# Lesson 22

- SIR2011
- Patterson Theory
- DIRDIF Patty
- Solve your structures

# A Quick Review

- Formula and proposed structure – User
- Unit cell constants and cell volume – Indexing
- Probable Bravais class – Indexing
- Data – Data collection
- Final Bravais class, Laue group, space group and Z -- XPREP

# Programs

- Charge Flipping—SUPERFLIP; PLATON, SHELXT
  - Ab initio does not need formula or space group
- Direct Methods—SIR2004, SHELXS(xs)
  - Uses E's
  - Needs approximate formula to calculate E and space group
- Dual Space Direct Methods—SHELXD(xm)
  - Like Direct Methods
  - Brute Force
  - Long Computing times

# Another Use of E's

- Calculate the mean value of $|E^2-1|$
- If this value is about 0.97 suggests the structure is centric
- If value is near 0.74 suggests structure is accentric
- Works best if all the atoms have about the same atomic number.
- Can be completely wrong!

# SIR2011

Released December 2011

Written by the SIR group at Bari, Italy.

So new that there is no published reference to it

It uses a new method which is neither charge flipping or direct methods but sort of a combination.

Uses Difference Fourier Maps

# Fourier Map

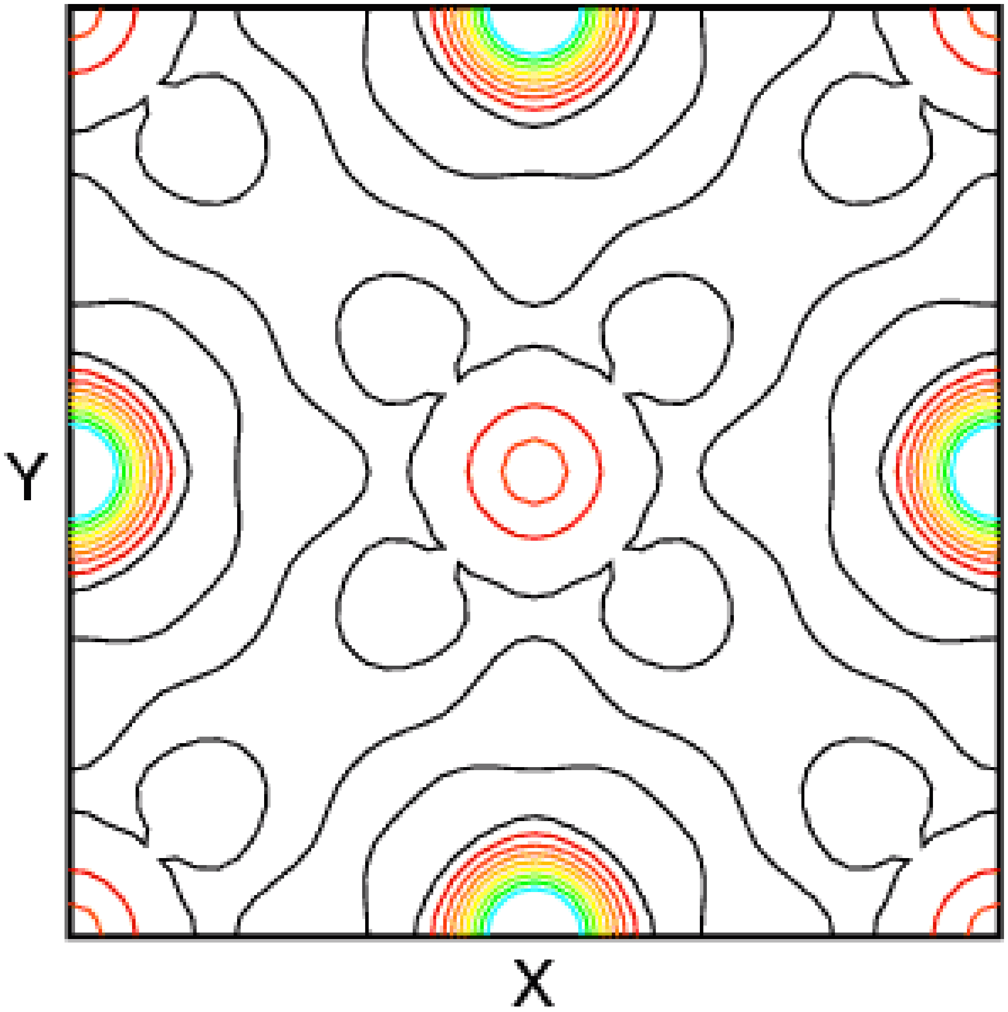
A Fourier map is a listing of electron density at any x,y,z in the unit cell.

