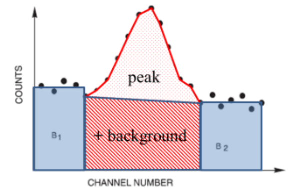
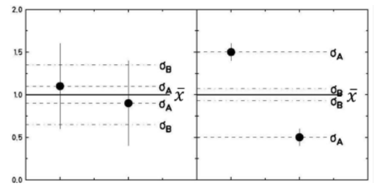
Error Analysis

- Quality of measurements (standard deviation, full with at half maximum FWHM)
- Statistical error: peak on top of a background

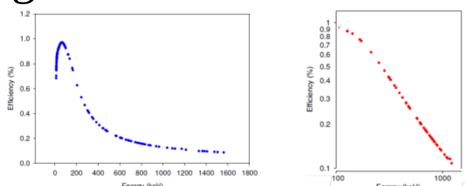


- Mean value and standard deviation (without and with errors of individual data points)

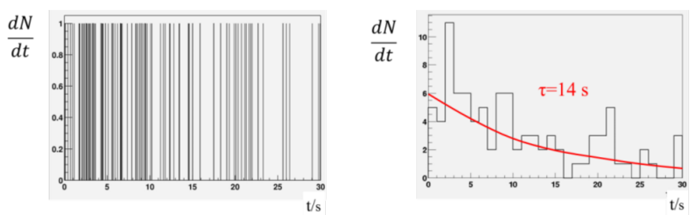
  Results for a limited number of measurements



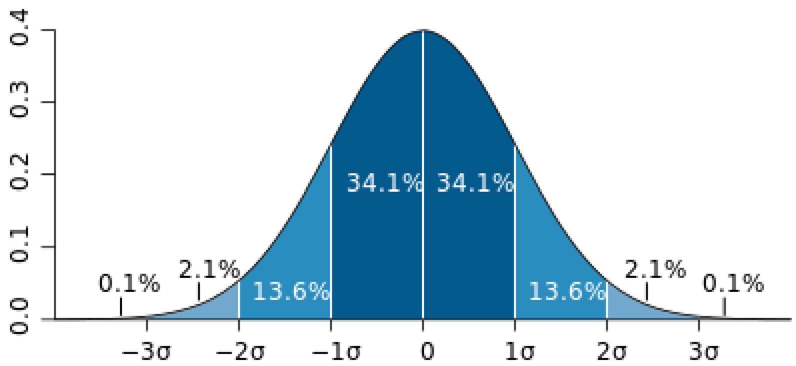
- Least-squares regression



- Decay curve measured with low statistics



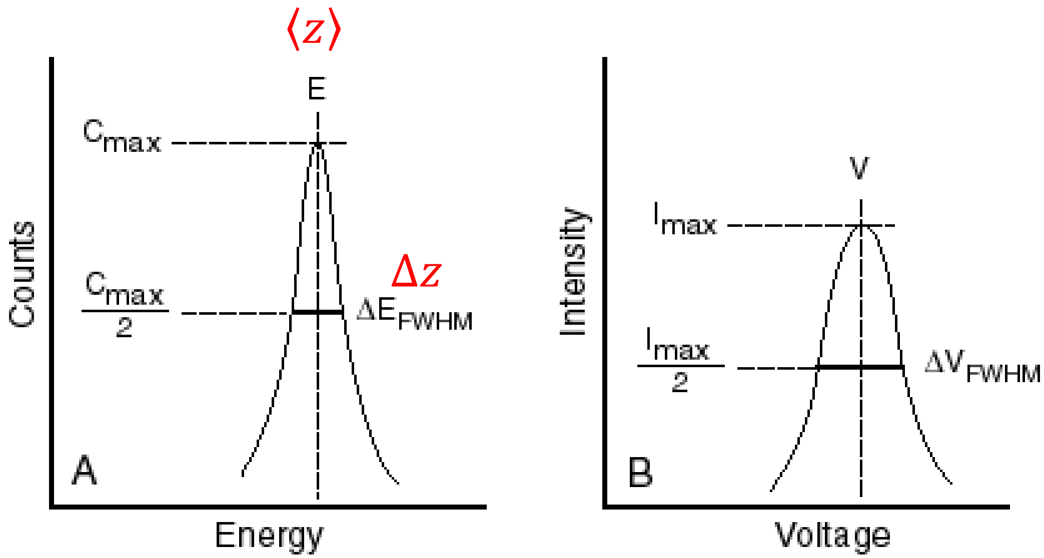
Gaussian Function



Gaussian function: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Full Width at Half Maximum: $FWHM = 2 \cdot \sqrt{2 \cdot ln2} \cdot \sigma$

Quality of Measurements: Resolution



Resolution is generally defined as 1 standard deviation (1σ) for a Gaussian distribution, or full width half maximum $(FWHM = \Delta z)$.

