Numerical Methods in Physics

Numerische Methoden in der Physik, 515.421.

Instructor: Ass. Prof. Dr. Lilia Boeri
Room: PH 03 090
Tel: +43-316-873 8191
Email Address: l.boeri@tugraz.at

Room: TDK Seminarraum

Time: 8:30-10 a.m.

Exercises: Computer Room, PH EG 004 F

http://itp.tugraz.at/LV/boeri/NUM_METH/index.html

(Lecture slides, Script, Exercises, etc).

TOPICS (this year):

- **Chapter 1: Introduction.** ✖
- **Chapter 2: Numerical methods for the solution of linear inhomogeneous systems.** ✖
- Chapter 3: Interpolation of point sets.
- **Chapter 4: Least-Squares Approximation.**
- **Chapter 5: Numerical solution of transcendental equations.**
- Chapter 6: Numerical Integration.
- Chapter 7: Eigenvalues and Eigenvectors of real matrices.
- **Chapter 8: Numerical Methods for the solution of ordinary differential equations: initial value problems.**
- Chapter 9: Numerical Methods for the solution of ordinary differential equations: marginal value problems.

Last week(22/10/2013)

- **Linear Systems:** Direct (**LU decomposition**) vs indirect methods (**Gauss-Seidel**).
- Direct methods**, iterative improvement** of the solution (reduce roundoff).
- Direct methods, special cases: **tridiagonal matrices**.
- **Indirect methods**: **iterative** solution for **sparse matrices**.
- Indirect methods: **Gauss-Seidel iteration rule**.
- **Band matrices**: definition and properties.
- **Gauss-Seidel** iteration for **band** matrices, storing information and efficiency**.**
- **Gauss-Seidel** method with over and under**-relaxation**.

Linear Systems

Methods of Solution (numerical):

$$\boxed{\hat{A}\mathbf{x} = \mathbf{b}}$$

- **Direct Methods:**
- **No methodological** error, BUT computationally expensive; **roundoff** errors can be **large**.

LU decomposition.

- **Iterative Methods:**
- **Simple algorithms; roundoff is easily controlled. The solution is approximate (truncation).**

Gauss-Seidel method.

Direct Methods:

$$\hat{A}\mathbf{x} = \mathbf{b} \quad \Rightarrow \quad \hat{U}\mathbf{x} = \mathbf{y}$$

Two-steps procedure:

1) LU decomposition (*Doolittle and Crout*): Reformulation of the Gaussian elimination. A real matrix can always be represented as the product of two real triangular matrices **L** and **U**, *i.e.*

$$\hat{A} = \hat{L} \cdot \hat{U}$$

2) The two auxiliary systems are solved through **back** and **forward** substitution:

$$\hat{U} \cdot \mathbf{x} = \mathbf{y}$$

Back

$$\hat{L} \cdot \mathbf{y} = \mathbf{b}$$

Forward

Gauss-Seidel Method:

iterative method for linear inhomogeneous sets of equations.

Advantages: simplicity; the matrix of coefficients is not changed during the iteration. Very **efficient** for systems with a **sparse** matrix of coefficients.

$$x_i^{(t+1)} = x_i^{(t)} - \Delta x_i^{(t)} \qquad (i = 1,\ldots,n)$$

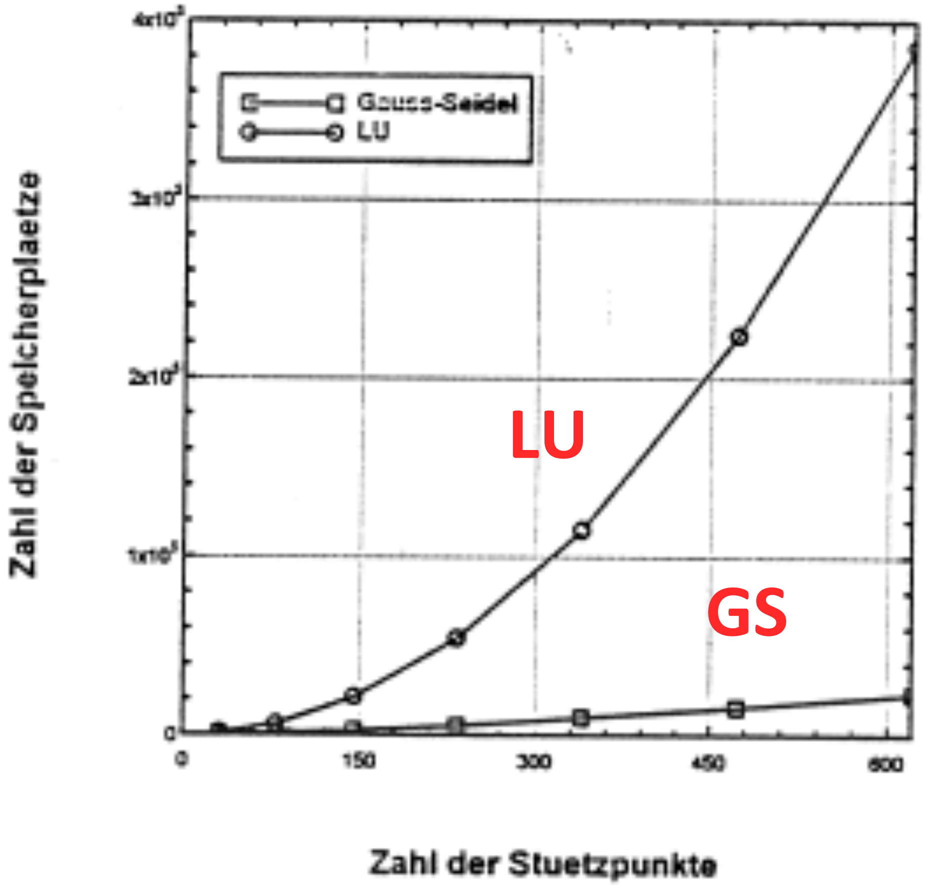
$$\Delta x_i^{(t)} = x_i^{(t)} + \frac{1}{a_{ii}}\left[\sum_{j=1(j\neq i)}^{n} a_{ij}x_j^{(t)} - b_i\right]$$

