate for a final conventional least-squares refinement on the basis of "observed" $e(h, n)$ extended to $n = 15$ (step 3). Frequently, step 2 can also be performed with, say, reflection orders 5 – 9 which improves discrimination and reduces the number of least-squares candidates. Generally, the winning X_{t2} , *i. e.* a satisfactory structure (projection) determination was found within the say 3 to 5 candidate vectors producing the best conventional agreement indices. It should, however, be noted in this context that the numbers of test reflections and the grid-point density are critical parameters that might be even better adjusted.

Three-dimensional Reconstruction of the Structure

Starting with the refined $x, y,$ and z parameters from the "best solutions", X_{t2} ($0 < x_i \leq 0.5; 0 < y_j \leq 1.0; 0 < z_k \leq 1.0; i, j, k = 1 \dots m$), all possible coordinate sums for the face and space diagonals, *e. g.* $\sum = x_i + y_j + z_k$ for the [111] direction or of $\sum = x_i - y_j$ for the [1,-1, 0] direction, were calculated and compared with the respective projection results. When the differences were below a preset threshold, *e. g.* 0.013, these coordinates were accepted as belonging to a possible atom position. From a total of $4m^3 = 5324$ "principally possible" positions in

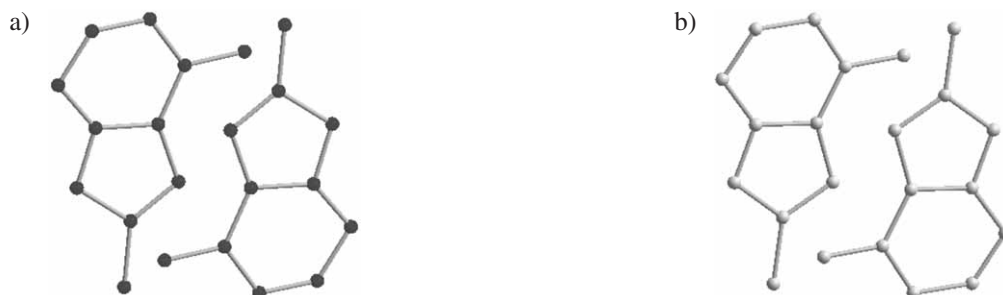


Figure 2. 11-atom part of the molecule's crystal structure. Left: true atom arrangement. Right: found as described in the text.

TABLE I. True and found fractional coordinates, their absolute differences on the basis of an assumed orthogonal lattice: $a_0 = 6.0$, $b_0 = 7.0$, $c_0 = 9.0$ Å. Dev is the distance between the true and the derived atomic position.

atom	coord.	true	found	diff	abs. diff / Å	dev / Å
1	x	0.4423	0.4423	0.0000	0.0	0.0037
	y	0.1610	0.1609	0.0001	0.0007	
	z	0.7398	0.7394	0.0004	0.0036	
2	x	0.4384	0.4345	-0.0039	-0.0234	0.0237
	y	0.1275	0.1277	-0.0002	-0.0014	
	z	0.6270	0.6266	0.0004	0.0036	
3	x	0.3722	0.3722	0.0000	0.0	0.0012
	y	0.2191	0.2190	0.0001	0.0007	
	z	0.4395	0.4394	0.0001	0.0009	
4	x	0.3693	0.3728	-0.0035	-0.0189	0.0194
	y	0.3069	0.3112	-0.0043	-0.0301	
	z	0.7903	0.7903	0.0000	0.0	
5	x	0.3691	0.3728	-0.0037	-0.0222	0.0224
	y	0.2565	0.2565	0.0000	0.0	
	z	0.5519	0.5516	0.0003	0.0027	
6	x	0.3040	0.3040	0.0000	0.0	0.0009
	y	0.4311	0.4310	0.0001	0.0007	
	z	0.7110	0.7111	-0.0001	-0.0009	
7	x	0.2993	0.3040	-0.0047	0.0282	0.0292
	y	0.4194	0.4205	-0.0011	-0.0077	
	z	0.5914	0.5914	0.0000	0.0	
8	x	0.2309	0.2309	0.0000	0.0	0.0033
	y	0.5967	0.5963	0.0004	0.0028	
	z	0.7360	0.7358	0.0002	0.0018	
9	x	0.2176	0.2176	0.0000	0.0	0.0329
	y	0.5842	0.5795	0.0047	0.0329	
	z	0.5419	0.5420	-0.0001	-0.0009	
10	x	0.1809	0.1809	0.0000	0.0	0.0343
	y	0.6891	0.6888	0.0003	0.0021	
	z	0.6304	0.6266	0.0038	0.0342	
11	x	0.1048	0.1048	0.0000	0.0	0.0432
	y	0.8774	0.8723	0.0051	0.0357	
	z	0.6293	0.6266	0.0027	0.0243	