$$\sigma_z = FWHM/2.355$$

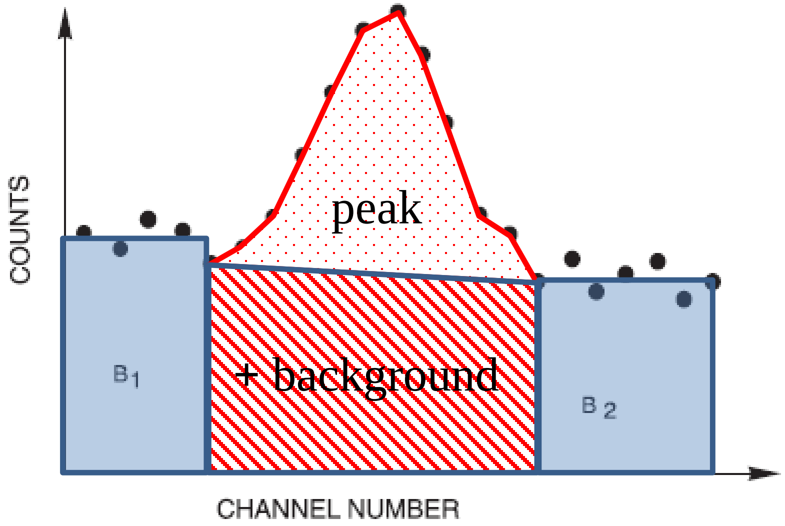
If the measurement is dominated by Poissonian fluctuations:

$$\frac{\sigma_z}{\langle z \rangle} = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$$

Fano factor F: fluctuations on N are reduced by correlation in the production of consecutive e-hole pairs. For Germanium detectors $F \sim 0.1$

$$\frac{\sigma_z}{\langle z \rangle} = \sqrt{\frac{F}{N}}$$

Statistical Error: Peak on top of Background



The area above the background represents the total counts between the vertical lines ***P*** minus the trapezoidal area ***B*** (red hatched). If the total counts are ***(P+B)*** and the end-points of the horizontal line are $\boldsymbol{B_1}$ and $\boldsymbol{B_2}$ (width of $B_1 + B_2$ = width of B), then the net area is given by:

$$P = (P + B) - B$$

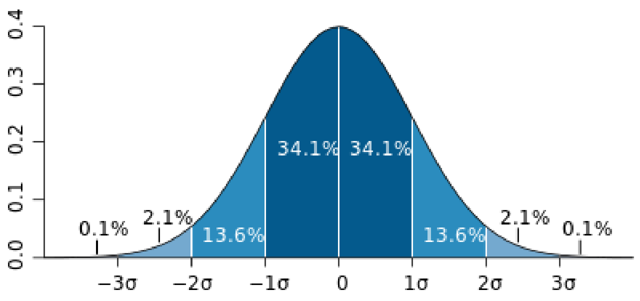
The ***standard deviation of ΔP*** is given by:

$$\Delta P = \sqrt{P + 2 \cdot B}$$

The mean value and the standard deviation

mean value $\bar{x} = \frac{1}{n} \sum_{1}^{n} x_i$

standard deviation $\sigma = \sqrt{\frac{\sum_1^n (x_i - \bar{x})^2}{n-1}}$



standard normal distribution

68.3% of data in the interval μ ± σ
94.4% of data in the interval μ ± 2σ
99.7% of data in the interval μ ± 3σ

weighted mean value $\bar{x} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$

weighted standard deviation $\sigma_w = \sqrt{\frac{\sum_{i=1}^n w_i \cdot (x_i - \bar{x})^2}{(n-1)/n \cdot \sum_{i=1}^n w_i}}$

The mean value and the standard deviation

weighted mean value $$\bar{x} = \frac{\sum_{i=1}^{n} w_i \cdot x_i}{\sum_{i=1}^{n} w_i}$$

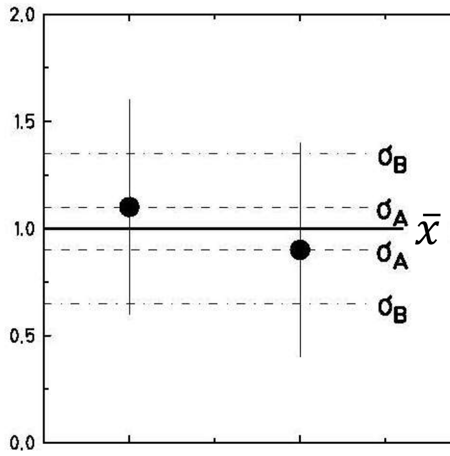
weighted standard deviation $$\sigma_w = \sqrt{\frac{\sum_{i=1}^{n} w_i \cdot (x_i - \bar{x})^2}{(n-1)/n \cdot \sum_{i=1}^{n} w_i}}$$

The weight $\mathbf{w_i}$ is given by the errors of the individual data $\mathbf{y_i \pm \sigma_i}$

weighted mean value $$\bar{x} = \frac{\sum_{i=1}^{n}(x_i/{\sigma_i}^2)}{\sum_{i=1}^{n}(1/{\sigma_i}^2)}$$

weighted standard deviation $$\sigma_A = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2/{\sigma_i}^2}{(n-1)/n \cdot \sum_{i=1}^{n}(1/{\sigma_i}^2)}}$$

Experimental standard deviation



Basic problems:

