Least-squares fitting of complex functions

$$y(x) = f(x; a_1, a_2, \ldots, a_{\mathcal{N}})$$

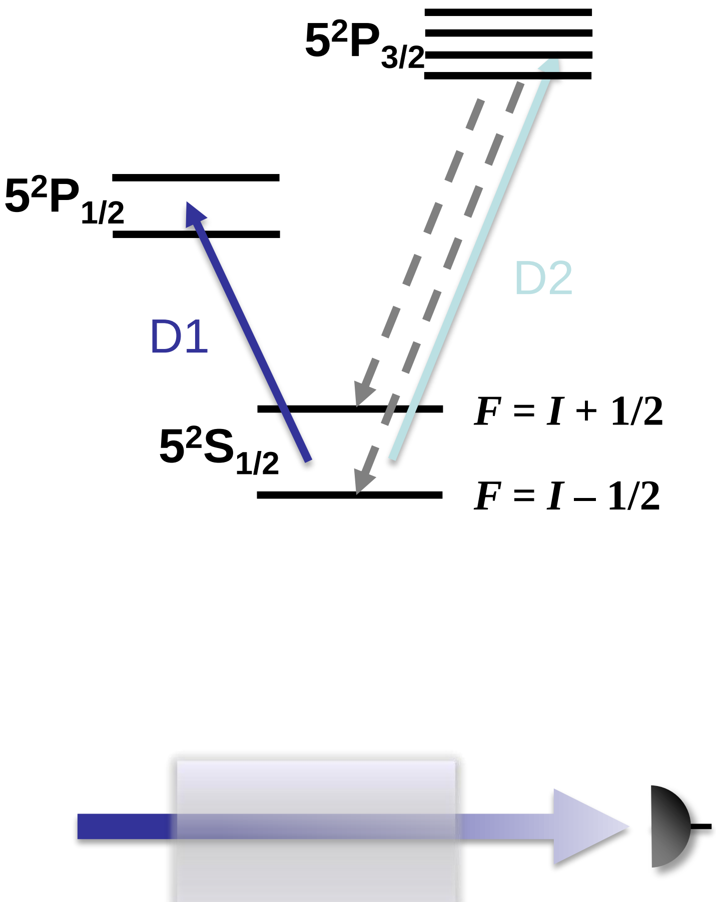
$\mathcal{N}$ parameters

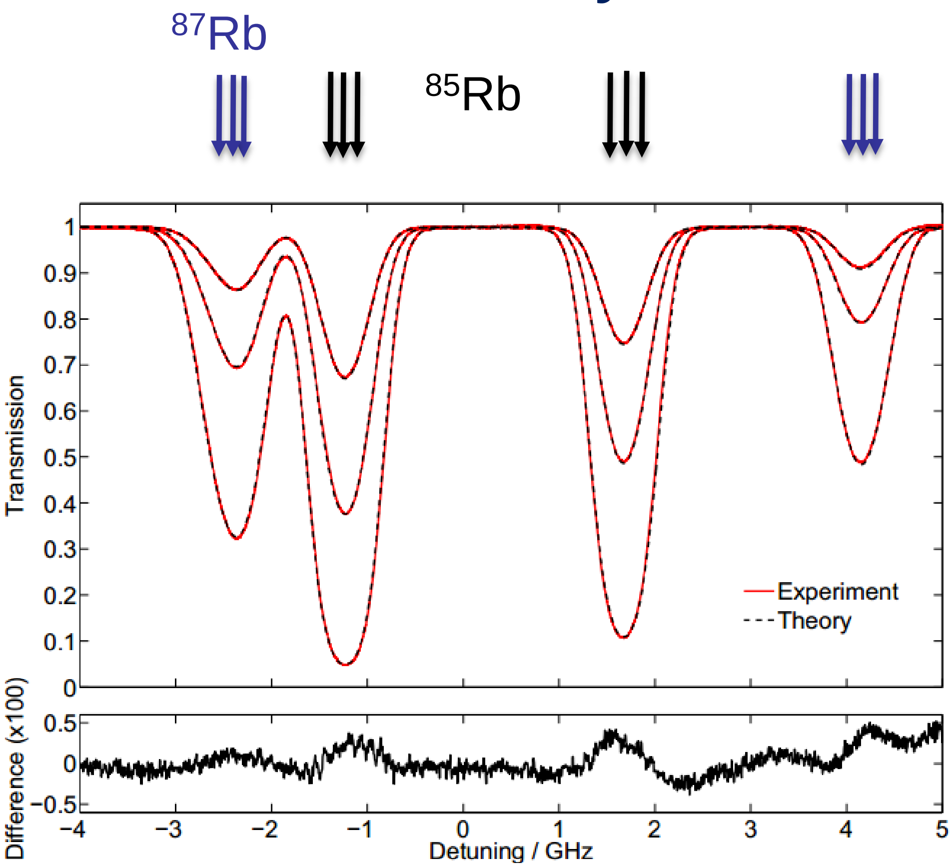
- For each value of the independent variable, x_i, calculate $y(x_i)$ from eqn (6.10) using an estimated set of values for the parameters.
- For each value of the independent variable calculate the square of the normalised residual, $\left[\frac{(y_i - y(x_i))}{\alpha_i}\right]^2$.
- Calculate χ^2 (by summing the square of the normalised residuals over the entire data set).
- Minimise χ^2 by optimising the fit parameters.

Use a computer for this step

Measurements and their Uncertainties, Pg. 72

Examples of least-squares fit to an arbitrary function

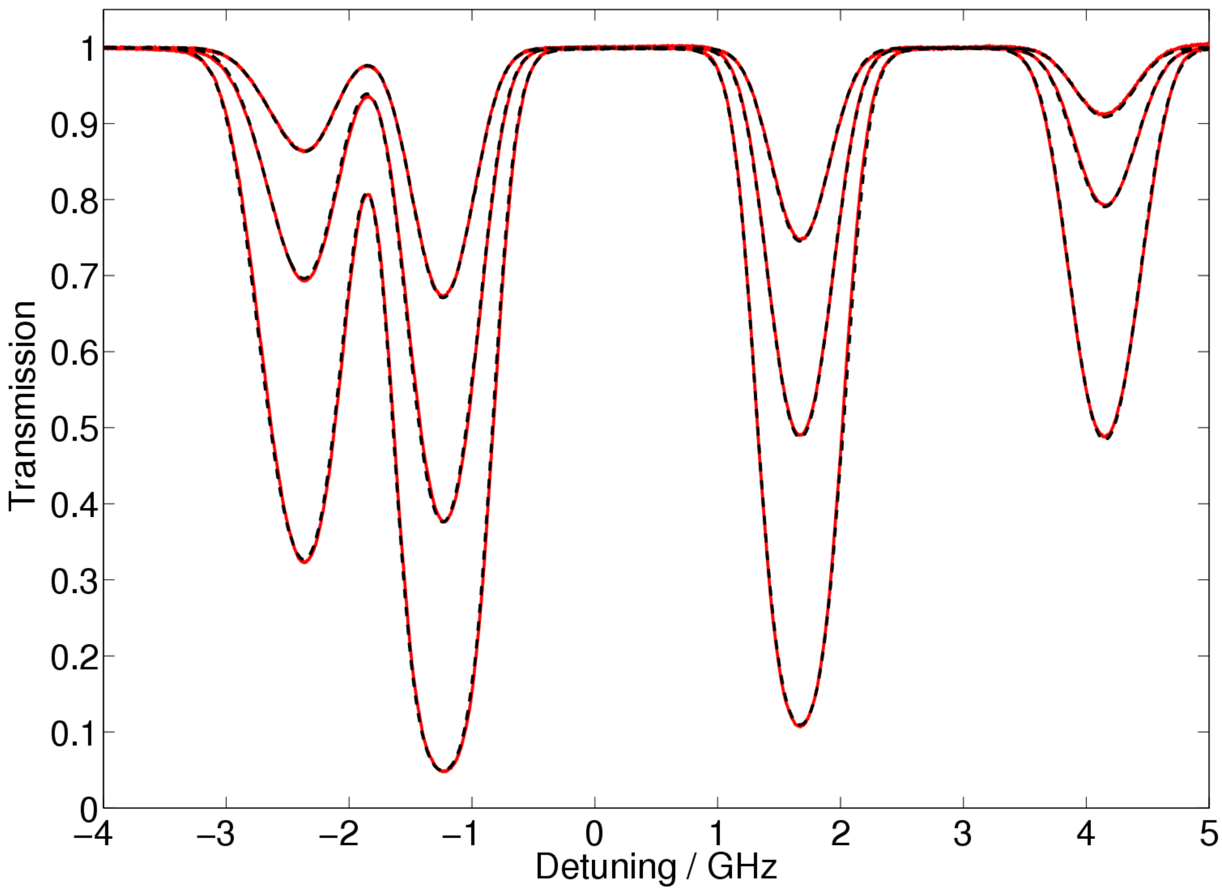




ElecSus

Program and example data available

Zentile et al, Computer Physics Communications **189** 162-174 (2015)



