

From beam time to structure factors

Including a concise guide to ANA and AVE

Elias Vlieg

Radboud University Nijmegen, Institute for Molecules and Materials, The Netherlands

e.vlieg@science.ru.nl

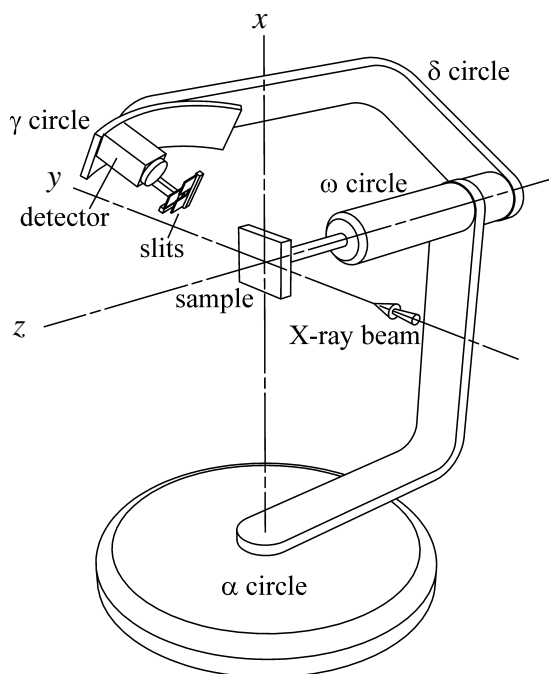
1	Introduction	2
2	Diffractometers.....	2
3	Sample alignment.....	3
4	Optimisation.....	4
5	Scan types.....	4
6	Collecting scans.....	5
7	Keeping track	6
8	ANA	6
8.1	Data input.....	7
8.2	Data adjustments	8
8.3	Scan integration	8
8.4	Peak fitting.....	10
8.5	Autoplot	10
8.6	Loop command	10
8.7	Summary of peak integration.....	11
9	AVE.....	11
9.1	Input	11
9.2	Symmetry-equivalent domains	12
9.3	Average	12
9.4	Weighting.....	13
9.5	Non-equal domain occupancy.....	14
9.6	L-merging of data	15
9.7	Other data files.....	15
10	Making the final data set	15
	References	16
	Appendix A: Geometrical correction factors	17
	Appendix B: Transforming between unit cells	24
	Appendix C: Line profiles in ANA	26
	Appendix D: Upper limit for structure factor value.....	27
	Appendix E: Required counting times	28
	Appendix F: Typical parameter settings in ANA.....	29
	Appendix G: ANA commands	31
	Appendix H: AVE commands.....	34

1 Introduction

This document contains some suggestions on how to perform a surface X-ray diffraction experiment, starting from the hardware and ending with the list of structure factors. The aim is to get a crystallography data set that is as accurate as possible in the least possible time. For a general description of the technique, several reviews are available [1, 2].[3]

2 Diffractometers

Diffractometers for surface X-ray diffraction come in many types with equally many different names: z-axis (see figure) [4], 5-circle [5] , 6-circle [6], 2+2-circles [7] and 2+3-circles [8], to name only the most popular. They come in both horizontal and vertical scattering geometries. The terminology horizontal and vertical can be confusing when discussing both scattering geometries, and therefore it can be convenient to use as reference a frame with respect to the surface. We will use the frame as shown in the figure, with the surface in the xy plane and z along the surface normal. The X-ray beam is along the y -axis (for all angles at zero). For the xy plane also frequently the terms in-plane and lateral direction will be used; similarly the z -direction will also be called the perpendicular or out-of-plane direction.



A z-axis diffractometer showing the definition of the x , y and z -axes with respect to the sample surface.

An accurate data set is only possible when the diffractometer is carefully aligned. In normal circumstances, the machine is in an aligned state, but it is often a good idea to check the alignment of the main circles. This will take a few hours, but finding an alignment error after several days of data acquisition is worse. Things that need to be checked are:

- Does beam go through the centre of the diffractometer?
- Are the zero's of the circles properly set?
- Are the detector slits and the beamline slits properly centred?

Every error may mean that scans are not properly centred (particularly at high l -values), which means that a lot of time has to be spend on offsetting the scan range. Alignment errors also lead to intensity differences between equivalent reflections.

Sometimes the incoming beam is not horizontal (due to a reflection from a mirror for example). The diffractometer should have the same tilt as the beam, but this is not always taken into account and will then be a source of errors.

Most of the time, the detector arm will have two sets of slits: one set close to the sample (slit1) and one in front of the detector (slit2). In the in-plane direction, slit1 en slit2 together define the active sample area. The value of slit1 is the most important one for the active area, while slit2 is more important for the angular acceptance. In order to minimise the effects of the coupling between real and angular space, it is a good idea to have slit2 slightly larger than slit1 [9]. Typical values are 2 mm for slit1, and 3 mm for slit2. In the out-of-plane direction (l -direction), only slit2 is important. Slit1 (if implemented) should be kept wide open in that direction [10].

3 Sample alignment

The alignment of the sample consists of two steps: physical surface and crystallographic (orientation matrix U). First the surface is aligned, typically with a laser. The surface normal should be aligned along the axis of the diffractometer. Two degrees of freedom are necessary for this. The implementation varies: cradle and circle, two cradles or hexapod. In the following, we will call the rotation axis of the sample ω . The laser alignment consists of the following steps:

- Reflect laser off the sample and mark the spot (e.g. piece of paper on the wall)
- Rotate ω by 180° and mark laser spot again (the laser spot describes an ellipse)
- The point exactly in between the two marked spots is the centre. Mark this point.
- Bring the laser spot on the centre point using the available degrees of freedom.

Sometimes a second iteration is necessary. Instead of using a laser, the same alignment can be done with the actual X-ray beam using total reflection. Due to the small angle of incidence, this is only accurate for the alignment angle in the scattering plane. If the sample is known to have a small miscut, one can skip the alignment of the optical surface and use a specular reflection to align the surface normal.

Since most surface science samples have a well defined crystallographic face with a small miscut, the laser alignment also largely determines the crystallographic alignment, except for the azimuthal angle. In the crystallographic alignment often only two reflections are used: the primary and the secondary (more sophisticated schemes to find the orientation

matrix are possible). The alignment starts with choosing an appropriate, out-of-plane bulk reflection. A reflection with several symmetry equivalent ones is easiest, because then there are many opportunities of finding one member of this set. The detector angles have to be set for this specific reflection. A rotation of ω is all that is required (hopefully!) to find the bulk reflection. Typical steps:

- With slits open and no attenuators, rotate ω to find the first reflection (use rate meter or oscilloscope for this).
- Close out-of-plane slit2 to ~ 0.5 mm, to avoid errors in this angle.
- Optimise all (non-degenerate) angles.
- Store the angle values as one of the orientation reflections.

Do the same thing for a second reflection and done!

When the laser and crystallographic alignments are done on a stable sample (e.g. no annealing in between), the miscut can be calculated. If this is done accurately, then after a new crystallographic alignment (done, for example, because the sample moved), no new laser alignment is required, because the alignment values can be derived from U and the miscut. This option, however, may not be implemented in the available software.

4 Optimisation

Once the sample is aligned, surface sensitive reflections can be measured. On a "representative" one, optimise the parameters that influence the signal and the signal-to-background ratio, like the detector slits (in particular slit2), the size of the incoming beam and the incoming or outgoing angle. The in-plane acceptance of slit2 should be sufficiently large in order to avoid systematic errors as a function of the scattering angle or between integer and fractional-order reflections [9]. For samples with broad peaks, this could mean that its value has to be more than the typical 3 mm.

A surface reflection can also be used to put the sample in the middle of the beam by doing a scan of its position while sitting on top of the reflection. The surface reflection has little chance of being contaminated with harmonics and is thus a save way to set the discriminator levels of the detector.

5 Scan types

In order to have a low background from the bulk, typically a low incoming or outgoing angle is used. Most of the time, the incoming angle is kept low, but for better line profiles (and possibly even better signal-to-background ratios) also a small exit angle can be considered [9].

The most common type of scan is a rocking scan in which the detector is fixed and the sample is rotated. Rather than rocking an angle, one can also scan in reciprocal space (diffraction index scan), but this is rarely done, since it requires more thinking about the scan direction.

Specular reflectivity scans can also be part of a data set. There are two ways to obtain reflectivity data. The first one is with a rocking scan of the surface normal within the scattering plane. Depending on the diffractometer, this can be done using an alignment cradle, the angle of incidence, or another appropriate degree of freedom.

The second type of reflectivity scan is a ridge scan. The detector and sample are rotated simultaneously such that the specular reflectivity condition is continuously maintained. One stays “on top of the peak”. In order to find the background in this case, the scan needs to be repeated with an offset on either side of the reflectivity profile. For a well-defined surface, sufficient detector acceptance and not too small angles, this type of scan gives reliable data [9]. If this is not the case and also in case of small angles, rocking scans are the only option. The ridge scan is attractive, because it quickly gives data with a fine spacing of l -values.

