Powder Data Analysis: Indexing & Structural Refinement based on the Rietveld Method

An introduction to the basic knowledge required for Project 2

X-ray Crystallography
Department of Chemistry & Biochemistry

cora.lind@utoledo.edu

THE UNIVERSITY OF TOLEDO

Analyzing Powder Data of New Material

- **Assumption: A high crystallinity, phase-pure sample has been prepared, and good data have been collected**
  - Refer to Powder Methods handout for sample prep and data collection!
- **First step is generally a search & match of a powder pattern database**
  - ICDD's PDF (Powder Diffraction File) if you have access
  - COD (Crystallography Open Database) allows export of a PDF as well
    - http://www.crystallography.net/cod/
- **If you find a reasonable match: Does the composition found make sense for the compound you measured?**
  - E.g., $AlScMo_3O_{12}$ gave a match to $Fe_2Mo_3O_{12}$, which makes sense
  - "Substitute" atoms and use as starting structural model
- **If no match is found: You need to find a cell by *indexing***

Indexing of Powder Data

- **Indexing of powder data is much harder than one may think**
  - Need to determine 6 or less variables (lattice parameters)
- **Problem: Only 1D data are available**
  - We are trying to solve a 3D problem
  - Peak overlap (exact or partial)
  - Weak peaks and systematic absences
  - Random and systematic errors
  - Usually more than 1 unique solution possible
- **We are trying to solve the following equation:**

$$\frac{1}{d^2} = h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2} + 2\,hk\, a^* b^* \cos\gamma^* + 2hl\, a^* c^* \cos\beta^* + 2kl\, b^* c^* \cos\alpha^*$$

$$\frac{1}{d^2} = h^2\mathbf{A} + k^2\mathbf{B} + l^2\mathbf{C} + hk\,\mathbf{D} + hl\,\mathbf{E} + kl\,\mathbf{F}$$

Indexing Approaches

- **There are several ways to solve the equation on the previous slide**
- **The most common algorithms are referred to as**
  - Zone indexing (used in ITO)
  - Successive dichotomy (used in DICVOL – exhaustive search)
  - Index heuristics (used in TREOR – “smart trial and error”)
  - There are also some more modern routines that are implemented in modern Rietveld codes
- **We will not talk about the underlying equations/details for each of these algorithms, but Project 2 will involve indexing powder data**
- **It can be a good idea to use multiple indexing programs for the same data**
  - When one fails, another one may succeed
  - If we get the same cell with multiple programs, it increases our confidence in that cell

Indexing – Practical Considerations

- **In contrast to automated indexing algorithms implemented in single crystal programs, YOU decide what peaks to include in indexing powder data!**
  - As a famous powder crystallographer has said, “Even the best program will not help you if you give it bad data!”
- **Two main options:**
  - Manual peak fitting, choose peaks from output to include in indexing list
  - Use indexing program that starts with powder pattern
- **It may be wise to exclude unreliable peaks (weak, overlapping)**
- **Closely inspect output**
  - Figure of merit is used to rank cells – the higher, the better
  - However, more important question is whether all strong/reliable peaks are accounted for!

Successfully Indexed – What's Next?

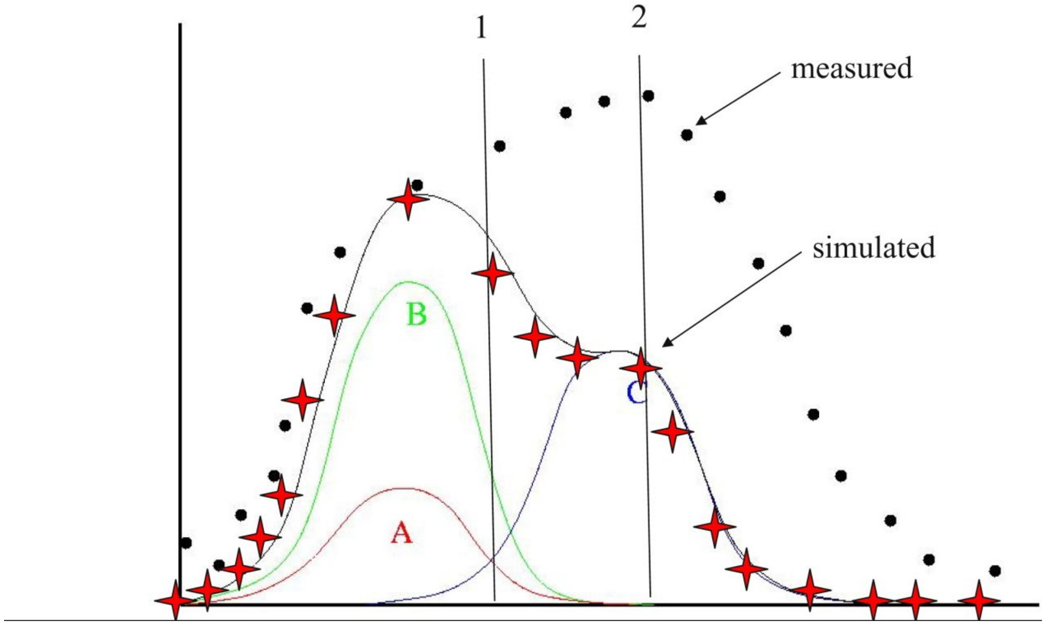
- **Once you have identified a cell that agrees well with your indexing peak list, you can confirm it by doing a Le Bail or Pawley fit**
  - Requires unit cell constants and space group, but no structural model
  - Look at systematic absences to choose a reasonable space group
  - Intensities are considered variables
  - Inspect your difference curve: Are all peaks accounted for?
- **Extracted intensities can be exported and used for structure solution**
  - We will not get into this in this class, as it is non-trivial!
- **Le Bail or Pawley fit shows what the best possible fit using a structural model could look like**
- **Once you have a structural starting model (from any source – search/match, structure solution etc.):**
  Carry out a Rietveld refinement!

Powder crystallography before Rietveld

- **Primary strategy: AVOID when it comes to structure determination!**
  - There was no straightforward way to deal with data
  - Had to manually integrate intensities
  - Overlapping reflections were a big problem
    - Usually discarded
    - Alternative: Rewriting of single crystal software to refine using sums of overlapped reflections
- **Powder pattern simulation was more common**
  - Relatively straightforward
  - Conclusions drawn based on similarities between patterns (e.g., isostructural compounds)
    - Visual comparison
- **Main use of powder diffraction was for phase identification**

Major breakthrough: Overlapped reflections

- Rietveld witnessed the power of introducing computers into crystallography during his dissertation (1961-1964)
- He realized that computers can handle individual datapoint intensities, which allowed calculation of $|F_{hkl}|$ even for overlapped reflection!



History of the Rietveld method

- **First introduced by Hugo Rietveld in 1966-1969**
    - He realized that if a pattern can be simulated/calculated, it should be possible to optimize the agreement between the calculated pattern and the collected data
    - Little response from the scientific community after first presentation of the method in 1966
    - Shared his program as well as knowledge freely
- **Became possible by increase in computing power**
    - Nowadays, a number of program codes are freely available for a variety of operating systems
- **Originally introduced for neutron data (Rietveld worked at a neutron reactor), first used for X-ray data in 1977**

What is the Rietveld method?

- **Least squares based minimization algorithm to obtain the best fit between a structural model and a powder pattern**
  - Demanding, as the algorithm is non-linear
  - User decides which parts of the model can be varied
- **Each point in the pattern can be regarded as an observation**
  - "No Bragg intensity" tells you something about your material, too!
- **Full pattern fitting**
  - In contrast to single crystal data, "experiment dependent parameters" must be fitted as well: Background, peak shape – sample and instrument contributions, lattice constants, …
- **Requires an approximate starting model**
  - Rietveld can NOT be used to "solve structures", it is a refinement method!

