Refinement

Fourier Methods
find heavy atom or some $F$'s phases using direct methods
locate new atoms, improve phases …
continue until all atoms found in more or less correct position
starting point of refinement
solvent or unexpected atoms next
harder to find in a Fourier map – do not know where to look
or what to look for; symmetry related atoms clutter map
use a difference map
$\Delta \rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} (|F_o| - |F_c|) \cos 2\pi(hx + ky + lz - \alpha)$

Fourier Methods

_atom_sites_solution_primary direct
_atom_sites_solution_secondary difmap
_atom_sites_solution secondary difmap
difmap difference Fourier map
vecmap realspace vector search
heavy heavy atom method
direct structure invariant direct methods
geom inferred from neighboring sites
disper anomalous dispersion techniques
isomor isomorphous structure methods

Difference Map
should just contain random noise
atoms not included in phase will stand out
H's often located this way – when structure is nearly complete

Difference Map – No C$_1$

Difference Map – Find H's
Final Difference Map
when structure complete (all atoms and H) \Rightarrow \text{final difference map}
largest, most negative and average e\textsuperscript{-} \ density given in cif (e/Å\textsuperscript{3})
- _refine_diff_density_max 0.375
- _refine_diff_density_min -0.331
- _refine_diff_density_rms 0.081
should be <1 e/Å\textsuperscript{3} and random
max/min often near heavy atom; ripples of Fourier summation

Atom Type Assignment
1. peak heights
2. magnitude of displacement parameters (temperature factors)
3. refinement with alternative atoms
4. chemistry (and Cambridge Structural Database) knowledge

Least-Square Methods
usually final refinement when all non-H atoms located
involves adjusting the 9N + 1 \ (N \text{ number of atoms in asymmetric unit}) (or more) parameters (p) to minimize the function:
\[ R = \sum w_{i} \left( |F_{o}| - |kF_{c}| \right)^{2} \]
sum of the squares of the difference
minimized by taking derivative w.r.t each parameter, setting equal to zero and approximating results with truncated Taylor series;
gives n equations (number of parameters) with n unknowns \((\Delta p_{j} \text{ parameter shifts})\)

Least-Square Methods
\[
\begin{bmatrix}
   a_{11} & a_{12} & \ldots & a_{1n} & x_{1} & v_{1} \\
   a_{21} & a_{22} & \ldots & a_{2n} & x_{2} & v_{2} \\
   \vdots & \vdots & & \vdots \\
   a_{n1} & a_{n2} & \ldots & a_{nn} & x_{n} & v_{n} \\
\end{bmatrix} = \begin{bmatrix} \mathbf{A} \mathbf{x} = \mathbf{v} \end{bmatrix}
\]
elements of \(\mathbf{A}^{-1}\) usually given as \(b_{ij}\)

SHELX
minimizes:
\[ R = \sum w_{i} \left( |F_{o}|^{2} - |kF_{c}|^{2} \right)^{2} \]
equations are similar:
\[ \Delta F_{i} \text{ replaced by } |F_{o}|^{2} - |kF_{c}|^{2} \]
\[
\frac{\partial |kF_{c}|}{\partial p_{i}} \text{ replaced by } 2|kF_{c}| \ \left( \frac{\partial |kF_{c}|}{\partial p_{i}} \right) \]
- _refine_ls_matrix_type full
- _atomblock

\[ \sum_{i=1}^{m} w_{i} \frac{\partial |kF_{c}|}{\partial p_{j}} \left( \frac{\partial |kF_{c}|}{\partial p_{j}} \right) x_{i} = \Delta p_{j} \]

\[ v_{i} = \sum_{i=1}^{m} w_{i} \left( |kF_{c}| \right) \frac{\partial |kF_{c}|}{\partial p_{j}} \]

\[ \mathbf{A} = \mathbf{A}^{-1} \mathbf{v} \]
Least-Square Methods

least-squares is nonlinear ⇒ run cycles until convergence
usually want at least 10 reflections/parameter:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>_refine ls number reflns</td>
<td>3753</td>
</tr>
<tr>
<td>_refine ls number parameters</td>
<td>263</td>
</tr>
</tbody>
</table>

shifting parameters allows an estimate of the standard deviation in p’s (σₚ); calculate Δp/σₚ (shift/error) of each parameter
structure converges when Δp/σₚ ≈ 0