An alternative for angular rocking scans is the use of an area detector. In this case an entire profile is collected at once. This is particularly suitable for out-of-plane data measured with a small incoming angle [9]. This stationary mode is becoming popular because of the availability of pixel detectors. The reflectivity ridge scans mentioned above are in fact an example of such a stationary data collection mode.

6 Collecting scans

A crystallographic data set consists of "as many reflections as possible". A few different CTR's are often sufficient to get accurate z -coordinates. Measuring many full rods is time consuming, while more information is gained by obtaining different in-plane reflections. These are, after all, different Fourier components, while extra rods repeat the same l -range. It is important to measure up to the highest l -value possible.

In order to get as much data as possible in the shortest possible time, consider the following points:

- The accuracy of the data is mostly determined by systematic errors, rather than counting statistics. When the integrated counts in a peak are about 1000, the counting error is only 3% (if the background is low, otherwise more counts are needed). Most data sets have systematic errors that are larger than 3%, so then there is no need to accumulate more than 1000 counts.
- If a CTR profile is smooth, the Δl interval can be taken quite large. No real information is contained in the intermediate points.
- A profile measured using a point detector can already be accurately integrated if the peak contains something like 7-10 points. Extra points make a nicer plot, but don't determine the error in the data set.

- Scans using a point detector should be sufficiently wide to allow the determination of the background. Time can be saved by using scan types that have a wider step spacing in the background region (e.g. 'xdscan' in SPEC).

Make sure that enough symmetry-equivalent reflections are measured. This is the best way to get a reliable estimate of the error. Doing a full set of equivalent reflections for a few cases can sometimes identify bad scan regions (ranges in ω), due to alignment errors or sample inhomogeneties. Such regions can then be avoided in subsequent scanning.

7 Keeping track

In order to get a data set of good quality, it's important to keep track of the data. *In situ* analysis is required, because then changes as a function of time, misalignments etc. can be identified and (hopefully) fixed. Two computer programs are helpful in this respect: ANA and AVE. ANA (ANALysis) can be used to integrate scans and to convert these into structure factors, while AVE (AVERage) can sort and average data, determine agreement factors and produce a data file for the program ROD for a structure analysis [11]. Comparing a preliminary rod profile with a calculation done in ROD may show that a sample is still too rough and needs further preparation.

All the programs use a command-line interpreter for their interactive use. They contain various commands and submenus (for ANA and AVE these are listed in Appendices B and C, respectively). Typing 'help' or '?' will list a menu or give additional information in many cases. In the list of available commands in a menu, only the letters shown in capitals need to be provided in order for a command to be recognised. Capitals and small case are treated the same. Often a command requires parameter values and it will prompt for a value after an <Enter> is given. The <Enter> is not necessary after each command or prompt: many commands can be typed on a single line. After typing <Enter>, the entire line will be executed. The relevant prompt will only appear when the command line does not contain the necessary item or when the line contains a mistake.

8 ANA

When starting ANA in a new directory for the first time, a number of parameter values are given default values. When quitting the program, the most relevant parameter values are saved in a file called ANA_INIT.MAC. This is a macro file, containing a list of commands available within the program, in the same way that the commands can also be typed interactively. When restarting ANA, this macro file will be automatically executed, so that the parameters resume the values they had in the previous session. For plotting, the parameters are saved in a similar way in the file PLOTINIT.MAC.

In this section the use of ANA for the integration of surface diffraction data will be briefly described. There are more options available, whose actions can be found by listing the various menus (see Appendix B).

ANA has a buffer for (normally) 5 spectra. Data can be read into such a spectrum and many operations on the data are possible. Typically the result of such an operation will be put into another spectrum. For the data analysis discussed here, ANA has to be in the 'datarun' mode, which is done by typing: ANA>SET GENERAL DATARUN YES.

8.1 Data input

First a scan has to be read into ANA. This is done by the command READ in the main menu. This prompts for a file type (like SPEC), a scan number and the spectrum number into which the scan is to be stored. The actual input from the file uses many defaults that are set in the menu ANA.SET.READ>. In that menu one needs to specify a number of parameters: column number of the x-values (typically 1), the normalisation value for monitor counts (typically 1e5), the column name of the y-value (typically "detector") and the file name (or beginning of the file name, depending on the data acquisition program used). Most parameters are explained in this submenu. When calculating structure factors, the incoming data should be normalised to a standard number of monitor counts. Thus both a value for the normalisation value and a value for the normalisation column have to be given. For data files of the type SPEC or SRS, ANA will automatically determine the correct column number for the detector and the monitor by finding the corresponding name.

A special case occurs when the xcolumn parameter is set to a negative value. Then the READ command will prompt each time for the column name.

After the read command, the data will be shown with some brief information in the specified spectrum. Also several parameters, like diffraction indices and various angle settings, will be extracted from the data file. ANA assumes that the detector column contains counts, and calculates the error simply as the square root of this total number of counts. In files that contain both normal detector counts and detector counts corrected for filter attenuation, ANA will use the relative error of the uncorrected detector counts.

There are a few special 'tricks' included in ANA that try to fix problems encountered with some set-ups. In particular there are:

- a FIX_L_INDEX mode that calculates the proper diffraction index l from the detector angle in case this is not correctly given in the SPEC file. This situation occurs for some Z-axis diffractometers in which the scan range is extended using a cradle and in which this is not included in the angle calculations. In this case also the corresponding value for $b_3\lambda$ needs to be provided, because l is calculated from

$$l = \frac{4\pi \sin \theta}{b_3\lambda},$$

with θ the angle of incidence (which equals the exit angle).

- a SCAN_REPAIR mode that uses the actual scan command to deduce the 'real' values during a scan. This is intended to repair the erroneous values that occur in DUBBLE spec files for small steps sizes due to communication problems with the diffractometer.

8.2 Data adjustments

Most of the time, the data as read in from the file will be suitable for immediate integration. Sometimes, however, adjustments need to be made, e.g. removing points from an aborted scan, deleting spikes or merging with other data. The ANA.OPERATE> submenu has many commands that allow such operations. E.g. CUT to remove data points below/above a specified value; DELETE to remove a specified range of points and LEVEL to straighten a sloping background.

8.3 Scan integration

There are basically three ways to integrate a scan:

- *Fully numerical.* The background is calculated from the intensity at the left and right in the scan and the peak is integrated numerically. This is quick and easy and thus convenient for the *in situ* data analysis. When scans are not well centred, or a singular high value is located in the designated background range, errors may occur.
- *Curve fitting.* The peak is fitted to a curve and the analytical integral from the fit is taken as integrated intensity. This works fine as long as the peak indeed has an analytical line shape (like Gaussian or Lorentzian). This is, however, often not the case.
- *Mixture.* A curve fit is used to estimate the background, and subsequently a numerical integration of the peak is done. This method is the most robust in practise, since it gives a fairly precise value for the background, while non-ideal line shapes are correctly integrated.

A spectrum is integrated by the command ANA.OPERATE>INTEGRATE. The precise action of this command depends on the parameter values in ANA.SET.INT> and in ANA.SET.COR>.

In ANA.SET.INT> a flag can be set for automatic (numerical) background estimation, using a specified number of points at the start and end of a spectrum. If automatic is off, the command will prompt for a background and background error, where it will use as defaults the values from the most recent fit. For convenience, typing a negative value is the same as accepting the default. In this case, the full command to integrate spectrum 1 with a predetermined background from a fit is: ANA>OP INT 1 -1 -1. The output of this command can be written to a file (with default extension ".int"), by specifying this in ANA.SET.INT>. The submenu also contains an integration range parameter. If a scan has a wide range of background and a tiny peak, it is better to ignore in the numerical integration the part of this background far away from the peak, because it only leads to a high error bar. When a scan has been fitted, the RANGE parameter specifies the range (as a multiplier of the fitted fwhm) over which the numerical integration will be performed.

When converting the integrated intensity into a structure factor, various correction factors need to be applied [8, 9]: Lorentz factor, polarisation factor, etc. The corrections

require the angle settings of the diffractometer, which are automatically extracted when a spectrum is read in. A number of other parameters is also required, and these need to be specified in SET.COR>. Most of these are explained in that submenu. The most important one is the type of scan used. For a normal angular rocking scan this is an RSCAN (default), for the rare case of a scan using diffraction indices one specifies HKLSCAN. In the latter case, no Lorentz factor applies. A reflectivity measurement using angular rocking scans is specified as REFROCK.

All rocking scans are normalised to a step size of 0.01° . The effective rotation speed of such a scan is thus 0.01° per the specified monitor normalisation. Taking for the latter quantity the total number equivalent to 1 second thus gives an integrated intensity corresponding to a rotation speed of $0.01^\circ/\text{sec}$.