Red – measured

Black – theory

Top 16.5°C

Middle 25.0°C

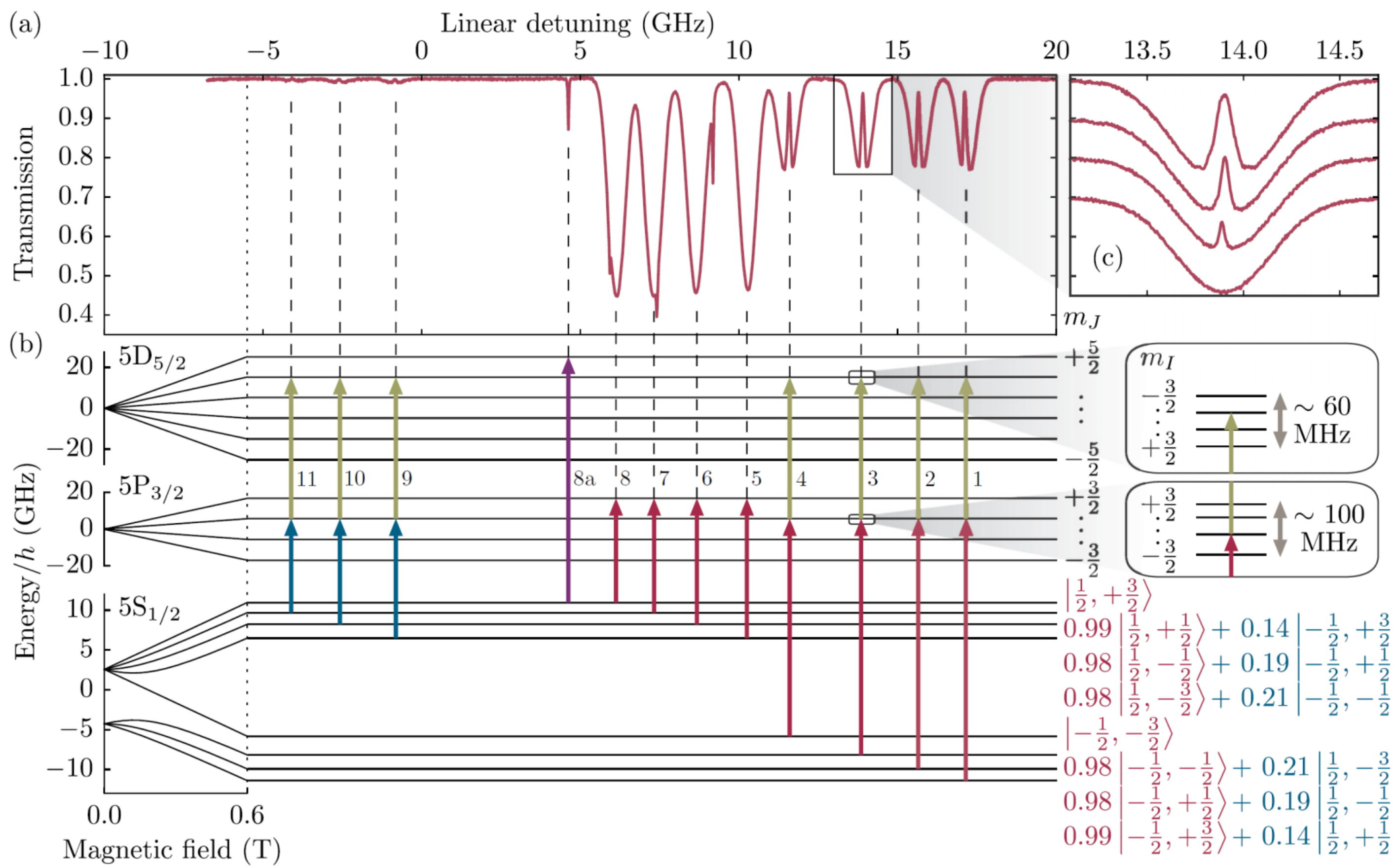
Bottom 36.6°C

Note:

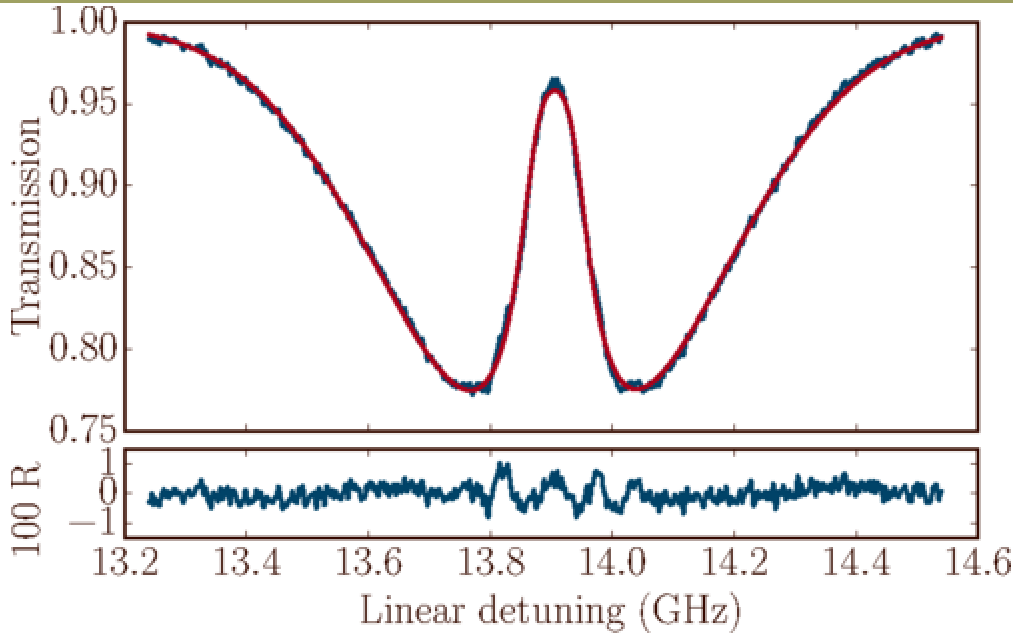
The temperature was measured using a thermocouple – there are no fitting parameters

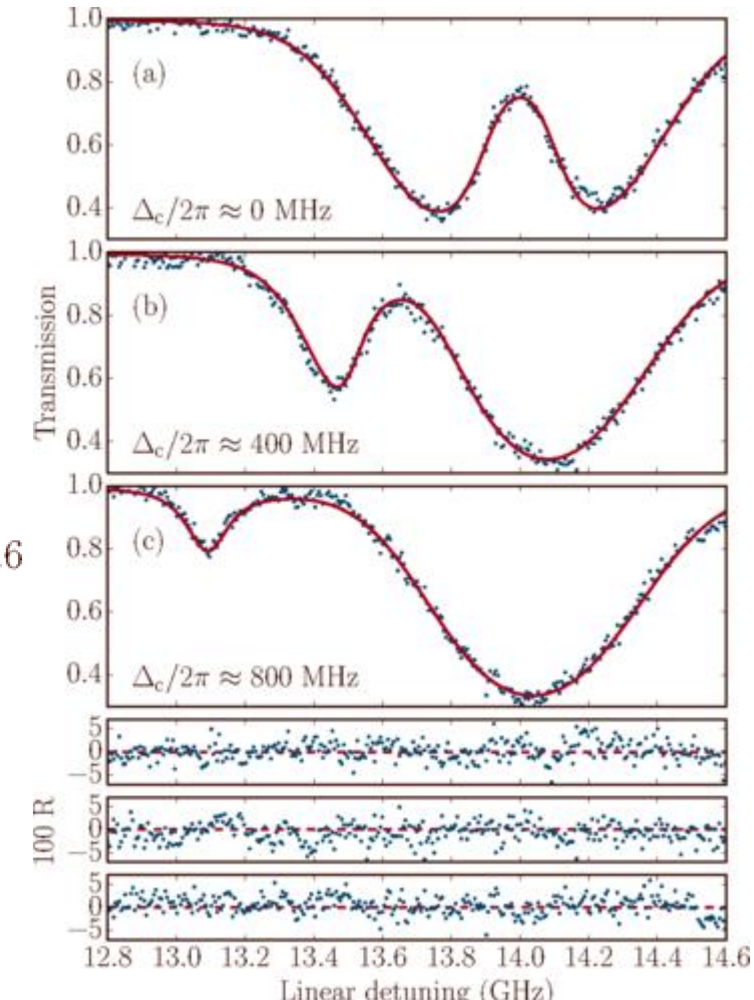
Siddons *et al. Absolute absorption on rubidium D lines: comparison between theory and experiment J. Phys. B: At. Mol. Opt. Phys.* **41** 155004 (2008)

Ben E Sherlock and Ifan G Hughes *How weak is a weak probe in laser spectroscopy? Am. J. Phys.* **77** 111–115 (2009)