Starting from an initial vector $\mathbf{x}_0$ (***starting vector***) one obtains a sequence of vectors $\mathbf{x}^{(t)}$, which **converges** to the **exact** solution: $\lim_{t\to\infty} \mathbf{x}^{(t)} \to \mathbf{x}$

In practice, the iteration stops when a given **precision** is reached: $|xi(t)-xi(t-1)| \leq \varepsilon$

For **sparse matrices**, only a few terms survive in the equation for the Δx_i's.

Efficiency of direct Methods:



Efficiency of the G-S method compared to LU decomposition for a Laplace equation (sparse matrix).

This week(29/10/2013)

- **Least Square Approximation:** Definition of the problem.
- **Statistical** distribution of **experimental data**: normal and Poisson distribution.
- **Statistical** properties of the **fitting parameters**.
- **Model** Functions with **Linear** parameters.
- **How is this implemented in practice?**

Least Squares Approximation

Optimal Fitting of a data set

Least-Squares Approximation:

Basic problem: Data collected from experiments form a ***discrete set***. Physical processes are typically described by mathematical expressions, which employ ***real numbers***. These are easily manipulated with the standard tools of ***algebra*** (integrals, derivatives, etc).

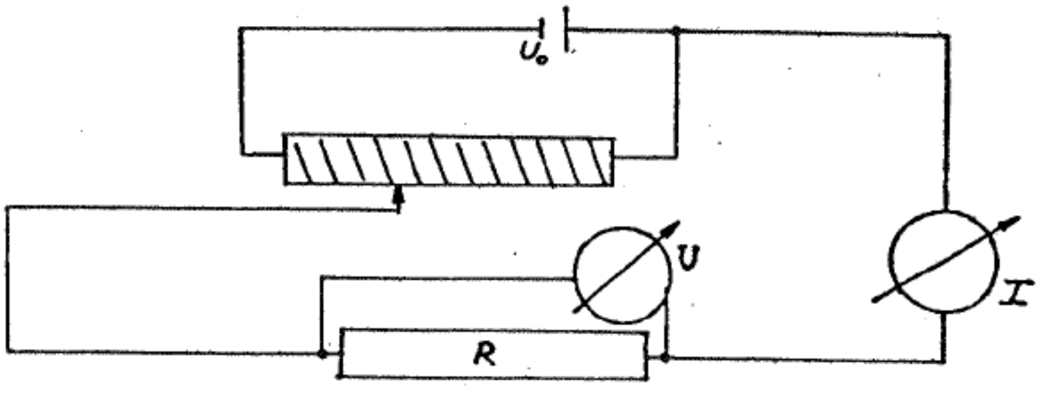
How to reduce a discrete set to a continuous function?

- **Interpolation:** Find an analytical expression (polynomial) which passes ***exactly*** through all experimental points.

- **Least-squares approximation:** Find the best possible analytical formula to approximate the behaviour of the n experimental points.