The type of diffractometer used also needs to be specified, because the correction factors vary from type to type (except for the REFROCK scan). ANA assumes a vertical scattering geometry. In case a diffractometer with a horizontal scattering plane ("vertical axis") is used, the meaning of horizontal and vertical simply needs to be inverted. For example, if the horizontal polarisation of the X-ray beam is 0.95, then for a diffractometer with a horizontal scattering plane, one should set this parameter to $1-0.95=0.05$.

For glancing incoming angles and not too small scattering angles, there is no need to worry about inhomogeneities in the profile of the X-ray beam or about the finite size of the sample. When this is not the case, however, the beam profile correction flag has to be set. The beam profile in this case is determined by four parameters: vertical and horizontal Gaussian widths (fwhm) and vertical and horizontal slit size, which are essentially cut-offs of the Gaussians. This allows one to describe Gaussian profiles, rectangular profiles and mixtures of the two. Note that the slit size is meant to describe the profile at the sample position, and does not necessarily equal the setting of the pre-sample slits in the beam line. The size of the in-plane detector slits (slit1) is also required, because this determines which fraction of the sample is observable. The finite sample size correction is currently only implemented in its simplest form: a circular sample.

A reflectivity ridge scan is converted to structure factors in a different way. In this case, one has to read in the peak and the two background scans in ANA in separate spectra. Using the ANA.OPERATE> submenu, the background can be subtracted. As an example, the command sequence, assuming the three spectra are in spectrum number 1, 2 and 3, is:

```
ANA>OP SUM 2 3 4 OP MUL -0.5 4 4 OP SUM 1 4 4.
```

This will put the background-subtracted data in spectrum 4. The conversion of this to structure factors requires the command ANA>OP REFTOF (REFlectivity TO F). This will apply the appropriate correction factors (from ANA.SET.COR>) and write the output to the *.int file (if specified). Note that for this operation it is assumed that the x-values in the spectrum represent the angle of incidence.

The correction factors for the various diffractometer types are summarised in the appendix.

8.4 Peak fitting

ANA has various line shapes that can be used to fit a curve using the command ANA>FIT. Typical curves are Lorentzian (for fractional-order reflections) or Gaussian (integer-order). The command prompts for the line shape, the input and the output spectrum. Next one enters the ANA.FIT.PAR> submenu, in which starting values can be specified, together with parameter ranges (if required) and flags to specify whether a particular parameter is included in the fit or not. For convenience, the command GUESS can be used to let the program calculate starting values for all parameters. The command RUN will start the actual least-squares fitting. The result is displayed on the screen, while the fitted curve is stored in the specified spectrum number. If desired, the output of a fit can be automatically written to a file, by specifying this in the ANA.SET.FIT> menu.

For the case we are discussing here, getting structure factors, the fitting is mainly used to get an accurate estimate of the background. For clear peaks there is never a problem, but when the signal-to-background ratio is more pathological, it may happen that during the fitting procedure the curve width explodes to a large value in order to accommodate a curving background. Such a 'fit' leads to a completely wrong estimate of the background, and therefore it is better to restrict the maximum value of the curve width by using the command UPPER in ANA.FIT.PAR>. Gaussian curves are less prone to exploding widths than Lorentzians, so Gaussians are often the best choice. The fit of the peak itself may not always be good, but the background value is the only thing that matters.

An example of a full command line for a fit is: ANA>F(it) G(auss) 1 2 G(uess) RU(n).

8.5 Autoplot

The command ANA>PLOT can be used to plot a particular spectrum. After specifying the spectrum, one enters the PLOT> menu where many options are available. For convenience, ANA also has an autoplot command in which a series of common plotting actions are combined. The most important ones are:

- ANA>AUTO PLOT DATA. Generates a plot of the data file that was most recently read in, together with a title.
- ANA>AUTO PLOT FIT. Produces a plot containing the data and its most recent fit, together with the fitting parameters.

8.6 Loop command

Integration of scans requires the same operations many times. The ANA>LOOP command is meant for such repetitive actions. It prompts for a first and a last scan number and then executes the commands specified in the ANA.SET.LOOP> menu. In this submenu

one specifies a string of commands to be executed before the scan number and a string to be executed after the scan number. In case of peak integration, this could be:

```
ANA.SET.LOOP>BEFORE "R SP" (Read SPEC data file)
```

```
ANA.SET.LOOP>AFTER "1 F G 1 2 G RU OP INT 1 -1 -1" (Fit and integrate scan)
```

(Note that the command BEFORE expects a string, which is either a series of commands in quotes on the same command line, or a series of commands typed in after the prompt. In the latter case, the quotes are not necessary.)

The command "ANA>LOOP 120 200" will then execute these commands for scans 120 until 200. The command is implemented in a fairly robust way. Normally, a sequence of commands (e.g. in a macro file) is terminated as soon as an error occurs. When loop encounters an empty scan or similar problems, it will keep running.

8.7 Summary of peak integration

Thus, in order to integrate a series of scans, the following things have to be done:

1. Set parameter values in ANA.SET.READ>.
2. Set parameter values in ANA.SET.CORRECT>.
3. Set parameter values in ANA.SET.INT>, including opening an output file.
4. Set parameter values in ANA.SET.LOOP>.
5. If desired, restrict fitting range of peak width in fitting routine.
6. Execute ANA>LOOP command.
7. Close the integration output file (or quit ANA).

The integration file consists of the following columns:

Scan#	<i>h</i>	<i>k</i>	<i>l</i>	<i>F</i>	σ_F	list of correction factors used.
-------	----------	----------	----------	----------	------------	----------------------------------

Reflectivity ridge scans are not integrated using the loop command, but by the command REFTOF.

9 AVE

The result of the operations in ANA is a *.int file with the full list of integrated data. The next step in getting the data in final form is to sort and average these. The program AVE is meant for this. The generation of the final data requires three steps: reading, averaging and weighting the data. In special cases, also a domain assignment step may be needed.

9.1 Input

The AVE>READ command is used to input data. The type of data file can be specified in the AVE.SET.READ> submenu, but all the defaults are set to read the integrate files generated by ANA (including the default extension .int). If more data files need to be read, the command READMORE has to be used.

For example, the command "AVE>READ NICEDATA" will read in the data file nicedata.int. A full cycle through AVE will generate a file with extension .dat, thus one should not start with a file with this extension because it will be overwritten.

Note that AVE will only accept *integer* values for h and k , because only in this way can it unambiguously find symmetry equivalent reflections. In AVE.SET.READ> a transformation matrix can be specified to transform the diffraction indices. This can e.g. be used to transform fractional coordinates to integer ones.

9.2 Symmetry-equivalent domains

A surface may contain symmetry-equivalent domains, e.g. on a Si(100) surface one will normally find both (1x2) and (2x1) domains. While there is a symmetry relation between these domains, it typically differs from the plane group symmetry of one domain. Depending on the symmetry relation between the domains, fractional-order reflections may not overlap, but integer-order reflections always will. If a data set contains fractional-order reflections from both domains, these need to be assigned to one domain only in order to get a unique data set that can be properly averaged over the equivalent domains.

In AVE this is done in the following way. The input data needs to have integer (hk) values, so in the (1x2) and (2x1) domain example, one would e.g. use a 2x2 unit cell to denote the reflections. It is straightforward to define matrices \mathbf{T}_n that transform the diffraction indices from the input data reference cell to that of domain n . If the indices after the transformation by matrix \mathbf{T}_n are integer, then the reflection indeed belongs to this domain. If the indices are fractional, then not.

In the AVE.SET.DOMAIN> menu the number of equivalent domains and their transformation matrices can be set. The command AVE>DOMAIN will start the search, in which each reflection is assigned a domain number and (hkl) values corresponding to domain 1. The assignment is listed in a file with extension .dom. In our example: nicedata.dom.

The domain assignment should be done before the averaging.

9.3 Average

AVE sorts and averages the data based on the plane group symmetry that needs to be specified in ANA.SET.SYMMETRY>. By default, AVE uses the Friedel rule to sort data into rods with both positive and negative l -values.

Once the plane group has been set, the command AVE>AVERAGE groups all symmetry-equivalent reflections together, calculates the average structure factor, the variance (called σ_1) and the statistical error (called σ_2) and sorts the data (starting with the smallest h and k values). The command generates a file with extension *.ave in which all this information is listed. In our example: nicedata.ave.

The command AVERAGE also calculates the agreement factor ϵ of the data [1]. This is done as follows. For each reflection of which more than one symmetry-equivalent

reflection is contained in the data set, the variance σ_1 is calculated. The ε_{hkl} of this reflection is the variance divided by the averaged structure factor. The ε_{hkl} 's of all reflections of sufficient quality are averaged into the overall agreement factor ε that is a measure of the systematic errors in the data set. The agreement factor of a data set typically varies from 0.03 to 0.15.