While it could be graphed usually it appears as a table. In most cases the peaks are selected and presented in a table

Since we do not know the phases for the Fo then usually the values calculated from the atom positions (Fc) are used.

It is the real space graph of the data

# Contour Plot



# Difference Fourier Map

A properly phased Fourier Map based on Fo shows the electron density in the unit cell

A Fourier map based on Fc shows the electron density accounted for by the current model

Can subtract the two maps or simply calculate a Fourier based on |Fo-Fc|

This is called a difference Fourier map.

# Interpreting the Difference Fourier

- If there is a large positive peak in the map
  - This means there is electron density in the data that is not account for in the model—i.e. A missing atom
  - The size of the peak is indicative of the number of electrons in the missing atom
- If there is a large negative peak in the map
  - This means there is electron density in the model that is not in the crystal –i.e. there is an atom in the model that should not be there
- Peaks very close to an atom are call ghosts and have to do with the scattering factors.

# Sir2011

Sir2011 works with difference Fourier maps to sort of do charge flipping.

This algorithm is called VLD which stands for *vive la différence*

It starts by randomly placing atoms from the formula into the unit cell

Therefore it needs the formula and cell symmetry unlike charge flipping

# VLD

1. Start with atom positions
2. Do a Fourier to reciprocal space to calculate Fc and phases. Calculate $r=\Sigma|Fo-Fc|/\Sigma Fo$.
3. Create a difference Fourier map (this will be incorrect because the phases are wrong and because many of the atoms are in the wrong position).
4. Get rid of atoms with large negative peaks in the difference Fourier and add atoms where there are large positive peaks.
5. Go to step 2

Continue until r-factor is acceptable

# Advantages of SIR2011

It is very fast.

It solves large structures that SIR2004 and charge flipping fail on

It has software to help determine the space group if needed.

# The heavy atom method

By interpretation of the Patterson function, it is possible to find the positions of a small number of heavy atoms in the cell. These enable the calculated structure factor $\mathbf{F}_c$ (a complex number) or the corresponding amplitude $F_c$ and phase $\phi_c$ to be calculated for each reflection. A difference Fourier synthesis with coefficients $|F_o - F_c|$ and phases $\phi_c$ enables further atoms to be found as peaks; the approximation $\phi = \phi_c$ becomes better as the structure becomes more complete. The full structure may be found by iterative application of such difference maps.

Although it is no longer considered necessary to prepare heavy atom derivatives such as bromobenzoates to solve small molecule structures (*direct methods* are so effective), heavy atom derivatives are still useful in macromolecular crystallography. Methionine can be replaced with selenomethionine and thymine with bromouracil; heavy ions such as $Hg^{2+}$, $Pb^{2+}$ and $I^-$ can be introduced by soaking the crystals in solutions of the appropriate salts, and even xenon derivatives may be prepared in special pressure chambers.

# The Patterson function

$$\begin{tikzcd} F^2 \arrow[r, shift left, "\text{FT}"] & \textbf{Patterson} \arrow[l, shift left, "\text{FT}"] \end{tikzcd}$$

**The Patterson function is a Fourier synthesis performed with all phases set to zero and the measured $F^2$-values as coefficients. The dimensions of the unit-cell and the lattice type of the Patterson are identical with those of the crystal structure. The Patterson is a vector map: each atom-atom vector in the structure corresponds to a peak in the Patterson. The peak height is proportional to the product of the atomic numbers of the two atoms and to the number *m* of vectors that are equal in magnitude and direction and so overlap:**

$$H \propto m Z_i Z_j$$

**Usually the origin peak (height proportional to $\sum Z_i^2$) is scaled to 999, so that the peak heights of the remaining peaks are given by:**

$$H = 999\, m Z_i Z_j / \sum Z_i^2$$

**This summation is over all atoms in the unit-cell.**

# One heavy atom in P$\bar{1}$

In this space group, for each atom at $x, y, z$ there is a symmetry equivalent atom at $-x, -y, -z$ and thus a Patterson peak at $2x, 2y, 2z$.