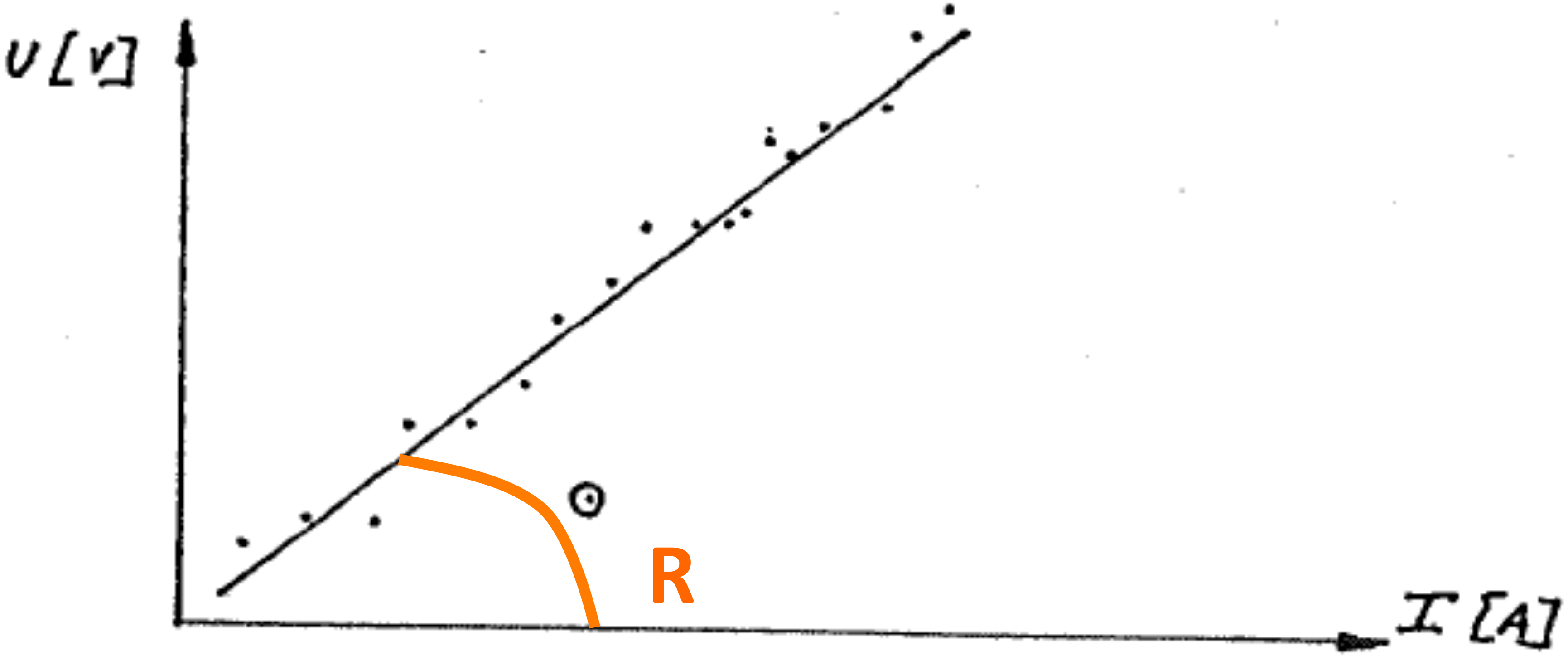
Example:

We wish to measure the resistivity of a **metallic conductor**:



The data points (Voltage: V; Current I) follow **Ohm's law**, but are affected by an **experimental error**.

$$V = R \cdot I$$



Due to experimental errors, the data do not lie ***exactly*** on top of a straight line (Ohm's law), but only ***approximately***.

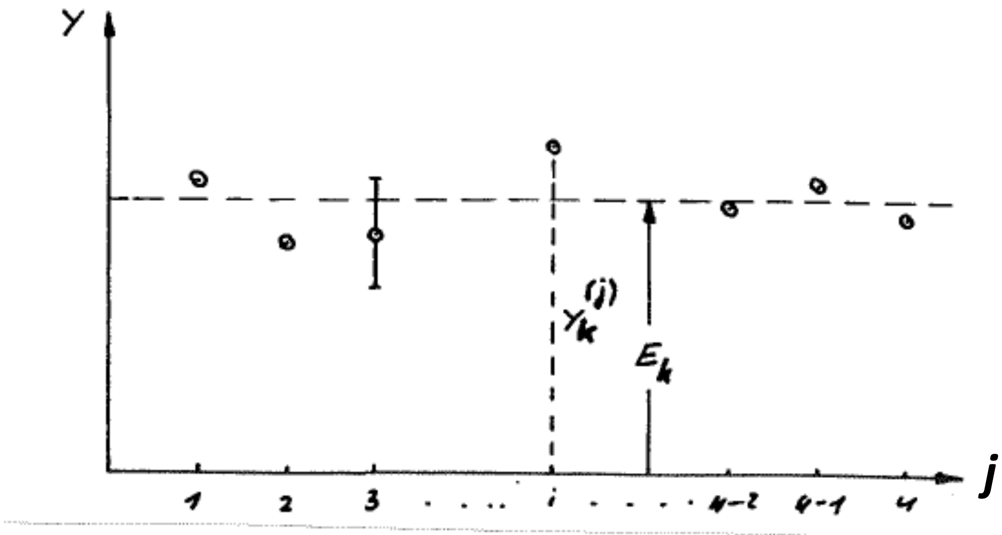
Least-Squares Approximation:

Mathematical Formulation: Given a set of n points (x_i/y_i), we wish to find a curve f(x) which approximates the points as closely as possible taking into account possible uncertainities due to measurement errors. We also would like to be able to assign different weights to the points through suitable weighting factors.

$$\chi^2 = \sum_{k=1}^{n} g_k \left[y_k - f(x_k; \mathbf{a}) \right]^2 \rightarrow \min$$

y_k	**Experimental values**
χ^2	**Weighted error sum**
$g_k > 0$	**Weighting factors (statistical)**
$f(x_k; \mathbf{a})$	**Model function**
$\mathbf{a} = a_1, \ldots, a_q$	**Model (fitting) parameters**

Statistics: The quality of the results of a LSQ fit depends on the statistical quality of the data to be fitted (y_k). The statistics of the y_k's is defined in terms of their expectation value (E_k) and standard deviation (σ_k).