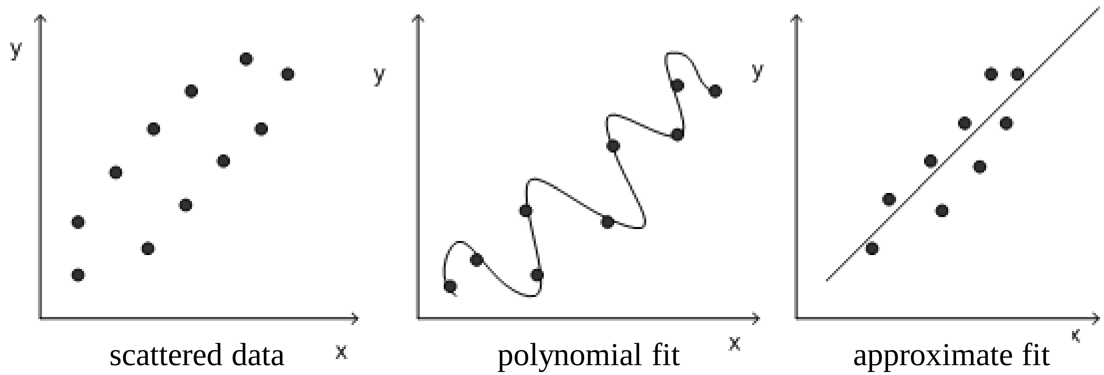
- The number of trial measurements are limited.
- The observed errors in these measurements include both random and systematic errors.
- For the error σ_A of the weighted mean $\bar{x}$ can yield unphysical values for very small samples.

weighted standard deviation $$\sigma_A = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2/\sigma_i{}^2}{(n-1)/n \cdot \sum_{i=1}^{n}(1/\sigma_i{}^2)}}$$

$$\sigma_B = \left[\sum_{i=1}^{n}(1/\sigma_i{}^2)\right]^{-1/2}$$

$$\boxed{\sigma = max(\sigma_A, \sigma_B)}$$

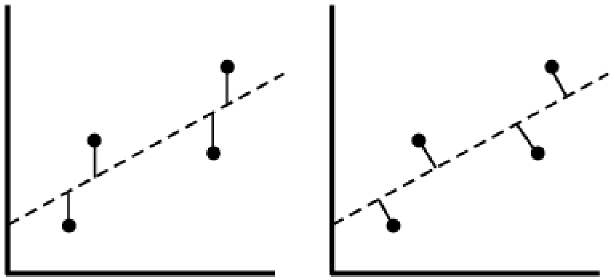
Least-Squares Regression



scattered data polynomial fit approximate fit

The principle of least squares is one of the popular methods for finding a curve fitting a given data.
Say (x_1,y_1), (x_2,y_2), ... (x_n,y_n), be ***n*** observations from an experiment. We are interested in finding a curve $y = f(x)$

Least-Squares Fit of a Straight Line



We are interested in finding a curve $\boxed{y = a + b \cdot x}$

We consider the sum of the squares of $$\chi^2(x_i, a, b) = \sum_{i=1}^{n} [y_i - (a + b \cdot x_i)]^2$$

We need to find a,b such that χ^2 is minimum. $\frac{d\chi^2}{da} = \frac{d\chi^2}{db} = 0$

$$\frac{d\chi^2}{da} = \sum_{i=1}^{n} 2[y_i - (a + b \cdot x_i)] = 0 \Longrightarrow n \cdot a + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\frac{d\chi^2}{db} = \sum_{i=1}^{n} 2 \cdot x_i [y_i - (a + b \cdot x_i)] = 0 \Longrightarrow a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i \cdot y_i$$

Least-Squares Fit of a Straight Line

$$\frac{d\chi^2}{da} = \sum_{i=1}^{n} 2[y_i - (a + b \cdot x_i)] = 0 \quad \Longrightarrow \quad n \cdot a + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\frac{d\chi^2}{db} = \sum_{i=1}^{n} 2 \cdot x_i[y_i - (a + b \cdot x_i)] = 0 \Longrightarrow \quad a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i \cdot y_i$$

In matrix form

$$\begin{vmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 \end{vmatrix} \begin{vmatrix} a \\ b \end{vmatrix} = \begin{vmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i \cdot y_i \end{vmatrix}$$

$$\begin{vmatrix} a \\ b \end{vmatrix} = \begin{vmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 \end{vmatrix}^{-1} \begin{vmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i \cdot y_i \end{vmatrix}$$

The 2x2 matrix inverse is

$$\begin{vmatrix} a \\ b \end{vmatrix} = \frac{1}{n \sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{vmatrix} \sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i \cdot y_i \\ n \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \end{vmatrix}$$

$$\boxed{\begin{aligned} \mathrm{a} &= \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i \cdot y_i}{n \sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \\ b &= \frac{n \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \end{aligned}}$$

Least-Squares Fit of a Straight Line

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i \cdot y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b = \frac{n \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

standard deviation:

$$\sigma_y^2 = \frac{1}{n-2} \sum_{i=1}^{n} [y_i - f(x_i)]^2$$

error propagation:

$$\sigma_a^2 = \sum \left(\frac{da}{dy_i}\right)^2 \sigma_y^2 = \frac{\sigma_y^2 \sum_{i=1}^{n} x_i^2}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\sigma_b^2 = \sum \left(\frac{db}{dy_i}\right)^2 \sigma_y^2 = \frac{n \cdot \sigma_y^2}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Example:

i	x_i	y_i
1	208	21,6
2	152	15,5
3	113	10,4
4	227	31,0
5	137	13,0
6	238	32,4
7	178	19,0
8	104	10,4
9	191	19,0
10	130	11,8
Σ	1678	184,1

