

## 3. CIF DATA DEFINITION AND CLASSIFICATION

radiation source and the specimen if no monochromator was present.

Two methods may be used to describe the slits limiting the divergence in the equatorial plane. The angular divergence allowed by the slits in degrees can be specified using the `_pd_instr_divg_eq_*` data items. Alternatively, the dimensions of the slits in the equatorial direction in millimetres may be specified using `_pd_instr_slit_eq_*`. The dimensions of these slits in the axial direction, *i.e.* the direction perpendicular to the equatorial plane and containing the incident or diffracted beam as appropriate, are specified in millimetres using `_pd_instr_slit_ax_*`. The axial slit lengths, along with the `_pd_instr_dist_*` distances, are useful for estimating the low-angle peak asymmetry (Finger *et al.*, 1994). Note that angular divergence in the axial plane is not a well defined concept for line-focus instruments, but can be specified, where appropriate, using `_pd_instr_divg_ax_*`.

The axial and equatorial directions are shown schematically in Fig. 3.3.4.1. The equatorial plane contains the equatorial direction vectors, as well as the incident beam, the diffracted beam and the scattering vector. The axial plane is perpendicular to the equatorial plane and contains the sample centre, which is the point where the incident and diffracted beams meet. For area-detection instruments, the designations of axial and equatorial directions may be arbitrary.

Soller collimators are described using `_pd_instr_soller_eq_*` data items rather than `_pd_instr_divg_eq_*` data items. It is common practice to specify the Soller collimation in arc-minutes (*e.g.* 30'). However, pdCIF defines these items to have units of degrees, so 30' would be recorded in the CIF as 0.5. It is not usual to limit the axial divergence, except to reduce low-angle asymmetry, but if this is done, the `_pd_instr_soller_ax_*` data items can be used to define this.

For constant-wavelength instruments, it is common to have a monochromator or filter either before the sample, or after the sample (an analyser), or sometimes both. This is described using `_pd_instr_monochr_pre_spec` and `_pd_instr_monochr_post_spec`. It is rare, but possible, to have both a filter and a monochromator in the same location. The BT-1 neutron powder diffractometer at NIST uses both a Cu(311) monochromator and a graphite filter to attenuate the  $\lambda/2$  component. In this case, the two elements would be placed in a loop:

```
loop _pd_instr_monochr_pre_spec
  '5 seg. vert. focusing Cu(311) monochromator'
  '3 cm graphite filter'
```

Note that the monochromator and analyser takeoff angles are given using `_pd_instr_2theta_monochr_pre` and `_pd_instr_2theta_monochr_post`. It is useful to record these values for X-ray studies, as they are needed for proper polarization corrections.

In a conventional Bragg–Brentano diffractometer, the divergence slits limit the illumination area at the sample. However, since the  $\varphi$  axis (the sample  $\theta$  axis) is usually set to bisect the  $2\theta$  angle of the detector, the actual length of the area of the sample that is illuminated changes with  $2\theta$ . One should choose divergence slits so that the beam does not illuminate areas outside the sample at the lowest diffraction angle used. An alternative method for data collection is to have a divergence slit that opens as  $2\theta$  increases, so that a constant area of the sample is illuminated. This is known as a  $\theta$ -compensating slit. Using a  $\theta$ -compensating slit provides a better signal-to-noise ratio at larger  $2\theta$  values, but means that the diffraction intensities have to be normalized to compensate for the change in illumination. It also introduces greater optical

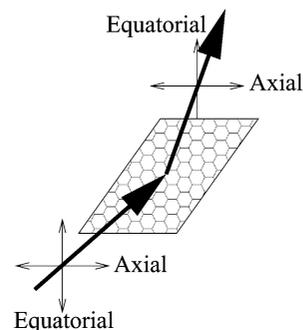


Fig. 3.3.4.1. The axial and equatorial directions in a powder-diffraction experiment.

aberrations with  $2\theta$ ; a flat plate becomes an increasingly worse approximation to the curved sample geometry in the true Bragg–Brentano geometry.