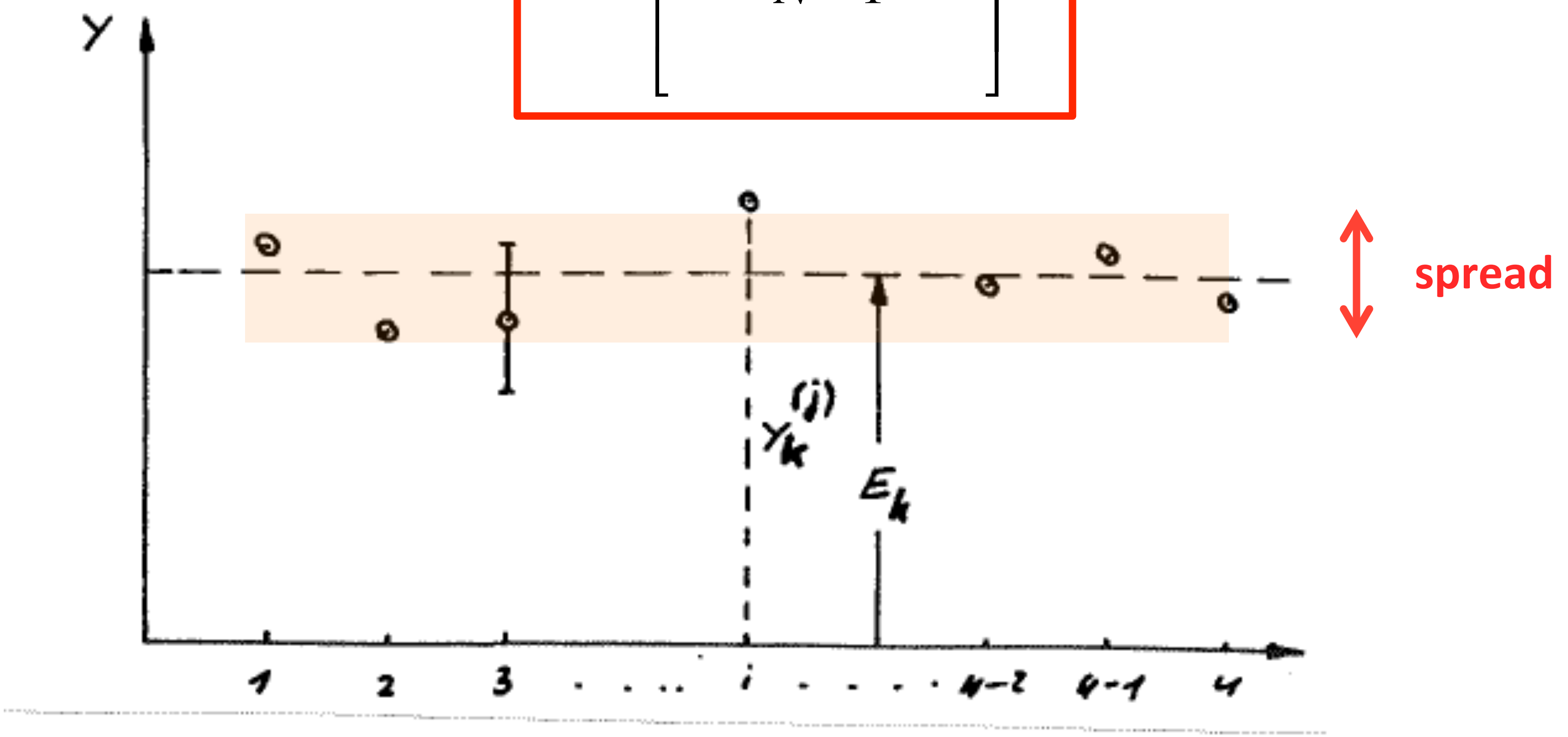


$$E_k = \frac{1}{N}\sum_{j=1}^{N} y_k^{(j)}$$ **Expectation value**

$$j = 1, \ldots, N$$ **Number of measurements**

Standard Deviation (σ_k): measures the spread of the y_k's around their expectation value E_k.

$$\sigma_k = \left[\frac{\sum_{j=1}^{N} \left(y_k^{(j)} - E_k \right)^2}{N-1} \right]^{1/2}$$



Gaussian normal distribution:

It is useful to describe experimental analogue measurements.

$$P(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

Parameters:

μ **Median** of the distribution

σ **Standard Deviation** σ^2 **Variance**

Central limit theorem: the mean of a random variable is distributed approximately normally, irrespective of the form of the original distribution -> very common!!!

Standard normal distribution: μ=0, σ=0.

Normalized Gaussian distribution:

We will assume that our measured (y) values have a normal distribution, with median = expectation value, and standard deviation=exp. standard deviation.

