## Precise lattice parameter

variance, weight, mean, weighted mean
> standard deviation ( $\sigma$ ) ; a measure of how spread out numbers are
$>$ variance $\left(\sigma^{2}\right)$; the average of the squared differences from the mean
(square of expected error)

$$
\sigma^{2}=\frac{1}{n-1} \sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)^{2}
$$

$>$ weight $\left(\omega_{\mathrm{i}}\right) \quad w_{i}=\frac{1}{\sigma_{i}^{2}}$.

$$
\sigma^{2}=\frac{n}{n-1} \frac{\sum_{i=1}^{n} w_{i}\left(x_{i}-\bar{x}\right)^{2}}{\sum_{i=1}^{n} w_{i}}
$$

$>$ weighted mean $\bar{x}=\frac{\sum_{i=1}^{n} w_{i} x_{i}}{\sum_{i=1}^{n} w_{i}}$,
$>$ Interpolation
$\checkmark$ connect the data-dots.
$\checkmark$ If data is reliable, we can plot it and connect the dots.

Depicting the trend in the data variation by assigning a single function to represent the data across its entire range



Interpolation


The goal is to identify the coefficients 'a' and ' $b$ ' such that $f(x)$ 'fits' the data well.

## Linear curve fitting, linear regression

A straight line function $f(x)=a x+b$
How can we pick the coefficients that best fits the line to the data?

First question: What makes a particular straight line a 'good' fit?
> Square the distance.
> Denote data values as ( $\mathrm{x}, \mathrm{y}$ ) and points on the fitted line as $(x, f(x))$.
$>$ Sum the error at the four data points.


The 'best' line has minimum error between line and data points.
least square minimization
the square of the error is minimized.


> Just as was the case for linear regression;
> How can we pick the coefficients that best fit the curve to the data?
$>$ The curve that gives minimum error between data $\rightarrow$ fit is 'best'.
>Quantify the error for these two second order curves...
$\checkmark$ Add up the length of all the red and blue vertical lines.
$\checkmark$ pick curve with minimum total error.




## Linear least square

4 measurements (observations)
2 unknown parameters
More equations than the \# of unknowns
$\rightarrow$ There are no values of $\beta_{1}$ and $\beta_{2}$ that satisfy the equations exactly.
$\rightarrow$ can get the $\beta_{1}$ and $\beta_{2}$ that satisfy the equations as much as possible (best straight line thru the points).
$\rightarrow$ best fit $\equiv$ values of $\beta_{1}$ and $\beta_{2}$ that minimizes $\sum \varepsilon_{i}^{2}$ when the residual (error) $\varepsilon_{i}=y-\beta_{1}-\beta_{2} x$.
$S\left(\beta_{1}, \beta_{2}\right)=\left[6-\left(\beta_{1}+1 \beta_{2}\right)\right]^{2}+\left[5-\left(\beta_{1}+2 \beta_{2}\right)\right]^{2}+\left[7-\left(\beta_{1}+3 \beta_{2}\right)\right]^{2}+\left[10-\left(\beta_{1}+4 \beta_{2}\right)\right]^{2}$
$\begin{aligned} & \text { For errors to } \\ & \text { be minimum }\end{aligned} \frac{\partial S}{\partial \beta_{1}}=2 \beta_{1}+5 \beta_{2}-14=0$

$$
\frac{\partial S}{\partial \beta_{2}}=10 \beta_{1}+30 \beta_{2}-77=0
$$

$\beta_{1}=3.5, \quad \beta_{2}=1.4 \rightarrow y=1.4 x+3.5$
$S(3.5,1.4)=1.1^{2}+(-1.3)^{2}+(-0.7)^{2}+(0.9)^{2}=4.2$

## Least square fitting


$>$ a way of finding the best curve to fit a given set of observations
$>$ it gives the best values of the constants in the equation selected.
$\checkmark \mathrm{q}$ is a function of 3 variables $\mathrm{x}, \mathrm{y}$ and z .
$\checkmark$ measurement of $q$ at various values of $x, y$ and $z$
$\checkmark 3$ unknown parameters $a, b$ and $c$
$\rightarrow$ With only 3 measurements at various $x, y$ and $z, 3$ equations can be uniquely solved for $\mathrm{a}, \mathrm{b}$ and c .
$>$ When number of measurements $>3$,
$\rightarrow$ (1) We can use only 3 measurements (equations) to solve for $a, b$, and $c$.
$\rightarrow$ (2) We can get more accurate values of $a, b$ and $c$ by taking advantage of the redundancy of the data; the best line that fits the experimental points.