The use of a variable divergence slit can be recorded in the form:

```
_pd_instr_cons_illum_flag    yes
_pd_instr_cons_illum_len    25.4
```

Note that if `_pd_instr_cons_illum_flag` is not specified, the value is assumed to be `no`, indicating that a fixed-width divergence slit has been used.

The beam size can be specified in two different ways: as the size at the source, using `_pd_instr_source_size_ax` and `_pd_instr_source_size_eq`, or as the size at the sample position, using `_pd_instr_beam_size_ax` and `_pd_instr_beam_size_eq`. Note that the size of the beam at the sample differs from the illumination length described above except when the sample is perpendicular to the beam. When a variable-divergence slit is in use, the beam size at the sample changes with  $2\theta$ , so if this size is known directly, the `_pd_instr_beam_size_eq` data item can be included in the loop containing the diffraction intensities. Similarly, in a constant-divergence instrument, where the illumination length changes with  $2\theta$ , the illumination length can be specified in the loop using `_pd_instr_var_illum_len`.

There are also several data items in the core CIF dictionary that should be present in the description of the instrument in a pdCIF. Use `_diffrn_radiation_probe` and `_diffrn_radiation_type` to specify the type of radiation used and `_diffrn_detector_type` to specify the detection type.

## 3.3.4.4. Observations and measurement conditions

The data items in these categories are as follows:

- (a) Part of PD\_DATA
- ```
_pd_data_point_id
_pd_meas_2theta_fixed
_pd_meas_2theta_range_min
_pd_meas_2theta_range_max
_pd_meas_2theta_range_inc
_pd_meas_2theta_scan
_pd_meas_angle_2theta
_pd_meas_angle_chi
_pd_meas_angle_omega
_pd_meas_angle_phi
_pd_meas_counts_total
_pd_meas_counts_background
_pd_meas_counts_container
_pd_meas_counts_monitor
_pd_meas_datetime_initiated
_pd_meas_detector_id
  → _pd_calib_detector_id
_pd_meas_point_id
_pd_meas_position
_pd_meas_rocking_angle
_pd_meas_step_count_time
_pd_meas_time_of_flight
```

### 3.3. CLASSIFICATION AND USE OF POWDER DIFFRACTION DATA

(b) PD\_MEAS\_INFO  
\_pd\_meas\_info\_author\_address  
\_pd\_meas\_info\_author\_email  
\_pd\_meas\_info\_author\_fax  
\_pd\_meas\_info\_author\_name  
\_pd\_meas\_info\_author\_phone  
\_pd\_meas\_intensity\_total  
\_pd\_meas\_intensity\_background  
\_pd\_meas\_intensity\_container  
\_pd\_meas\_intensity\_monitor

(c) PD\_MEAS\_METHOD  
\_pd\_meas\_number\_of\_points  
\_pd\_meas\_rocking\_axis  
\_pd\_meas\_scan\_method  
\_pd\_meas\_special\_details  
\_pd\_meas\_units\_of\_intensity

The arrow (→) is a reference to a parent data item.

The item `_pd_data_point_id` identifies each entry in the list of measured, processed or simulated intensities. It is the only item in the PD\_DATA category that actually begins with the string `_pd_data_`, and is included here for convenience. If the list of intensities is split across several distinct loops, the role of this identifier may be adopted by other identifiers, such as `_pd_meas_point_id` in an isolated list of measured intensities.

The `_pd_meas_*` data items contain unprocessed measurements and documentation on the instrumental settings used for the measurements. Note that the choice of the data items used to represent this information is determined by the type of diffraction instrument, as well as how the measurement was conducted. This will be discussed further in Section 3.3.8. However, some `_pd_meas_*` data items are independent of the instrument type. For example, the use of `_pd_meas_datetime_initiated` is good practice, as is use of the `_pd_meas_info_author_*` data items. It is probably good practice to record the number of data points in `_pd_meas_number_of_points` for the benefit of people who might read the CIF, but there is no requirement that this item be present. This means that software should determine the number of data points directly when reading the CIF, rather than relying on the presence of a value for `_pd_meas_number_of_points`.