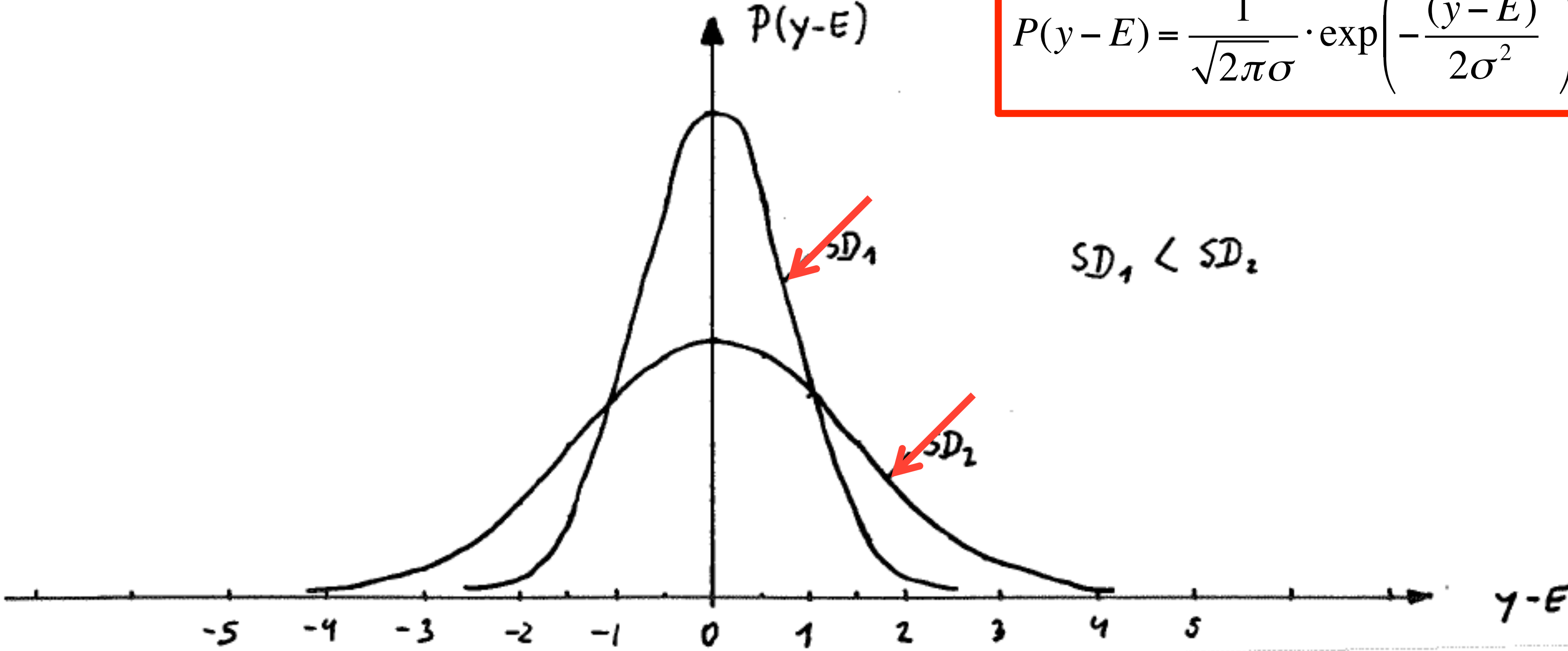
$$P(y-E)=\frac{1}{\sqrt{2\pi}\sigma}\cdot\exp\left(-\frac{(y-E)^2}{2\sigma^2}\right)$$

P(y-E) describes the probability with which a specific distance (y-E) between the measured value y and its expectation value occurs.

The standard deviation indicates the **spread** with which the measured values are distributed around their expectation value E.

Data-sets with a smaller σ are “more accurate”.

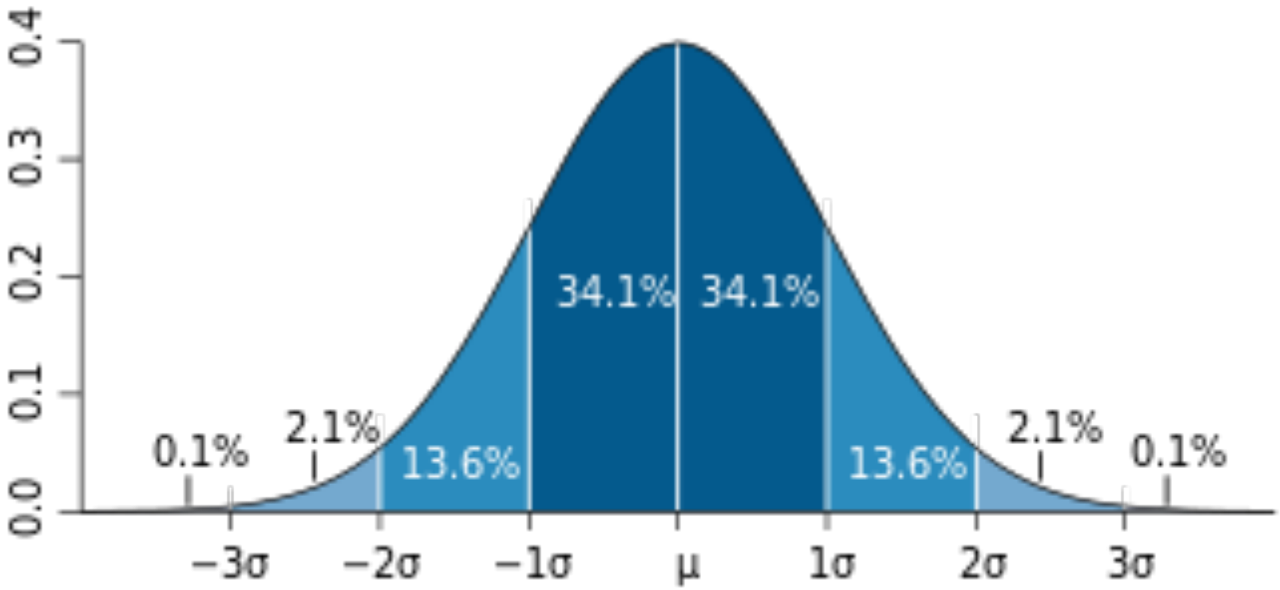
$$P(y-E) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(y-E)^2}{2\sigma^2}\right)$$



$$y = E \pm \sigma$$

Inflection points (P’’(y)=0).