Least-Squares Regression

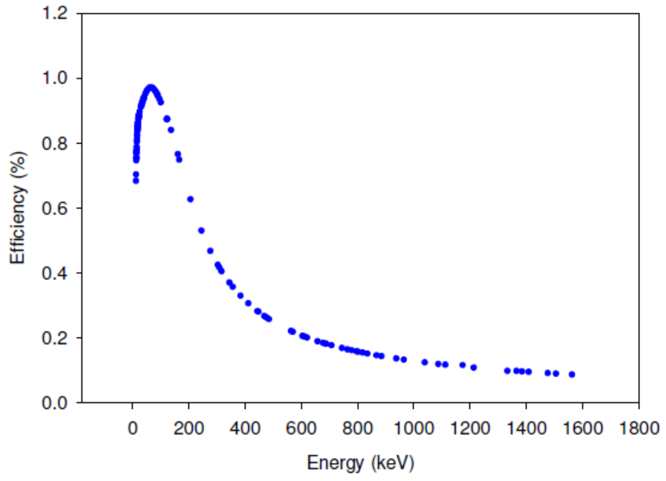
Linear models are not limited to being straight lines, but include wide range of shapes. For example

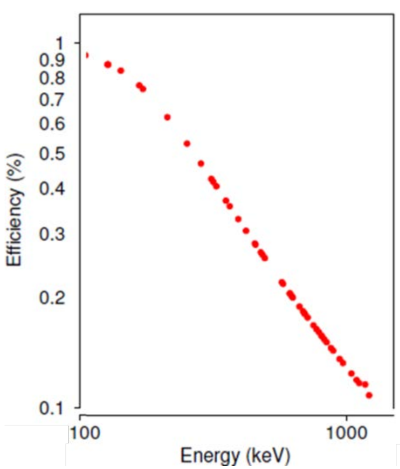
$$f(x;\vec{a}\,) = a_0 + a_1 \cdot ln(x)$$

Linearization of non-linear relation:

$$f(x) = a \cdot x^b \quad \Longrightarrow \quad log[f(x)] = log(a) + b \cdot log(x)$$

$$f(x) = a \cdot e^{b \cdot x} \quad \Longrightarrow \quad ln[f(x)] = ln(a) + b \cdot x$$





Weighted Least-Squares

$$\chi^2 = \sum_{i=1}^{n} w_i [y_i - f(x_i)]^2$$

Example: $\boxed{y = a + b \cdot x}$

$$\chi^2 = \sum_{i=1}^{n} w_i [y_i - (a + b \cdot x_i)]^2$$ with the weight given by the individual errors $w_i = {}^{1}/_{\sigma_i{}^2}$

Appendix

Assume we have data points that lie on a straight line:

$$y = \alpha + \beta x$$

Previously we showed that the solution for α and β is:

$$\alpha = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \text{ and } \beta = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

use the Propagation of Errors technique to estimate σ_α and σ_β.

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2 + \sigma_y^2 \left(\frac{\partial Q}{\partial y}\right)^2 + 2\sigma_{xy}\left(\frac{\partial Q}{\partial x}\right)\left(\frac{\partial Q}{\partial y}\right)$$

Assumed that each measurement is independent of each other:

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2 + \sigma_y^2 \left(\frac{\partial Q}{\partial y}\right)^2$$

$$\sigma_\alpha^2 = \sum_{i=1}^{n} \sigma_{y_i}^2 \left(\frac{\partial \alpha}{\partial y_i}\right)^2 = \sigma^2 \sum_{i=1}^{n} \left(\frac{\partial \alpha}{\partial y_i}\right)^2$$

$$\frac{\partial\alpha}{\partial y_i} = \frac{\partial}{\partial y_i}\frac{\sum_{i=1}^{n} y_i \sum_{j=1}^{n} x_j^2 - \sum_{i=1}^{n} x_i y_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{\sum_{j=1}^{n} x_j^2 - x_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$\sigma_\alpha^2 = \sigma^2 \sum_{i=1}^{n}\left(\frac{\sum_{j=1}^{n} x_j^2 - x_i \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\right)^2 = \sigma^2 \sum_{i=1}^{n}\left(\frac{(\sum_{j=1}^{n} x_j^2)^2 + x_i^2(\sum_{j=1}^{n} x_j)^2 - 2x_i \sum_{j=1}^{n} x_j \sum_{j=1}^{n} x_j^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2}\right)$$

$$\sigma_\alpha^2 = \sigma^2 \frac{n(\sum_{j=1}^{n} x_j^2)^2 + \sum_{i=1}^{n} x_i^2(\sum_{j=1}^{n} x_j)^2 - 2(\sum_{j=1}^{n} x_j)^2 \sum_{j=1}^{n} x_j^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2} = \sigma^2 \frac{n(\sum_{j=1}^{n} x_j^2)^2 - \sum_{i=1}^{n} x_i^2(\sum_{j=1}^{n} x_j)^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2}$$

