

Tutorial 5: Anomalous Phasing

James Foadi
email j.foadi@bath.ac.uk

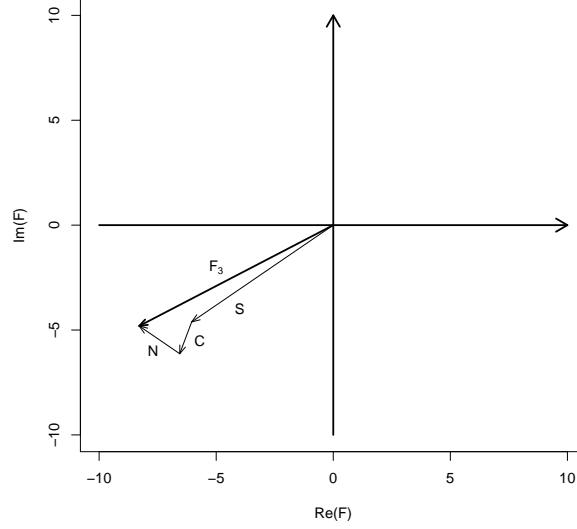
July, 2019

1 Anomalous scattering

This is a very cursory presentation of anomalous scattering in crystallography. A very accessible but more in-depth introduction to the same subject can be found on the web site maintained by Ethan Merritt at the Biomolecular Structure Centre, University of Washington [[merritt](#)]. Atoms scatter anomalously when the energy, and thus the wavelength, used is close to their resonant values. This means that in general crystals do not scatter anomalously or, better, that the magnitude of anomalous diffraction is negligible when compared to ordinary scattering. At specific wavelengths, though, certain atoms in the crystal resonate and the resulting diffracted beam includes significant anomalous contributions. Their particular nature is exploited for phasing, i.e. to find phases for the scattering factors. Ordinary diffraction is mathematically reflected in the usual factor appearing in the following equation,

$$F_h = \sum_{j=1}^N f_j \exp \left(2\pi i h \frac{x_j}{a} \right),$$

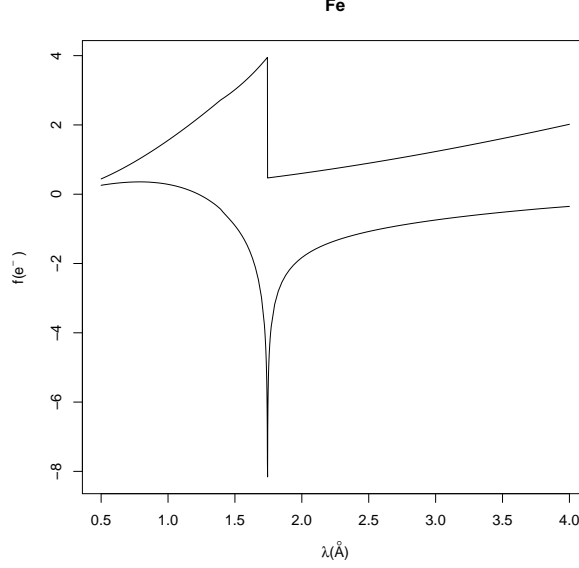
where N is the number of atoms in the unit cell. From the above expressions it emerges clearly that the structure factors of crystallographic structures composed of several atoms in the unit cell are the sum of individual complex numbers, each one of them having length equal to the atom's scattering factor f_j and phase built out of the atom's position. A representation of this sum in the Argand plane for the thiocyanate structure is shown here.



The expression for anomalous diffraction makes use of imaginary scattering factors in the following way:

$$F_h = \sum_{j=1}^N (f_j + f'_j + if''_j) \exp\left(2\pi i h \frac{x_j}{a}\right)$$

The new scattering factor is equal to the ordinary, real, component plus a complex one, $f'_j + if''_j$. Both real and imaginary parts of this additional anomalous factor are virtually independent of resolution and only depend on wavelength. Values of f'_j and f''_j for all atomic species and for several wavelengths have been calculated theoretically and can be easily tabulated. The values used in **crone** have been extracted from Ethan Merritt's site [**merritt**]. An example of f'_j and f''_j as functions of wavelength for the iron atom can be obtained using **crone** function `plot_absorption_curves`, The second argument of this function also allows to zoom in or out the plot at a specified window with minimum and maximum wavelength λ .