Geometrical Meaning of the Standard Deviation:



$$\int_{-\sigma}^{+\sigma} d(y-E)\frac{1}{\sqrt{2\pi\sigma}}\cdot\exp\left(-\frac{(y-E)^2}{2\sigma^2}\right)=0.68\ldots \quad \sigma$$

$$\int_{-2\sigma}^{+2\sigma} d(y-E)\frac{1}{\sqrt{2\pi\sigma}}\cdot\exp\left(-\frac{(y-E)^2}{2\sigma^2}\right)=0.95\ldots \quad 2\sigma$$

$$\int_{-3\sigma}^{+3\sigma} d(y-E)\frac{1}{\sqrt{2\pi\sigma}}\cdot\exp\left(-\frac{(y-E)^2}{2\sigma^2}\right)=0.997\ldots \quad 3\sigma$$

Poisson's statistics:

The **Poisson**'s distribution is a **discrete probability distribution** that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.

$$P_{pois}(X = k) = f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Properties:

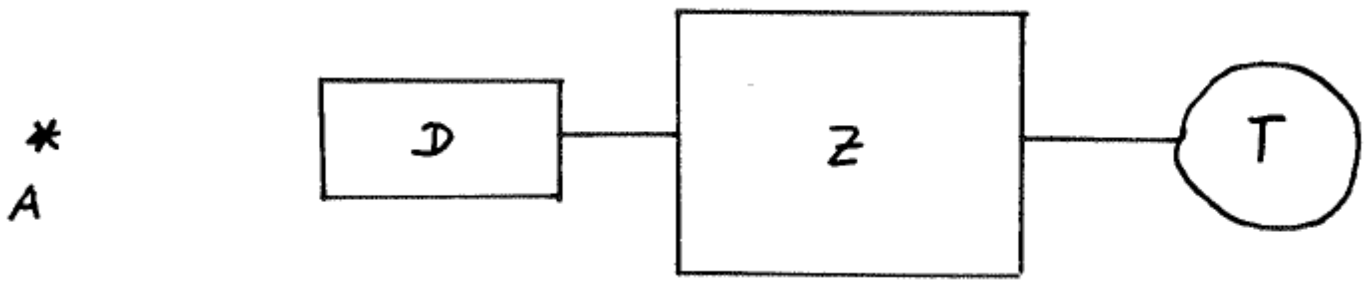
$E(X) = \lambda$ **Expectation Value**

$\text{var}(X) = \lambda$ **Variance**

In a Poisson's distribution the expectation value and variance are equal.

The **Poisson**'s distribution can be applied to systems with a large number of possible events, each of which is rare.

Example: Counting Experiment (radioactive decay):



where A could be a radioactive substance, D a radiation detector, Z a digital counter and T a clock.

D is a **digital** counter; the number of counts has a Poisson distribution.

Statistics in the LSQ method:

$$\chi^2 = \sum_{k=1}^{n} g_k \left[y_k - f(x_k;\mathbf{a}) \right]^2 \rightarrow \min$$

The **weighting factors** g_k depend on the statistics of the experimental data sets y_k:

$$g_k = \frac{1}{\sigma_k^2}$$

Normal distribution (analogic)

$$g_k = \frac{1}{y_k}$$

Poisson distribution (digital)

Normal matrix of the LSQ fit:

For the weighted error sum χ^2 we can define a **normal matrix** as:

$$[N]_{ij} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial \alpha_i \partial \alpha_j}$$

Its ***inverse*** is the ***covariance matrix***, which gives important information on the statistical properties of the fitting parameters:

$$C = N^{-1}$$

Covariance Matrix

$$\sigma_{a_i} = \sqrt{c_{ii}}$$

Standard deviation of the a_i's (fitting parameters)

$$r_{ij} = \frac{c_{ij}}{\sqrt{c_{ii}c_{jj}}}$$

Correlation coefficients for the a_i's: ($|cij| \leq 1$)

Measure how strongly the i-th and j-th parameter influence each other.

Variance and Standard Deviation:

$$V = \frac{\chi^2}{N-q}$$ **Variance**

$$\sigma_V = \sqrt{\frac{2}{N-q}}$$ **Standard Deviation**

q Number of **fitting parameters**

$N-q$ Number of **degrees of freedom**

Quality of the LSQ fit: In case of an **ideal model** for N>>1, V has approximately a **normal distribution** with E=1 and $\sigma=\sigma_V$. If V lies **significantly** outside the interval: $[1-\sigma_V, 1+\sigma_V]$ the fit is bad!

Model Functions with Linear Parameters

Model functions with linear parameters:

A model function with **linear parameters** has the form:

$$f(x;\mathbf{a}) = \sum_{j=1}^{m} a_j \cdot \varphi_j(x)$$

Here, $\phi_j(x)$ are **arbitrary** (linearly independent) **basis functions**. f(x,**a**) is linear in the fitting parameters, not in x. In the special case in which one uses as basis function a straight line: $\phi_j(x)=f_{lin}$ (x; a_1,a_2) we have the ***linear regression***.