$$
q_{j}=a x_{j}+b y_{j}+c z_{j} \quad(\mathrm{j}>3)
$$

For every measurement $\mathrm{q}_{j}$ the error $\mathrm{E}_{\mathrm{j}}$ is given by $E_{j}=a x_{j}+b y_{j}+c z_{j}-q_{j}$

The sum of the squares of errors for all $q_{j}$ must be minimum w.r.t. unknowns.

$$
\sum_{j} E_{j}^{2}=\sum_{j}\left(a x_{j}+b y_{j}+c z_{j}-q_{j}\right)^{2}
$$

## Linear least square analysis

$\sum_{j} E_{j}^{2}=\sum_{j}\left(a x_{j}+b y_{j}+c z_{j}-q_{j}\right)^{2} \quad$ must be minimum w.r.t. unknowns.
At minimum, $\frac{\partial \sum_{j} E_{j}^{2}}{\partial a}=\frac{\partial \sum_{j} E_{j}^{2}}{\partial b}=\frac{\partial \sum_{j} E_{j}^{2}}{\partial c}=0$
$\frac{\partial \sum_{j} E_{j}^{2}}{\partial a}=2 \sum_{j} x_{j}\left(a x_{j}+b y_{j}+c z_{j}-q_{j}\right)=0$

$$
\begin{aligned}
& a \sum_{j} x_{j}^{2}+b \sum_{j} x_{j} y_{j}+c \sum_{j} x_{j} z_{j}-\sum_{j} q_{j} x_{j}=0 \\
& a \sum_{j} x_{j} y_{j}+b \sum_{j} y_{j}^{2}+c \sum_{j} y_{j} z_{j}-\sum_{j} q_{j} y_{j}=0 \\
& a \sum_{j} x_{j} z_{j}+b \sum_{j} y_{j} z_{j}+c \sum_{j} z_{j}^{2}-\sum_{j} q_{j} z_{j}=0
\end{aligned}
$$

3 equations can be solved for 3 unknowns.

```
n measurements
2 unknown parameters
```

$$
\begin{aligned}
& y_{i}=m x_{i}+c \\
& y_{1}=m x_{1}+c \\
& y_{2}=m x_{2}+c \\
& \text {----- } \\
& y_{n}=m x_{n}+c \\
& \begin{array}{c}
\left(\begin{array}{cc}
x_{1} & 1 \\
x_{2} & 1 \\
. . & . \\
x_{n} & 1
\end{array}\right) \\
\mathbf{A} \\
\end{array}\binom{m}{c}=\underset{\left(\begin{array}{c}
y_{1} \\
y_{2} \\
. \\
y_{n}
\end{array}\right)}{\mathbf{b}} \quad \begin{array}{c}
\mathbf{A x}=\mathbf{b} \\
\downarrow \\
\left(\mathbf{A}^{\top} \mathbf{A}\right) \mathbf{x}=\mathbf{A}^{\top} \mathbf{b}
\end{array} \\
& \text { observation equation }
\end{aligned}
$$

Least square solution is that which minimizes the sum of squares of residuals of the observation equations.
$w_{i}$; inversely proportional to the (expected error) ${ }^{2}$ of each observation equation
W weight matrix

$$
\begin{gathered}
\text { WAx }=W b \\
\downarrow \\
\left(A^{\top} W A\right) x=\left(A^{\top} W\right) b
\end{gathered}
$$

n linear equations \& m unknown parameters

| $a_{11} x_{1}+a_{12} x_{2}+\ldots+a_{1 m} x_{m}=y_{1}$ |
| :--- |
| $a_{21} x_{1}+a_{22} x_{2}+\ldots+a_{2 m} x_{m}=y_{2}$ |
| $\ldots$ |
| $a_{n 1} x_{1}+a_{n 2} x_{2}+\ldots+a_{n m} x_{m}=y_{n}$ |\(\quad \mathbf{A}=\left(\begin{array}{cccc}a_{11} \& a_{12} \& ··· \& a_{1 m} <br>