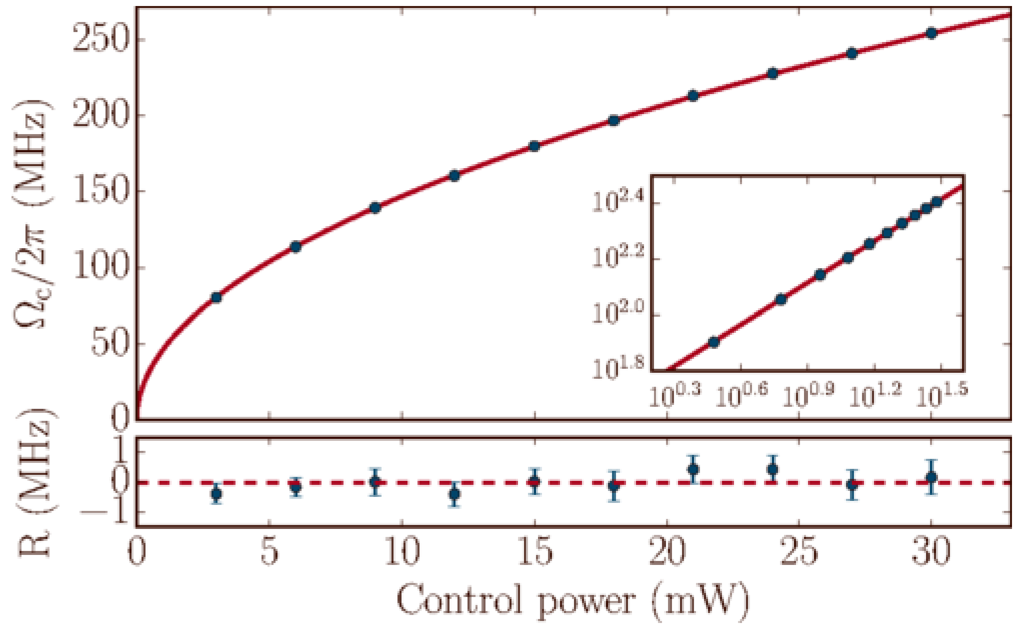
Whiting *et al.* Direct measurement of excited-state dipole matrix elements using electromagnetically induced transparency in the hyperfine Paschen-Back regime
Phys. Rev. A **93** 043854 (2016)





$$\chi(v) = \frac{4i\hbar d_{\mathrm{p}}^2/\epsilon_0}{\gamma_{\mathrm{p}} - i\delta_1(v) + \frac{\Omega_{\mathrm{c}}^2/4}{\gamma_{\mathrm{c}} - i\delta_2(v)}} N$$

$$T = \exp\left(-k_{\mathrm{p}} l \int_{-\infty}^{\infty} \mathrm{Im}[\chi(v)] n(v) dv \right)$$



We measure $|\langle 5P_{3/2}||er||5D_{5/2}\rangle| =$

$$(2.290 \pm 0.002_{\rm stat} \pm 0.04_{\rm syst})ea_0$$

Whiting *et al.* Direct measurement of excited-state dipole matrix elements using electromagnetically induced transparency in the hyperfine Paschen-Back regime
Phys. Rev. A **93** 043854 (2016)

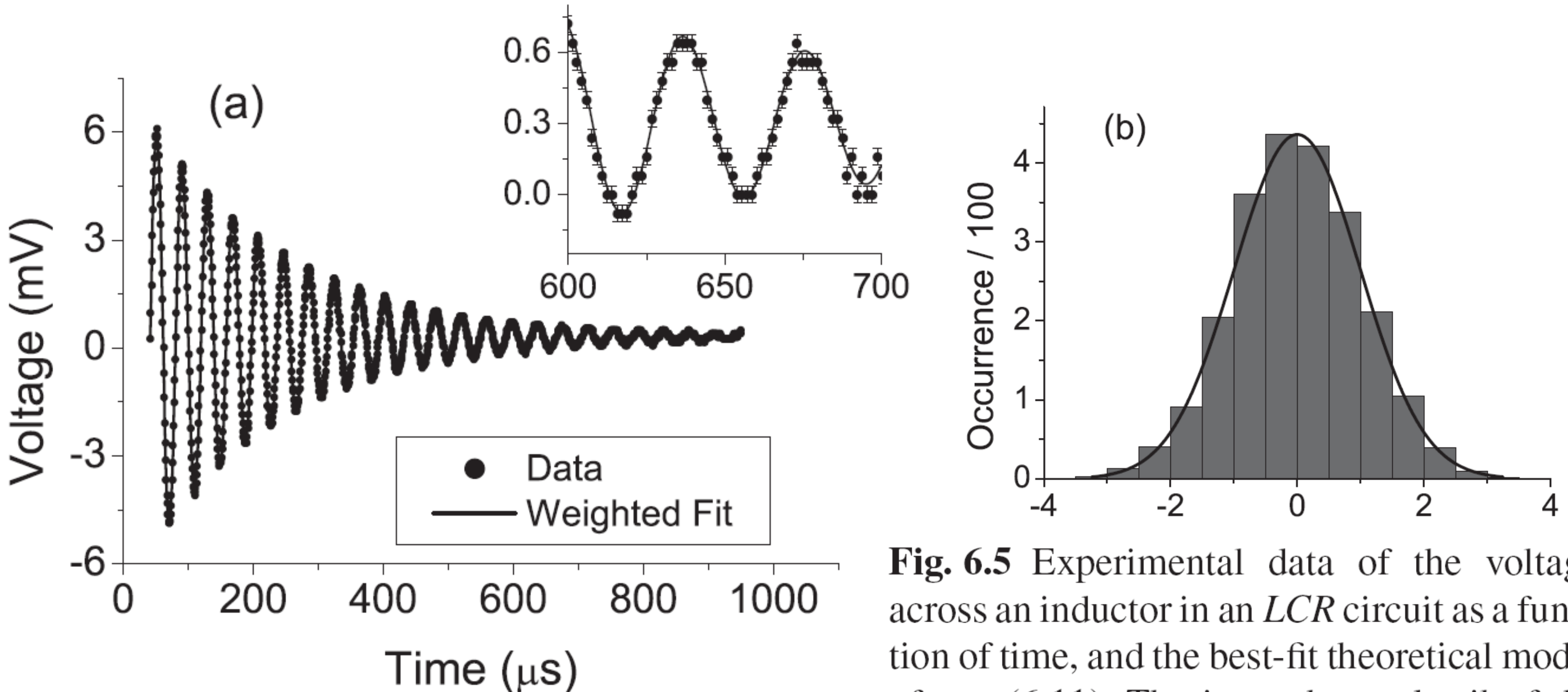
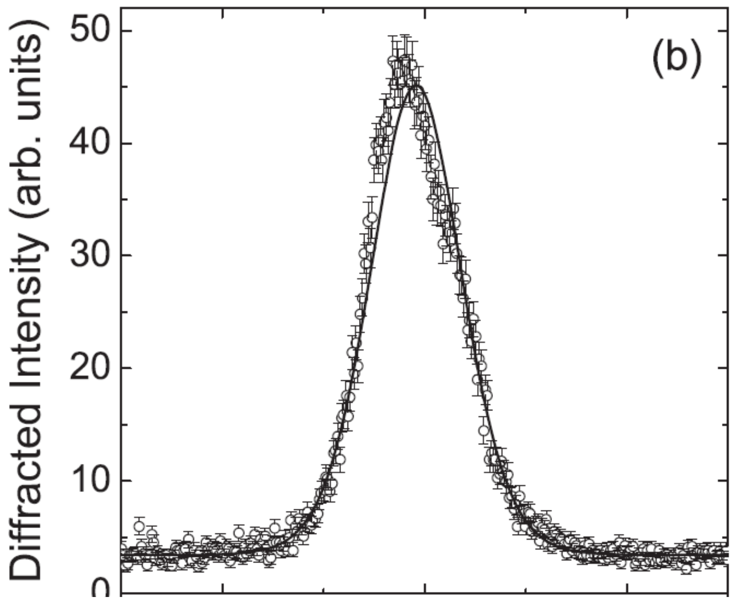
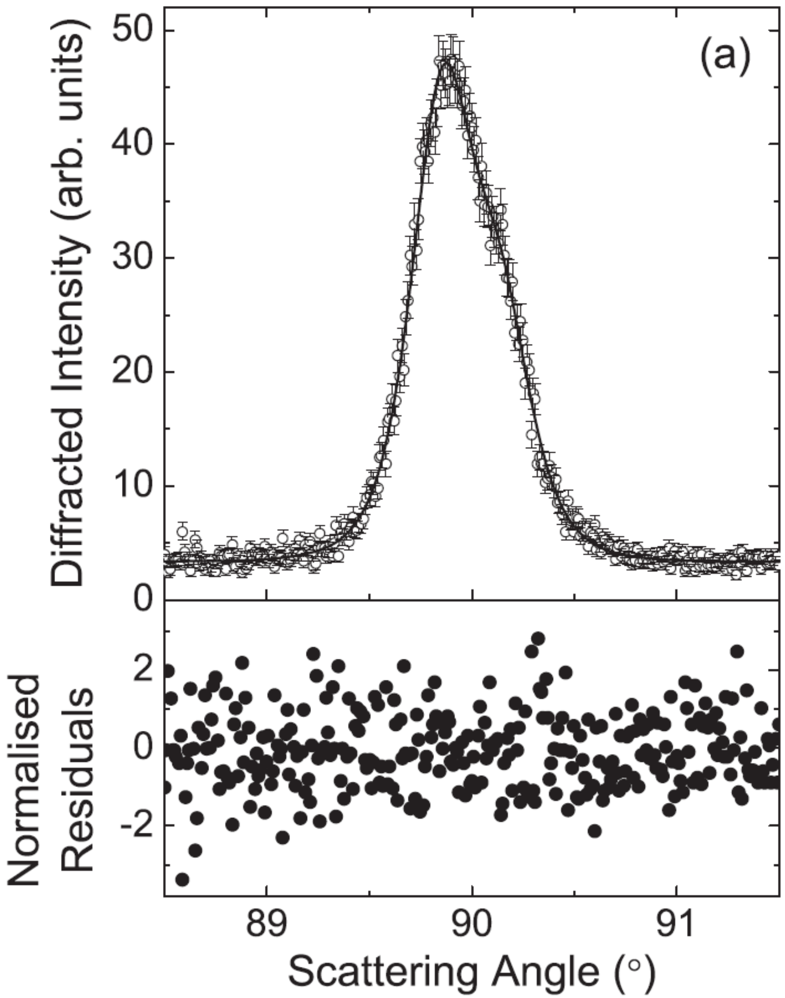


Fig. 6.5 Experimental data of the voltage across an inductor in an *LCR* circuit as a function of time, and the best-fit theoretical model of eqn (6.11). The inset shows detail of the oscillations, and the magnitude of the experimental error bars. The histogram shows the distribution of normalised residuals. Nearly all of the data lie within ±2 error bars of the theoretical model, which is consistent with a good fit.