Linear Regression:

$$f(x;a_1,a_2) = a_1 + a_2 x$$

Model functions with linear parameters, LSQ formula:

Inserting the expression of the model function with linear parameters in the expression for the weighted error sum we have:

$$\chi^2 = \sum_{k=1}^{n} g_k \left[y_k - \sum_{j=1}^{m} a_j \varphi_j(x_k) \right]^2 \rightarrow \min$$

Deriving with respect to the fitting parameters we get:

$$\frac{\partial \chi^2}{\partial a_i} = 2 \sum_{k=1}^{n} g_k \varphi_i(x_k) \left[y_k - \sum_{j=1}^{m} a_j \varphi_j(x_k) \right] = 0 \qquad i = 1,\ldots,m$$

Which can be recast in linear set of *m* equations:

$$\sum_{j=1}^{m} a_j \sum_{k=1}^{n} g_k \varphi_i(x_k) \varphi_j(x_k) = \sum_{k=1}^{n} g_k y_k \varphi_i(x_k) \qquad \hat{A}\alpha = \beta$$

This inhomogeneous linear set of *m* equations is (in matrix form):

$$\hat{A}\mathbf{a} = \beta$$

$$\hat{A} \equiv \left[\alpha_{ij}\right] \qquad \alpha_{ij} = \sum_{k=1}^{n} g_k \varphi_i(x_k)\varphi_j(x_k)$$

$$\beta_i = \sum_{k=1}^{n} g_k y_k \varphi_i(x_k)$$

One can define a **normal matrix,** given by:

$$\hat{N} \equiv \left[n_{ij}\right]$$

$$n_{ij} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial a_i \partial a_j} = \sum_{k=1}^{n} g_k \varphi_i(x_k)\varphi_j(x_k) = \alpha_{ij}$$

Standard Deviation of the fitting parameters:

We consider the ***linear regression*** formula (*q=2*):

$$f(x; a_1, a_2) = a_1 + a_2 x$$

The ***least-squares*** formula gives:

$$\chi^2 = \sum_{k=1}^{n} g_k \left[y_k - a - bx_k \right]^2 \rightarrow \min \Leftrightarrow a = a_{opt}, b = b_{opt}$$

The ***normal matrix*** of the problem is given by:

$$N = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

with:

$$\alpha_{11} = \sum_{k=1,n} g_k \qquad \alpha_{12} = \sum_{k=1,n} g_k x_k \qquad \alpha_{22} = \sum_{k=1,n} g_k x_k^2$$

The ***optimized parameters*** (a_{opt} and b_{opt}) can be obtained from:

$$N\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

with:

$$\beta_1 = \sum_{k=1}^{n} g_k y_k, \qquad \beta_2 = \sum_{k=1}^{n} g_k x_k y_k$$

This gives:

$$a = a_{opt} = \frac{\beta_1 \alpha_{22} - \beta_2 \alpha_{12}}{D}, \qquad b = b_{opt} = \frac{\beta_2 \alpha_{11} - \beta_1 \alpha_{21}}{D}$$

D is the **determinant** of the **normal matrix**:

$$D = \alpha_{11}\alpha_{22} - \alpha_{12}^2$$

The **covariance** matrix is the **inverse** of the **normal** matrix:

$$C = N^{-1} = \frac{1}{D}\begin{pmatrix} \alpha_{22} & -\alpha_{21} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}$$

The ***diagonal elements*** give the standard deviations of the fitting parameters:

$$\sigma_a^2 = \frac{\alpha_{22}}{D}, \qquad \sigma_b^2 = \frac{\alpha_{11}}{D}$$

$$a = a_{opt} \pm \sigma_a, \qquad b = b_{opt} \pm \sigma_b$$

We will now show that the **same result** can be obtained using the **Error Propagation Rule** (EPR).

Error Propagation Rule:

The fitting parameters are function of the *x*, *y* components of the data points:

$$a = a(x_1,\ldots,x_n;y_1,\ldots,y_n) \qquad b = b(x_1,\ldots,x_n;y_1,\ldots,y_n)$$

We can then use the standard **formula for the error propagation**. Given a function f(**x**) of n variables ($x_1,\ldots,x_n$), with errors $\Delta x_1,\ldots, \Delta x_n$ the **error on the function** is given by:

$$f = f(x_1,\ldots,x_n) = f(\mathbf{x})$$

$$f(\mathbf{x}) = \overline{f(\mathbf{x})} \pm \Delta f(\mathbf{x})$$

$$\left(\Delta f(\mathbf{x})\right)^2 = \left[\left(\frac{\partial f}{\partial x_1}\right)^2 (\Delta x_1)^2 + \ldots + \left(\frac{\partial f}{\partial x_n}\right)^2 (\Delta x_n)^2\right]$$

Applying this to the expression for the standard deviations gives:

$$\sigma_a^2 = \sum_{l=1}^{n}\left[\left(\frac{\partial a}{\partial x_l}\right)^2 \sigma(x_l)^2 + \left(\frac{\partial a}{\partial y_l}\right)^2 \sigma(y_l)^2\right]$$

$$\sigma_b^2 = \sum_{l=1}^{n}\left[\left(\frac{\partial b}{\partial x_l}\right)^2 \sigma(x_l)^2 + \left(\frac{\partial b}{\partial y_l}\right)^2 \sigma(y_l)^2\right]$$

If we assume that there is no statistical error on the x_l's, and that the y_l's have a **normal distribution**, we have:

$$\sigma(x_l) = 0 \qquad g_l = \frac{1}{\sigma^2(y_l)}$$

And thus:

$$\sigma_a^2 = \sum_{l=1}^{n}\left[\frac{1}{g_l}\left(\frac{\partial a}{\partial y_l}\right)^2\right]; \qquad \sigma_b^2 = \sum_{l=1}^{n}\left[\frac{1}{g_l}\left(\frac{\partial b}{\partial y_l}\right)^2\right]$$