Goal of Rietveld refinement

- **Obtain the best possible fit between a structural model and a powder pattern**
    - User decides what parameters can be varied
- **Can be used to confirm that a sample has a specific structure**
    - Especially for single phase samples, a good fit is usually accepted as confirmation of phase purity/structural identity
- **Useful for quantitative analysis**
    - This is by definition a multi-phase refinement!
- **Frequently used to investigate structural changes as a response to external stimuli**
    - Thermal expansion (lattice constants as f(T)), compressibility (lattice constants as f(P), oxygen deficiency (site occupancy as f(P(O_2)) etc.

Calculated powder patterns

- **Steps involved in calculating a powder pattern:**
  - Generate list of reflections
    - From unit cell data and space group
  - Compute F_{hkl}
    - Requires atomic positions from model
  - Generate peak heights from $|F_{hkl}|^2 \cdot$multiplicity
  - Convolute peaks
    - Peak shape function
  - Add background
    - Could be a constant (rare), simple function (most common), or a function that accounts for amorphous materials' contributions
- **Rietveld: Optimize parameters to match experimental data**

Possibilities

- **Works for simple and complicated structures**
  - Thanks to today's computing power, even complicated structures can be refined rather rapidly
- **Can be used to refine several phases as well as mixed occupancies**
  - Use of internal standard possible
  - Quantitative analysis of mixture or versus a standard
  - Non-stoichiometry/partial occupancy can be refined
- **Each point in the pattern can be regarded as an observation**
  - "No Bragg intensity" tells you something about your crystal, too!
- **Refinement of several data sets together**
  - X-ray and neutron data
  - Several different wavelengths => changes scattering contrast between atoms

Limitations

- **Parameters can sometimes be correlated**
  - e.g. zero point and sample height
- **For big structures, constraints or restraints can be necessary**
  - Restrain bond distances or bond angles (define rigid bodies)
  - Constrain composition if known
- **The method only works if you have a good starting model!**
  - Divergence might be observed
  - A local instead of a global minimum may be found

The most important part: Data collection!

- **Without high quality data, even the best program can’t help you!**
- **Choice of instrument**
  - Resolution
  - Accessible angular range
  - “Tricks” you can play
- **Sample preparation (powder methods handout!)**
  - Particle size
  - Surface roughness
  - Homogeneity
  - Preferred orientation
- **Instrument alignment, sample height and other errors**

Choosing a Rietveld program

- **Let's assume that you have access to a good diffractometer, prepared an excellent sample, chose an appropriate setup, and thus collected refinement quality data. So, what's next?**
- **You have to choose a Rietveld program, export your data in the appropriate format, and set up an input file for your refinement**
    - Most diffractometer software gives you multiple export options
    - We will use GSAS-II in this class
- **But how do you decide which Rietveld program you should use? And does it matter?**

First Consideration: Commercial or Freeware?

- **Pros**
  - Freeware:
    - Well, it's free!
    - You can use it anywhere
    - It usually deals well with data from many different sources
  - Commercial:
    - If it came with the instrument you are using, it is probably very well programmed to handle data from that instrument!
    - You usually have access to (paid, professional) customer support & documentation
    - Especially in recent years, some of the codes have become very advanced
- **Cons**
  - Freeware:
    - No professional customer support, you have to often figure things out on your own
    - Code may have limitations, and people may or may not care about your concerns!
  - Commercial:
    - Comes with a price tag
    - If you move, you may no longer have access to a program & need to learn another one

A partial list of Rietveld programs

- **There are quite a few Rietveld programs out there!**
  - GSAS, GSAS-II
  - FullProf
  - Rietan
  - BGMN
  - Highscore Plus
  - PDXL
  - Topas, Topas Academic
  - Jade Plus
- **This is not a comprehensive list – and Rietveld codes are also included in many structure solution codes like EXPO, FOX etc.**

Least squares minimization

- All Rietveld programs try to minimize the following sum:

$$\Sigma\, w_i[y_i - Y(x_i, \mathbf{p})]^2 \text{ where}$$

Data

y_i = observed powder diffraction intensities

Model

$Y(x_i, \mathbf{p})$ = computed diffraction intensities

from F_{hkl}(calc), background model, profile convolution, correction terms…

Parameters

p = parameters

Parameters

- **Structural variables**
  - Atom positions, fractional occupancies, atomic displacement parameters (ADPs)
  - Only these parameters are refinable in most single crystal software
- **Profile parameters**
  - Background
  - Peak shape, including width and asymmetry
  - Unit cell constants
  - Wavelength
  - Diffractometer zero point
  - Sample height and transparency
- **Correction terms**
  - Absorption
  - Extinction
  - Surface roughness
  - Preferred orientation

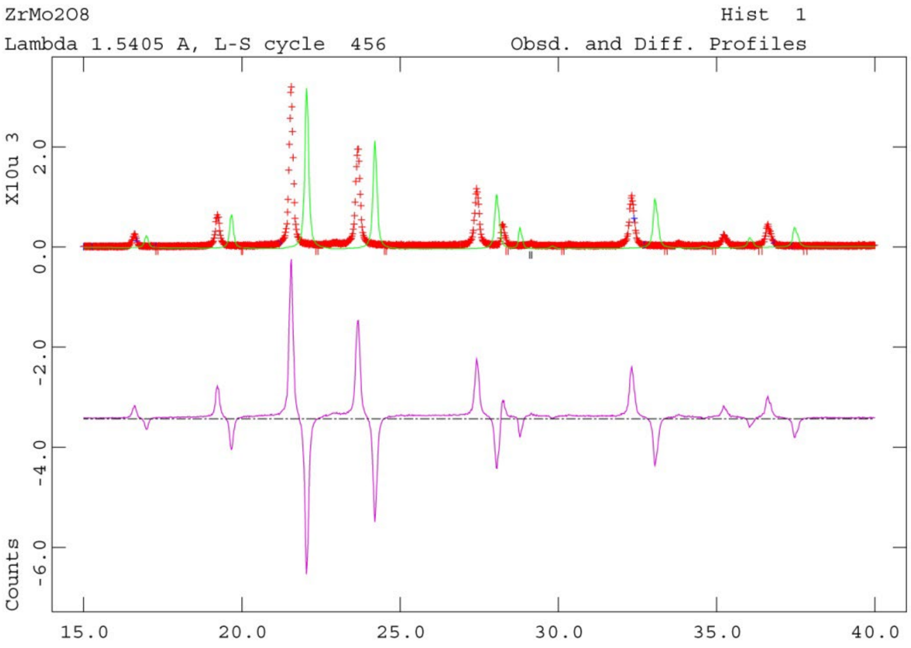
Order of refining parameters

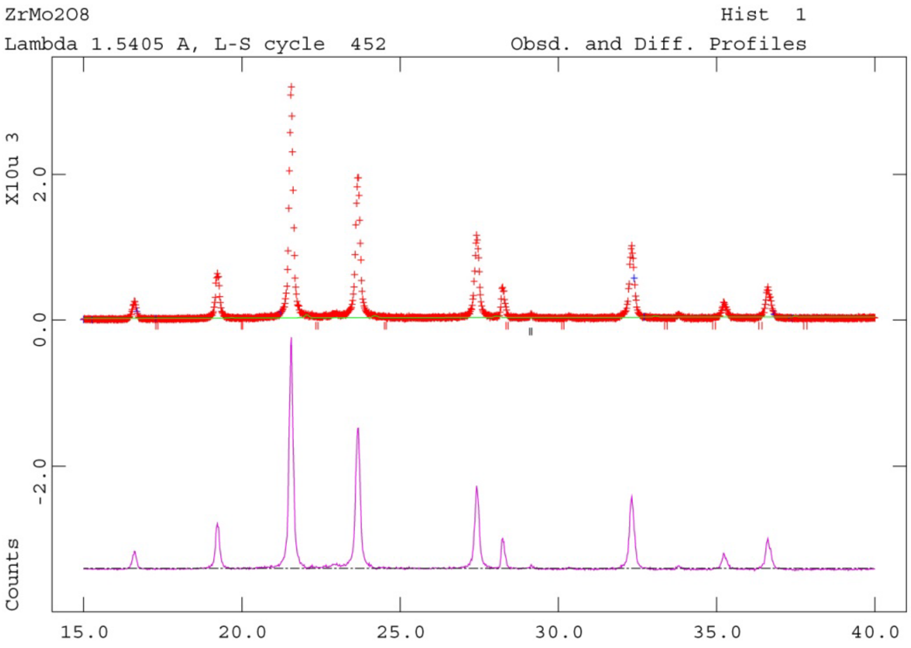
- **In the vast majority of cases, variables have to be added to a Rietveld refinement in a stepwise fashion to avoid divergence**
  - Program may otherwise try to account for some features by changing the wrong variable!
- **There are rules of thumb for the order of turning on variables, however, there is NO “one-fits-all” approach!**
  - Inspect your difference curve to identify what causes the worst discrepancy – this variable should be tackled first!
- **Generally, the structure (atom positions, occupancies, ADPs) are varied last**
- **Make sure background, peak positions, and approximate peak shapes are accounted for before attempting to vary atom positions**
  - Note that none of these are “variables” in single crystal refinements!