Measurements and their Uncertainties, Pg. 73

Testing the fit using residuals

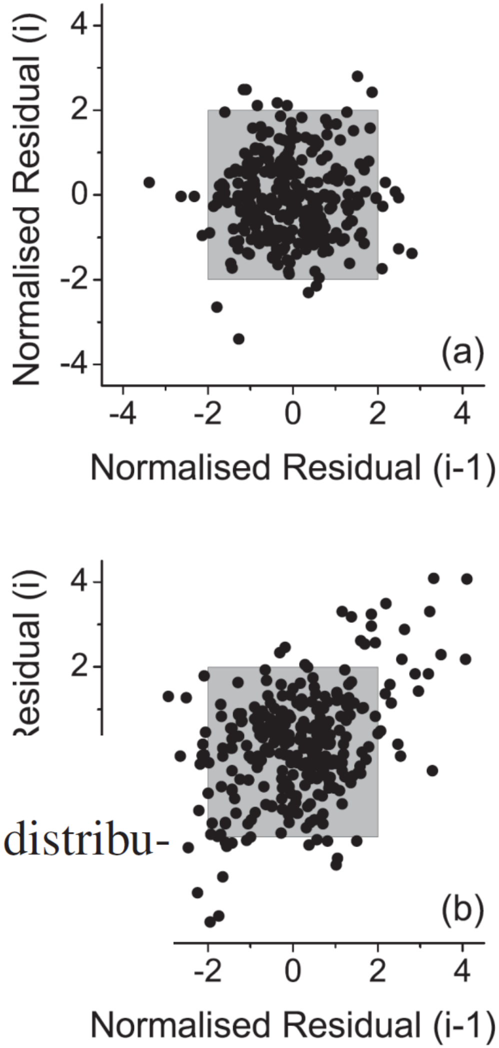




Testing the fit using residuals: Durbin-Watson statistic

Structure in the residuals can be visualised by use of a lag plot.

The degree of correlation in the lag plot can be reduced to a single numerical value by evaluating the **Durbin-Watson** statistic, $\mathcal{D}$.



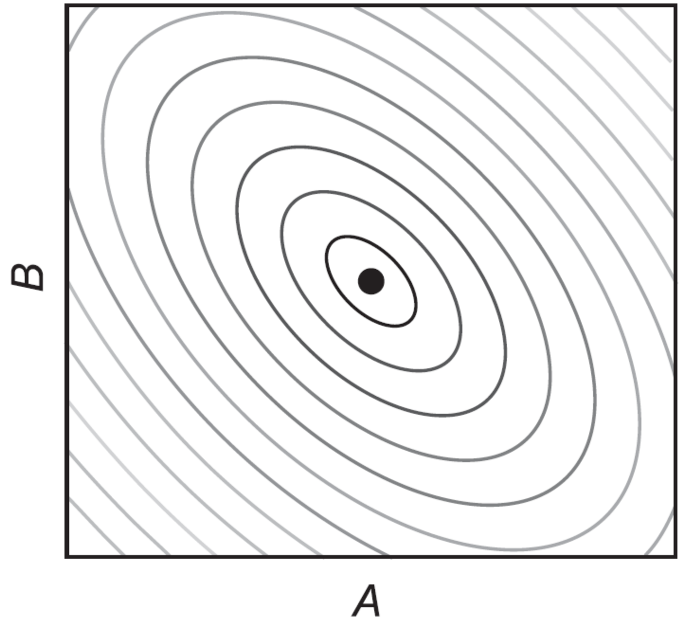
$$\mathcal{D} = \frac{\sum_{i=2}^{N} \left[R_i - R_{i-1}\right]^2}{\sum^{N} \left[R_i\right]^2}.$$

(1) $\mathcal{D} = 0$: systematically correlated residuals;

(2) $\mathcal{D} = 2$: randomly distributed residuals that follow a Gaussian distribution;

(3) $\mathcal{D} = 4$: systematically anticorrelated residuals.

Measurements and their Uncertainties, Pg. 81



$$\chi^2\left(\overline{a}_j + \Delta a_j\right) = \chi^2\left(\overline{a}_j\right) + \frac{1}{2}\left.\frac{\partial^2\chi^2}{\partial a_j^2}\right|_{\overline{a}_j}\left(\Delta a_j\right)^2$$

$$\chi^2 \rightarrow \chi^2_{\mathrm{min}} + 1$$

Fig. 6.6 Contours of constant χ^2 in th $A-B$ plane for a general two-paramete nonlinear function $f(A, B)$. The minimun value of χ^2, χ^2_{min}, is obtained with th best-fit values of the parameters, $\overline{A}$ and $\overline{B}$ shown by the dot in the centre. The contour increase in magnitude as one departs from th best-fit parameters.

$$\alpha_j = \sqrt{\frac{2}{\left(\frac{\partial^2\chi^2}{\partial a_j^2}\right)}}$$

Measurements and their Uncertainties, Pg. 74

Calculating the error in a least-squares fit

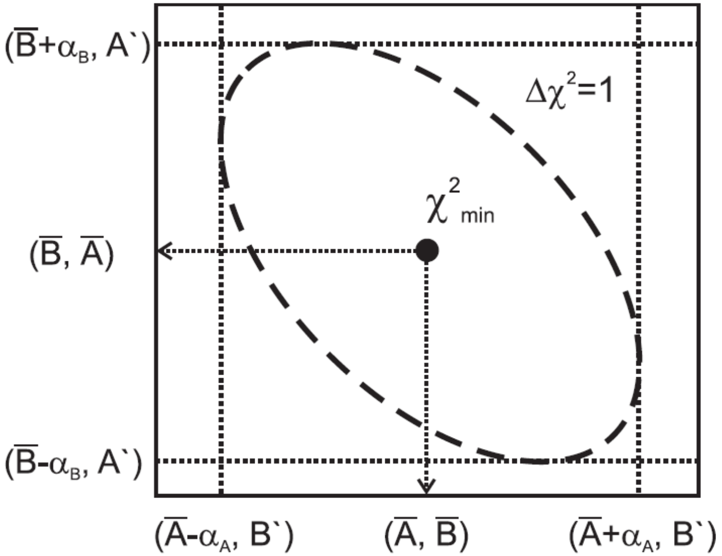


Fig. 6.7 The $\Delta\chi^2 = 1$ contour in the A–B plane. The horizontal and vertical tangent planes define the uncertainties in the parameters.

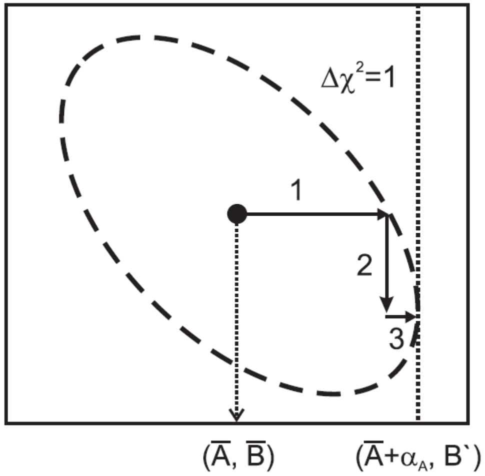


Fig. 6.8 The procedure for obtaining the uncertainties in the best-fit parameters. χ^2_{min} is achieved when the parameters have the values $\overline{A}, \overline{B}$. If A is increased with B held constant one follows the trajectory labelled 1 to arrive at the desired $\Delta\chi^2$ contour. Then, B must be allowed to vary to reduce χ^2, which results in motion across the error surface along the path labelled 2. This iterative procedure is repeated until an extremum of the desired $\Delta\chi^2$ contour is achieved, where the coordinate of the abscissa yields $\overline{A} + \alpha_A$. Repeating this procedure from $\overline{A}, \overline{B}$ by keeping A constant and increasing B will give the error bar for B.

Measurements and their Uncertainties, Pg. 76

Confidence levels for multiparameter fits

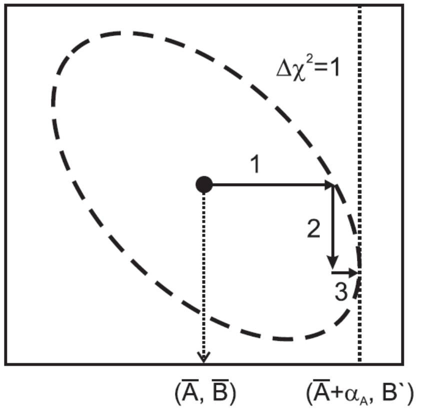


Fig. 6.8 The procedure for obtaining the uncertainties in the best-fit parameters. $\chi^2_{\min}$ is achieved when the parameters have the values $\overline{A}, \overline{B}$. If A is increased with B held constant one follows the trajectory labelled 1 to arrive at the desired $\Delta\chi^2$ contour. Then, B must be allowed to vary to reduce χ^2, which results in motion across the error surface along the path labelled 2. This iterative procedure is repeated until an extremum of the desired $\Delta\chi^2$ contour is achieved, where the coordinate of the abscissa yields $\overline{A}+\alpha_A$. Repeating this procedure from $\overline{A}, \overline{B}$ by keeping A constant and increasing B will give the error bar for B.