To calculate the **partial derivatives** we have to recall:

$$a = a_{opt} = \frac{\beta_1\alpha_{22} - \beta_2\alpha_{12}}{D}, \qquad b = b_{opt} = \frac{\beta_2\alpha_{11} - \beta_1\alpha_{21}}{D}$$

$$\alpha_{11} = \sum_{k=1,n} g_k \quad , \alpha_{12} = \sum_{k=1,n} g_k x_k \quad , \alpha_{22} = \sum_{k=1,n} g_k x_k^2 \qquad \frac{\partial \alpha_{ij}}{\partial y_l} = 0;$$

$$\beta_1 = \sum_{k=1}^{n} g_k y_k, \qquad \beta_2 = \sum_{k=1}^{n} g_k x_k y_k \qquad \frac{\partial \beta_1}{\partial y_l} = g_l; \qquad \frac{\partial \beta_2}{\partial y_l} = g_l x_l$$

Substituting in the expressions for the standard deviations we get:

$$\sigma_a^2 = \sum_{l=1}^{n} \frac{1}{g_l}\left[\frac{1}{D}(\alpha_{22} g_l - \alpha_{12} g_l x_l)\right]^2$$

$$\sigma_b^2 = \sum_{l=1}^{n} \frac{1}{g_l}\left[\frac{1}{D}(-\alpha_{12} g_l + \alpha_{11} g_l x_l)\right]^2$$

We obtain (for example for a):

$$\sigma_a^2 = \sum_{l=1}^{n} \frac{1}{g_l}\left[\frac{1}{D}(\alpha_{22} g_l - \alpha_{12} g_l x_l)\right]^2 =$$

$$= \sum_{l=1}^{n} \frac{1}{g_l}\left(\frac{1}{D} g_l\right)^2 \left[(\alpha_{22} - \alpha_{12} x_l)\right]^2 =$$

$$= \frac{1}{D^2}\sum_{l=1}^{n} g_l \left[(\alpha_{22} - \alpha_{12} x_l)\right]^2 =$$

$$= \frac{1}{D^2}\sum_{l,k,k'} g_l g_k g_{k'} \left[x_k^2 x_{k'}^2 - 2 x_l x_k^2 x_{k'} + x_l^2 x_k x_{k'}\right]$$

Renaming the indexes of the sum in the last two terms we obtain:

$$\sigma_a^2 = \frac{1}{D^2}\sum_{l,k,k'} g_l g_k g_{k'} x_k^2 \left[x_{k'}^2 - x_l x_{k'}\right] = \frac{1}{D^2}\sum_{k=1}^{n} g_k x_k^2 \left[\sum_{l=1}^{n} g_l \sum_{k'=1}^{n} g_{k'} x_{k'}^2 - \left(\sum_{l=1}^{n} g_l x_l\right)^2\right]$$

$$\sigma_a^2 = \frac{1}{D^2}\sum_{k=1}^{n} g_k x_k^2 \left[\sum_{l=1}^{n} g_l \sum_{k'=1}^{n} g_{k'} x_{k'}^2 - \left(\sum_{l=1}^{n} g_l x_l\right)^2\right]$$

$$D = \alpha_{11}\alpha_{22} - \alpha_{12}^2$$

$$\alpha_{11} = \sum_{k=1,n} g_k \quad ,\alpha_{12} = \sum_{k=1,n} g_k x_k \quad ,\alpha_{22} = \sum_{k=1,n} g_k x_k^2$$

We thus obtain:

$$\sigma_a^2 = \frac{1}{D^2} D \sum_{k=1}^{n} g_k x_k^2 = \frac{\alpha_{22}}{D}$$

i.e. the **diagonal terms** of the **covariance matrix** give the standard deviation for the fitting parameters!!!!

In summary (LSQ with linear model parameters):

1) Input the experimental data set: x_k, y_k and the statistical weights g_k

2) Choose a set of **basis functions** $\phi_j(x)$: $f(x;\mathbf{a}) = \sum_{j=1}^{m} a_j \cdot \varphi_j(x)$

3) Construct the auxiliary linear problem: $\hat{A}\mathbf{a} = \beta$

$$\hat{A} \equiv [\alpha_{ij}] \qquad \alpha_{ij} = \sum_{k=1}^{n} g_k \varphi_i(x_k)\varphi_j(x_k), \qquad \beta_i = \sum_{k=1}^{n} g_k y_k \varphi_i(x_k)$$

4) Solve the linear problem (LU decomposition); Find optimal fitting parameters.

$$\mathbf{a}^{opt} = (a_1^{opt}, \ldots, a_q^{opt})$$

5) Calculate the covariance matrix (standard deviations of the fitting parameters).

6) Calculate the value of the optimal fitting function on the given data points: $f(x_k; \mathbf{a}^{opt})$

7) Evaluate the weighted error sum. $\chi^2 = \sum_{k=1}^{n} g_k \left[y_k - f(x_k; \mathbf{a}) \right]^2$

This week(29/10/2013)

- **Least Square Approximation:** Definition of the problem.
- **Statistical** distribution of **experimental data**: normal and Poisson distribution.
- **Statistical** properties of the **fitting parameters**.
- **Model** Functions with **Linear** parameters.
- **How is this implemented in practice?**