$$= \sigma^2 \sum_{j=1}^{n} x_j^2 \frac{n\sum_{j=1}^{n} x_j^2 - (\sum_{j=1}^{n} x_j)^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2}$$

$$\sigma_\alpha^2 = \sigma^2 \frac{\sum_{j=1}^{n} x_j^2}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

variance in the intercept

We can find the variance in the slope (β) using exactly the same procedure:

$$\sigma_\beta^2 = \sum_{i=1}^{n} \sigma_{y_i}^2 \left(\frac{\partial\beta}{\partial y_i}\right)^2 = \sigma^2 \sum_{i=1}^{n} \left(\frac{\partial\beta}{\partial y_i}\right)^2 = \sigma^2 \sum_{i=1}^{n} \left(\frac{\partial}{\partial y_i} \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \right)^2 = \sigma^2 \sum_{i=1}^{n} \left(\frac{n x_i - \sum_{j=1}^{n} x_j}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \right)^2$$

$$= \sigma^2 \frac{n^2 \sum_{j=1}^{n} x_j^2 + n(\sum_{j=1}^{n} x_j)^2 - 2n \sum_{i=1}^{n} x_i \sum_{j=1}^{n} x_j}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2} = \sigma^2 \frac{n^2 \sum_{j=1}^{n} x_j^2 - n(\sum_{j=1}^{n} x_j)^2}{(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)^2}$$

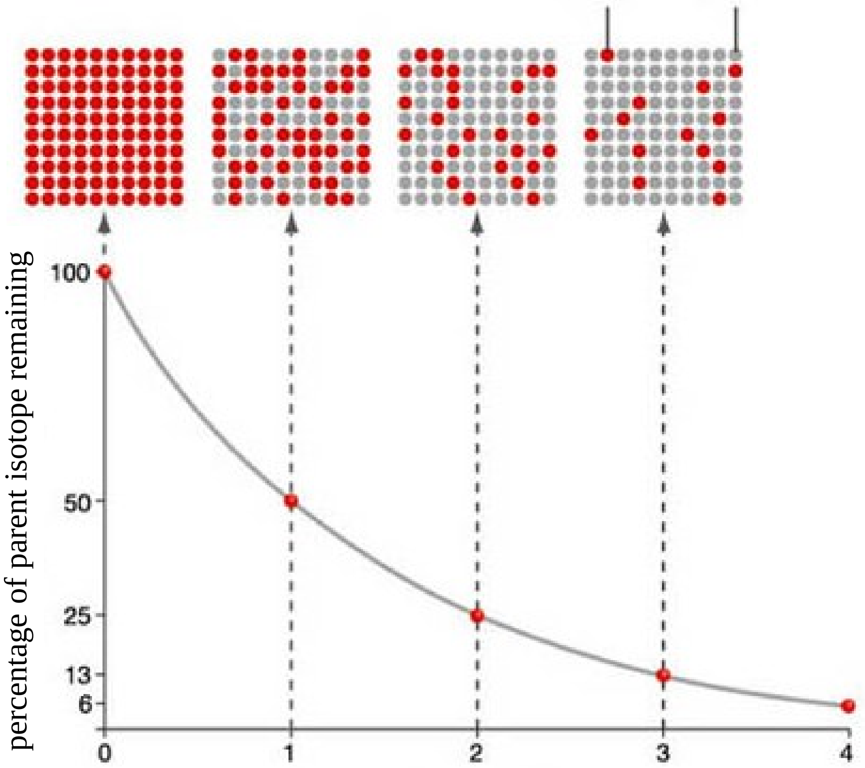
$$\sigma_\beta^2 = \frac{n\sigma^2}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

variance in the slope

Radioactive Decay Activities

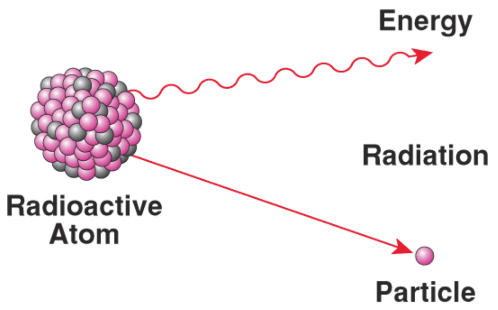
parent isotope (red)

daughter isotope (grey)