a_{21} \& a_{22} \& ··· \& a_{2 m} <br>
··· \& ··· \& ··· \& ··· <br>
a_{n 1} \& a_{n 2} \& ··· \& a_{n m}\end{array}\right) ; \mathbf{x}=\left($$
\begin{array}{c}x_{1} \\
x_{2} \\
\ldots \\
x_{m}\end{array}
$$\right) ; \mathbf{y}=\left($$
\begin{array}{c}y_{1} \\
y_{2} \\
\ldots \\
y_{n}\end{array}
$$\right)\)

When $n>m$, vector $x$ can be found, which will be the best solution for all n existing equations using the least square technique.

$$
\begin{aligned}
& a_{11} x_{1}+a_{12} x_{2}+\ldots+a_{1 m} x_{m}-y_{1}=\varepsilon_{1} \\
& a_{21} x_{1}+a_{22} x_{2}+\ldots+a_{2 m} x_{m}-y_{2}=\varepsilon_{2} \\
& \ldots \\
& a_{n 1} x_{1}+a_{n 2} x_{2}+\ldots+a_{n m} x_{m}-y_{n}=\varepsilon_{n}
\end{aligned}
$$

Find the minimum of

$$
\Phi\left(x_{1}, x_{2}, \ldots x_{m}\right)=\sum_{i=1}^{n} \varepsilon_{i}^{2}
$$

The best solution is found by calculating the minimum condition of

$$
\Phi\left(x_{1}, x_{2}, \ldots x_{m}\right)=\sum_{i=1}^{n} \varepsilon_{i}^{2}
$$

$\boldsymbol{\sim} |$| $a_{11} x_{1}+a_{12} x_{2}+\ldots+a_{1 m} x_{m}-y_{1}=\varepsilon_{1}$ |
| :--- |
| $a_{21} x_{1}+a_{22} x_{2}+\ldots+a_{2 m} x_{m}-y_{2}=\varepsilon_{2}$ |
| $\ldots$ |
| $a_{n 1} x_{1}+a_{n 2} x_{2}+\ldots+a_{n m} x_{m}-y_{n}=\varepsilon_{n}$ |

$$
\begin{array}{|l}
\begin{array}{l}
\frac{\partial \Phi\left(x_{1}, x_{2}, \ldots, x_{m}\right)}{\partial x_{1}}=0 \\
\frac{\partial \Phi\left(x_{1}, x_{2}, \ldots, x_{m}\right)}{\partial x_{2}}=0
\end{array} \\
\frac{\partial \Phi\left(x_{1}, x_{2}, \ldots, x_{m}\right)}{\partial x_{m}}=0
\end{array} \begin{aligned}
& \begin{array}{l}
x_{1} \sum_{i=1}^{n} a_{i 1}^{2}+x_{2} \sum_{i=1}^{n} a_{i 1} a_{i 2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i 1} a_{i n}=\sum_{i=1}^{n} a_{i 1} y_{i} \\
x_{1} \sum_{i=1}^{n} a_{i 2} a_{i 1}+x_{2} \sum_{i=1}^{n} a_{i 2}^{2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i 2} a_{i m}=\sum_{i=1}^{n} a_{i 2} y_{i} \\
\ldots \\
x_{1} \sum_{i=1}^{n} a_{i m} a_{i 1}+x_{2} \sum_{i=1}^{n} a_{m} a_{i 2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i m}^{2}=\sum_{i=1}^{n} a_{i m} y_{i}
\end{array} \\
& \text { ex) } \mathrm{m}=2, \frac{\partial S}{\partial x_{1}}=a_{11}^{2} x_{1}+a_{11} a_{12} x_{2}-a_{11} y_{1}=0
\end{aligned}
$$