We separate our χ^2 fitting problem into two parts:

- A fit of N data points to $\mathcal{N}$ -1 parameters with N -$\mathcal{N}$ -1 degrees of freedom
- A variation of χ^2 with the parameter of interest about the minimum with 1 degree of freedom.

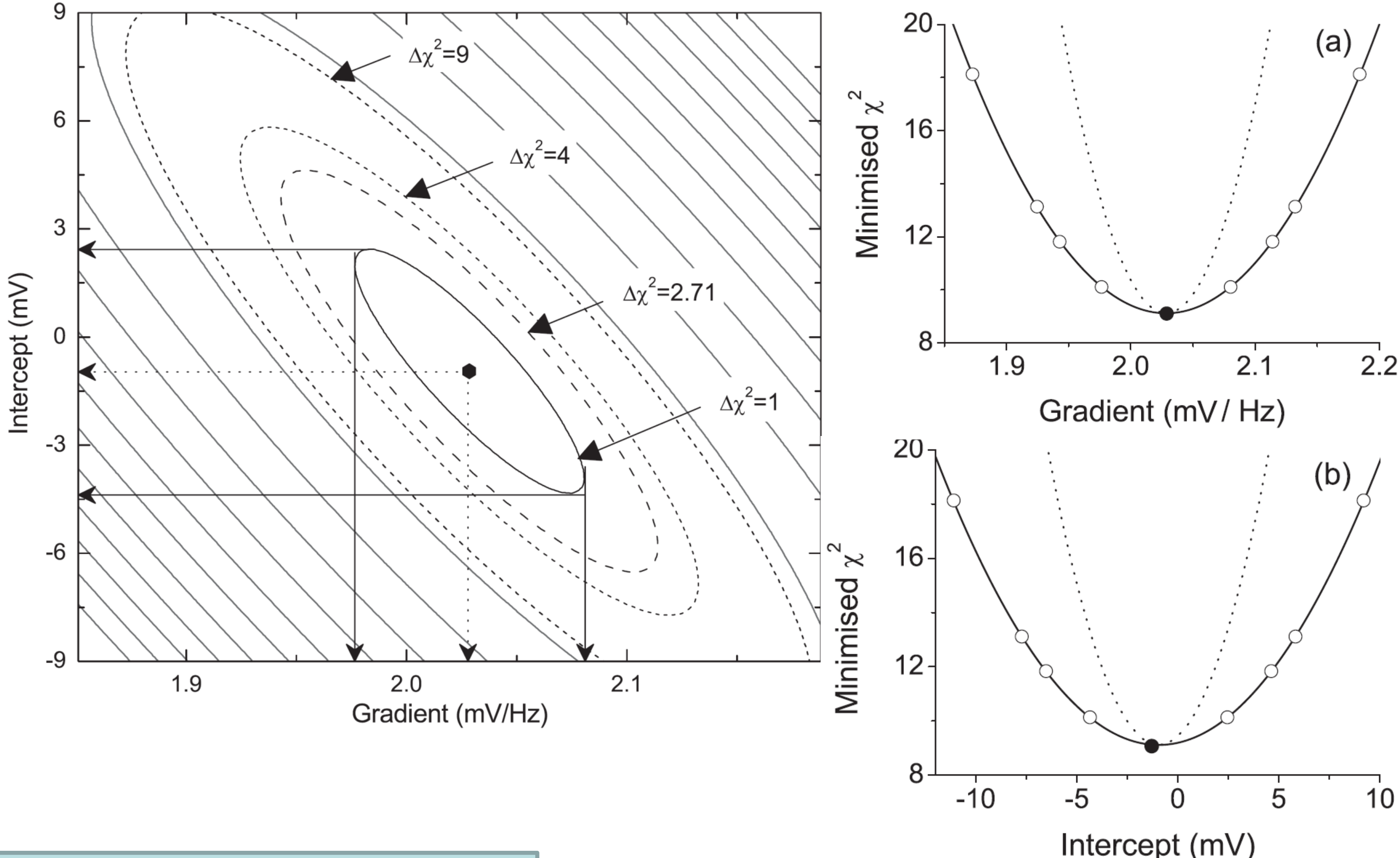
From the integrated probability distribution $\Delta\chi^2 \geq 1$ corresponds to 31.7% of the probability, or $\Delta\chi^2<1$ corresponds to 68.3%.

$$\chi^2 \longrightarrow \chi^2_{\min} + 1$$

Measurements and their Uncertainties, Pg. 76

Data Reduction and Error Analysis
P R Bevington, Pg. 211

Calculating the error in a least-squares fit



Measurements and their Uncertainties, Pg. 78

Curvature matrix for a straight-line fit

AtMol. Durham Atomic & Molecular Physics

For a **straight-line fit** the χ^2 surface is perfectly parabolic with respect to both variables. Therefore:

- There is only one minimum
- It is easy to find the minimum
- Analytic results exist for the **curvature matrix** which allow for easy determination of the errors:

$$A_{cc} = \sum_i \frac{1}{\alpha_i^2},$$

$$A_{cm} = A_{mc} = \sum_i \frac{x_i}{\alpha_i^2},$$

$$A_{mm} = \sum_i \frac{x_i^2}{\alpha_i^2}.$$

Measurements and their Uncertainties, Pg. 93

Topical example of chi-squared contours

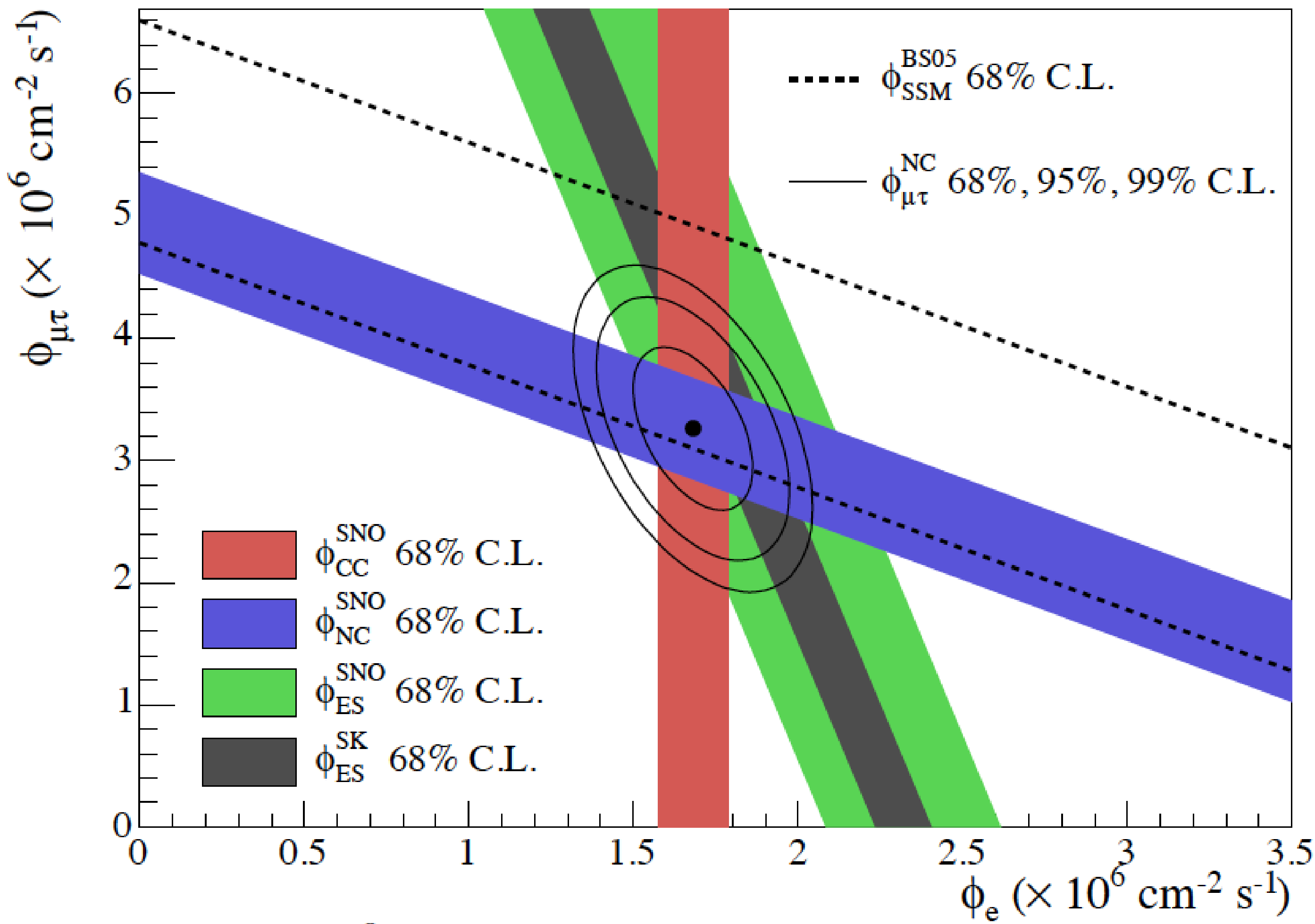


Figure 4: Fluxes of ^{8}B solar neutrinos from SNO and Super-Kamiokande. The SSM BS05 prediction is shown as a range between the dashed lines. C.L. stands for confidence level.
From: Scientific Background on the Nobel Prize in Physics 2015 NEUTRINO OSCILLATIONS