percentage of parent isotope remaining

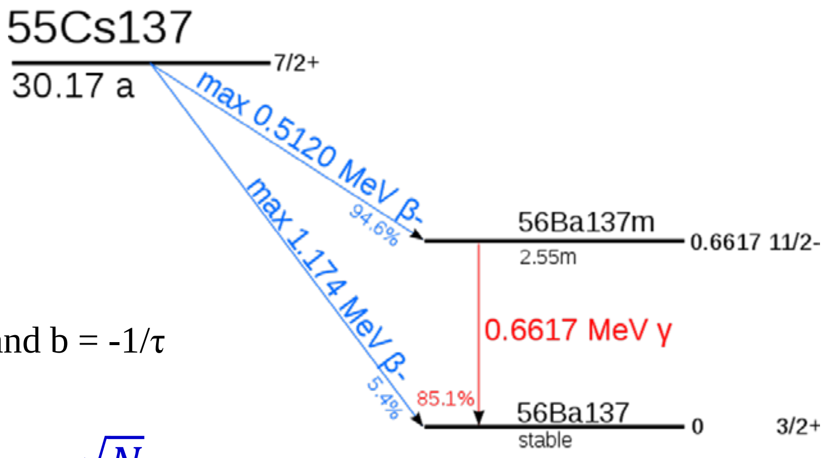
number of half-lives



Radioactive Decay Activities

$$N = N_0 \cdot e^{-t/\tau}$$

$$\ln N = \ln N_0 - {}^{t}/_{\tau}$$



$y = a + bt$ with y = ln(N), a = ln(N0) and b = -1/τ

uncertainty of decay counts N (Poisson): $\sigma_N = \sqrt{N}$

At time progresses, N, is getting smaller and smaller.
What is the uncertainty of ln(N) ?

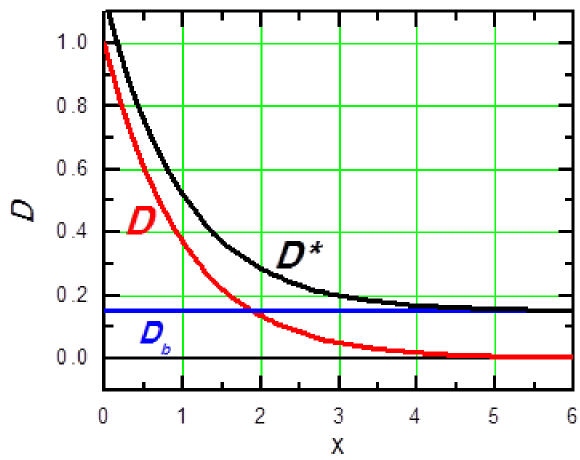
$$\sigma_{lnN} = \left|\frac{dln(N)}{dN}\right| \sigma_N = \frac{\sigma_N}{N} = \frac{1}{\sqrt{N}}$$

The Background Radiation

Background radiation is the radiation constantly present in the natural environment of the Earth which is emitted by natural and artificial sources (**D = N**)

$$D^{*}=D+D_{B} \Rightarrow D=D^{*}-D_{B}$$

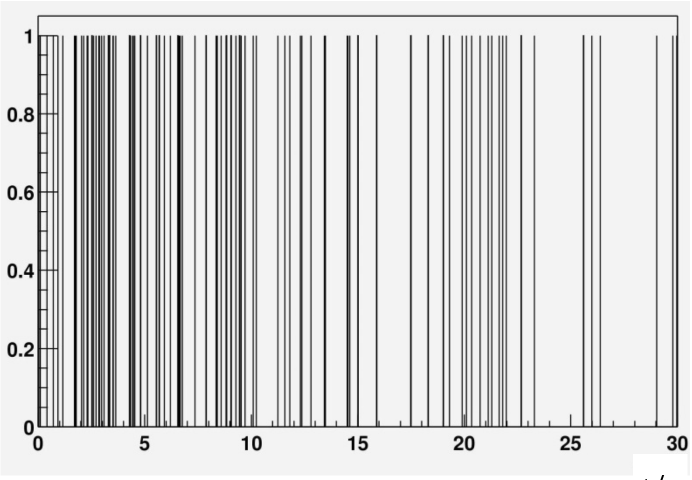
- Sources in the Earth
- Sources from outer space, such as cosmic rays
- Sources in the atmosphere, such asthe radon gas released fromthe Earth's crust



Attention: Error Propagation !