the asymmetric half of the unit cell, only 15 survived the checking with respect to the 10 non-axial projections (four pairs were so close to each other that they could be considered identical). These final results reproduced all atom positions, and they were essentially the same independent of the sequence of the projection checks. The differences between the "true" and the derived coordinates are rather small yielding coordinate and "true" deviations as given in Table I. For an assumed lattice of $6.0 \times 7.0 \times 9.0$ Å, on average of 0.02 Å with a maximum of 0.04_3 Å. These results demonstrate again the resolution potential of the method (see Figure 2 for comparison).

In addition, the example conveys an important result: less than m independent $e(h, n)$ data suffice (at least in this case) to select a set of solution vectors allowing for a successful one-dimensional structure determination upon "refinement" on $n > m$ data.

(Examination of other types of reconstruction algorithms are under progress and will be reported elsewhere.)

DISCUSSION AND SUMMARY

Not quite unexpectedly, the method outlined has both advantages and disadvantages. Some of the "pro's" have already been mentioned in the above section "principal advantages". In addition, it is noteworthy that the use of one-dimensional projections reduces the problem of space group handling in the parameter space to the two one-dimensional cases, $P1$ and $P\bar{1}$. The actual space group of the structure becomes only relevant in the reconstruction process ("point tomography") which takes place in the 3-dimensional fractional coordinate space, optionally supported by monitoring the results in the direct crystal space which allows for checking against chemical constraints, *i. e.* reasonable bond lengths and angles. Other interesting advantages result from the high resolution, because:

- i) any well-defined solution(s) of one- or two-dimensional structure projection(s) should be able to supply Direct Methods with useful phase information;
- ii) the determination of the partial structure of one or more anomalous scatterers (the phases of their partial structure contributions thus known) may provide a better basis for completing the structure than does the same partial structure obtained from Patterson methods, and
- iii) selected data may be used to unravel structural details as illustrated by a feasibility study on a split atom problem.²⁷

Inherent "contra's", on the other hand, are readily defined by:

- i) two types of problems caused by data errors in the "experimental" $g(hkl)$ data, namely the above mentioned errors in the data reduction process, and the vulnerability of the method against individual, undetected gross errors in the measurements (whereas a few of them do not necessarily compromise a Fourier summation);
- ii) the use of one-dimensional projections which give rise to a partial loss of "atomicity", an increased probability to obtain (quasi-)homometric solutions, and a loss of "easy" origin definition in acentric projections;
- iii) the large computing demand: increasing almost proportional to p^m .

Adopting, however, an optimistic point of view, the last disadvantage is the least serious considering the still rapid progress in both numerical methods and computing technology. Calculations seeming hopeless to be performed on a to-date PC may well become routine task in the near future. Therefore, the presently achieved knowledge of the method should be consolidated (*e. g.* by improvements of the algorithms) but simultaneously, it is of pressing importance to focus on the compatibility of the "parameter space approach" with "real" experimental data.