First things first – peak positions

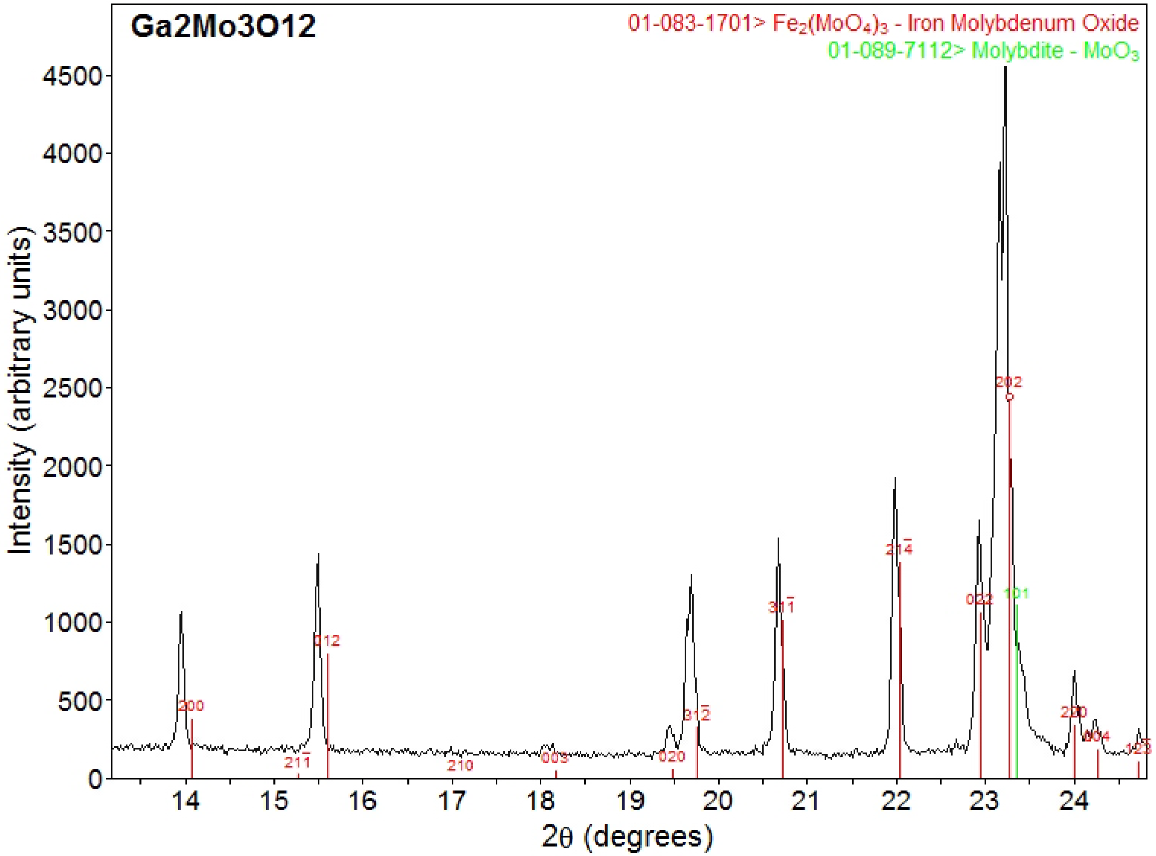
- **Several variables affect peak positions**
  - Wavelength (unlikely culprit for lab instruments)
  - Zero error (usually small)
  - Sample displacement/transparency (usually small)
  - Lattice constants
- **Make sure starting lattice constants are close to what they should be**
  - Often obtained from PDF search/match
  - May require indexing – either from scratch or “manually” (from known structure/indices)
- **Calculated peak positions must be close to observed peaks for a refinement to work!**
  - Otherwise, a “flat line fit” may result

First things first – peak positions





“Manual indexing”



“Manual indexing”

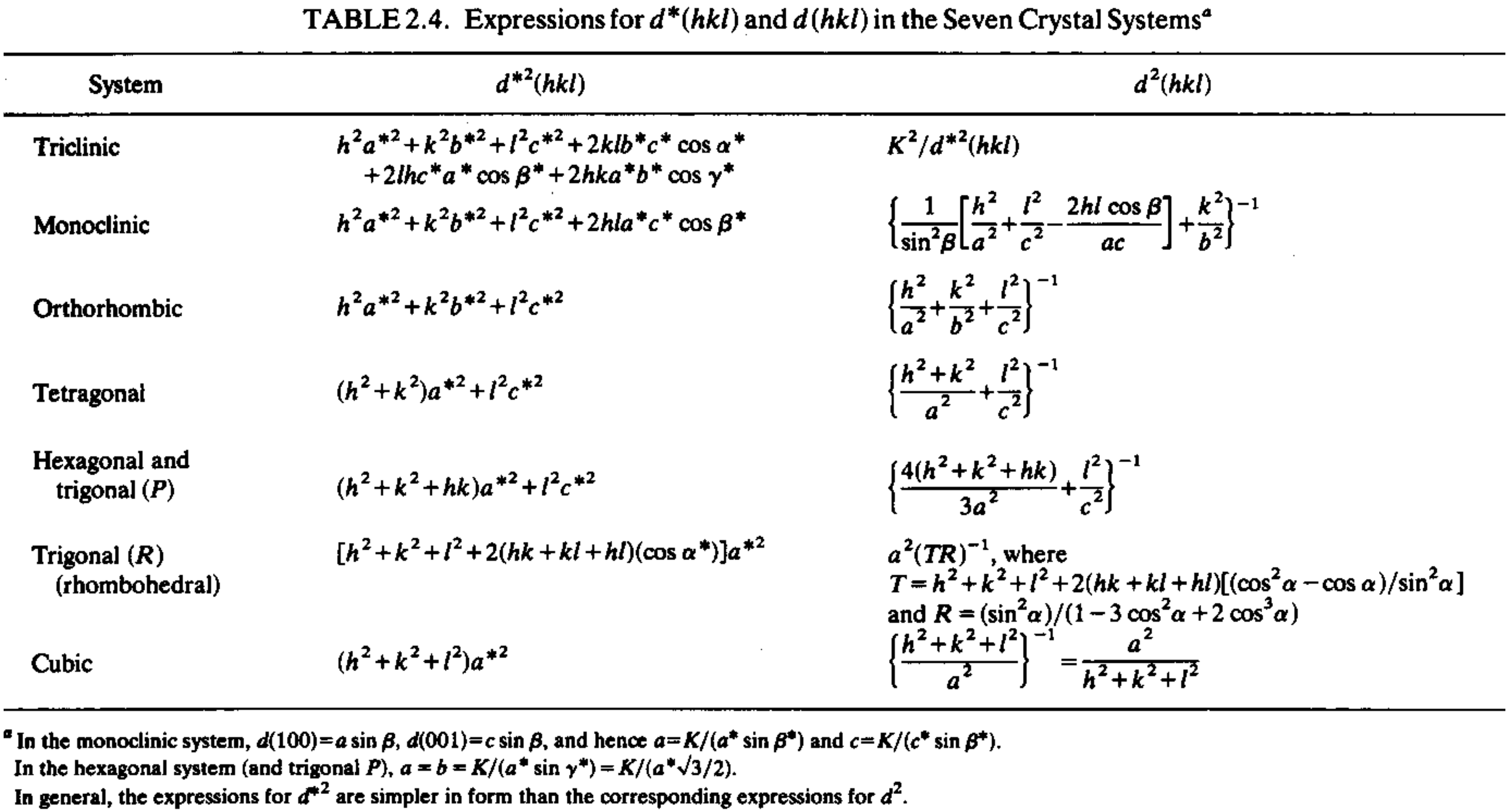


TABLE 2.4. Expressions for $d^*(hkl)$ and $d(hkl)$ in the Seven Crystal Systems[a]