Traversing the χ^2 surface for an arbitrary function

AtMol
Durham Atomic & Molecular Physics

For an arbitrary function the χ^2 surface can be much more complicated! Therefore:

- There can be numerous **local** minima
- Good choice of **initial parameters** are necessary to avoid being trapped in a local minimum
- The $\Delta\chi^2$ contours can be asymmetric – Gaussian confidence limits might become questionable!
- There are unlikely to be analytic results for the curvature, therefore **numerical techniques** are necessary

Iterative approaches

- Newton-Raphson

$$x_{s+1} = x_s - \frac{f\,(x_s)}{f'\,(x_s)}$$

Update Rule

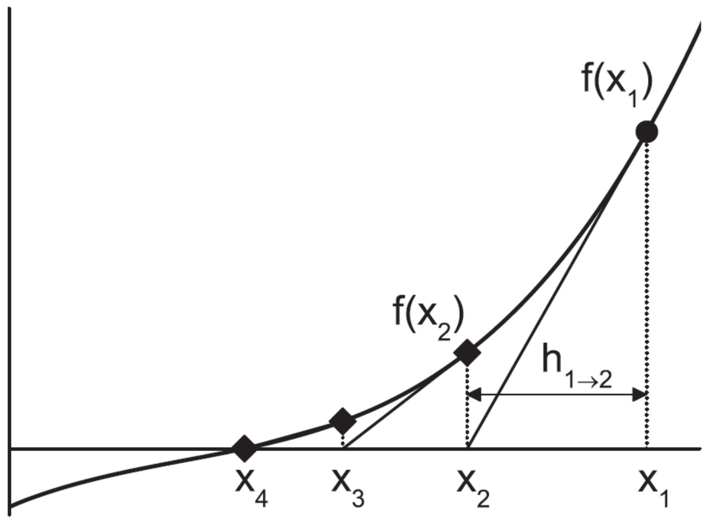


Fig. 7.1 The Newton–Raphson method for finding the zero crossing of a function. After s iterations the function is $f\,(x_s)$, and its derivative at x_s is $f'\,(x_s)$. The tangent to the curve is drawn at x_s, and the zero crossing of the tangent found. This point is taken as the updated coordinate of the zero-crossing point of the function, x_{s+1}. The iterative procedure continues until convergence at the desired tolerance.

Measurements and their Uncertainties, Pg. 86

Iterative approaches

- Grid Search

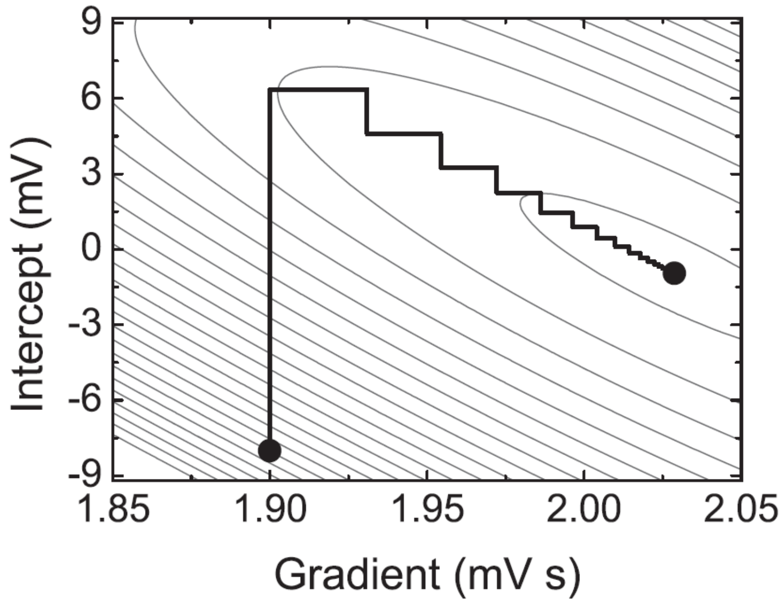


Fig. 7.2 An illustration of a grid search on a two-dimensional error surface. Initial trial values of the gradient and intercept were chosen (gradient 1.90 mV s, intercept −8 mV), and each parameter was optimised in turn until convergence and the minimum χ^2 was obtained. Note the inefficient zig-zag motion along the flat valley bottom.

Measurements and their Uncertainties, Pg. 87

AtMol. Durham Atomic & Molecular Physics

Iterative approaches

- Gradient descent

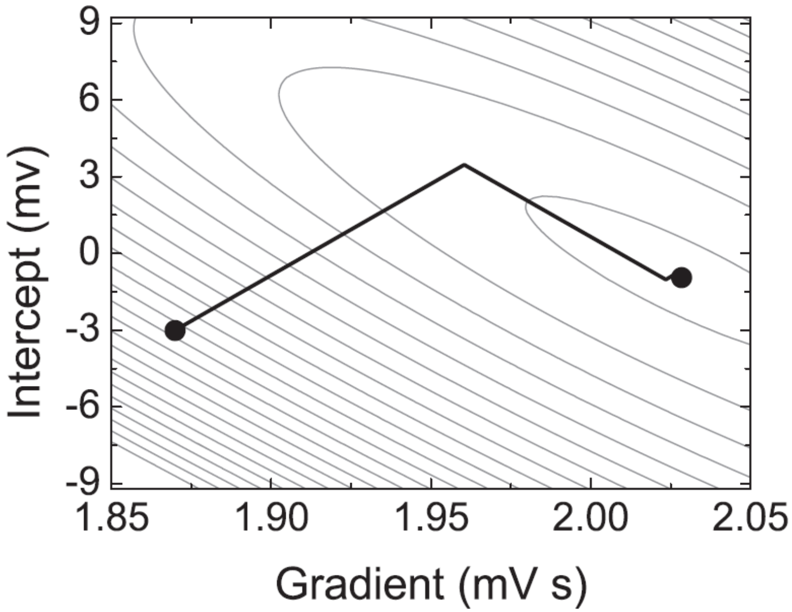


Fig. 7.4 An illustration of a gradient-descent method on a two-dimensional error surface. Each parameter is optimised simultaneously along a vector defined using the gradient.

Update Rule

$$\mathbf{a}_{s+1} = \mathbf{a}_s - \beta \nabla \chi^2 (\mathbf{a}_s)$$

$$\left(\nabla \chi^2\right)_j = \frac{\partial \chi^2}{\partial a_j} \approx \frac{\chi^2 (a_j + \delta a_j) - \chi^2 (a_j)}{\delta a_j}$$

Measurements and their Uncertainties, Pg. 88

Second-order expansion: the Newton method

$$\chi^2(\mathsf{a}_s + \mathsf{h}) \approx \chi^2(\mathsf{a}_s) + \mathsf{g}_s^{\mathrm{T}}\mathsf{h} + \frac{1}{2}\mathsf{h}^{\mathrm{T}}\mathsf{H}_s\mathsf{h}$$

the **gradient vector** $\mathsf{g}_s = \nabla\chi^2(\mathsf{a}_s) = \left[\frac{\partial\chi^2}{\partial a_1}, \cdots, \frac{\partial\chi^2}{\partial a_{\mathcal{N}}}\right]^{\mathrm{T}}$

the $\mathcal{N} \times \mathcal{N}$ **Hessian matrix** H_s is

$$\mathsf{H}_s = \mathsf{H}(\mathsf{a}_s) = \begin{bmatrix} \frac{\partial^2\chi^2}{\partial a_1^2} & \cdots & \frac{\partial^2\chi^2}{\partial a_1 \partial a_{\mathcal{N}}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2\chi^2}{\partial a_1 \partial a_{\mathcal{N}}} & \cdots & \frac{\partial^2\chi^2}{\partial a_{\mathcal{N}}^2} \end{bmatrix}$$