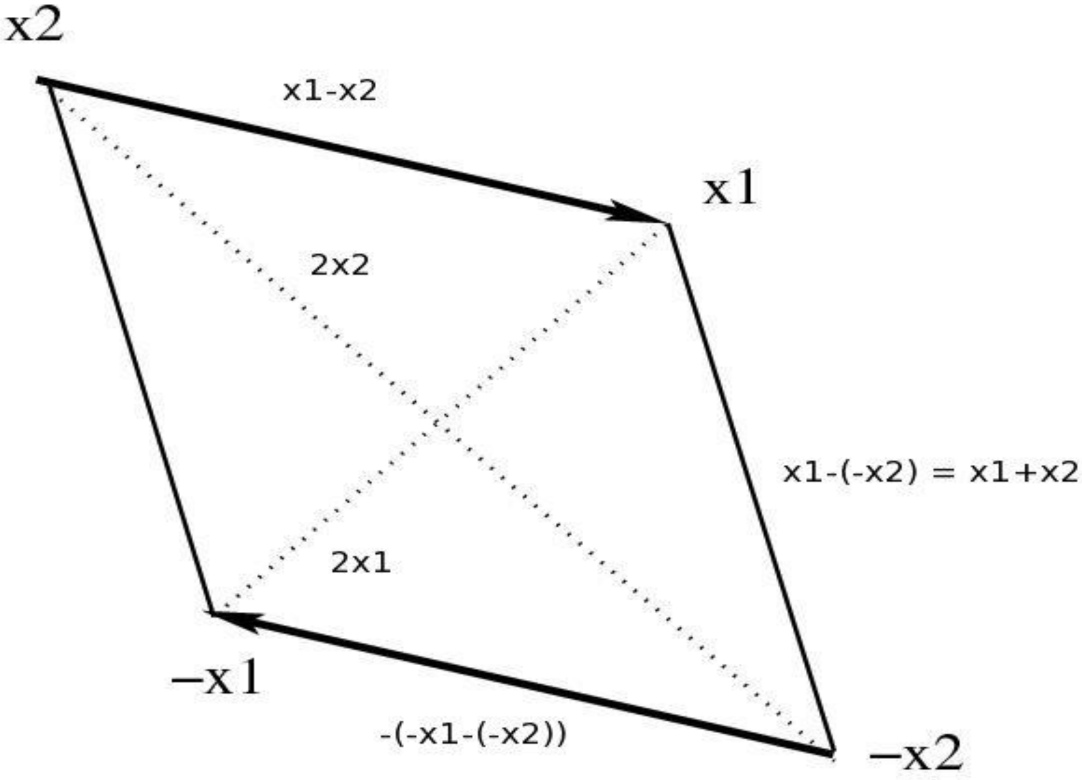
Crystals of $C_{32}H_{24}AuF_5P_2$ have two formula units in the unit-cell. The two gold atoms are related to each other by the inversion center. In addition to the origin peak (999) there is a peak of height 374 with $X$=0.318, $Y$=0.471, $Z$=0.532, and of course a peak at $-X, -Y, -Z$, that are much higher than the rest (≤145). This peak height is in good agreement with the calculated value of 377 for a Au—Au vector.

In order to calculate $x$ from $2x$ it would be far too simple just to divide by 2! We have to take into account that there must also be peaks in the next unit-cell of the Patterson with $X = 2x+1$ etc. Thus we calculate the possible gold coordinates as follows:

$$x = 0.318/2 \text{ or } 1.318/2 = 0.159 \text{ or } 0.659$$
$$y = 0.471/2 \text{ or } 1.471/2 = 0.236 \text{ or } 0.736$$
$$z = 0.532/2 \text{ or } 1.532/2 = 0.266 \text{ or } 0.766$$

These *eight* possible positions correspond to the *eight* possible ways of putting the origin on an inversion center in space group P1.