"Sufficient quality" is something that can be specified in AVE.SET.AVERAGE>. A reflection is of sufficient quality if its average is larger than CUTOFF times the statistical error σ_2 , where the default CUTOFF is 2. In this way only well measured and sufficiently strong reflections are used in the agreement factor calculation. Reflections that pass the 'sufficient quality' criterion are indicated by a '*' in the .ave file. The reflections that are used thus depend on the value of CUTOFF and on the way in which σ_2 is calculated. By default, the standard deviation of the statistical average is used in the calculation of σ_2 (SIGMA2MODE 0). Two other modes are available in AVE, which can be used if the standard deviation is considered to be unrealistically small. This can easily occur, because the data contain more errors than just statistics. The main systematic errors, however, are taken into account in the weighting scheme, and in practise using the default SIGMA2MODE is often satisfactory.

9.4 Weighting

The final step is to calculate appropriate errors and to write all the unique structure factors to a file. The error that is assigned to a reflection is calculated as follows:

$$\sigma_{hkl} = \sqrt{\varepsilon^2 F_{hkl}^2 + \sigma_2^2}$$

The error of a particular structure factor F_{hkl} is thus the square sum of the average agreement factor of the entire data set and its own statistical error. In this way, also reflections of which no symmetry-equivalent ones have been measured, get a realistic error. Often the systematic error dominates.

The command WEIGHT prompts for a comment line and then produces a data file with extension *.dat (nicedata.dat) with the following format:

```
1st line:      comment
next lines:   h      k      l      F       $\sigma$ .
```

This is the format required for a ROD data file.

If the data set contains no symmetry-equivalent reflections of sufficient quality, no ε is calculated during the AVERAGE command. Weighting produces then data files with anomalously low error bars. This can be avoided by setting a 'realistic' value for ε in AVE.SET.WEIGHT> before executing the WEIGHT command; for example 0.1.

It is possible not to use the average normalized variance (i.e. agreement factor ε), but to use the variance of the specific data point measured. This can be useful if a data point is more extensively measured than the typical data and has a variance that deviates from the average value. The command AVE.SET.WEIGHT>MINIMUM allows setting the minimum

number of 'equivalent' reflections (at least 2!) that needs to be measured in order to do this alternative weighting of the data. With only limited data points (2 for example), the variance can be an overestimate of the actual error. Therefore a multiplier for the variance can be specified.

The 'approximate beta' roughness model in ROD requires the l -value of a near Bragg peak along a crystal truncation rod. If the flag LBRAGG is set in AVE.SET.WEIGHT>, WEIGHT will prompt for such a value for each rod. This value is added as an extra column in the data file.

ROD is prone to numerical errors for l -values that are integer. Therefore, by default a small offset is added to such integer values.

9.5 Non-equal domain occupancy

It may occur (but only rarely) that a surface consists of several domains that are not equally occupied. In that case, the symmetry-equivalence that is expected for equal domain occupancy is absent. It is still possible to obtain an estimate for the structure factor expected for equal occupancy by doing the averaging based on intensities, rather than on the structure factor. This option is activated by setting the flag AVE.SET.AVERAGE>INTENSITY. In this case the unweighted average is calculated, both for the structure factor and the statistical error. The calculated variance for a specific data point is a reasonable estimate for the error at that data point: if the non-equivalence is large, the error will be large, if the data points are nearly equivalent, the error will be small. For the calculated error of a data point, the variance of that particular point should be used, not the global agreement factor. This is possible by setting the parameter MINIMUM to the number of domains (e.g. 2) in AVE.SET.WEIGHT>. If the variance leads to error bars that are too large, a fraction of the variance can be used instead by setting the parameter VARIANCE in AVE.SET.WEIGHT>.

In addition to the systematic error due to the non-equal domain occupancy, the data set will still contain the normal systematic errors. These are difficult to estimate in this case, but some rods will likely be relatively insensitive to the domain occupancy and can thus be used to estimate the global value of ϵ . Two iterations of error estimates are now needed: first for the non-equal domains, and second for the overall agreement factor.

The procedure to obtain a properly weighted data file in this case is (with examples values shown):

1. read example.int
2. set INTENSITY mode to YES in AVE.SET.AVERAGE>
3. calculate AVERAGE
4. set EPS = 0; MINIMUM = 2; VARIANCE = 0.5 in AVE.SET.WEIGHT>
5. calculate SWEIGHT (special form of weight that generates file example.wgt)
6. read example.wgt

7. set INTENSITY mode to NO in AVE.SET.AVERAGE>
8. calculate average (no equivalence should be found)
9. set EPS = 0.06 (or other value) in AVE.SET.WEIGHT>
10. calculate WEIGHT

Instead of this averaging, the occupancies of the domains can be used as a fitting parameter when refining a model against the data.

9.6 *L-merging of data*

When acquiring data with an area detector, often data is taken that is very closely spaced along the rod. Such data points contain little extra information, and thus speed in the subsequent data analysis can be gained by merging the data that is very close in l . In the AVE.SET.WEIGHT> menu parameters for this can be set. Default points that are within 0.1 units in l and within 15% relative intensity are merged. After a normal WEIGHT is executed first, the command LMERGE will generate a data file with l -merged data. The filename has an 'L' attached to it.

9.7 *Other data files*

The entire data set is written to the *.dat file. Even though ROD now has options that make this no longer necessary, during analysis it may be convenient to have the different rods available as separate data sets, The command AVE>RODFILE does this. After the weighting, this command produces a file for each rod that has sufficient data points. The name of each file is generated by adding the (h,k) values to the filename. Thus for the (1,-2) rod, one would get nicedata1_2.dat.

It is also often convenient to have a data set with only the in-plane data. This is, after all, how surface X-ray diffraction started! The command AVE>INPLANE will generate such a file from the full data set, trying to interpolate or extrapolate from nearby l -values to $l \approx 0$. It generates a file with extension *.inp.

10 Making the final data set

One analysis cycle through ANA and AVE is normally insufficient to get the data in its best form. Particularly for large data sets in which many scans are integrated automatically, without visual inspection, some of the scans may be completely wrong. This could for example be due to a beam dump, a monochromator instability, a scan abortion, etc. After the first cycle through ANA and AVE, it is therefore a good idea to take a careful look at the .ave file. All scans are listed there, and it is easy to find a singular strong or weak reflection in a range of equivalent ones that is otherwise more or less constant. To aid a little in finding such reflections, AVE puts a 'w' (warning) next to a reflection that appears very different from the rest (based on its σ_2 value).

When a particular structure factor is off, it is necessary to take a more careful look at it using ANA. Making a plot of the data may identify the problem (peak completely off-centre, insufficient points, one or more spikes). The various operations that ANA offers can be used to solve the problem, if possible. The error may also arise from a strange fit, because for example a spike misguided the GUESS command. By checking such a scan, and trying to integrate it in different ways, a more reliable structure factor may be obtained. This value and error bar can then replace the values in the original .int file using a normal text editor. If the problem with the scan cannot be repaired, the structure factor should be removed from the .int file.

If a reflection is very weak compared to the background, it is difficult to integrate it accurately. By using information from other scans in the same region of reciprocal space about the peak shape, width and/or position, it is easier to determine the integrated intensity. In the fitting procedure, one or more of these parameters can then be fixed. Providing an upper bound for the structure factor is very valuable in a structure analysis. If the structure factor is, for example, determined to be less than 3, then this can be included in the data file with value 1.5 and error bar 1.5.

This 'manual' cleaning of the data set may require a few iterations. The .int file is modified and new .ave files are generated until all the data is considered satisfactory. In this process one may find that the agreement factor also significantly improves. The final weighting produces the data set that is used in ROD. The cleaning procedure is typically done based on the consistency of the data set itself (and the information in the experiment logbook about possible changing experimental conditions). Only when this is completed, a serious analysis using ROD should be done. However, it may occur that in the structural analysis one specific reflection is always off. This can be a reason to reconsider the integration of that reflection.