Linear least square analysis

$$
\begin{aligned}
& x_{1} \sum_{i=1}^{n} a_{i 1}^{2}+x_{2} \sum_{i=1}^{n} a_{i 1} a_{i 2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i 1} a_{i n}=\sum_{i=1}^{n} a_{i 1} y_{i} \\
& x_{1} \sum_{i=1}^{n} a_{i 2} a_{i 1}+x_{2} \sum_{i=1}^{n} a_{i 2}^{2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i 2} a_{i m}=\sum_{i=1}^{n} a_{i 2} y_{i} \\
& x_{1} \sum_{i=1}^{n} a_{i m} a_{i 1}+x_{2} \sum_{i=1}^{n} a_{m} a_{i 2}+\ldots+x_{m} \sum_{i=1}^{n} a_{i m}^{2}=\sum_{i=1}^{n} a_{i m} y_{i} \\
& \left(\begin{array}{cccc}
a_{11} & a_{21} & \ldots & a_{n 1} \\
a_{12} & a_{22} & \ldots & a_{n 2} \\
\ldots & \ldots & \ldots & \ldots \\
a_{1 m} & a_{2 m} & \ldots & a_{n m}
\end{array}\right)\left(\begin{array}{cccc}
a_{11} & a_{12} & \ldots & a_{1 m} \\
a_{21} & a_{22} & \ldots & a_{2 m} \\
\ldots & \ldots & \ldots & \ldots \\
a_{n 1} & a_{n 2} & \ldots & a_{n m}
\end{array}\right) \mathrm{X}=\left(\mathbf{A}^{\top} \mathbf{y}\right)
\end{aligned}
$$

$1 / \mathrm{d}^{2}=\left(\mathrm{h}^{2}+\mathrm{k}^{2}\right) / \mathrm{a}^{2}+\mathrm{l}^{2} / \mathrm{c}^{2}$
d-value of a tetragonal elementary cell

(100)

$$
\begin{equation*}
1 / d^{2}=\left(h^{2}+k^{2}\right) / a^{2}+l^{2} / c^{2} \tag{110}
\end{equation*}
$$

$1 / d^{2}=\left(h^{2}+k^{2}\right) / a^{2}+l^{2} / c^{2}$
$1 / d^{2}=\left(h^{2}+k^{2}\right) / a^{2}+l^{2} / c^{2}$
(200)
$1 / d^{2}=\left(h^{2}+k^{2}\right) / a^{2}+l^{2} / c^{2}$
(210)
$1 / d^{2}=\left(h^{2}+k^{2}\right) / a^{2}+l^{2} / c^{2}$


## Evaluation of $\mathrm{F}_{\mathrm{N}}$

- 50 possible lines
- 42 have observable intensity
- $2 \theta_{\text {calk }}$; calculated $2 \theta$ values based on the known lattice parameters
- $\Delta 2 \theta=2 \theta_{\text {calc }}-2 \theta_{\text {obs }}$
$F_{N}=\frac{1}{\Delta \Delta 2 \theta} \frac{N}{N_{\text {poss }}}$
SS figure of merit

$N$; \# experimental lines (peaks) considered
$N_{\text {possi }}$ \# possible, space group-allowed diffraction lines
(Cont'd.)

| No. | $2 \theta_{\text {calc }}$ | $I^{\text {rel }}$ | $d(\AA)$ | $2 \theta_{\text {obs }}$ | $\Delta 2 \theta$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 40 | 80.253 | 5 | 1.195 | 80.238 | -0.015 |
| 41 | 81.772 | 9 | 1.177 | 81.776 | 0.004 |
| 42 | 82.025 | 1 | 1.174 | - | - |
| 43 | 84.002 | 1 | 1.151 | - | - |
| 44 | 84.923 | 1 | 1.141 | 84.916 | -0.007 |
| 45 | 85.773 | 3 | 1.132 | 85.770 | -0.003 |
| 46 | 8.246 | 6 | 1.106 | 88.249 | 0.003 |
| 47 | 89.114 | 2 | 1.098 | 89.110 | -0.004 |
| 48 | 90.002 | 1 | 1.089 | - | - |
| 49 | 90.734 | 4 | 1.082 | 90.720 | -0.014 |
| 50 | 91.720 | 1 | 1.073 | 91.726 | 0.006 |
| Avg $42 \theta$ | 0.0066 | 0.0066 | 0.0083 | 0.0104 | 0.0084 |
| $N_{\text {poss }}$ | 50 | 50 | 40 | 30 | 20 |
| $N_{\text {obs }}$ | 50 | 42 | 35 | 27 | 18 |
| FOM | 151.6 | 127.4 | 104.9 | 86.2 | 107.3 |