The top curve describes f_j'' , while the bottom curve describes f_j' . In order to phase structures using anomalous scattering it is necessary to choose wavelengths close to the resonant energies (close to the dip of f_j'), because in its neighbourhood the f'' anomalous contribution will be largest, causing structure factors to change appreciably from their ordinary values.

2 The breaking of Friedel's law

The most important consequence that anomalous scattering has on structure factors is the breaking of the so-called *Friedel's law*, according to which amplitudes of structure factors with opposite Miller indices are identical, and phases are opposite of each other. For structures not scattering anomalously in any appreciable way, such amplitudes will not be identical, due to experimental errors, but very similar. When some atoms in the structure scatter anomalously, the same related amplitudes are expected to be systematically different.

Given a Miller index, h , and its opposite, $-h$, Friedel's Law states that the structure factor associated with h is the complex conjugate of the structure factor associated with $-h$:

$$F_{-h} = F_h^* \quad \Leftrightarrow \quad |F_{-h}| = |F_h| \quad \text{and} \quad \varphi_{-h} = -\varphi_h$$

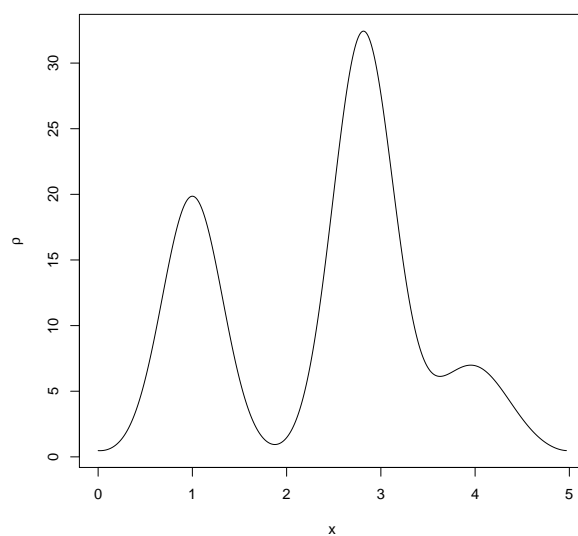
When contributions from the anomalous scattering become important, though,

Friedel's Law is not valid any longer and in general the following will be true:

$$|F_{-h}| \neq |F_h|$$

A demonstration of Friedel's law breaking down can be carried out using function `strufac`, exploiting more specifically the parameter `'anoflag'`.

```
# Replace the carbon of thiocyanate with Iron, as  
# this scatters anomalously in a significant way  
sdata$Z  
  
## [1] 16 6 7  
  
sdata$Z[2] <- 26  
  
# Reduce thermal vibration  
sdata$B  
  
## [1] 5.000 13.333 11.429  
  
sdata$B[2] <- 3  
  
# Plot electron density  
rtmp <- structure_gauss(sdata=sdata,N=1000)  
plot(rtmp$x,rtmp$rr,type="l",xlab=expression(x),  
      ylab=expression(rho))
```



The wavelength at which the difference between f_j'' and f_j' is largest, and at which the Friedel's law is mostly violated, can be found using the function `fluorescent_scan` which mimics the physical process carried out at synchrotrons to find the optimal wavelength for anomalous phasing.

```
# Wavelength used when anomalous scattering is significant
out <- fluorescent_scan("Fe")
my_lambda <- out$lambda
my_lambda

## [1] 1.742797

# Miller indices used
hidx <- -2:2

# No breaking of Friedel's law (lambda = 0.5). Differences
# in Bijvoet pairs are minimal
ftmp <- strufac(hidx=hidx,sdata=sdata,lbda=0.5,
               anoflag=TRUE,f1f2out=FALSE) # No message printing
ftmp$Fmod[c(2,4)] # h=-1,1

## [1] 15.77355 15.72275

ftmp$Fmod[c(1,5)] # h=-2,2

## [1] 18.02202 18.17212

# Breaking of Friedel's law
# (lambda = 'my_lambda', roughly 1.743)
ftmp <- strufac(hidx=hidx,sdata=sdata,lbda=my_lambda,
               anoflag=TRUE,f1f2out=FALSE) # No message printing
ftmp$Fmod[c(2,4)] # h=-1,1

## [1] 13.75429 13.25682

ftmp$Fmod[c(1,5)] # h=-2,2

## [1] 15.77879 17.23431
```