# Dimer in P-1



# Interpretation of the Patterson

| | *X* | *Y* | *Z* | height | interpretation |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 999 | origin |
| • | 0.765 | 0.187 | 0.974 | 310 | x(Ag)+x(As) |
| • | 0.392 | 0.099 | 0.325 | 301 | x(Ag)–x(As) |
| 4 | 0.159 | 0.285 | 0.298 | 250 | 2x(Ag) |
| . . | . . | . . | . . | . . | . . |
| . . | . . | . . | . . | . . | . . |
| • | 0.364 | 0.077 | 0.639 | 102 | 2x(As) |

**A consistent solution would be:**

Ag: $x = 0.080$, $y = 0.143$, $z = 0.149$
As: $x = 0.682$, $y = 0.039$, $z = 0.820$

**For the first atom (Ag) there are 8 equivalent solutions, so one can simply divide 2x by 2. The second atom (As) must however be consistent with the Ag, so x(Ag) is subtracted from x of peak 2 and then checked whether this As position agrees with the other vectors.**

# The symmetry of the Patterson function

**Because for each vector i → j there is always a vector j → i, the Patterson must possess an *Inversion center*.**

**In general the symmetry of the Patterson is determined by the symmetry of the diffraction pattern, which also possesses an inversion center. Glide planes of the space group become, as in reciprocal space, normal mirror planes, and screw axes become rotation axes without translation. In other words, the Patterson exhibits the same symmetry as the Laue group.**

# Harker sections

The symmetry of the Patterson leads to a concentration of peaks in particular planes (or along particular lines). E.g.

**$P2_1$**: Atoms at $x\ y\ z$; $-x\ \tfrac{1}{2}+y\ -z$. Self-vectors at $2x\ \ \tfrac{1}{2}\ \ 2z$, ***Harker section*** **at** $Y = \tfrac{1}{2}$.

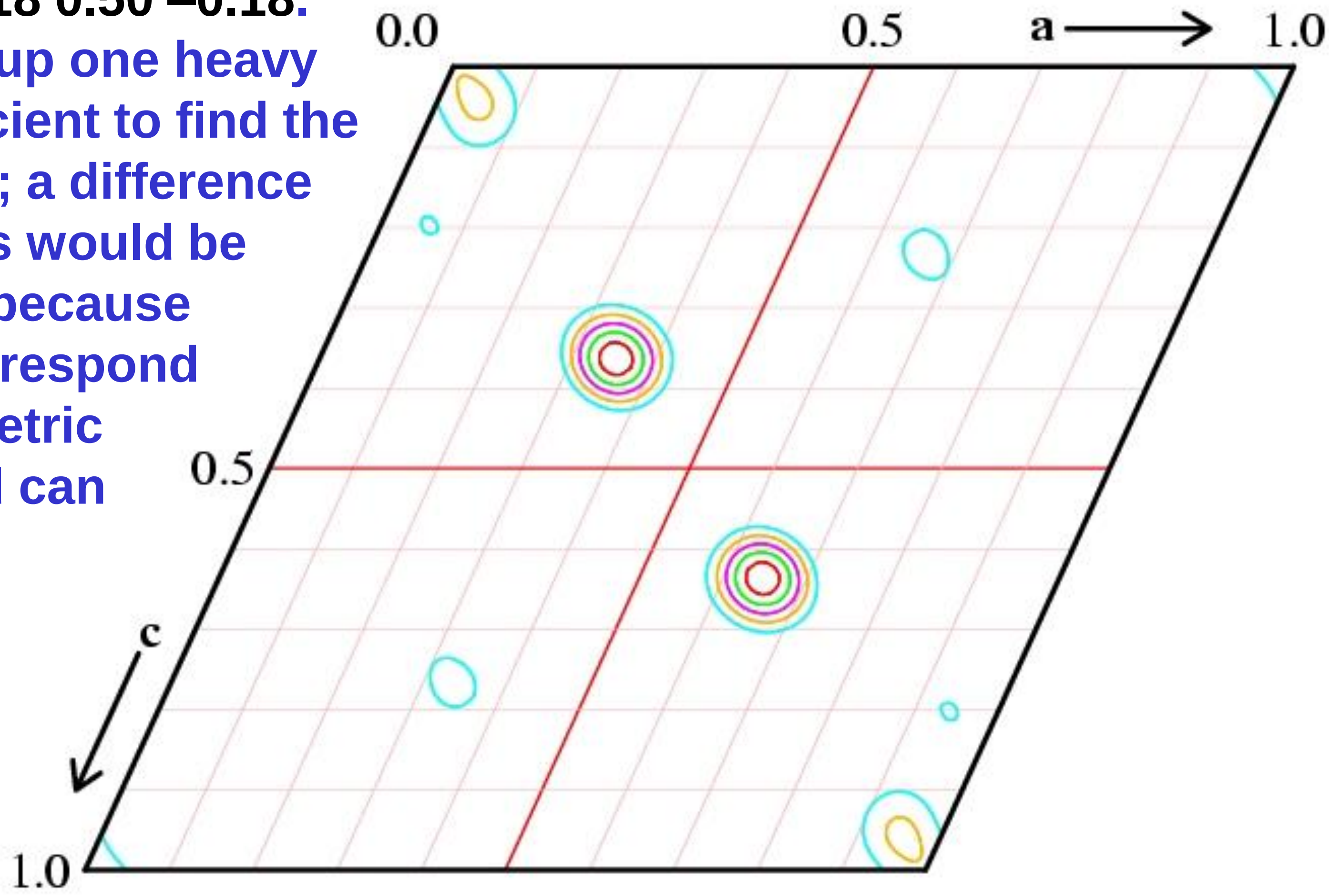
**P2**: Atoms at $x\ y\ z$; $-x\ y\ -z$. Self-vectors at $2x\ \ 0\ \ 2z$, ***Harker section*** **at** $Y = 0$.