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>_refine ls shift/su max</td>
<td>0.000</td>
</tr>
<tr>
<td>_refine ls shift/su mean</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Estimate of Standard Deviations (esd’s)

\[
σ_{pi} = \frac{b_0 \sum w (|F_i| - |F_o|)^2}{N_{\text{reflections}} - N_{\text{parameters}}}^{1/2}
\]

α 1/Nref; more reflections = better esd’s

α σᵢ (w); good reflections = better esd’s

Correlation Coefficients

off diagonal elements of A⁻¹, bᵢⱼ, correlation of parameters i and j

\[δ_{ij} = \frac{b_{ij}}{(b_{ii})^{1/2}(b_{jj})^{1/2}}\]

0 to ±1

no correlation completely correlated
indicates refinement problems: disorder or missed symmetry

Largest correlation matrix elements

<table>
<thead>
<tr>
<th>Correlation Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.676 U13 F1 / U11 F1</td>
<td></td>
</tr>
<tr>
<td>0.563 U13 F3 / U33 F3</td>
<td></td>
</tr>
<tr>
<td>0.502 U22 Cu / OSF</td>
<td></td>
</tr>
<tr>
<td>0.440 U13 F1 / U33 F1</td>
<td></td>
</tr>
<tr>
<td>0.562 U12 F1 / U23 F1</td>
<td></td>
</tr>
<tr>
<td>0.502 U33 Cu / OSF</td>
<td></td>
</tr>
<tr>
<td>0.583 x F1 / x F1</td>
<td></td>
</tr>
<tr>
<td>0.559 U13 F3 / U11 F3</td>
<td></td>
</tr>
</tbody>
</table>

Goodness of Fit (GOF, GooF or S)

\[S = \left( \frac{\sum w (|F_o| - |kF_c|)^2}{N_{\text{reflections}} - N_{\text{parameters}}} \right)^{1/2}\]

fit of distribution of ΔF with distribution expected from w
if w are correct and model good, S = 1.0

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>_refine ls goodness of fit ref</td>
<td>1.042</td>
</tr>
<tr>
<td>_refine ls restrained S all</td>
<td>1.042</td>
</tr>
</tbody>
</table>

used to adjust w: forced = 1

<table>
<thead>
<tr>
<th>Correlation Coefficient Calculation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>calc</td>
<td></td>
</tr>
<tr>
<td>_refine ls weighting_scheme</td>
<td>calc</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>_refine ls weighting_details</td>
<td></td>
</tr>
</tbody>
</table>

used to adjust w: forced = 1

Parameters

9N + 1 parameters are:

1. scale factor: K
2. x, y, z of each non-H atom of asymmetric unit
3. temperature factor (thermal parameters)
   a. isotropic: one parameter
      first, refine all atoms isotropic to identify any problems:
      atoms with Uₘₐₓ > 0.10 or < 0.01 may be in error

Temperature Factors

too large: 1. atom position incorrect: smeared out to eliminate
          2. atom type wrong (more e⁻ than should have):
             spreads e⁻ over a larger volume to lower e⁻ density

O (8 e⁻) but should be N (7 e⁻)

3. Disorder

too small: atom type wrong (fewer e⁻ than should have):
          concentrates e⁻ in smaller volume to raise e⁻ density

N (7 e⁻) but should be O (8 e⁻)
## Disorder

1. **static disorder** atom in two (or more) position with about same energy; “see” units cells with atom in different positions

2. **dynamic disorder** atom moving relatively large distances from rest position

### Parameters

- **9N + 1 parameters are:**
  - 1. scale factor: \( K \)
  - 2. \( x, y, z \) of each non-H atom of asymmetric unit
  - 3. temperature factor (thermal parameters)
    - **anisotropic:** six parameters
      - atoms refined anisotropically before locating H’s if:
        - there is sufficient data; otherwise just heavy atoms
        - the atom is not disordered