It is also possible to use the default output of `strufac` and re-establish the ideal Friedel's law using `anoflag=FALSE` (default).

```

ftmp <- strufac(hidx=hidx,sdata=sdata,lbda=my_lambda)
ftmp$Fmod[c(2,4)] # h=-1,1

## [1] 15.68228 15.68228

ftmp$Fmod[c(1,5)] # h=-2,2

## [1] 18.02887 18.02887

```

3 Anomalous phasing

The pair corresponding to opposite Miller indices is called Bijvoet pair. In many crystallographic computer programs the Bijvoet pair is typically divided in its *plus* and a *minus* part, indicated as F^+ and F^- . One of the methods used to phase crystal structures with anomalous phasing employs Bijvoet pairs to create a special density map called *anomalous Fourier difference* map (or anomalous difference in short). This density map has amplitudes equal to the absolute square difference of each Bijvoet pair, and all phases equal to 0. If the anomalous difference map is indicated as $\rho_{ano}(x)$, then:

$$\rho_{ano}(x) = \frac{1}{a} \sum_{h=-\infty}^{+\infty} \Delta_{ano}^2(h) \exp\left(-2\pi i h \frac{x}{a}\right)$$

where,

$$\Delta_{ano}(h) \equiv ||F^+| - |F^-||$$

It can be shown that the anomalous difference map is equivalent to a corrupted Patterson map in which some of the prominent peaks correspond to distances between the anomalous scatterers, e.g. peaks between heavy atoms [**woolfson**]. Once these distances have been measured, an initial, approximate model of the structure, made of anomalous scatterers, can be suggested. An updated Fourier transform calculated with the experimental structure factors' amplitudes and phases from the initial model, should show improved positions for the anomalous scatterers, but also peaks corresponding to atoms not yet determined. The updated density map will be closer to the true density map if the selected peaks from the anomalous Fourier difference are the correct ones. The updated density can, in turn, be used to confirm or reject the positions of the anomalous scatterers and to select new peaks for the atoms missing from the initial approximate model of the structure. Old and new peaks will form a new structural model used to carry

out another Fourier cycle (observed amplitudes and calculated phases). This sort of Fourier recycling provides atomic positions converging to the correct structure, if the initial anomalous scatterers have been properly selected. A specific example will make the procedure clearer.

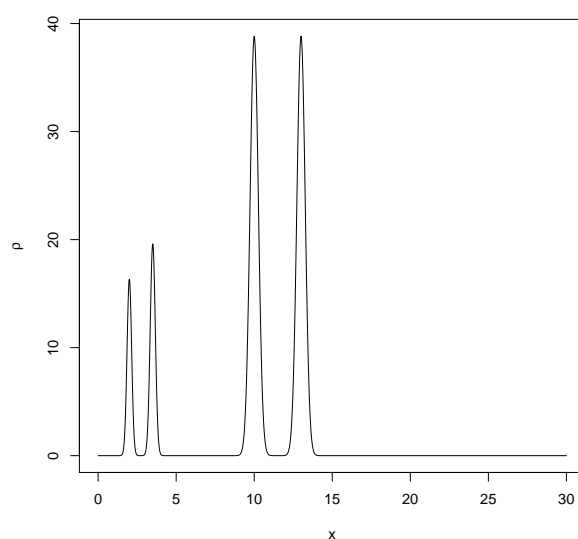
Consider a P1 structure made of one carbon, one oxygen and two iron atoms at positions:

$$x_1 = 2.0, x_2 = 3.5, x_3 = 10.0, x_4 = 13.0$$

in a unit cell of side $a = 30 \text{ \AA}$. All atoms are kept at a very low temperature, corresponding to a B-factor equal to 0.5 \AA^2 .

```
# New structure (use 'standardise_sdata')
sdata <- standardise_sdata(a=30,
                           SG="P1",
                           x0=c(2,3.5,10,13),
                           Z=c(6,8,26,26),
                           B=c(0.5,0.5,0.5,0.5),
                           occ=c(1,1,1,1))

# Electron density
rtmp <- structure_gauss(sdata=sdata,N=1000)
plot(rtmp$x,rtmp$rr,type="l",xlab=expression(x),
     ylab=expression(rho))
```