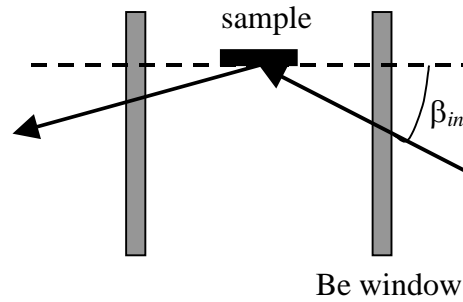
References

- [1] I.K. Robinson, in *Handbook on synchrotron radiation*, edited by G.S. Brown and D.E. Moncton (North-Holland, Amsterdam, 1991), Vol. 3, p. 221.
- [2] R. Feidenhans'l, *Surf. Sci. Rep.* 10 (1989) 105.
- [3] E. Vlieg, in *Surface and Interface Science*, edited by K. Wandelt (Wiley, 2011), p. to be published.
- [4] J.M. Bloch, *J. Appl. Cryst.* 18 (1985) 33.
- [5] E. Vlieg, J.F. van der Veen, J.E. Macdonald and M. Miller, *J. Appl. Cryst.* 20 (1987) 330.
- [6] M. Lohmeier and E. Vlieg, *J. Appl. Cryst.* 26 (1993) 706.
- [7] K.W. Evans-Lutterodt and M.T. Tang, *J. Appl. Cryst.* 28 (1995) 318.
- [8] E. Vlieg, *J. Appl. Cryst.* 31 (1998) 198.
- [9] E. Vlieg, *J. Appl. Cryst.* 30 (1997) 532.
- [10] O. Robach, Y. Garreau, K. Aïd and M.B. Véron-Jolliot, *J. Appl. Cryst.* 33 (2000) 1006.
- [11] E. Vlieg, *J. Appl. Cryst.* 33 (2000) 401.

Appendix A: Geometrical correction factors

The z-axis and the 5-circle geometries are special cases of the 6-circle geometry (e.g., by taking $\gamma = 0$ in the 6-circle geometry, one arrives at the 5-circle geometry) [9]. In ANA the correction factors for these three cases are therefore calculated in one subroutine for the general 6-circle case. At the moment, this is (rather confusingly) listed as z-axis geometry in ANA.

Most of the correction factors listed here can be found in refs. [8, 9]. The transmission is not discussed in those references, so we briefly explain it here. There are two different situations, that may occur together, depending on whether the dimensions in the parallel or perpendicular direction determine the total path of the X-ray beam through the absorbing medium.



The first case occurs when the glancing X-ray beam traverses a sample chamber, liquid column etc. in the perpendicular direction. The figure shows how the path length traversed by the X-rays depends on the incoming or outgoing angles. We discuss here only the incoming beam, the exit beam is completely analogous. For incoming angle equal to zero ($\beta_{in} = 0$), the transmission is equal to:

$$T_{0,par} = e^{-\mu d},$$

with μ the linear absorption coefficient and d the thickness of the Be window (or the liquid until the sample centre). $T_{0,par}$ is thus the transmission with the beam in the inplane direction, but with the absorbing material perpendicular to the beam. For a finite value of β_{in} , we have:

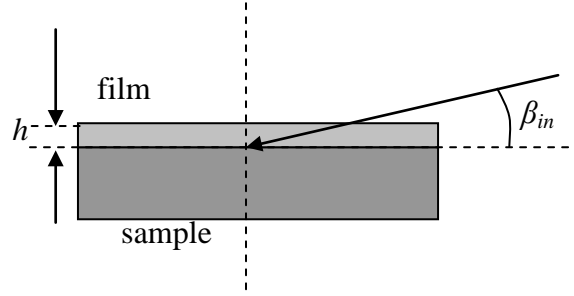
$$T_{par} = e^{-\mu d / \cos \beta} = (T_{0,par})^{1/\cos \beta}.$$

The absorption correction factor is the change in transmission with respect to zero angle of incidence, thus:

$$C_{abs,par} \equiv \frac{T_{\beta}}{T_0} = (T_{0,par})^{(1/\cos \beta - 1)}.$$

The expression for β_{in} depends on the particular diffractometer used. To apply the correction, the value for $T_{0,par}$ needs to be provided. We assume that the geometry is cylindrically

symmetric, and thus the inplane transmission is assumed to be the same for the incoming and outgoing directions.



For the second case, the path length traversed is determined by the perpendicular thickness of a film (or other material) on top of surface/interface (see figure). The transmission through the film in the perpendicular direction is:

$$T_{0,perp} = e^{-\mu h}.$$

For a glancing incoming angle, the transmission is:

$$T_{perp} = e^{-\mu h / \sin \beta_{in}} = (T_{0,perp})^{1/\sin \beta_{in}} \equiv C_{abs,perp}.$$

Since the transmission is zero for $\beta_{in} = 0$, we cannot use this as the reference situation here. Assuming again cylindrical symmetry, the parameter $T_{0,perp}$ will apply for both the incoming and outgoing angle.

For $T_{0,par} = T_{0,perp} = 0$ no absorption correction is applied.

Z-axis diffractometer

Lorentz factor rocking scan	$\frac{1}{\sin \delta \cos \alpha \cos \gamma}$
Lorentz factor stationary mode	$\frac{1}{\sin \gamma}$
Lorentz factor reflectivity rocking scan	$\frac{1}{\sin 2\alpha}$
Rod interception	$\cos \gamma$
Linear γ table (= changing slit distance)	$\cos \gamma$
Polarisation factor	$P = p_h P_{hor} + (1-p_h) P_{ver}$
Horizontal polarisation P_{hor}	$1 - (\sin \alpha \cos \delta \cos \gamma + \cos \alpha \sin \gamma)^2$
Vertical polarisation P_{ver}	$1 - \sin^2 \delta \cos^2 \gamma$
Transmission incoming beam, parallel direction.	$(T_{0,par})^{(1/\cos \alpha - 1)}$
Transmission outgoing beam, parallel direction.	$(T_{0,par})^{(1/\cos \gamma - 1)}$
Transmission incoming beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \alpha)}$
Transmission outgoing beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \gamma)}$
Area correction (ignoring footprint and sample size)	$\frac{1}{\sin \delta}$
Beam profile and finite sample size	Calculated numerically in ANA. See ref. [9].

6-circle diffractometer

Lorentz factor rocking scan	$\frac{1}{\sin \delta \cos \beta_{in} \cos \gamma}$
Lorentz factor stationary mode	$\frac{1}{\sin \beta_{out}}$
Lorentz factor reflectivity rocking scan	$\frac{1}{\sin 2\beta_{in}} = \frac{1}{\sin 2\beta_{out}}$
Rod interception	$\frac{2 \cos \beta_{in} \cos \gamma}{\cos \alpha + \cos(\alpha - 2\beta_{in}) + \sin(2\alpha - 2\beta_{in}) \sin \gamma + 2 \sin^2(\alpha - \beta_{in}) \cos \delta \cos \gamma}$
Linear γ table (= changing slit distance)	$\cos \gamma$
Polarisation factor	$P = p_h P_{hor} + (1-p_h) P_{ver}$
Horizontal polarisation P_{hor}	$1 - (\sin \alpha \cos \delta \cos \gamma + \cos \alpha \sin \gamma)^2$
Vertical polarisation P_{ver}	$1 - \sin^2 \delta \cos^2 \gamma$
Transmission incoming beam, parallel direction.	$(T_{0,par})^{(1/\cos \alpha - 1)}$
Transmission outgoing beam, parallel direction.	$(T_{0,par})^{(1/\cos \gamma - 1)}$
Transmission incoming beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \alpha)}$
Transmission outgoing beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \gamma)}$
Area correction (ignoring footprint and sample size)	$\frac{1}{\sin \delta \cos(\alpha - \beta_{in})}$
Beam profile and finite sample size	Calculated numerically in ANA. See ref. [9].

5-circle diffractometer

Lorentz factor rocking scan	$\frac{1}{\sin \delta \cos \beta_{in}}$
Lorentz factor stationary mode	$\frac{1}{\sin \beta_{out}}$
Lorentz factor reflectivity rocking scan	$\frac{1}{\sin 2\beta_{in}} = \frac{1}{\sin 2\beta_{out}}$
Rod interception	$\frac{2 \cos \beta_{in}}{\cos \alpha + \cos(\alpha - 2\beta_{in}) + 2 \sin^2(\alpha - \beta_{in}) \cos \delta}$
Polarisation factor	$P = p_h P_{hor} + (1-p_h) P_{ver}$
Horizontal polarisation P_{hor}	$1 - (\sin \alpha \cos \delta)^2$
Vertical polarisation P_{ver}	$1 - \sin^2 \delta$
Transmission incoming beam, parallel direction.	$\left(T_{0,par}\right)^{\left(1/\cos\alpha^{-1}\right)}$
Transmission outgoing beam	1
Area correction (ignoring footprint and sample size)	$\frac{1}{\sin \delta \cos(\alpha - \beta_{in})}$
Beam profile and finite sample size	Calculated numerically in ANA. See ref. [9].

2+3-axis diffractometer

Lorentz factor rocking scan	$\frac{1}{\sin \delta \cos \alpha}$
Lorentz factor stationary mode	
Lorentz factor reflectivity rocking scan	$\frac{1}{\sin 2\alpha}$
Rod interception	$\cos \beta_{out}$
Polarisation factor	$P = p_h P_{hor} + (1-p_h) P_{ver}$
Horizontal polarisation P_{hor}	$1 - (\cos \delta \sin \gamma)^2$
Vertical polarisation P_{ver}	$1 - \sin^2 \delta$
Transmission incoming beam, parallel direction.	$(T_{0,par})^{(1/\cos \alpha - 1)}$
Transmission outgoing beam, parallel direction.	$(T_{0,par})^{(1/\cos \beta_{out} - 1)}$
Transmission incoming beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \alpha)}$
Transmission outgoing beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \beta_{out})}$
Area correction (ignoring footprint and sample size)	$\frac{\cos \beta_{out}}{\sin \delta}$
Beam profile and finite sample size	Calculated numerically in ANA. See ref. [9].