## Temperature Factors

\[
f = \sum e^{i \sum_{h,k,l} U_{11} h^2 a^* + U_{22} k^2 b^* + U_{33} \ell^2 c^* + 2 U_{12} hka^*b^* + 2 U_{13} h\ell a^*c^* + 2 U_{23} k\ell b^*c^*}}
\]

- cross terms related to orientation
- if very elliptical, maybe split into two atoms: **partial occupancy**

## Constraints

**exact** mathematical condition: one or more variables expressed exactly in terms of other variables, and eliminated for example, coordinates and anisotropic parameters for atoms on special positions:

- **geometrical constraints:** rigid-group refinements (benzene ring) and riding models for hydrogen atom refinement
- calculated H: 1.0 Å from C at 120°
- as attached C shifts (least-squares), H moves with it

## Restraints

- additional information, not **exact** but subject to probability; inserted into least-squares as additional parameters
  1. **two or more atoms sharing the same site:** sof’s expressed as variables so sum is fixed as a constant (e.g. 1.0); disorder
    - here \(-\frac{1}{2}\) the time sof = 0.5374(2)
    - here \(-\frac{1}{2}\) the time sof = 0.4626(2)
  2. make chemically, but not crystallographically, equivalent distances equal (within an esd)
  3. **anisotropic parameters:** isolated or disordered atoms made approximately isotropic to prevent them **blowing-up**
Solving Structure

1. **heavy atom** or direct methods to get approximate phases
2. Fourier synthesis with phases and $F_o$ to find atoms
3. calculate better phases and repeat step 2; watch $R$ value
4. when all atoms ‘found’, check difference map
5. least-squares refine scale factor, $x$, $y$, $z$ and isotropic
   temperature factors until $R$ converged
6. refine anisotropically until convergence, if you have data;
   look for small shift/error
7. difference map to find H’s or calculate positions
8. refine all non-H parameters to 0 shift/error

Finally

Is structure chemically reasonable!??!

Structure Validation - An Emotive Issue.
D.J. Watkin & R.A. Cooper, Chemical Crystallography
Laboratory, Oxford, UK

In the days of photographic data collection, validation was implicit because of the painful-taking labour needed to resolve a structure, and external validation was more-or-less limited to checking for typing errors in the long lists of numbers. Serial diffractometers and computers changed this. Improvements in software made it possible for non-specialists to undertake analyses. However, it became evident that do-it-yourself analyses were not all of a high quality. It fell to the referees not only to assess the scientific merit of a paper, but also to verify that the work was technically sound.


Validation of a reported single crystal structure determination is a long-standing issue requiring an independent structure evaluation system. Dick Marsh (Space group corrections) and Dick Harlow (ORTEP-of-the-Year) clearly demonstrated the need for this already for many years. It is becoming even more urgent in view of the current practice where non-professionals carry out structure determinations using 'user-friendly-firmware' and referees with limited expertise or given insufficient access to the supporting data for independent evaluation.

Example of Bad Crystallography

paper submitted for publication in *Acta Cryst. C*

Cu$^{2+}$ L$^2^-$, confirmed with routine X-ray. $R < 7\%$

Br$^-$ L$^+$

Cu$^{2+}$ → Br$^-$ (HBr used in synthesis)

H-atoms needed/found on sp$^3$ N and $-\text{CO}_2$

An ORTEP-of-the year Award Winner

Pca2$_1$

should have been Pcam
Problem Structures

- Described as dichloride, actually disordered CO and Cl.
- True dichloride structure.

Problem Structures

- Unusual coordination number of 1.
- Correct structure.

Problem Structures

- Wrong space group.
- Correct space group.

Problem Structures

- Missing H's.
- NO₃⁻ actually CO₃²⁻.
  - Note: relatively large N temperature factor.
Wrong Structures

usually exhibit one or more of three symptoms:
- unusual bond distances
- unreasonable thermal ellipsoids
- impossible intra- and interatomic nonbonded contacts