Representation of Decay Curve

$\frac{dN}{dt}$



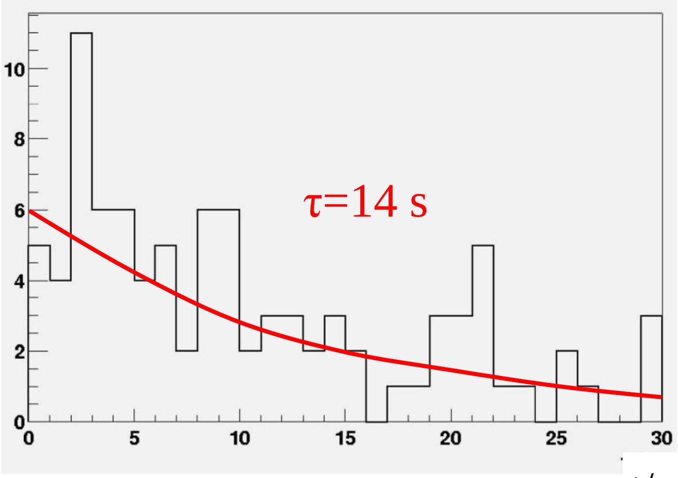
t/s

$$N = N_0 \cdot e^{-\lambda \cdot t}$$

$$\frac{dN}{dt} = -N_0 \cdot \lambda \cdot e^{-\lambda \cdot t}$$

lifetime: $\tau = {}^{1}/_{\lambda}$

$\frac{dN}{dt}$



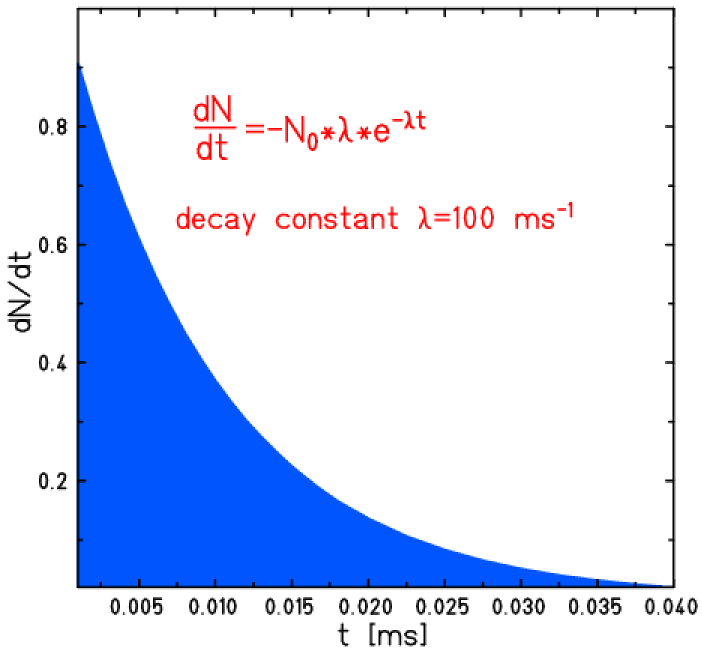
t/s

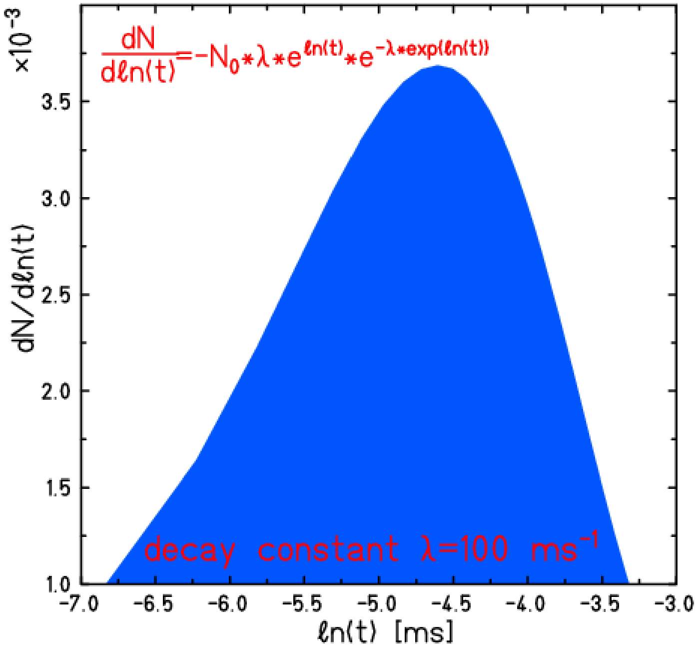
Binning of a data set is a pre-processing technique for data smoothing .

The content of **10 original channels** is now contained **in 1 new channel**

Representation of Decay Curve

The maximum likelihood estimate of the lifetime τ





example: $\ln\tau = -4.61$

$$\Theta \equiv \ell n\, t \;\rightarrow\; \frac{d\Theta}{dt} = \frac{1}{t} \qquad dt = t\cdot d\Theta$$

$$\frac{dN}{d\Theta} = -N_0 \cdot \lambda \cdot t \cdot e^{-\lambda t}$$

$$\frac{dN}{d\Theta} = -N_0 \cdot \lambda \cdot e^{\Theta} \cdot e^{-\lambda \cdot e^{\Theta}}$$

$$\frac{d^2N}{d\Theta^2} = -N_0 \cdot \lambda \cdot \left[e^{\Theta} \cdot e^{-\lambda \cdot e^{\Theta}} + e^{\Theta} \cdot e^{-\lambda \cdot \Theta} \cdot \left(-\lambda \cdot e^{\Theta} \right) \right] = 0$$

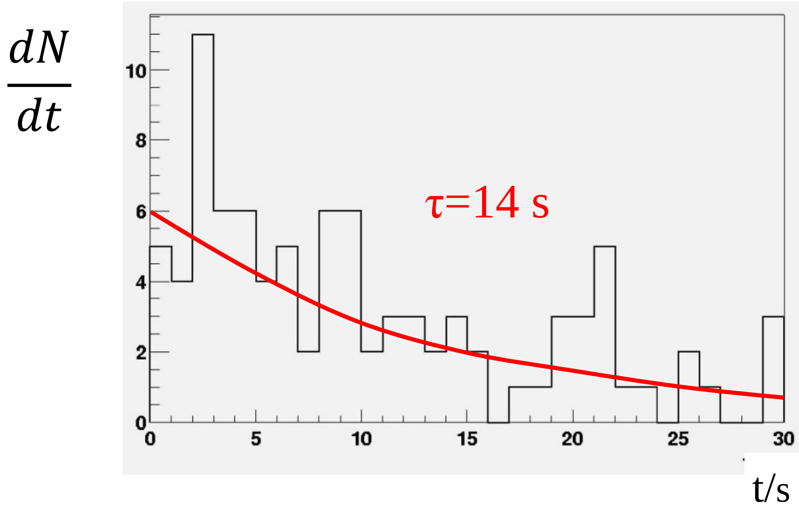
$$\frac{d^2N}{d\Theta^2} = -N_0 \cdot \lambda \cdot e^{\Theta} \cdot e^{-\lambda \cdot e^{\Theta}} \cdot \left[1 - \lambda \cdot e^{\Theta} \right] = 0$$

$$\Rightarrow \; 1 - \lambda \cdot e^{\Theta} = 0 \quad \Rightarrow \quad \Theta_{max} = \ln\tau$$