Structure factors for this structure can be computed for positive and negative Miller indices up to, say, $h = 80$. In the following code snippet their amplitudes for the first 5 Bijvoet pairs and corresponding averages, taken to be equal to the observed amplitudes, are also displayed.

```
# h from 1 to 80
hidx <- -80:80

# Structure factors with Friedel's law violated
ftmp <- strufac(hidx=hidx,sdata=sdata,anoflag=TRUE,
                lbda=my_lambda)

##      Z          f1          f2
## 1   6   0.02247688 0.01190210
## 2   8   0.06101792 0.04200023
## 3  26  -7.71243900 3.95011200

# F-
Fminus <- ftmp$Fmod[1:80]
Fminus <- rev(Fminus)

# F+
Fplus <- ftmp$Fmod[82:161]

# Averages
Fobs <- 0.5*(Fplus+Fminus)

# Delta_ano
Dano <- abs(Fplus-Fminus)

# Display
line1 <- sprintf("%10.3f %10.3f %10.3f %10.3f",
                 Fplus[1],Fminus[1],Fobs[1],Dano[1])
line2 <- sprintf("%10.3f %10.3f %10.3f %10.3f",
                 Fplus[2],Fminus[2],Fobs[2],Dano[2])
line3 <- sprintf("%10.3f %10.3f %10.3f %10.3f",
                 Fplus[3],Fminus[3],Fobs[3],Dano[3])
line4 <- sprintf("%10.3f %10.3f %10.3f %10.3f",
                 Fplus[4],Fminus[4],Fobs[4],Dano[4])
line5 <- sprintf("%10.3f %10.3f %10.3f %10.3f",
                 Fplus[5],Fminus[5],Fobs[5],Dano[5])
cat(c(line1,line2,line3,line4,line5),sep="\n")

##      43.191      44.542      43.867      1.351
```


##	27.420	26.568	26.994	0.852
##	37.172	36.480	36.826	0.692
##	22.517	23.099	22.808	0.581
##	9.880	9.890	9.885	0.009

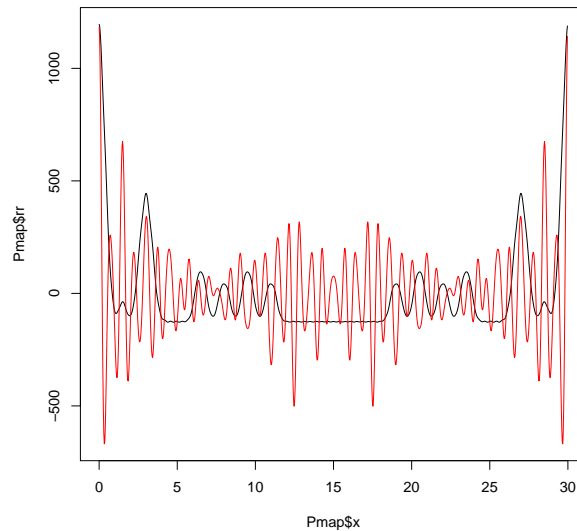
Let's now compare the anomalous difference map with the Patterson map. We have scaled up the anomalous map in order to compare it visually with the Patterson map.

```
# Patterson map
Fmod <- Fobs^2
Fpha <- rep(0,length=80)
Pmap <- fousynth(a=sdata$a,Fmod=Fmod,Fpha=Fpha,
                 hidx=1:80,N=1000)

# Anomalous difference map
Fmod <- (Fplus-Fminus)^2
Anomap <- fousynth(a=sdata$a,Fmod=Fmod,Fpha=Fpha,
                  hidx=1:80,N=1000)

# Scale for anomalous map
KK <- sd(Pmap$rr)/sd(Anomap$rr)

# Compare maps
ym <- min(Pmap$rr,KK*Anomap$rr)
yM <- max(Pmap$rr,KK*Anomap$rr)
plot(Pmap$x,Pmap$rr,type="l",ylim=c(ym,yM))
points(Anomap$x,KK*Anomap$rr,type="l",col=2)
```