${ }^{a}$ Wavelength $=1.54056$.

$$
F_{N}=\frac{1}{|\Delta 2 \theta|} \frac{N}{N_{\text {poss }}}
$$

CHAN PARK, MSE, SNU Spring-2022 Crystal Structure Analyses

Accurate vs. Precise


High accuracy High precision


Low accuracy
High precision


High accuracy Low precision


Low accuracy Low precision


Measured mean Real value

## Precision - reproducibility

Accuracy - approach to the "true" value
> Precision ; the degree to which further measurements show the same or similar results
> Accuracy; the degree of conformity of a measured quantity to its true value.
> How to prepare powder?
$\checkmark$ Grind in mortar \& pestle (wet or dry)
$\checkmark$ Crush (percussion mill)
$\checkmark$ Cryo-grind
$\checkmark$ Micronising mill
$\checkmark$ Treatments/separations

## Mounting specimen

$\checkmark$ Front, side, back-loaded powders
$\checkmark$ Films \& disks
$\checkmark$ Diluents \& dispersants
$\checkmark$ Adhesives
$\checkmark$ Reactive samples (windows)
$\checkmark$ Capillaries
$\checkmark$ Odd shapes
$\checkmark$ Zero-background holders (ZBH)

Panalytyical


BRUKER


BRUKER

www2.arnes.si/~sgszmera1/html/xrd/preparation2.html

Spray drying - can eliminate preferred orientation


## - Specimen preparation

$\checkmark$ Specimen should represent the bulk.
$\checkmark$ 1um < Particle size < 15um
$\checkmark$ Method should not distort the lattice
$\checkmark$ Avoid solid state reactions.

Diffractometer specimen
requirements
$\checkmark$ Flat specimen surface
$\checkmark$ Smooth specimen surface
$\checkmark$ Area greater than that irradiated by beam
$\checkmark$ Specimen support gives zero diffraction or zero contribution.

## Common problems in

 specimen preparation$\checkmark$ Preferred orientation
$\checkmark$ Crystallite statistics
$\checkmark$ Crystallite \& particle size
$\checkmark$ Particle morphology
$\checkmark$ Specimen configurations
$\checkmark$ Crystallite perfection
$\checkmark$ Absorption effects

## 5 fingers of quartz



## > Typical error windows

$\checkmark$ Debye Scherrer camera $\pm \Delta 2 \theta=0.1^{\circ}$
$\checkmark$ diffractometer $\pm \Delta 2 \theta=0.05^{\circ}$
$\checkmark$ diffractometer (internal standard corrected) $\pm \Delta 2 \theta=0.01^{\circ}$
$\checkmark$ diffractometer (internal standard corrected \& peaks profile fitted) $\pm \Delta 2 \theta=0.005^{\circ}$
$\Delta 2 \theta$ - d relationship is non-linear
$\checkmark$ Low angle (low $2 \theta$, large $d$-value) lines have large error.

| Table 12.2. Errors in $\boldsymbol{d}$-Values Resulting from Fixed 2 $\boldsymbol{\theta}$ Errors |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| $d$ | $2 \theta$ | $\pm \Delta 2 \theta$ <br> (degrees) | $\pm \Delta d$ <br> $(\AA)$ | $\pm \Delta 2 \theta$ <br> $($ degrees $)$ | $\pm \Delta d$ <br> $(\AA)$ |
| $(\AA)$ | 17.73 | 0.1 | 0.04 | 0.05 | 0.014 |
| 5 | 22.20 | 0.1 | 0.02 | 0.05 | 0.008 |
| 4 | 29.76 | 0.1 | 0.01 | 0.05 | 0.005 |
| 3 | 45.30 | 0.1 | 0.004 | 0.05 | 0.002 |
| 2 | 61.80 | 0.1 | 0.002 | 0.05 | 0.0011 |
| 1.5 | 100.76 | 0.1 | 0.0007 | 0.05 | 0.0004 |