Measurements and their Uncertainties, Pg. 85

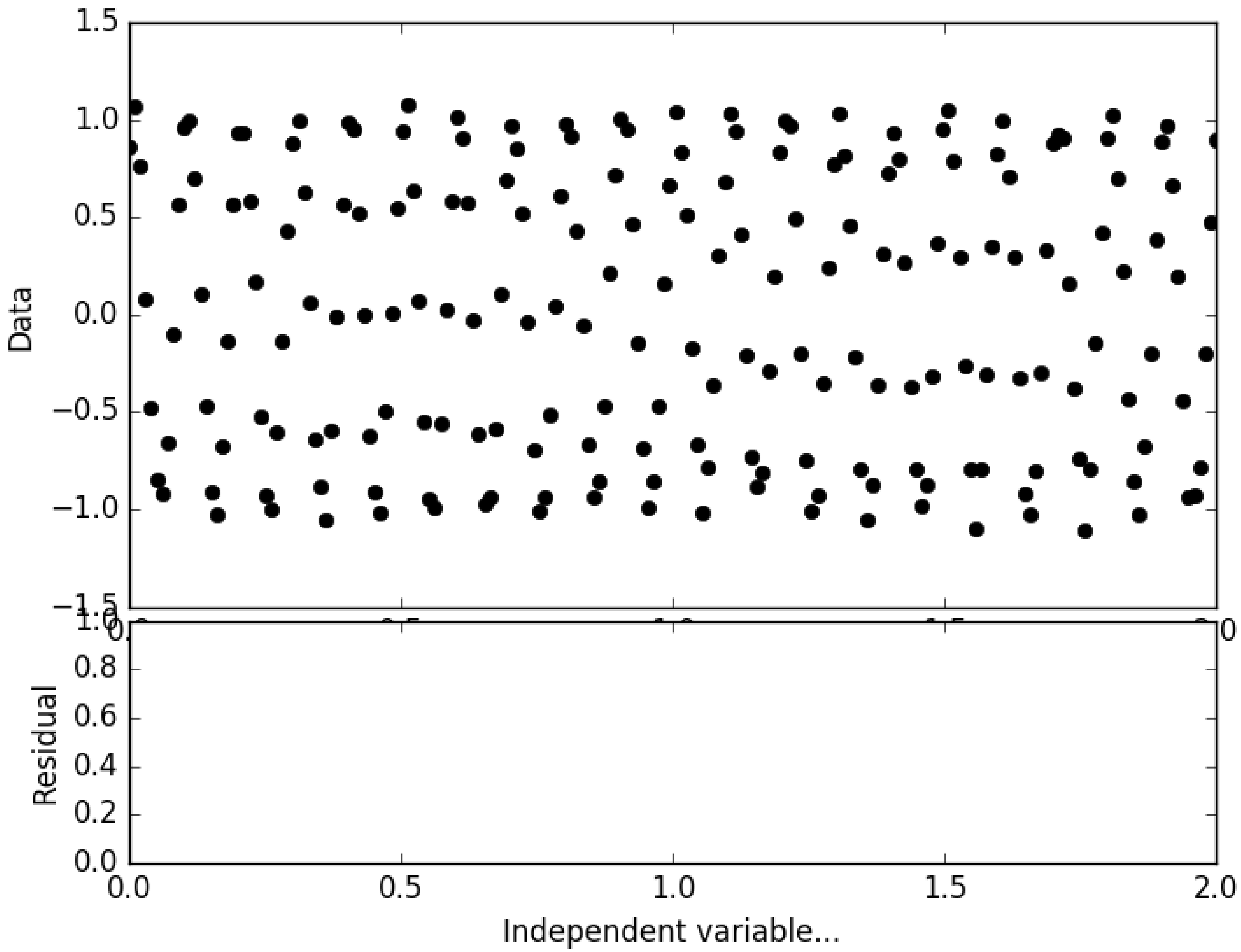
Marquardt-Levenberg Method

Combines best features of gradient and expansion approaches

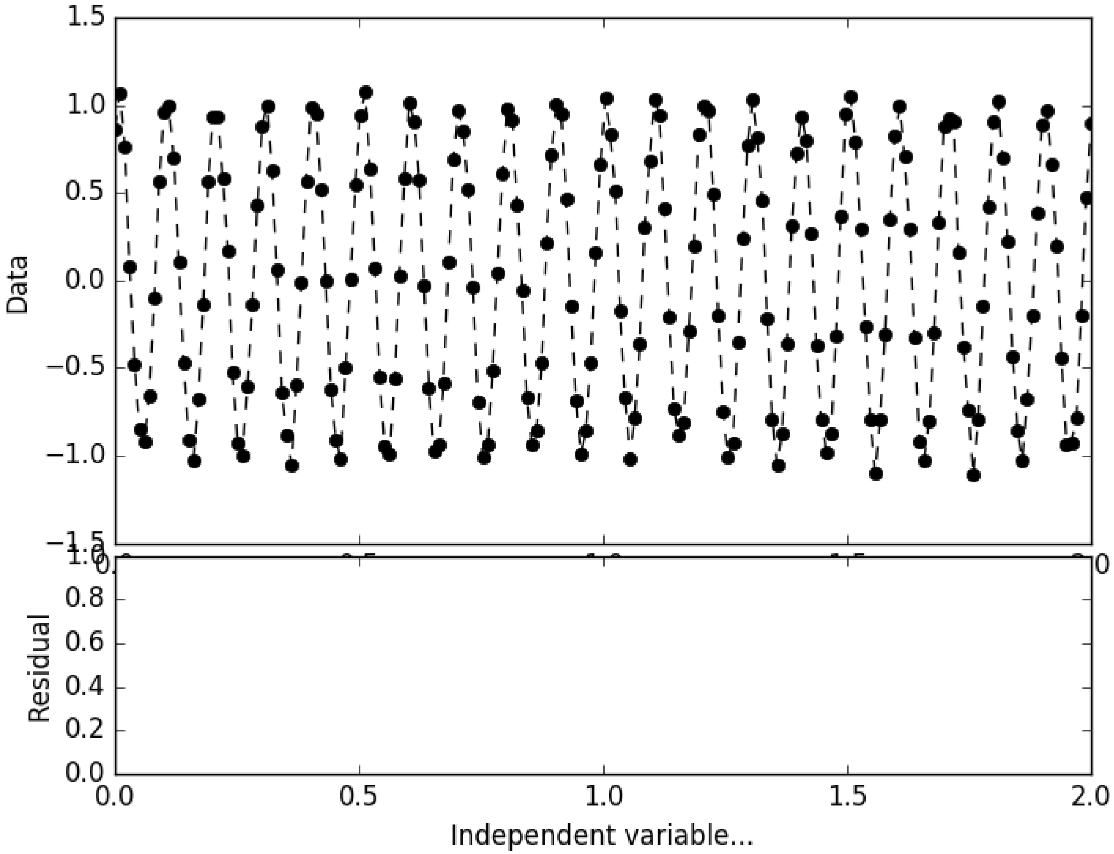
- **Steepest descent** when far from optimum values
- **Expansion method** when surface becomes parabolic

Update Rule

$$\mathsf{a}_{s+1} = \mathsf{a}_s - \left(\mathsf{H}_s + \lambda\,\mathrm{diag}\left[\mathsf{H}_s\right]\right)^{-1} \mathsf{g}_s$$

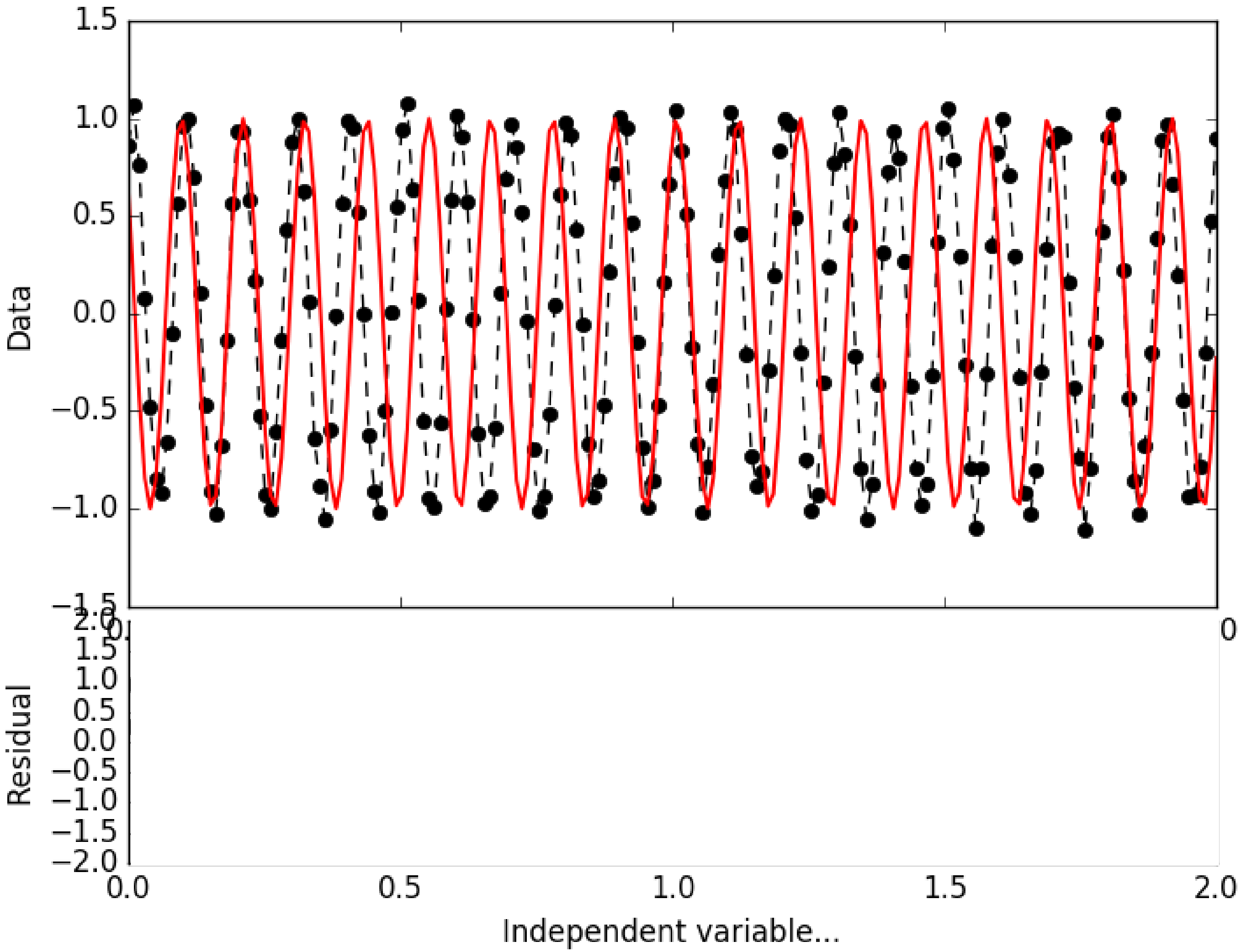


Raw data. Without a theory very hard to know what it might be!