Some of the highest peaks in the anomalous difference map (with the exclusion of the origin peak, which is always the highest peak) correspond to the interatomic distance between anomalous scatterers. Let's find out what positions these peaks fall at.

```
# Patterson map peaks
idxP <- local_maxima(Pmap$rr)
jdxP <- order(Pmap$rr[idxP],decreasing=TRUE)
for (i in 1:16) {
  line <- sprintf("%10.3f\n",Pmap$rr[idxP[jdxP[i]]])
  cat(line)
}

## 1195.505
## 444.795
## 444.795
## 95.872
## 95.872
## 95.571
## 95.571
## 42.191
## 42.191
## 42.095
## 42.095
```

```
##      -37.117
##      -37.117
##     -123.438
##     -123.438
##     -123.519
```

Aside from the origin peak, there are only five other peaks (each peak has a symmetry-related mate) that should indicate inter-atomic distances. The same process carried out above can be repeated for the anomalous difference.

```
# Anomalous map peaks
idxA <- local_maxima(Anomap$rr)
jdxA <- order(Anomap$rr[idxA],decreasing=TRUE)
for (i in 1:16) {
  line <- sprintf("%10.3f\n",Anomap$rr[idxA[jdxA[i]]])
  cat(line)
}

##      7.370
##      4.207
##      4.207
##      2.139
##      2.139
##      1.981
##      1.981
##      1.933
##      1.933
##      1.614
##      1.614
##      1.544
##      1.544
##      1.283
##      1.283
##      1.250
```

Here we also have an origin peaks and symmetry-related mates. The first higher peak aside from the origin peak should correspond to the inter-atomic distance between the two iron atoms. A comparison of the positions between the peaks in the Patterson and anomalous difference maps should locate such a distance.

```
# First three anomalous-difference peaks
```

```
Anomap$x[idxA[jdxA[c(2,4,6)]]]
```

```
## [1] 1.50 3.03 17.19
```

```
# First three Patterson peaks
```

```
Pmap$x[idxP[jdxP[c(2,4,6)]]]
```

```
## [1] 3.00 6.48 9.48
```

For this specific case, the second highest peak (both for the Patterson and the anomalous difference) is the one corresponding to the inter-atomic distance of the two irons which is, then, fixed at 3. We can place two iron atoms anywhere in the unit cell (because of the arbitrariness of cell's origin), as long as their distance is 3.00 Å, i.e. the position of the chosen peak in the anomalous difference map. The two iron atoms arranged in this way form an initial model for the complete structure. In order to avoid placing atoms in the cell origin, let's choose positions $x = 1$ and $x = 4$ for the iron atoms in our initial model. The updated electron density map will consist of the observed structure factors and the phases coming from this model.

```
# Updated model
```

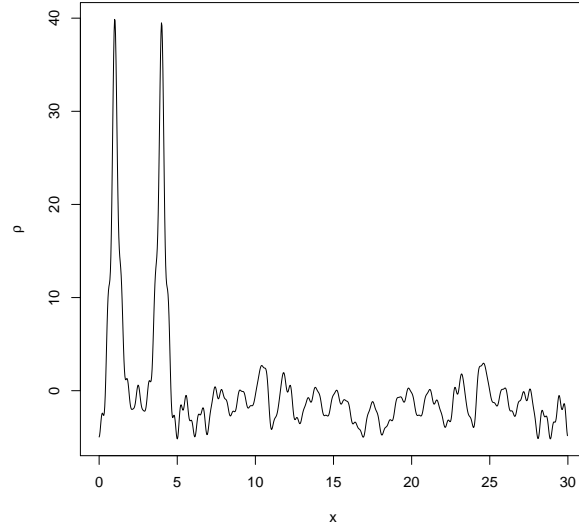
```
sdataU <- standardise_sdata(a=sdata$a,  
                           SG="P1",  
                           x0=c(1,4),  
                           Z=c(26,26),  
                           B=c(0,0),  
                           occ=c(1,1))
```

```
# Calculated phases
```

```
ftmpU <- strufac(hidx=1:80,sdata=sdataU)  
Fpha <- ftmpU$Fpha
```

```
# New electron density map
```

```
rtmpU <- fousynth(a=sdata$a,Fmod=Fobs,Fpha=Fpha,hidx=1:80,N=1000)  
plot(rtmpU$x,rtmpU$rr,type="l",xlab=expression(x),  
     ylab=expression(rho))
```



Aside from the main two peaks corresponding to iron atoms, there are many more smaller peaks, mostly corresponding to noise. Some of higher peaks, though, should correspond to the remaining carbon and oxygen. Trial and error maps with attempted C and O at these peaks position will eventually reveal the complete and correct structure.