**P*m***: Atoms at $x\ y\ z$; $x\ -y\ \ z$. Self-vectors at $0\ \ 2y\ \ 0$, ***Harker line*** **at** $X = 0$, $Z = 0$.

**P2** and **P*m*** would have identical systematic absences, but could be distinguished from one another using the ***Harker section*** and ***line***. The **$2_1$** axis in **$P2_1$** results in peaks concentrated in the Harker section at $Y = \tfrac{1}{2}$; however the systematic absences $k \neq 2n+1$ for reflections $0\ k\ 0$ would give a more reliable indication of the the space group.

# One heavy atom in $P2_1$

An organoselenium compound with one unique Se atom (two in the unit-cell) in **$P2_1$** gave the **$Y = ½$** Harker section in the diagram. The peaks at **$\pm(2x\ ½\ 2z) = \pm(0.36\ 0.50\ 0.36)$** correspond to a Se atom at **0.18 0 0.18** and its symmetry equivalent at **–0.18 0.50 –0.18**. In this space group one heavy atom is not sufficient to find the remaining atoms; a difference Fourier synthesis would be a double image, because the Se atoms correspond to a centrosymmetric substructure and can not resolve the enantiomorph!



# Two independent heavy atoms in $P2_1$

|  | $x_1\ y_1\ z_1$ | $x_2\ y_2\ z_2$ | $-x_1\ \tfrac{1}{2}+y_1\ -z_1$ | $-x_2\ \tfrac{1}{2}+y_2\ -z_2$ |
|---|---|---|---|---|
| $x_1\ y_1\ z_1$ | 0 | $-x_1+x_2\ -y_1+y_2\ -z_1+z_2$ | $-2x_1\ \tfrac{1}{2}\ -2z_1$ | $-x_1-x_2\ \tfrac{1}{2}-y_1+y_2\ -z_1-z_2$ |
| $x_2\ y_2\ z_2$ | $x_1-x_2\ y_1-y_2\ z_1-z_2$ | 0 | $-x_1-x_2\ \tfrac{1}{2}+y_1-y_2\ -z_1-z_2$ | $-2x_2\ \tfrac{1}{2}\ -2z_2$ |
| $-x_1\ \tfrac{1}{2}+y_1\ -z_1$ | $2x_1\ \tfrac{1}{2}\ 2z_1$ | $x_1+x_2\ \tfrac{1}{2}-y_1+y_2\ z_1+z_2$ | 0 | $x_1-x_2\ -y_1+y_2\ z_1-z_2$ |
| $-x_2\ \tfrac{1}{2}+y_2\ -z_2$ | $x_1+x_2\ \tfrac{1}{2}+y_1-y_2\ z_1+z_2$ | $2x_2\ \tfrac{1}{2}\ 2z_2$ | $-x_1+x_2\ y_1-y_2\ -z_1+z_2$ | 0 |