2+2-axis diffractometer

Lorentz factor rocking scan	$\frac{1}{\sin \delta \cos \alpha}$
Lorentz factor stationary mode	
Lorentz factor reflectivity rocking scan	$\frac{1}{\sin 2\alpha}$
Rod interception	$\frac{\cos \alpha}{\sin^2(\alpha - \gamma) \cos \delta + \cos \alpha \cos(\alpha - \gamma)}$
Polarisation factor	$P = p_h P_{hor} + (1-p_h) P_{ver}$
Horizontal polarisation P_{hor}	$1 - (\cos \delta \sin \gamma)^2$
Vertical polarisation P_{ver}	$1 - \sin^2 \delta$
Transmission incoming beam, parallel direction.	$(T_{0,par})^{(1/\cos \alpha - 1)}$
Transmission outgoing beam, parallel direction.	$(T_{0,par})^{(1/\cos \beta_{out} - 1)}$
Transmission incoming beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \alpha)}$
Transmission outgoing beam, perpendicular direction.	$(T_{0,perp})^{(1/\sin \beta_{out})}$
Area correction (ignoring footprint and sample size)	$\frac{\cos \beta_{out}}{\sin \delta}$
Beam profile and finite sample size	Calculated numerically in ANA. See ref. [9].

Appendix B: Transforming between unit cells

Suppose we want to change from a unit cell with lattice vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 to a new one with lattice vectors \mathbf{a}'_1 , \mathbf{a}'_2 and \mathbf{a}'_3 . The transformation between these is given by a transformation matrix \mathbf{T} , such that.

$$\mathbf{a}'_i = \sum_j \mathbf{T}_{ij} \mathbf{a}_j \quad (\text{B1})$$

The inverse matrix will transform the lattice vector back:

$$\mathbf{a}_i = \sum_j \mathbf{T}_{ij}^{-1} \mathbf{a}'_j. \quad (\text{B2})$$

Assume that position \mathbf{r} has coordinates α_i in the starting unit cell:

$$\mathbf{r} = \sum_i \alpha_i \mathbf{a}_i. \quad (\text{B3})$$

We want to know the coordinates in the new frame:

$$\mathbf{r} = \sum_i \alpha'_i \mathbf{a}'_i. \quad (\text{B4})$$

The solution for this can be found by substituting (B2) in (B3):

$$\mathbf{r} = \sum_i \alpha_i \mathbf{a}_i = \sum_i \alpha_i \left(\sum_j \mathbf{T}_{ij}^{-1} \mathbf{a}'_j \right) = \sum_j \left(\sum_i \alpha_i \mathbf{T}_{ij}^{-1} \right) \mathbf{a}'_j. \quad (\text{B5})$$

Thus

$$\alpha'_j = \sum_i \alpha_i \mathbf{T}_{ij}^{-1} = \sum_i (\mathbf{T}^{-1})_{ji}^t \alpha_i, \quad (\text{B6})$$

where t stands for transpose. Thus the matrix that transforms the coordinates from the old to the new unit cell is given by the transpose of the inverse of matrix \mathbf{T} .

The matrix \mathbf{T} itself transforms the diffraction indices $\mathbf{H} = (hkl)$ to the ones corresponding to the new unit cell (see e.g. B.E. Warren, *X-ray diffraction*). In summary, we thus have the following transformations between the two unit cells:

$$\begin{aligned} \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \mathbf{a}'_3 \end{pmatrix} &= \mathbf{T} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix} & \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix} &= \mathbf{T}^{-1} \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \mathbf{a}'_3 \end{pmatrix} \\ \begin{pmatrix} h' \\ k' \\ l' \end{pmatrix} &= \mathbf{T} \begin{pmatrix} h \\ k \\ l \end{pmatrix} & \begin{pmatrix} h \\ k \\ l \end{pmatrix} &= \mathbf{T}^{-1} \begin{pmatrix} h' \\ k' \\ l' \end{pmatrix} \\ \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \alpha'_3 \end{pmatrix} &= (\mathbf{T}^{-1})^t \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} & \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} &= (\mathbf{T})^t \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \alpha'_3 \end{pmatrix}. \end{aligned} \quad (\text{B7})$$

Example: from fcc cubic to (111) surface unit cell

$$\mathbf{H}_{(111)} = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix} \mathbf{H}_{fcc} \quad \mathbf{H}_{fcc} = \frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 2 & 4 & 1 \\ -4 & -2 & 1 \end{pmatrix} \mathbf{H}_{(111)}$$

$$\alpha_{(111)} = \frac{1}{3} \begin{pmatrix} 2 & 2 & -4 \\ -2 & 4 & -2 \\ 1 & 1 & 1 \end{pmatrix} \alpha_{fcc} \quad \alpha_{fcc} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} & 1 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix} \alpha_{(111)}$$

Example: from (111) surface unit cell to ($\sqrt{3} \times \sqrt{3}$)R30° surface unit cell

$$\mathbf{H}_{\sqrt{3}} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{H}_{(111)} \quad \mathbf{H}_{(111)} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{H}_{\sqrt{3}}$$

$$\alpha_{\sqrt{3}} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \alpha_{(111)} \quad \alpha_{(111)} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \alpha_{\sqrt{3}}$$

Appendix C: Line profiles in ANA

A number of line profiles are used in ANA, the two most popular are the Lorentzian and the Gaussian profiles. In all cases also a sloping background is part of the theoretical line shape. In the line shapes, the following parameters are used:

- b constant background
- s slope in background
- p peak position
- h height of the peak
- w width of the peak (typically fwhm)
- η exponent

The following line shapes are used:

$$\text{Lorentzian: } L(x) = b + s(x - p) + \frac{hw^2}{w^2 + 4(x - p)^2}.$$

The integrated intensity of the peak is: $\frac{1}{2}\pi wh$.

$$\text{Gaussian: } G(x) = b + s(x - p) + he^{-4\ln 2(x-p)^2/w^2}.$$

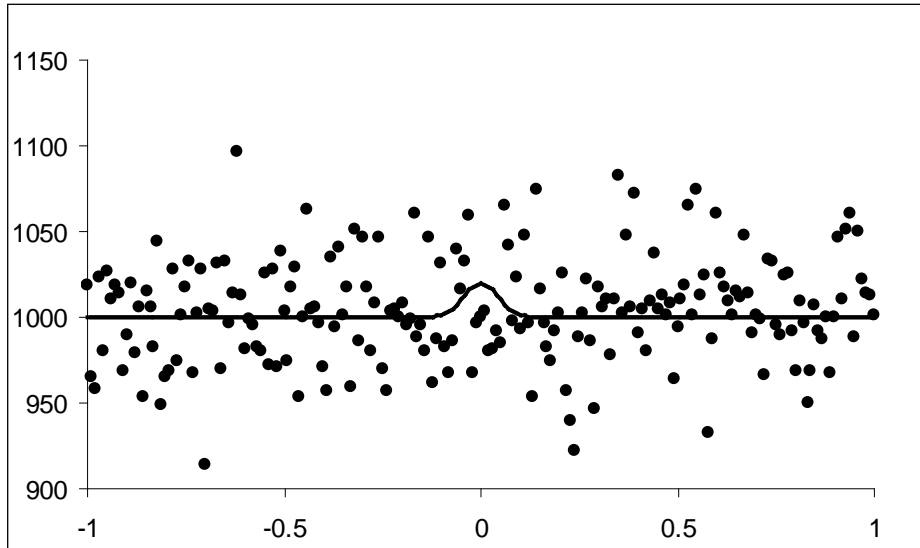
The integrated intensity is: $\frac{1}{2}\sqrt{\pi/\ln 2}hw$.

$$\text{Approximate power law: } P(x) = b + s(x - p) + \frac{h}{\left[4\left(\frac{x-p}{w}\right)^2 + 1\right]^{\eta/2}}.$$

For $\eta = 2$ this is the same as a Lorentzian profile. Note that in this case w is not the fwhm; that quantity is given by: $w\sqrt{2^{2/\eta} - 1}$.

Appendix D: Upper limit for structure factor value

If a scan does not (appear to) contain a peak, it is still important to estimate an upper limit for the structure factor. The situation is illustrated in the figure. The normal integration in ANA does not provide a good estimate in this case, because the procedure assumes a real peak.