System	$d^{*2}(hkl)$	$d^2(hkl)$
Triclinic	$h^2a^{*2}+k^2b^{*2}+l^2c^{*2}+2klb^*c^*\cos\alpha^*$ $+2lhc^*a^*\cos\beta^*+2hka^*b^*\cos\gamma^*$	$K^2/d^{*2}(hkl)$
Monoclinic	$h^2a^{*2}+k^2b^{*2}+l^2c^{*2}+2hla^*c^*\cos\beta^*$	$\left\{\frac{1}{\sin^2\beta}\left[\frac{h^2}{a^2}+\frac{l^2}{c^2}-\frac{2hl\cos\beta}{ac}\right]+\frac{k^2}{b^2}\right\}^{-1}$
Orthorhombic	$h^2a^{*2}+k^2b^{*2}+l^2c^{*2}$	$\left\{\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right\}^{-1}$
Tetragonal	$(h^2+k^2)a^{*2}+l^2c^{*2}$	$\left\{\frac{h^2+k^2}{a^2}+\frac{l^2}{c^2}\right\}^{-1}$
Hexagonal and trigonal (P)	$(h^2+k^2+hk)a^{*2}+l^2c^{*2}$	$\left\{\frac{4(h^2+k^2+hk)}{3a^2}+\frac{l^2}{c^2}\right\}^{-1}$
Trigonal (R) (rhombohedral)	$[h^2+k^2+l^2+2(hk+kl+hl)(\cos\alpha^*)]a^{*2}$	$a^2(TR)^{-1}$, where $T=h^2+k^2+l^2+2(hk+kl+hl)[(\cos^2\alpha-\cos\alpha)/\sin^2\alpha]$ and $R=(\sin^2\alpha)/(1-3\cos^2\alpha+2\cos^3\alpha)$
Cubic	$(h^2+k^2+l^2)a^{*2}$	$\left\{\frac{h^2+k^2+l^2}{a^2}\right\}^{-1}=\frac{a^2}{h^2+k^2+l^2}$

[a] In the monoclinic system, $d(100)=a\sin\beta$, $d(001)=c\sin\beta$, and hence $a=K/(a^*\sin\beta^*)$ and $c=K/(c^*\sin\beta^*)$.
In the hexagonal system (and trigonal P), $a=b=K/(a^*\sin\gamma^*)=K/(a^*\sqrt{3}/2)$.
In general, the expressions for d^{*2} are simpler in form than the corresponding expressions for d^2.

“Structure Determination by X-ray Crystallography”, Ladd and Palmer, Plenum, 1994.

Errors/parameters that affect peak positions

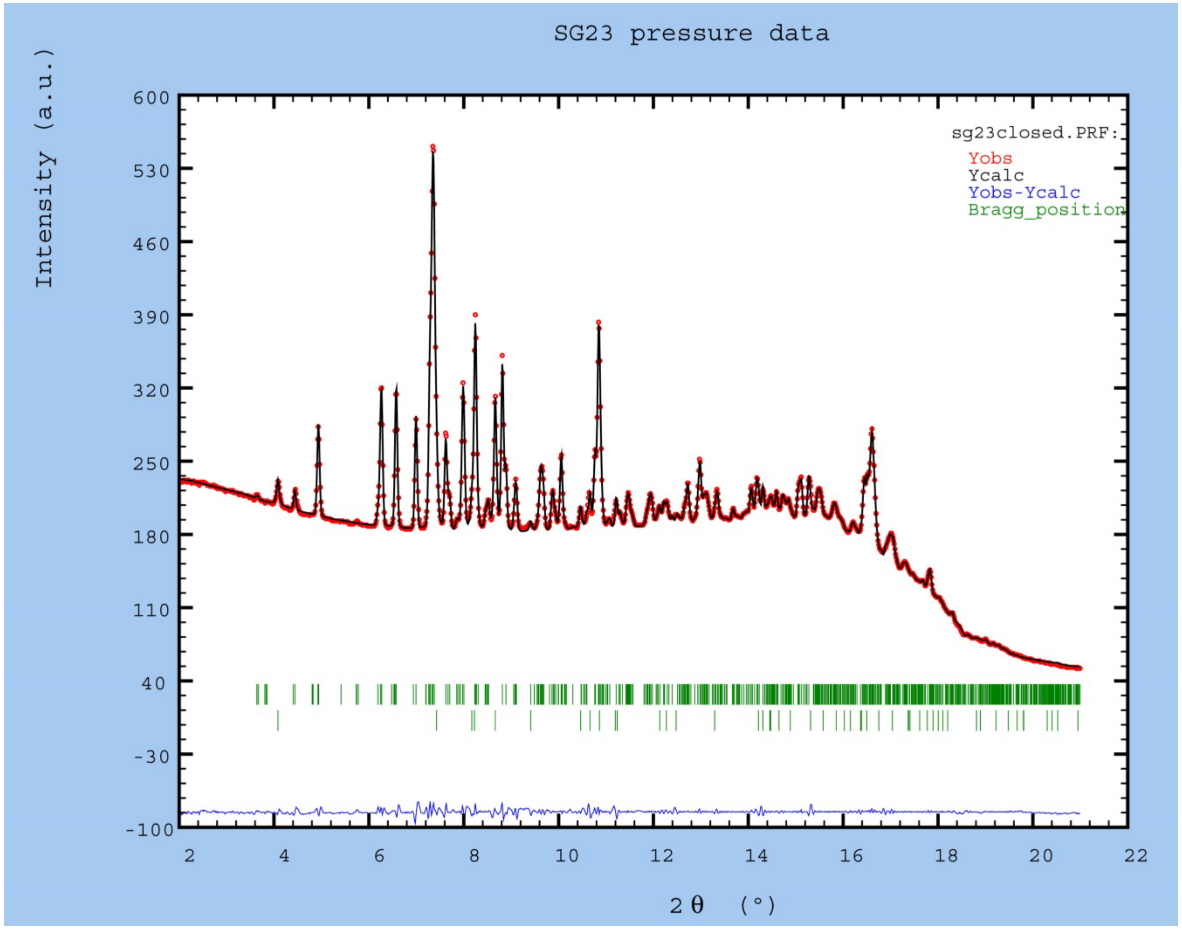
- **Wavelength (fixed for lab instrument!)**
- **Zero error – check with standard**
- **Sample displacement and transparency**
  - Transparency is effectively a "sample displacement" of part of the sample!
  - Common error for Bragg-Brentano setup
  - Do NOT vary for parallel beam setup or capillary experiments!
- **Note that sample displacement and zero error are usually strongly correlated!**
  - Unless you have a large data range, excellent data, and an internal standard, do NOT attempt to refine both simultaneously!
- **Again: YOU decide what should reasonably be varied**