K.H.Schmidt, Z.Phys. A316 (1984), 19

Representation of Decay Curve

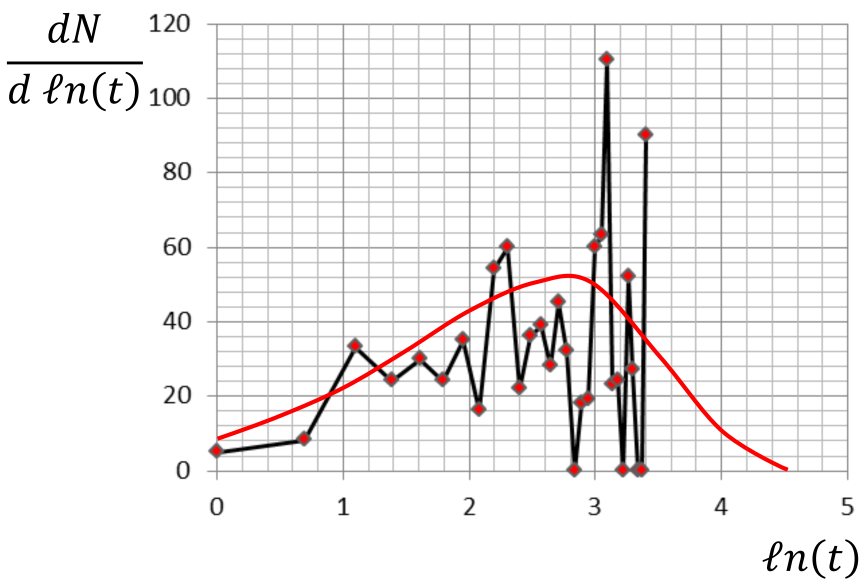
The maximum likelihood estimate of the lifetime τ



$$N = N_0 \cdot e^{-\lambda \cdot t}$$

$$\frac{dN}{dt} = -N_0 \cdot \lambda \cdot e^{-\lambda \cdot t}$$

lifetime: $\tau = {}^{1}/_{\lambda}$



$\Theta_{max} = \ell n\,(\tau) = 2.64 \rightarrow \tau = 14.1$ s

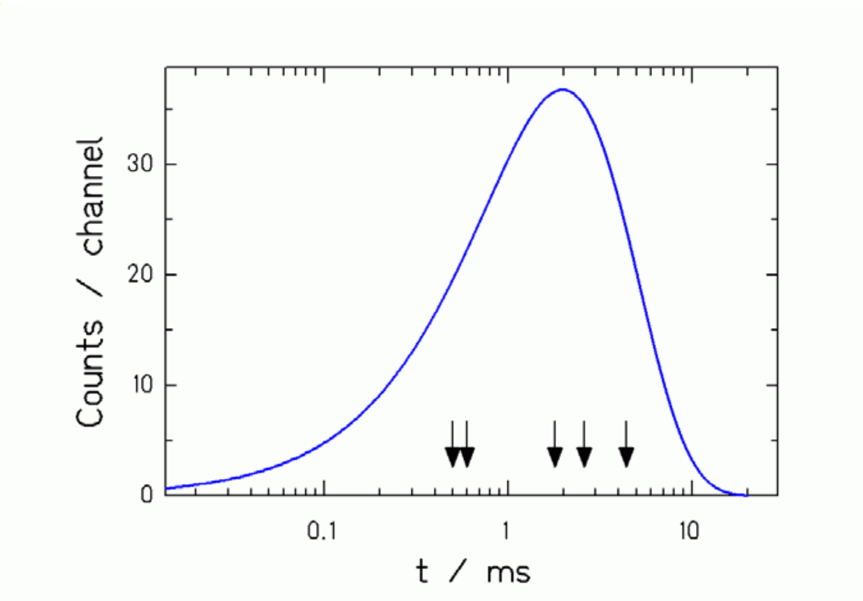
K.H.Schmidt, Z.Phys. A316 (1984), 19

Representation of Decay Curve

The maximum likelihood estimate of the lifetime τ

no.	t (ms)
1	0.6
2	1.8
3	4.4
4	0.5
5	2.6

average value: 2.0 ms



Logarithmic decay-time distribution of 5 events observed in the α decay of $^{271}110$ with alpha energies close to 10.74 MeV. Data are taken from S.Hofmann, Rep.Prog.Phys. 61(1998), 639. The curve shows the logarithmic decay-time distribution (τ=2.3 ms). The units on the ordinate are arbitrary.