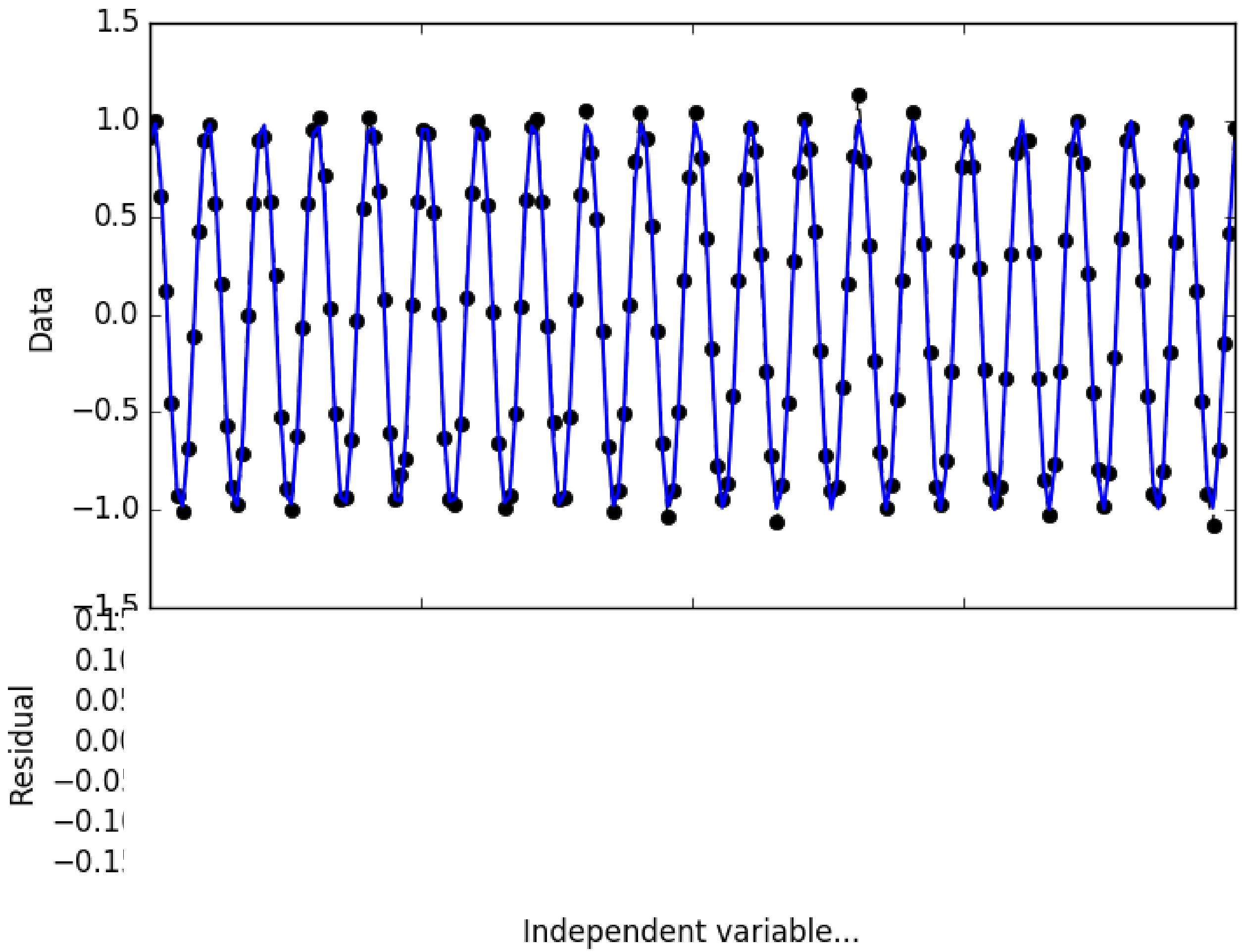


A sinusoidal fit.

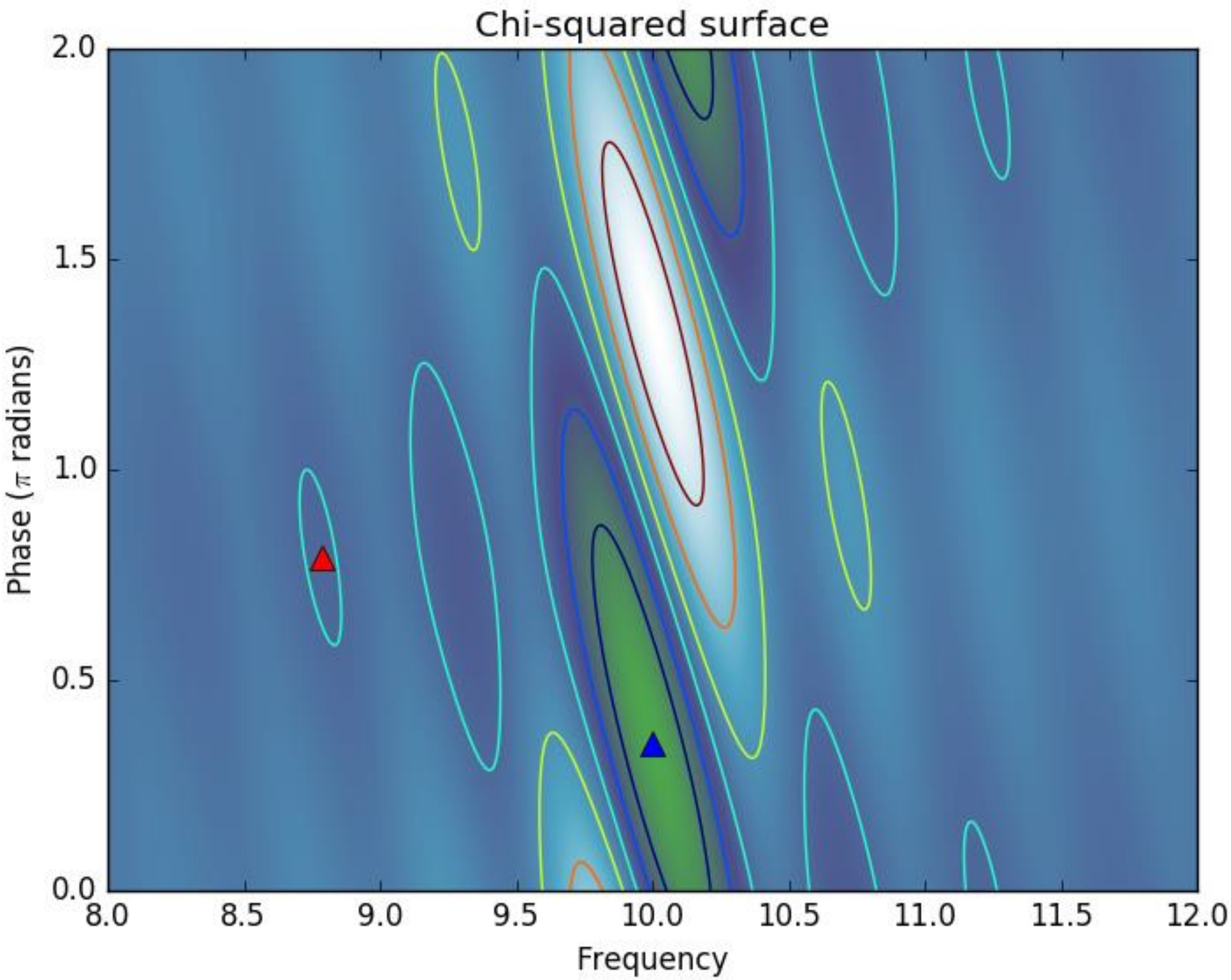
That looks better!



A “local minimum” fit – many peaks fit, but WRONG period – hence residuals nice beat pattern. Shows danger of not getting good enough initial guess for period.



Now with fit! Clearly a “good fit”, structure-less residuals.


Chi-squared surface
2.0
1.5
1.0
0.5
0.0
Phase (π radians)
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
Frequency

The **curvature matrix A** is one half of the Hessian matrix

The **covariance** or **error matrix C** is the inverse of the curvature matrix

$$[\mathrm{C}] = [\mathrm{A}]^{-1}$$

The variance in the j^{th} parameter is given by C_{jj}, the j^{th} diagonal element of the error matrix evaluated with the best-fit paramters.

$$\boxed{\alpha_j = \sqrt{C_{jj}}}$$

Measurements and their Uncertainties, Pg. 92

The off-diagonal elements of the covariance matrix are the **correlation coefficients**

Worked example for *V=mf+c*

What is the voltage and its error for *f*=75 Hz?

Ignoring the correlation the answer is

$$\alpha_V^2 = f^2\alpha_m^2 + \alpha_c^2 = f^2 C_{22} + C_{11}$$

Incorporating the correlation the answer is

$$\alpha_V^2 = f^2 C_{22} + C_{11} + 2 f C_{12}$$

Measurements and their Uncertainties, Pg. 94

AtMol.
Durham Atomic & Molecular Physics

“…error analysis is a participation, rather than a spectator, sport.”

Please attempt homework before 5pm next Monday

(28th October).

For MiSCaDA students via https://notebooks.dmaitre.phyip3.dur.ac.uk/miscada-da/hub/login

For PhD students either the notebook server OR email me a document