If no peak is observed, it means that the intensity in the peak is not above the background in a statistical significant sense. Thus the signal is less than the statistical error in the background over the peak region. The approximate peak width will often be known (e.g. from neighbouring scans with a proper peak). In that case, the error can be found by using all data points in the peak region:

$$\frac{1}{\sigma_{bgr}^2} = \sum_i^{region < FWHM} \frac{1}{\sigma_i^2}.$$

An estimate for the upper limit of the integrated peak intensity is then:

$$I_{upper} = 2 \times \sigma_{bgr} \times FWHM,$$

from which an estimate of the corresponding structure factor F_{upper} can be calculated. (The factor 2 is an estimate relating integrated intensity to the peak width.) It turns out that using:

$$F = F_{upper}$$

$$\sigma_F = 0.9 \times F_{upper}$$

yields values that work well in ROD. Using this estimate it pays to do long counting, even without observing a peak, because this will lower the value for F_{upper} .

Appendix E: Required counting times

The counting time that is required to reach a specific accuracy depends strongly on the signal-to-background ratio. If this ratio is high, the accuracy depends only on the statistical error in the signal, but for low ratios, the counting time can become significantly longer.

We define the following parameters

s peak signal rate

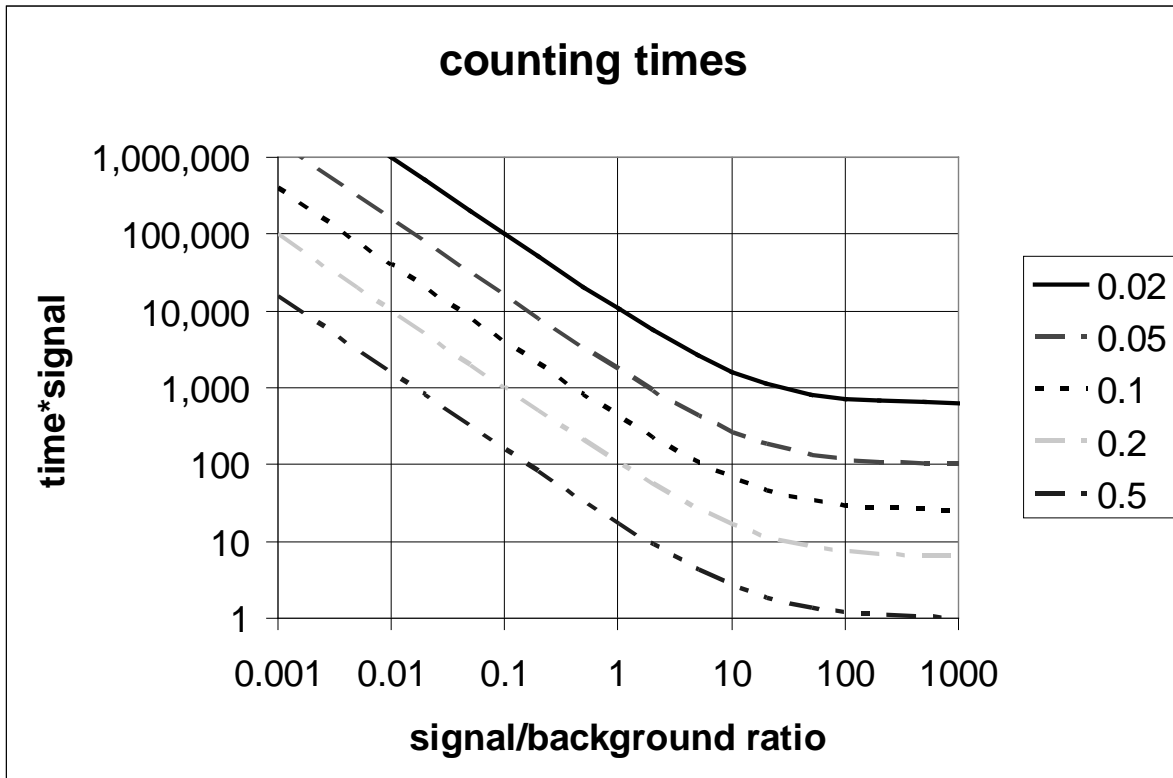
b background rate

ε relative error in integrated signal

An approximate value for the required counting time t is found to be:

$$t = \frac{1}{\varepsilon^2} \frac{\sqrt{1 + 4\left(\frac{s}{b}\right)^{-1}}}{s}.$$

This assumes that approximately 10 points are measured in the peak. Below the function ts is plotted versus the s/b ratio for four values of the required accuracy.



Appendix F: Typical parameter settings in ANA

(with thanks to Chris Walker)

set read menu

	ID3	ID32	DUBBLE	9.4	comments
xcolumn	1	1	1	1	
ycolumn	-	-	-	5	at ESRF obtained from 'yname'
mcolumn	≠ 0	≠ 0	≠ 0	3	at ESRF obtained from column name 'monitor'
norm	1e5	1e5	1e5	1e5	Convenient is average monitor counts per second
filename	spec filename containing the scan number	spec filename containing the scan number	spec filename containing the scan number	<i>path/Rnn</i> with nn most significant part of scan number	
solidstate	no	no	no	yes	patch for Daresbury counters
yname	det_corr	detector	detector	-	
<i>type of file</i>	ID3	ID32	DUBBLE	SRS	get correct values for fixed angles

set integration menu

	ID3	comments
open/close		open/close file name with integration output
transmat	identity matrix	can be used to change coordination system
nbackground	3	
autoback	yes	use 'nbackground' points on beginning and end of scan
autoback	no	use user specified background (value -1 gives the values from most recent fit)
range	0	full range
range	<i>n</i>	use <i>n</i> times fwhm of fit (useful for very broad scans)

set correct menu

	ID3-EH1	ID3-EH2	ID32	DUBBLE	9.4	comments
<i>rocking scan (rscan)</i>	zaxis	zaxis	twoptwo	twopthree	zaxis	zaxis includes 6 and 5-circles.
<i>reflectivity rocking scan</i>	refrock	refrock	refrock	refrock	refrock	doesn't care about diffractometer type
beamcor	yes	yes	yes	yes	yes	Only important for large incoming angles. If no, slit and beam parameters are ignored.
gammafix	-	-	-	-	0/15	
lingamma	no	yes	no	no	no	

dslit	2	2	2	ds1v	2	in-plane opening of detector slit near sample
hslit	3					in-plane cut-off of beam profile
vslit	1					out-of-plane cut-off of beam profile
hbeam						in-plane Gaussian fwhm of beam profile
vbeam						out-of-plane Gaussian fwhm of beam profile
radius	4	4	4	4	4	circular samples, <i>not</i> diameter
hpolar	0	1	1	0.95	0.95	
tin	1	1	1	1	1	only $\neq 1$ if serious beam absorption
tout	1	1	1	1	1	only $\neq 1$ if serious beam absorption