REFERENCES

1. G. Oszlány and A. Sütö, *Acta Cryst.* **A60** (2004) 134–141.
2. R. Rothbauer, *Neues Jahrbuch Mineral. Mon.-Hefte* (1974) 14–21.
3. R. Rothbauer, *Neues Jahrbuch Mineral. Mon.-Hefte* (1975) 121–134.
4. R. Rothbauer, *Z. Kristallogr.* **209** (1994) 578–582.
5. R. Rothbauer, *Z. Kristallogr.* **210** (1995) 255–259.
6. R. Rothbauer, *Z. Kristallogr.* **215** (2000) 157–168.
7. J. Navaza and A. Silva, *Acta Cryst.* **A35** (1979) 266–275.
8. A. Cervellino and S. Ciccariello, *Riv. Nuov. Cim.* **19** (1996) 1–59.
9. A. Cervellino and S. Ciccariello, *Z. Kristallogr.* **214** (1999) 739–750.
10. A. Cervellino and S. Ciccariello, *J. Phys. A, Math. Gen.* **34** (2001) 731–755.
11. K. Fischer and K. Pilz, *Acta. Cryst.* **A53** (1997) 475–483.
12. W. Knof (*PhD Thesis, Saarbruecken*, 1989).
13. K. Fischer, A. Kirfel, and H. Zimmermann, *Z. Kristallogr.* **220** (2005) 643–656.
14. A. Kirfel, K. Fischer, and H. Zimmermann, *Z. Kristallogr.* **221** (2006) 673–680.
15. A. Navaza and J. Navaza, *Acta Cryst.* **A48** (1992) 695–700.
16. S. Ciccariello and A. Cervellino, *Z. Kristallogr. Suppl.* **10** (1995) 59.
17. A. Kirfel and K. Fischer, *Abstr. DGK Meeting 199-07-id61* (2007).
18. H. Zimmermann, A. Kirfel, and K. F. Fischer, *Z. Kristallogr. Suppl.* **22** (2005) 122.
19. H. Zimmermann, A. Kirfel, and K. F. Fischer, *Z. Kristallogr. Suppl.* **23** (2006) 107.
20. B. Buras, J. S. Olsen, and L. Gerward, *Nucl. Instrum. Meth. Phys. Res.* **152** (1978) 293–296.
21. H. G. Krane (*PhD Thesis, Saarbruecken*, 1991).
22. K. Fischer, H. G. Krane, and W. Morgenroth, *Nucl. Instrum. Meth. A* **369** (1996) 306–311.
23. J. Karle, *Int. J. Quantum Chem., Quantum Biology Symposium* **7** (1980) 357–367.
24. D. Harker and J. Kasper, *Acta Cryst.* **1** (1948) 70–75.
25. H. Ott, *Z. Kristallogr.* **66** (1927) 136–153.
26. H. Zimmermann *et al.*, in preparation.
27. A. Kirfel and K. Fischer, *Z. Kristallogr. Suppl.* **21** (2004) 101.

SAŽETAK

Koncept određivanja kristalne strukture bez uporabe Fourierove inverzije: neki koraci prema primjeni

Karl F. Fischer, Armin Kirfel i Helmuth Zimmermann

Određivanje kristalne strukture bez proračuna raspodjele elektronske gustoće Fourierovom transformacijom (čime se izbjegava i tzv. fazni problem) ostvaruje se u $3m$ dimenzijskom prostoru relativnih prostornih parametara, gdje je m broj neovisnih atoma, svedenih na jednake centre (točke) raspršenja u mirovanju. Kao demonstracija osnovnih ideja, primijenjuju se dvodimenzijски parametarski prostori, koji predočuju na pr. jednodimenzijске strukture od dva atoma. Središnji nizovi refleksa u recipročnoj rešetki omogućavaju rješavanje jednodimenzij-skih projekcija strukture, uz manji broj refleksa u svakom nizu, te uz ostvarenje boljeg razlučivanja nego pomoću Fourierovih redova. Rješenje se može ostvariti ili pomoću zajedničkog presjeka hiper-ploha u m dimenzijskom parametarskom prostoru definiranom odabranim amplitudama raspršenja, ili inspekcijom dozvoljenih, mogućih, područja rješenja što slijedi iz jednostavnog nizanja amplituda. Nalaze se sva moguća rješenja koja zadovoljavaju ulazne podatke, uključujući i "lažne minimume". Primjerom, koji se osniva na teorijskoj pretpostavki da su ulazni podatci točni, ilustrira se određivanje hipotetičke trodimenzijске strukture od 11 atoma na osnovi jednodimenzij-skih projekcija. Budući da je teorijska osnova opisana u dva prethodno objavljena rada, u ovom radu naglasak je na različitim opcijama koje se usklađuju s računalnim zahtjevima u praktičnoj primjeni. Raspravljaju se prednosti i nedostaci predloženog koncepta.