Important early variable: Background

- **Unless you spot a significant other “biggest issue”, make sure your background is reasonable before refining peak shapes or structure!**
- **Background can be determined manually**
  - May be necessary for significant background contributions from complex experimental setups or amorphous sample components
  - Program uses x,y pairs given by user and linear interpolation
- **Background can be calculated using a number of mathematical functions**
  - Available in most codes: Fourier cosine series, radial distribution functions, Chebychev polynomial
- **For a calculated background, the number of background terms can be chosen by user**
  - Usually start with 3 to 6, adjust as necessary – adjustments could be done later as background errors become more obvious

Patterns requiring manual background

- **High pressure data collected inside a diamond anvil cell – scattering from diamonds, pressure transmitting liquid etc.**



Approaches to peak shape models

- **Empirical models**
  - Functions are chosen because they are simple and fit the peaks
  - Parameters cannot be interpreted as there is no relationship to underlying physics
- **Physically-based models**
  - Functions are based on physical phenomena
  - Often found empirically originally, but parameters have meaning
- **Fundamental parameter approach**
  - Functions and even many parameter values are determined based on the underlying physics
  - Only sample property related parameters are refined

Peak shape models

- **Several different peak shape models are available in most Rietveld packages, as the peak shape is determined by**
  - sample characteristics (size, strain)
  - instrument characteristics
- **Convolution of several different factors, but in many cases, one factor will be dominant**
- **Laboratory x-rays will usually give Lorentzian peak shape**
- **Medium resolution CW neutron diffractometers give Gaussian peak shape**
- **Neutron TOF instruments give highly asymmetric peak shapes**
- **In high resolution data, sample characteristics determine the peak shape**
  - can be very challenging to model the peak shape properly

Physics behind peak shapes

Common factors

- Instrument resolution (intrinsic peak width due to instrument)
- Axial divergence (low angle asymmetry)
- Sample position, transparency
- Crystallite size broadening*
- Strain broadening*

Less common factors#

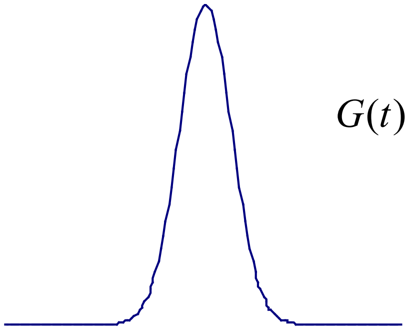
- Stacking faults
- Modulated structures
- Compositional inhomogeneity

* Could be anisotropic (only affects certain classes of reflections then)

Often hard or impossible to model

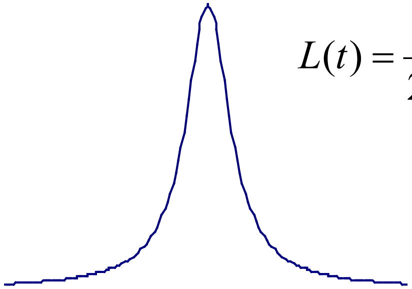
Examples of peak shapes

Gaussian peak shape



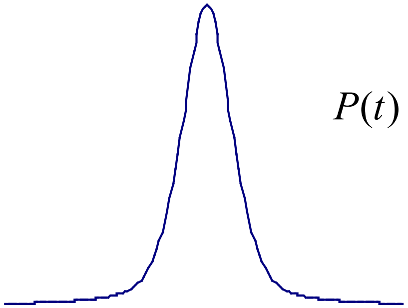
$$G(t)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-t^2}{2\sigma^2}\right)$$

Lorentzian peak shape



$$L(t)=\frac{\Gamma}{2\pi}\frac{1}{(\Gamma/2)^2+t^2}$$

Pseudo-Voigt peak shape



$$P(t)=\eta\cdot L(t,\Gamma)+(1-\eta)\cdot G(t,\sigma)$$

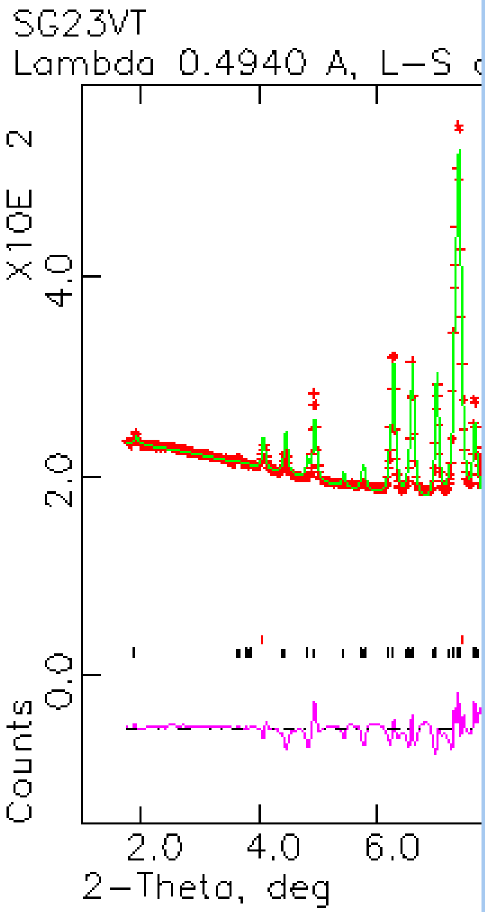
Finally – getting ready to vary the structure!

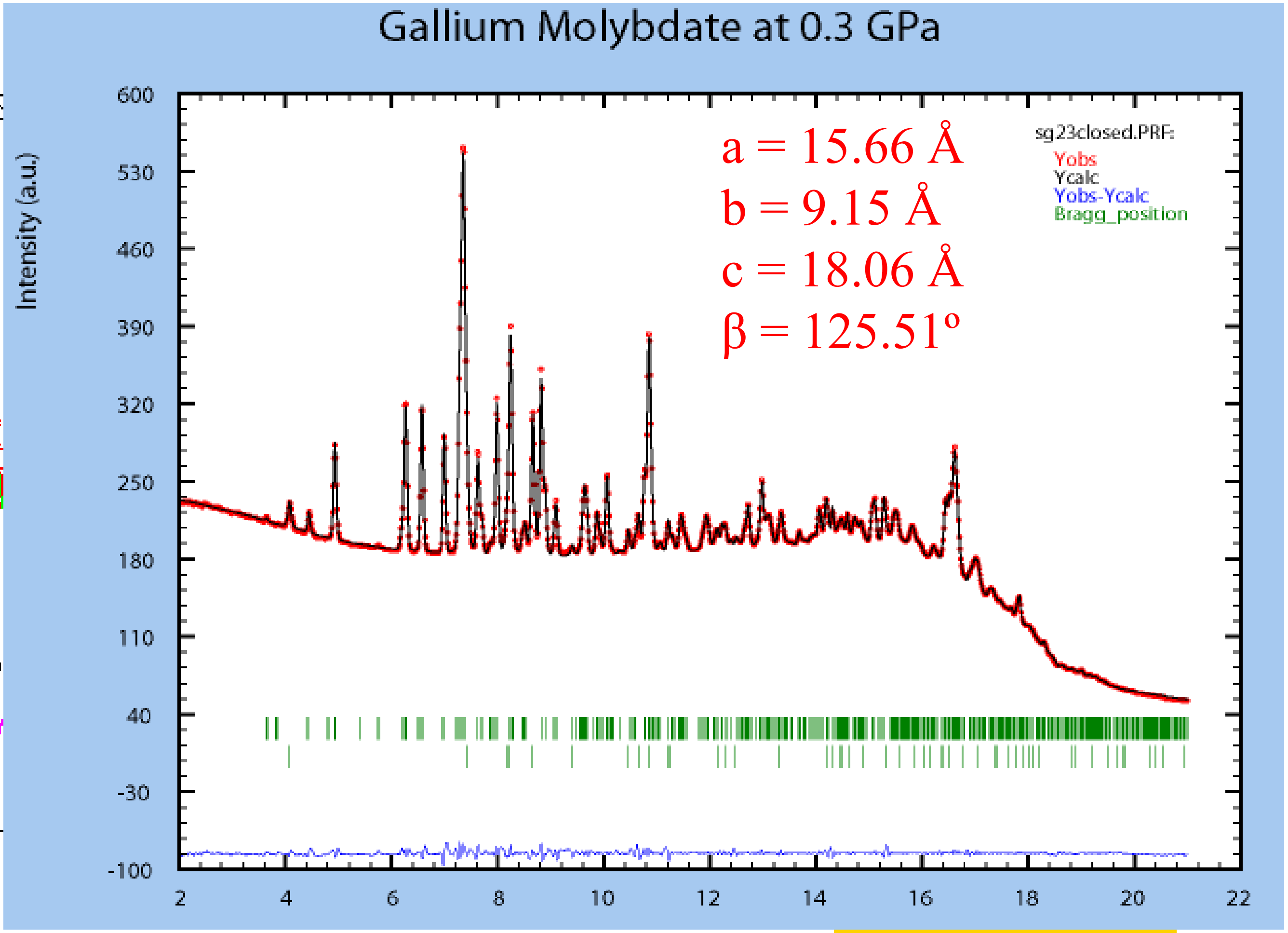
- **Before you start varying your structural model, *<u>closely inspect your refinement/difference curve!</u>***
  - Is background properly fitted? Is the peak shape reasonable?
  - Are all peaks accounted for (tick marks) and in the right positions?
- **Consider your data quality and the complexity of the structure**
  - Can you expect a stable refinement if you independently vary all parameters?
    - For lab X-ray data, the answer is often "no"!
  - Would constraints or restraints make sense?
    - Now is the time to implement them before your structure may fall apart during refinement!
    - Composition, thermal parameters, atomic distances, rigid bodies…
  - Look carefully at your model
    - e.g., refining the positional parameters of a half-occupied site just off an inversion center (due to disorder) will likely cause trouble
    - In some programs, special positions will still vary if you release their positional parameters!