Appendix G: ANA commands

```
***** MAIN MENU *****
Read      : Read a spectrum from file
List      : List a spectrum to terminal or file
Clear     : Clear all spectra
Set       : Goto set parameter menu
Operate   : Goto operations menu
Theory    : Goto theory menu
Fit       : Fit a spectrum
Plot      : Plot a spectrum
Autoplot  : Make automated plot
P3d       : Plot spectrum as 3D curve
LooP      : Execute loop command
Macro     : Run macro file
:         : Execute an operating system command
Help      : Display menu
QUIT     : Quit program

ANA>
***** FILE TYPES *****
SRs       : SRS data file
SPec      : SPEC data file
SUPER     : Super data file (single scan)
Osuper    : Old super scan (in bigdata file)
Xye       : File with x y (and opt. error) columns
Column    : File with arbitrary amount of columns
Ftascom   : Single column tascom data file
Help      : Display menu
Return    : Return to main menu

File type:
***** SET PARAMETERS *****
REAd      : Set read parameters
Fit       : Set fit parameters
General   : Set general parameters
LooP      : Set loop strings
Integrate : Set numerical integration parameters
Correct   : Set correction parameters
Help      : Display menu
Return    : Return to main menu

ANA.SET>
***** SET READ PARAMETERS *****
Xcolumn   : Column number of x values
Ycolumn   : Column number of y values
Mcolumn   : Column number of monitor values
Ecolumn   : Column number of error values
Norm      : Normalization value for monitor counts
HEAdER    : Number of header lines (obsolete ...)
Filename  : Name of data file
Solidstate: Solid-state detector correction (SRS)
YName     : Name of column of y values in data file
Pathname  : Pathname of SUPER file
NAutosub  : # of files in automatic subdir. (SUPER)
SRs       : Standard srs scan settings
AEs       : Standard VSW-AES scan settings
ID03      : Spec file of ID03 type
ID32      : Spec file of ID32 type
Dubble    : Spec file of DUBBLE type
BL13xu    : Spec file of BL13XU type
REPair    : Repair angle scans of DUBBLE diffract.
FIXl      : Fix l-index for spec. scans on z-axis
B3lambda  : Product of b3 and wavelength
Corfilter : Correct for filter attenuation
FILLTer   : Transmission of filter n
List      : List parameter values
Help      : Display menu
Return    : Return to main menu

ANA.SET.READ>
***** SET FIT PARAMETERS *****
Open      : Open file for fit results
Close     : Close fit output file
Standard  : Use standard deviation weights in fit
Unit      : Use unit weights in fit
List      : List parameter values
```

```

Help      : Display menu
Return    : Return to main menu
ANA.SET.FIT>
***** SET GENERAL PARAMETERS *****
ONscreen  : Screen output on
OFFscreen : Screen output off
Datarun   : Set data run mode
List      : List parameter values
Help      : Display menu
Return    : Return to main menu
ANA.SET.GENERAL>
***** SET LOOP STRINGS *****
Before    : Set string before file number
After     : Set string after file number
Shownumber: Show scan number during looping
Increment : Run number increment in loop
List      : List parameters
Help      : Display menu
Return    : Return to main menu
ANA.SET.LOOP>
***** SET INTEGRATION PARAMETERS *****
Open      : Open output file for integration results
Close     : Close output file with integr. results
Transmat  : Set elements of transformation matrix
NBackgr   : # left/right background points
Autoback  : Automatic background calculation yes/no
Range     : Integration range (*fitted FWHM)
List      : List parameter values
Help      : Display menu
Return    : Return to main menu
ANA.SET.INT>
***** SET CORRECTION PARAMETERS *****
NOcorrect : No correction factors applied
ZAxis     : Z-axis diffractometer
TWOptwo   : (2+2)-type diffractometer
TWOPTHree : (2+3)-type diffractometer
REFrock   : Reflectivity rocking scan
Beamcor   : Beam profile correction yes/no
Gammafix  : Detector angle gamma at the SRS
LIngamma  : Linear gamma table yes/no
RScan     : Spectrum is angular scan
HKlscan   : Spectrum is diffraction-index scan
Dslit     : In-plane width of detector slits
HSlit     : Horizontal slit width of incoming beam
VSlit     : Vertical slit width of incoming beam
HBeam     : Horizontal fwhm of incoming beam
VBeam     : Vertical fwhm of incoming beam
RADIUS    : Sample radius (only round samples)
HPolar    : Horizontal polarisation fraction
TPar      : Transmission in parallel direction
TPerp     : Transmission in perpendicular dir.
List      : List current values
Help      : Display menu
Return    : Return to main menu
ANA.SET.COR>
***** Ordinate/spectrum operation menu *****
Merge     : Merge two spectra
APpend    : Append two spectra
SUM       : Sum y-values of two spectra
Shift     : Shift x-values of spectrum
Center    : Shift x-values to put maximum at x=0
Scale     : Multiply x-values by constant
Delete    : Delete range of points from spectrum
Wrap      : Wrap x-range to avoid 360 degr jumps
MULTIPLY  : Multiply y-values by constant
Add       : Add constant to y-values
CUTOFF    : Cut off points; +/- = above/below value
Lump      : Lump data points
Integrate : Numerically integrate a spectrum
AUTOMERGE : Try merging with next-higher run number
DIFFERENT : Differentiate spectrum
FITSUB    : Subtract last fit from spectrum
SPlit     : Split spectrum into peak and background
DIV       : Divide two buffers by each other
LOG10     : Take log10 of y values
REFTOF    : Convert refl. ridge scan to str. factors
LEvel     : Level spectrum with sloping background

```



```

DESpike : Remove spikes from spectrum
FLimit  : Estimate upper limit integr. intensity
Help    : Display menu
Return  : Return to main menu
ANA.OPERATE>
***** THEORY MENU *****
Theory  : Choose type of theory
LOWerx  : Lower x-value
Upperx  : Upper x-value
Stepx   : Step size in x
SPecx   : Use x-range of spectrum (if 0, use range)value
Value   : Set parameter value
Compute : Compute theoretical spectrum
List    : List parameter values
Help    : Display menu
Return  : Return to main menu
ANA.THEORY>
***** THEORY TYPES *****
LOrentz : Single Lorentzian
LLOrentz : Double Lorentzian
Gauss   : Single Gaussian
GGauss  : Double Gaussian
LGauss  : Lorentzian plus Gaussian
GLOrentz : Gaussian plus Lorentzian
Power   : Power-law lineshape
Voigt   : Pseudo-Voigt lineshape
Slit    : Slit function
Help    : Display menu
Give type of theory:
***** SET FITTING PARAMETERS *****
Value   : Set parameter value
LOWer   : Set lower bound on parameter range
Upper   : Set upper bound on parameter range
Fix     : Fix a parameter value
LOose   : Make parameter free
Guess   : Estimate fitting parameters
Midpoint : Set peak position to center of scan
List    : List parameter values
Help    : Display menu
RUn     : Start fit
Return  : Return to main menu
ANA.FIT.PAR>
***** AUTO PLOT MENU *****
Data    : Plot data only
Fit     : Plot data and fit
Title   : Set standard data title
Scannumber: Make scan number the title
HKl     : Put Diffraction indices in header
Errors  : Error bar plotting on/off
Help    : Display menu
Return  : Return to main menu
ANA.AUTO PLOT>

```

Appendix H: AVE commands

```
***** MAIN MENU *****
Set      : Set parameters
Read     : Read in new data
READMore : Read additional data
MANipulate: Manipulate data
DOfains  : Search for domains
Average  : Do the averaging
Weight   : Make file with weighted data
Lmerge   : As weight, merged l values
SWeight  : Same as weight, plus scan #'s
ROdfile  : Make separate file for each rod
LRodfile : Make separate file for l-merged data
Inplane  : Make file with inplane data
DEfault  : Set all parameters to default
Macro    : Run macro file
:        : Execute an operating system command
Help     : Display menu
QUIT     : Quit program
AVE>
***** SET MENU *****
Read     : Parameters for reading input file
SYmmetry : Set symmetry group
Average  : Averaging parameters
Weight   : Weighting parameters
Domains  : Domain parameters
Help     : Display menu
Return   : Return to main menu
AVE.SET>
***** SET FILE READ PARAMETERS *****
Header   : Number of header lines
Idcol    : Column with identification (scan) number
HCol     : Column with diffraction index h
KCol     : Column with diffraction index k
LCol     : Column with diffraction index l
Fcol     : Column with structure factor
Sigcol   : Column with error in structure factor
Tmat     : Parameters of transformation matrix
LAsonly  : Only use last of consecutive data
List     : List parameters
Help     : Display menu
Return   : Return to main menu
AVE.SET.READ>
***** SET DOMAIN PARAMETERS *****
Ndomain  : Number of domains
Matrix   : Transformation matrix parameters
List     : List parameters
Help     : Display menu
Return   : Return to main menu
AVE.SET.DOMAIN>
***** SET PLANE GROUP SYMMETRY *****
P1       : Plane group no. 1
P2       : Plane group no. 2
PM       : Plane group no. 3
PG       : Plane group no. 4
CM       : Plane group no. 5
P2MM    : Plane group no. 6
P2MG    : Plane group no. 7
P2GG    : Plane group no. 8
C2MM    : Plane group no. 9
P4       : Plane group no. 10
P4MM    : Plane group no. 11
P4GM    : Plane group no. 12
P3       : Plane group no. 13
P3M1    : Plane group no. 14
P31M    : Plane group no. 15
P6       : Plane group no. 16
P6MM    : Plane group no. 17
FRiedel : Friedel pairs yes/no
List     : List current plane group
Help     : Display menu
Return   : Return to main menu
AVE.SET.SYMMETRY>
***** SET AVERAGING PARAMETERS *****
```

```

Sigma2mode: Mode of sigma_2 calculation
CUtloff   : Cutoff in agreement factor evaluation
Intensity : Special mode to average intensity
List      : List parameters
Help      : Display menu
Return    : Return to main menu
AVE.SET.AVERAGE>
***** SET WEIGHT PARAMETERS *****
Epsilon   : Average agreement factor
Minimum   : Minimum # of equivalent reflections
LMult     : Multiplication factor of l
Integerl  : Add offset to integer l yes/no
LBragg    : Input nearest Bragg peak l-values
ROdmin    : Minimum # points on rod for listing
LStep     : Stepsize in l for l-merging of data
Delf      : Maximum rel. diff. in f for l-merging
Variance  : Multiplier of variance (indiv. refl.)
List      : List parameters
Help      : Display menu
Return    : Return to main menu
AVE.SET.WEIGHT>
***** MANIPULATE MENU *****
Scselect  : Multiply f for range of scan #'s
HKselect  : Multiply f for specific {hk}
LSelect   : Multiply f for specific |l|
Help      : Display menu
Return    : Return to main menu
AVE.MANIPULATE>

```