#### 3.3.5. Analysis

The categories relating to the information derived from the measurements are as follows:

*Processed intensities, positions and data processing* (§3.3.5.1)

PD\_DATA group  
PD\_DATA (items beginning with `_pd_proc_*`)  
PD\_PROC group  
PD\_PROC\_INFO  
PD\_PROC\_LS

*Simulated intensities and their positions* (§3.3.5.2)

PD\_DATA group  
PD\_DATA (items beginning with `_pd_calc_*`)  
PD\_CALC group  
PD\_CALC

*Diffraction peak table* (§3.3.5.3)

PD\_PEAK group  
PD\_PEAK  
PD\_PEAK\_METHOD

*Reflection assignments and intensities* (§3.3.5.4)

REFLN

In Rietveld and other studies, processed or simulated intensities are presented alongside measured values. This leads to the presence of both derived and measured values in the same category (PD\_DATA). However, the purposes of the data items that

refer to processed and simulated data points are made clear by the way they are named. Overall descriptions of processed and simulated intensity data are covered by the categories PD\_PROC\_INFO, PD\_PROC\_LS and PD\_CALC. The two categories PD\_PEAK and PD\_PEAK\_METHOD are used to describe lists of peak positions, which would typically be used to search and match powder profiles. Some additional data items relevant to the table of Bragg reflections are defined as additions to the existing REFLN category in the core CIF dictionary.

#### 3.3.5.1. Processed intensities, their positions and processing information

The data items in these categories are as follows:

(a) Part of PD\_DATA  
\_pd\_proc\_2theta\_corrected  
\_pd\_proc\_2theta\_range\_min  
\_pd\_proc\_2theta\_range\_max  
\_pd\_proc\_2theta\_range\_inc  
\_pd\_proc\_d\_spacing  
\_pd\_proc\_energy\_incident  
\_pd\_proc\_energy\_detection  
\_pd\_proc\_intensity\_net  
\_pd\_proc\_intensity\_total  
\_pd\_proc\_intensity\_bkg\_calc  
\_pd\_proc\_intensity\_bkg\_fix  
\_pd\_proc\_intensity\_incident  
\_pd\_proc\_intensity\_norm  
\_pd\_proc\_point\_id  
\_pd\_proc\_recip\_len\_Q  
\_pd\_proc\_wavelength

(b) PD\_PROC\_INFO  
\_pd\_proc\_info\_author\_address  
\_pd\_proc\_info\_author\_email  
\_pd\_proc\_info\_author\_fax  
\_pd\_proc\_info\_author\_name  
\_pd\_proc\_info\_author\_phone  
\_pd\_proc\_info\_data\_reduction  
\_pd\_proc\_info\_datetime  
\_pd\_proc\_info\_excluded\_regions  
\_pd\_proc\_info\_special\_details

(c) PD\_PROC\_LS  
\_pd\_proc\_ls\_background\_function  
\_pd\_proc\_ls\_peak\_cutoff  
\_pd\_proc\_ls\_pref\_orient\_corr  
\_pd\_proc\_ls\_prof\_R\_factor  
\_pd\_proc\_ls\_prof\_wR\_factor  
\_pd\_proc\_ls\_prof\_wR\_expected  
\_pd\_proc\_ls\_profile\_function  
\_pd\_proc\_ls\_special\_details  
\_pd\_proc\_ls\_weight

The pdCIF dictionary distinguishes between values that are measured directly and values that are derived from these observations. For example, in a constant-wavelength instrument, diffraction intensities are recorded as a function of  $2\theta$ . One may derive  $d$ -space values from the  $2\theta$  values using the value of the wavelength and corrections for the  $2\theta$  zero-point error and the sample displacement. One may also derive a new set of data points from the observations, for example by summing adjacent data points when the increment between the data points is much smaller than is warranted by the peak widths. For peak searching and other non-quantitative purposes, the diffraction intensities may be smoothed or otherwise modified. Note that the unprocessed measurement values are retained using the data items `_pd_meas_*`. Since the original measurements are still available, modifications like these do not result in the loss of the original data. In fact, by placing processed values in multiple blocks, a single CIF may contain measurements that have been processed in more than one way.

It is good practice to use the `_pd_proc_info_author_*` and `_pd_proc_info_datetime` data items. It is also a good idea to