**In the table –½ is replaced by +½ because they are equivalent.**

**These Patterson peaks can be summarized as follows:**

|  |  |  |
|---|---|---|
| $0\quad 0\quad 0$ | $m = 4$ | **origin** |
| $\pm\{\ 2x_1\ \ \tfrac{1}{2}\ \ 2z_1\ \}$ | $m = 1$ | **Harker section at y = ½** |
| $\pm\{\ 2x_2\ \ \tfrac{1}{2}\ \ 2z_2\ \}$ | $m = 1$ |  |
| $\pm\{\ x_1-x_2\ \ y_1-y_2\ \ z_1-z_2\ \}$ | $m = 1$ |  |
| $\pm\{\ x_1-x_2\ \ -y_1+y_2\ \ z_1-z_2\ \}$ | $m = 1$ | **cross vectors** |
| $\pm\{\ x_1+x_2\ \ \tfrac{1}{2}+y_1-y_2\ \ z_1+z_2\ \}$ | $m = 1$ |  |
| $\pm\{\ x_1+x_2\ \ \tfrac{1}{2}-y_1+y_2\ \ z_1+z_2\ \}$ | $m = 1$ |  |

**The effective space group of this Patterson is $P2/m$, which can be deduced from the last four (general) vectors.**

# Problems with the heavy atom method

**In many cases, one heavy atom suffices to find the remaining atoms by (possibly iterative) application of difference Fourier syntheses with amplitudes $|F_o – F_c|$ and Phases $\phi_c$. Exceptions are those non-centrosymmetric space groups (such as $P2_1$ and C2) in which one unique heavy atom generates a centrosymmetric arrangement of heavy atoms in the cell; then all phases are 0º or 180º and the resulting difference electron density retains the inversion center, i.e. corresponds to a double image consisting of the two possible enantiomorphs.**

**A different problem – which also arises when the structure is determined by direct methods – is the assignment of the element types; electron density is always blue! The density is proportional to the atomic number, but isoelectronic species are particularly difficult to distinguish, as will be seen in the first exercise...**

# DIRDIF Patty

- DIRDIF is a program package written by Paul Beurskens at the University of Nijmengen, Netherlands.
- Patty is the Patterson solver.
- Has one weakness—insists on using the formula for heavy atoms even if it is obviously wrong
- Look at r-factor
  - >0.50 no solution;  <0.20 solution; intermediate may be solution with mis-assigned atoms.

# The .ins FIle

```
TITL xl in P2(1)/n
CELL 1.54178  12.8613  13.7235  24.4278  90.000  91.856  90.000
ZERR    8.00   0.0004   0.0005   0.0009   0.000   0.002   0.000
LATT  1
SYMM 0.5-X, 0.5+Y, 0.5-Z
SFAC C H N I U
UNIT 256 304 16 8 8
TEMP -123.15
SIZE 0.10 0.14 0.16
TREF
HKLF 4
END
```

# Homework

- Assign all possible space groups to data with the following systematic presences
- Monoclinic—hkl, h+k=2n; h0l, l=2n
- Orthorhombic—0kl, k+l=2n;  h0l,h=2n
- Orthorhombic-- hkl, h+k=2n, k+l=2n, h+l=2n; 0kl, k+l=4n; h0l, h+l=4n
- hexagonal/trigonal-- h-hl, l=2n; 00l, l=2n