What if…

- **…you did everything you could to fulfill the demands on the previous slides but things still don't look as good as you would like them to**
- **If you start varying "non-standard variables" (correction terms), make a backup of your current input file – could save you A LOT of time!**
  - Depending on program, copy the whole folder. GSAS makes backup copies automatically, so no need to worry about this!
- **If the major problem lies in judging intensity mismatches, consider doing a Le Bail or Pawley refinement**
  - Structureless refinement, intensities are considered variables
  - Can be useful to get a good idea of "best possible fit", too
  - Could show problems due to incorrect space group assignment
    - missing peaks or lots of peak tick marks with zero calculated intensity
  - Also used in SDPD to extract integrated intensities from powder data

Structural fit versus Le Bail refinement





a = 15.66 Å
b = 9.15 Å
c = 18.06 Å
β = 125.51°

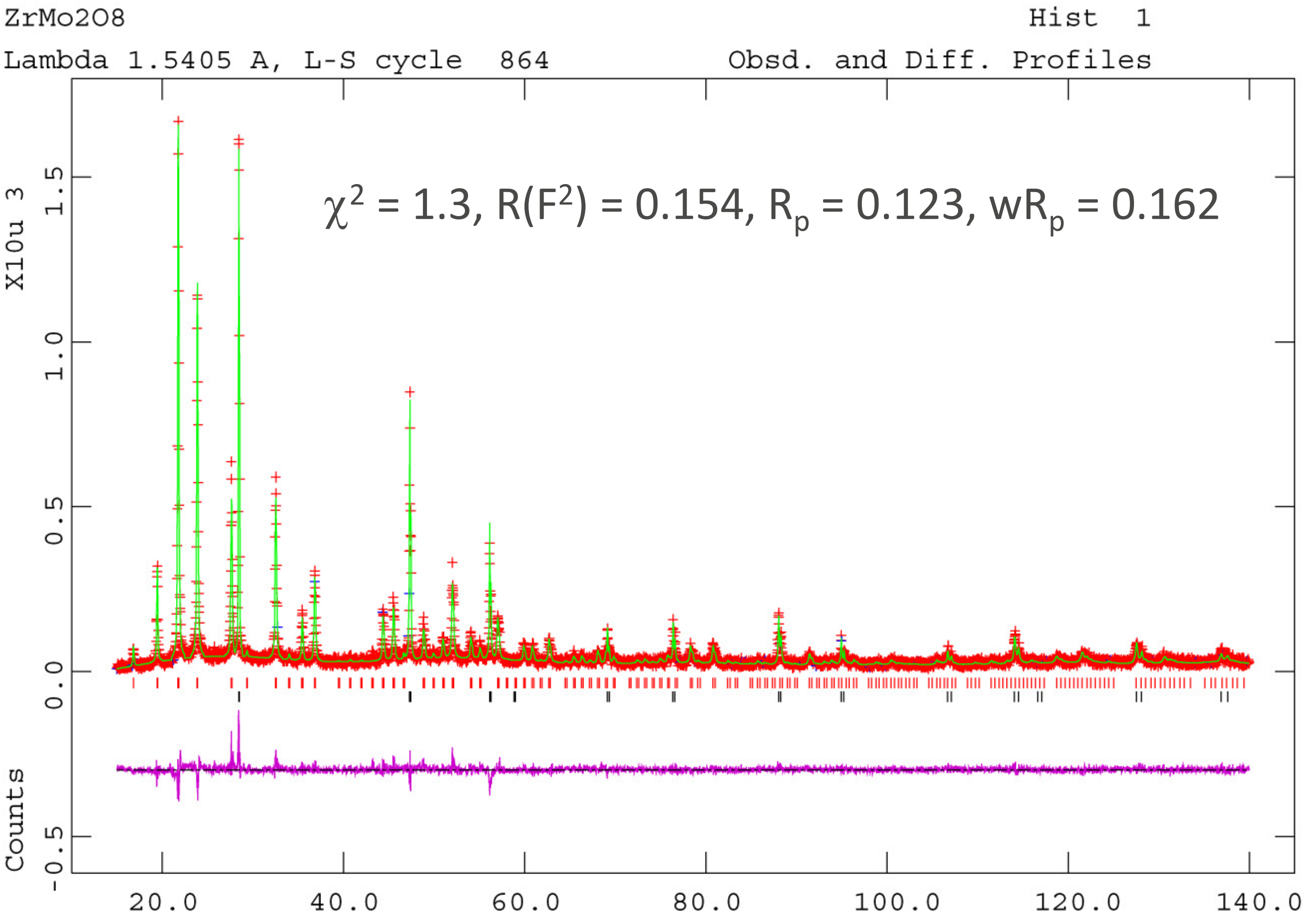
Varying the structure

- **For lab data, *never* attempt to refine anisotropic displacement parameters!**
  - Usually reserved for high quality neutron data
  - Occasionally attempted for synchrotron data
- **For materials with mixed site occupancy, constrain positional parameters and ADPs for the two element sites!**
- **For lab X-ray data, it is quite common to only refine one ADP value per atom type (or possibly one for all atoms in a phase!)**
  - Again, use constraints
- **Damping of variables can help or hinder refinement**
- **Usually start with ADPs, then positions of heavy atoms, then positions of light atoms, occupancies last (if necessary, at all – do not vary if known!)**
  - *Watch* evolution of refinement/statistics!
  - Go back to previous copy if necessary!
  - Restraints on bond distances can help stabilize structures during refinement!