$\Delta 2 \theta \& \Delta \mathrm{~d}$


Spectral dispersion; peak breadth increases with $2 \theta$


Line (peak) profile analysis


Smith \& Snyder FOM

$$
F_{N}=\frac{1}{|\Delta 2 \theta|} \frac{N}{N_{\text {poss }}}
$$

De Wolff FOM

$$
M_{20}=\frac{\mathbf{d}_{20}^{* 2}}{2\left|\Delta \mathbf{d}^{* 2}\right|} \frac{1}{N_{\text {poss }}}
$$

Jenkins \& Snyder, page 316

Intensity FOM

$$
R_{I}=\sum \frac{I_{\mathrm{obs}}-I_{\mathrm{calc}}}{I_{\mathrm{obs}}}
$$

## Precise lattice parameter

> Composition of a solid solution
> Thermal expansion coefficients
> Solubility limit
> -----

In cubic,
$>$ measure $\theta$ for $\mathrm{hkl} . \rightarrow$ determine $\mathrm{d} . \rightarrow$ calculate a .

$>$ Precision in $d$ or a depends on precision in $\sin \theta$, not on $\theta$.
$>$ Accurate value of $\sin \theta \leftarrow$ measurement of $\theta$ near $90^{\circ}$
$>$ Differentiation of Bragg's law w.r.t. $\theta \rightarrow \Delta \mathrm{d} / \mathrm{d}=-\Delta \theta \cot \theta=\Delta \mathrm{a} / \mathrm{a}$
$>\Delta$ a caused by a given $\Delta \theta \rightarrow$ zero as $\theta \rightarrow 90^{\circ}$.
$\Rightarrow$ precision when using peaks in $2 \theta \sim 180^{\circ}$
$\Rightarrow$ precision of 0.001 Å possible

Internal standard method
$\checkmark \Delta 2 \theta=0.01$ Sample +
Standard powder

$\checkmark$ Can eliminate both displacement and transparency errors.
$\checkmark$ Both instrument \& specimen errors are corrected.

## External standard method

$\checkmark \Delta 2 \theta=0.25$
$\checkmark$ Do not correct displacement errors.

$\checkmark$ With ZBK holder, can correct all errors $\rightarrow \Delta 2 \theta=0.01$.

Zero background holder method
Standard powder
Sample


## Profile fit peak positions

## Zero Background Holder

$>$ A single crystal of quartz that is cut and polished in an orientation such that it produces no diffraction peaks.


Thick specimens
$\checkmark$ Good intensity...but problems defining depth (position)
$\square$
Thin specimens
$\checkmark$ No penetration depth effect (good position)... but low intensity


## Internal Standard Calibration

> Mix sample powder and silicon SRM.

Sample + Silicon SRM




이 Calibration curve

CHAN PARK, MSE, SNU Spring-2022 Crystal Structure Analyses



Errors removed by calibration

Effectiveness of Standards for the Correction of $2 \theta$ Errors

| Use of Standard | Type of Standard |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | None | External (2 $\theta$ ) | Internal (2 $\theta$ ) | $\begin{gathered} \text { ZBH } \\ (2 \theta) \end{gathered}$ | External (Intensity) |
| Instrument misalignment | No | Yes | Yes | Yes | (Yes) |
| Inherent aberrations | No | Yes | Yes | Yes | No |
| Specimen transparency | No | No | Yes | Yes | No |
| Specimen displacement | No | No | Yes | Yes | No |
| Instrument sensitivity | No | No | No | No | Yes |

$>$ None on a random instrument $=0.1^{\circ}$
> None on a well-aligned instrument $=0.05^{\circ}$
External standard method $=0.025^{\circ}$
$>$ Internal standard method $=0.01^{\circ}$
$>$ Zero background holder method $=0.01^{\circ}$
Profile fit peak positions $=0.005^{\circ}$

Calibration - quartz crystal

TABLE II. ZBH (mounted quartz crystal) calibrated data.