How to judge your refinement

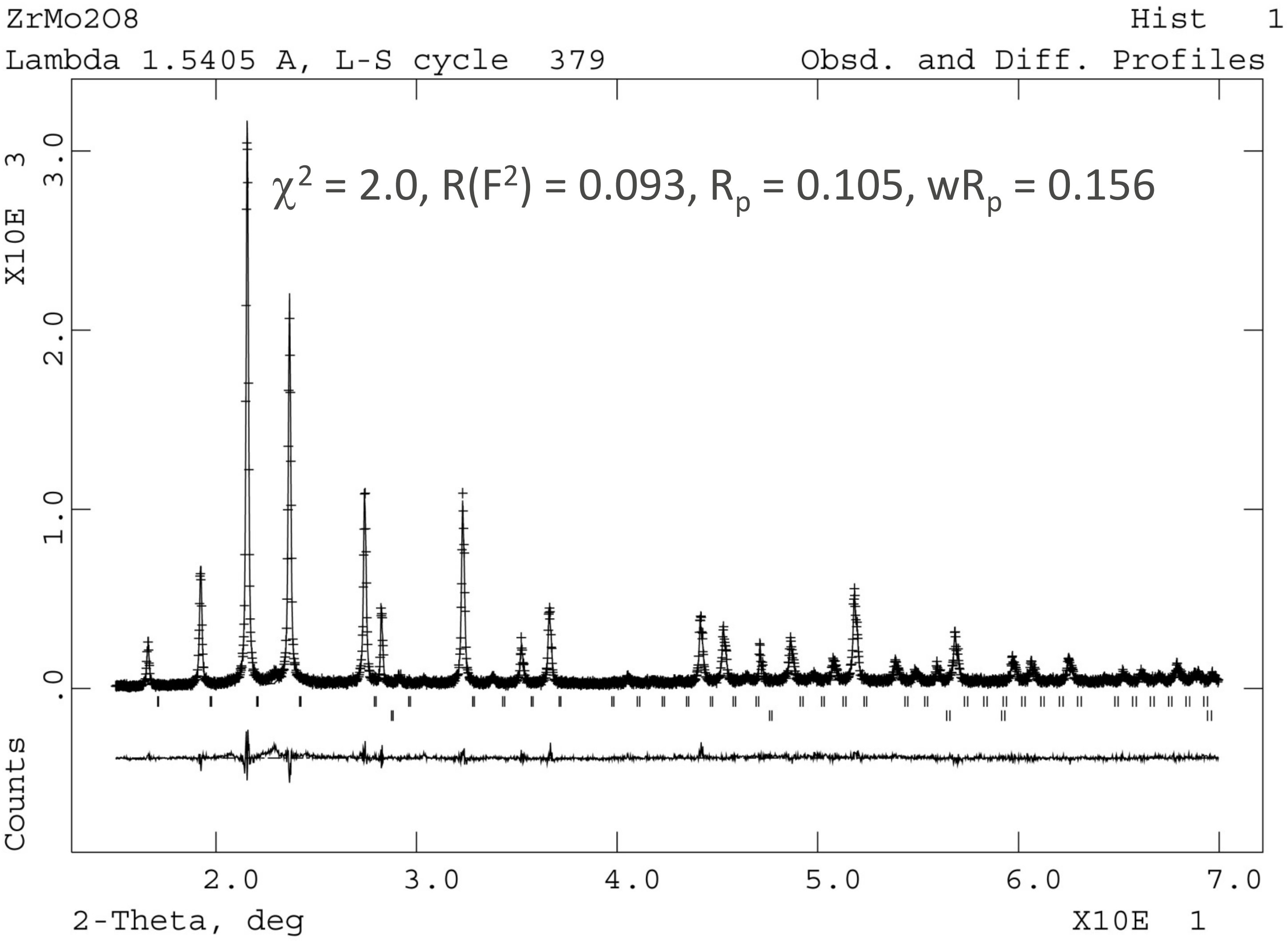
- **Several different indices:**
  - R_{wp} $$\left[\frac{\sum w_i(y_i(obs)-y_i(calc))^2}{\sum w_i(y_i(obs))^2}\right]^{1/2}$$
  - R_w $$\frac{\sum |y_i(obs)-y_i(calc)|}{\sum y_i(obs)}$$
  - R_F $$\frac{\sum |(I_K(obs))^{1/2}-(I_K(calc))^{1/2}|}{\sum (I_K(obs))^{1/2}}$$
  - Goodness of fit Chi^2
- **The easiest way to get low Chi^2 values is to collect noisy data**
- **R_F is not minimized during the refinement, but gives information about the agreement between the structural model and the pattern**
- **The most important judgment is in all cases the visual judgment!**
  - Low indices mean nothing if the fit does not look convincing

A good fit



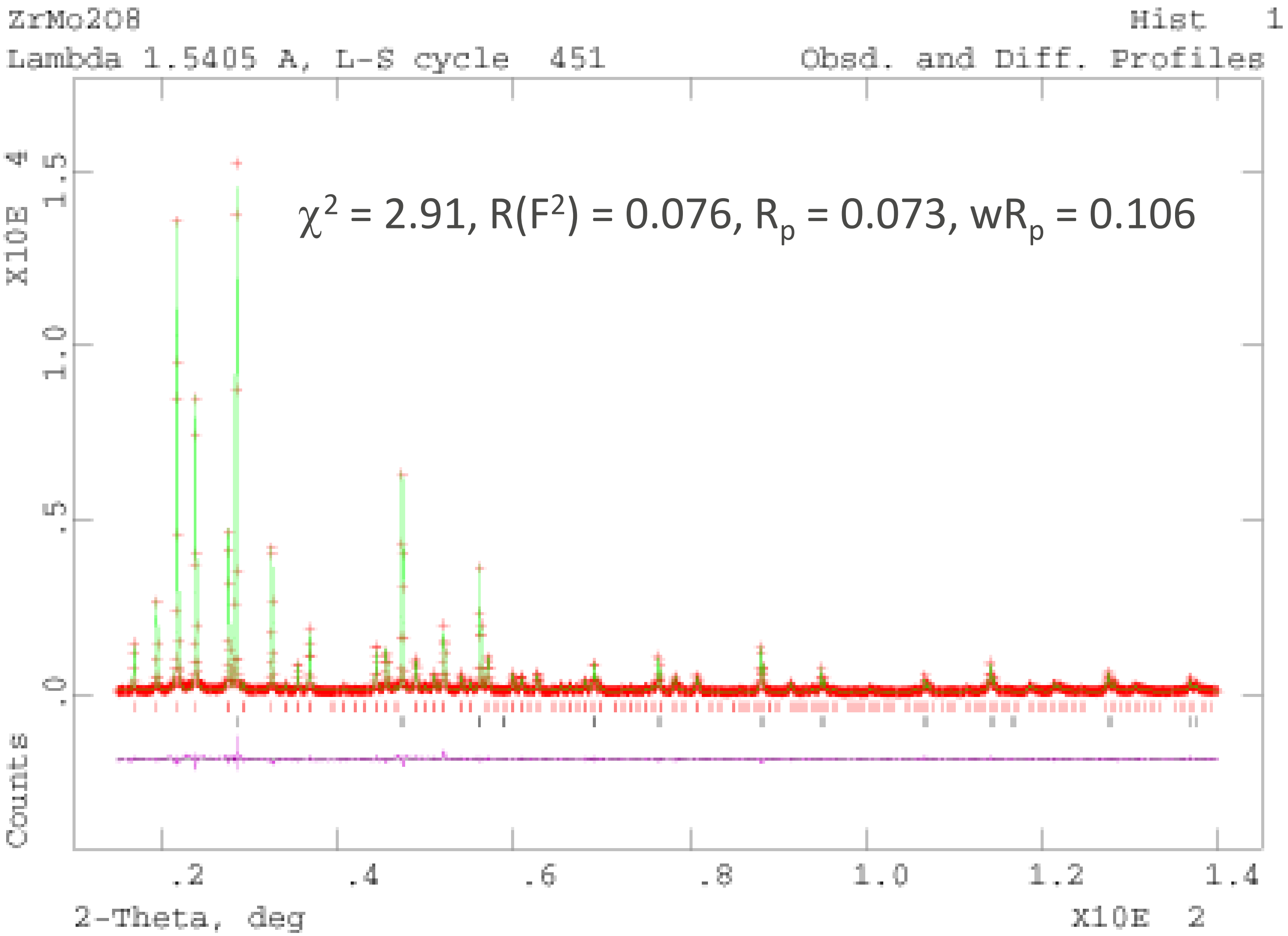
χ^2 = 1.3, $R(F^2)$ = 0.154, R_p = 0.123, wR_p = 0.162

Also a good fit



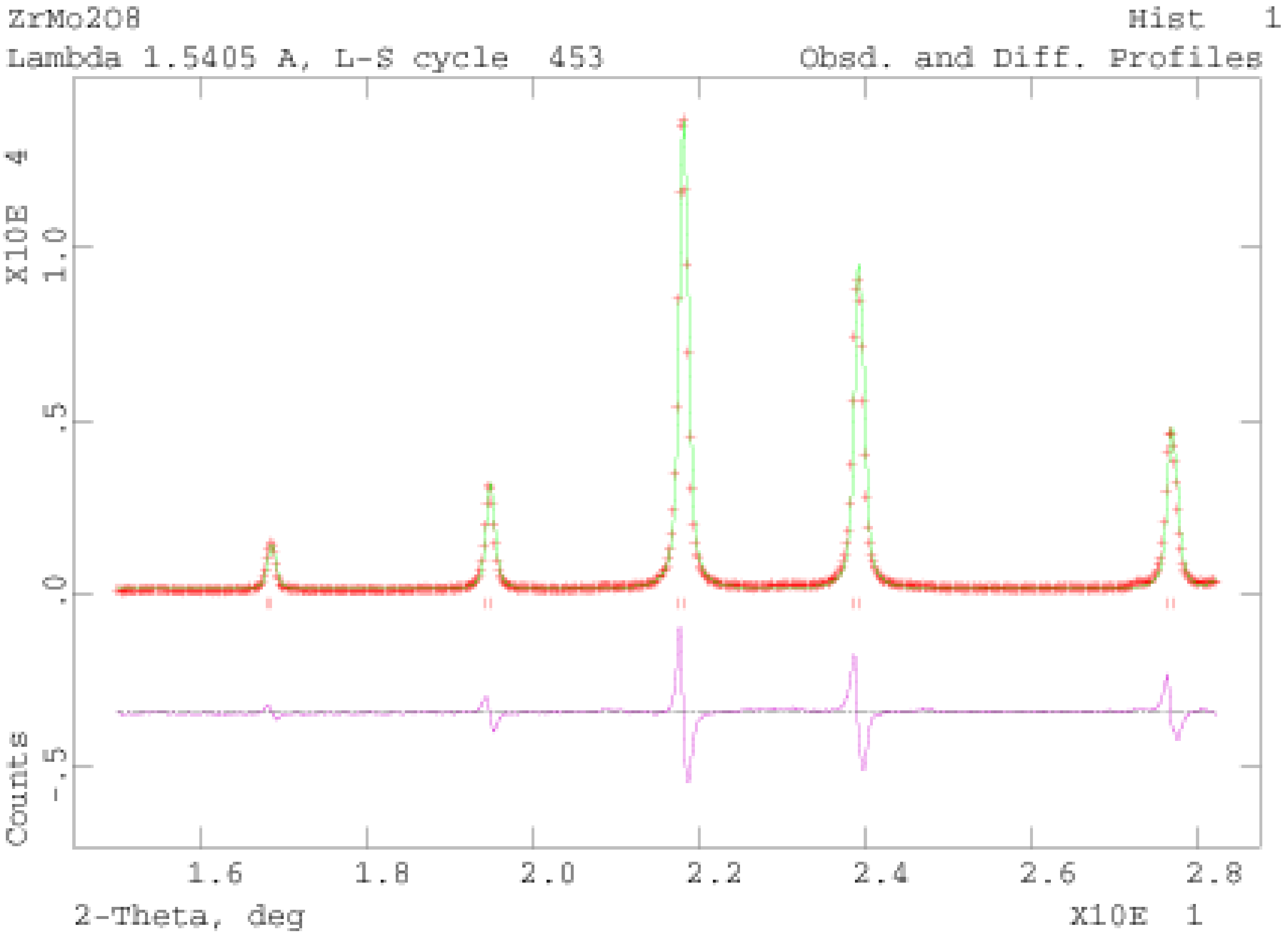
$\chi^2 = 2.0$, $R(F^2) = 0.093$, $R_p = 0.105$, $wR_p = 0.156$

Another good fit



$\chi^2 = 2.91$, $R(F^2) = 0.076$, $R_p = 0.073$, $wR_p = 0.106$

A fit that is clearly not there yet



$\chi^2 = 8.54$, $R(F^2) = 0.115$, $R_p = 0.139$, $wR_p = 0.182$

Peaks are shifted due to incorrect lattice constant

Modern instrumentation

- **High resolution diffractometers and synchrotron data can provide extra information**
  - Better peak resolution (FWHM can be as low as 0.005° in some cases)
  - High peak to background ratios
- **However, there is a price to pay for this:**
  - Peak shape often determined by the sample
  - TOF diffractometers have asymmetric peaks
- **Modeling of high-resolution data can be very demanding**

Refining several patterns simultaneously

- **Refinement of several patterns simultaneously can provide additional information**
- **X-ray and neutron data**
    - X-rays usually give better resolution, but the intensity falls off at small d-spacings
    - Good for resolving small lattice distortions
    - Neutron data provide much higher intensities at small d-spacings
    - More reliable atomic positions and thermal parameters, esp. for light atoms
- **X-ray data of several different wavelengths**
    - Can be used to change the scattering contrast between atoms
    - More reliable refinement of partial occupancies

What to watch out for

- **Some parameters can be highly correlated**
  - Unit cell constants and wavelength
  - Thermal parameters and partial occupancy
  - Sample height and zero point
  - Look at correlation matrix!
- **Check whether your results are physically meaningful!**
  - If you know the wavelength, fix it
  - Does your occupancy agree with your chemical composition?
  - Bond lengths?
- **Use an internal standard if possible**
  - Known lattice constant
  - Constrain sample height, zero-point, asymmetry etc. to be identical for standard and sample

Final steps

- **Closely inspect your fit/difference curve**
- **Any bad correlations?**
    - Look at correlation matrix!
- **Check whether your results are physically meaningful!**
    - If you know the wavelength, fix it
    - Does your occupancy agree with your chemical composition?
        - Make sure you take into account multiplicity of positions!
    - ADPs from laboratory powder data are often not very meaningful, so don't stress too much if that is the only problem
    - But bond distances and packing <u>must</u> be reasonable!

Summary: Some basic “rules of thumb” for Rietveld refinement

- **Collect good data**
- **Your eyes are the most important judgement tool: Inspect the difference curve!**
- **If you refine several phases, constrain changes in parameters like sample height or asymmetry to be identical for all phases before varying them!**
- **If you are using an internal standard, do NOT vary its lattice constants!**
- **Suggested order for turning on parameters unless you spot a “biggest culprit”:**
  - automatically varied in GSAS: 3 background parameters, histogram scale
  - if necessary, insert more background parameters
  - lattice constants (except for standard; make sure the starting point isn’t too far off!)
  - profile parameters LX, LY, SHFT, ASYM; if necessary, also GW, TRNS
  - ADPs (keep isotropic, possibly constrain), atom positions
  - preferred orientation, diffractometer constants, absorption, extinction etc. should only be varied if absolutely necessary!

Useful resources

- CCP14: Free software including tutorials and examples http://www.ccp14.ac.uk/
  - Unfortunately, no longer maintained due to lack of funding
- Rietveld mailing list http://crystsun1.unige.ch/stxnews/riet/welcome.htm
- GSAS-II download https://subversion.xray.aps.anl.gov/trac/pyGSAS
- R. A. Young; "The Rietveld method
  - Comprehensive text including history, description of several Rietveld programs, as well as details about certain parameters (e.g., background modeling, peak shapes, pattern decomposition…)