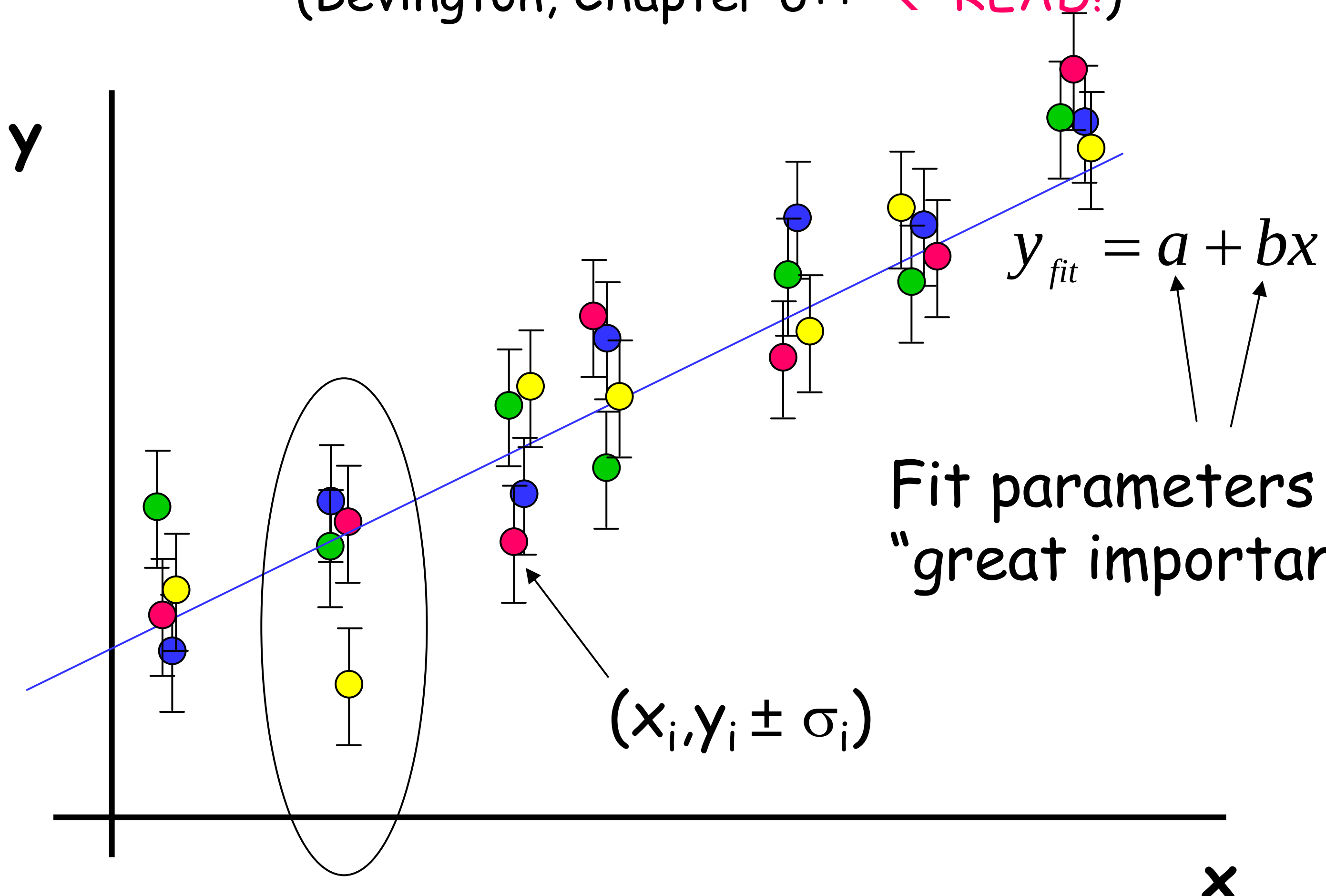
Week 4

The deeper meaning of χ^2

Recall Least-Squares Fitting

(Bevington, Chapter 6++ ← READ!)



Fit parameters are of "great importance"

Expect repeat measurement of i^{th} point to fall within $\pm 1\sigma$ about 2/3 times.

Vary parameters to minimize χ^2

$$\boxed{\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - y_{fit}(x_i)}{\sigma_i} \right)^2}$$

General form

$$\chi^2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left(y_i - (a + bx_i) \right)^2$$

This case (straight-line fit)

What are the "best" values?

Must estimate these correctly beforehand.

Sometimes an analytic solution exists… like in this case

$$\frac{\partial \chi^2}{\partial a} = 0 = \sum_{i=1}^{N} \frac{1}{\sigma_i^{\ 2}} \left(y_i - (a + b x_i) \right)$$

$$\frac{\partial \chi^2}{\partial b} = 0 = \sum_{i=1}^{N} \frac{1}{\sigma_i^{\ 2}} x_i \left(y_i - (a + b x_i) \right)$$

2 equations
2 unknowns

Solve via method of determinants

$$a = (\text{big formula involving sum over data points})$$

$$b = (\text{another big formula})$$

See Bevington Eq 6.13 !

The fit parameters themselves have uncertainty

$$\sigma_a^2 = \sum_{i=1}^{N} \left(\frac{\partial a}{\partial y_i} \right)^2 \sigma_i^2$$

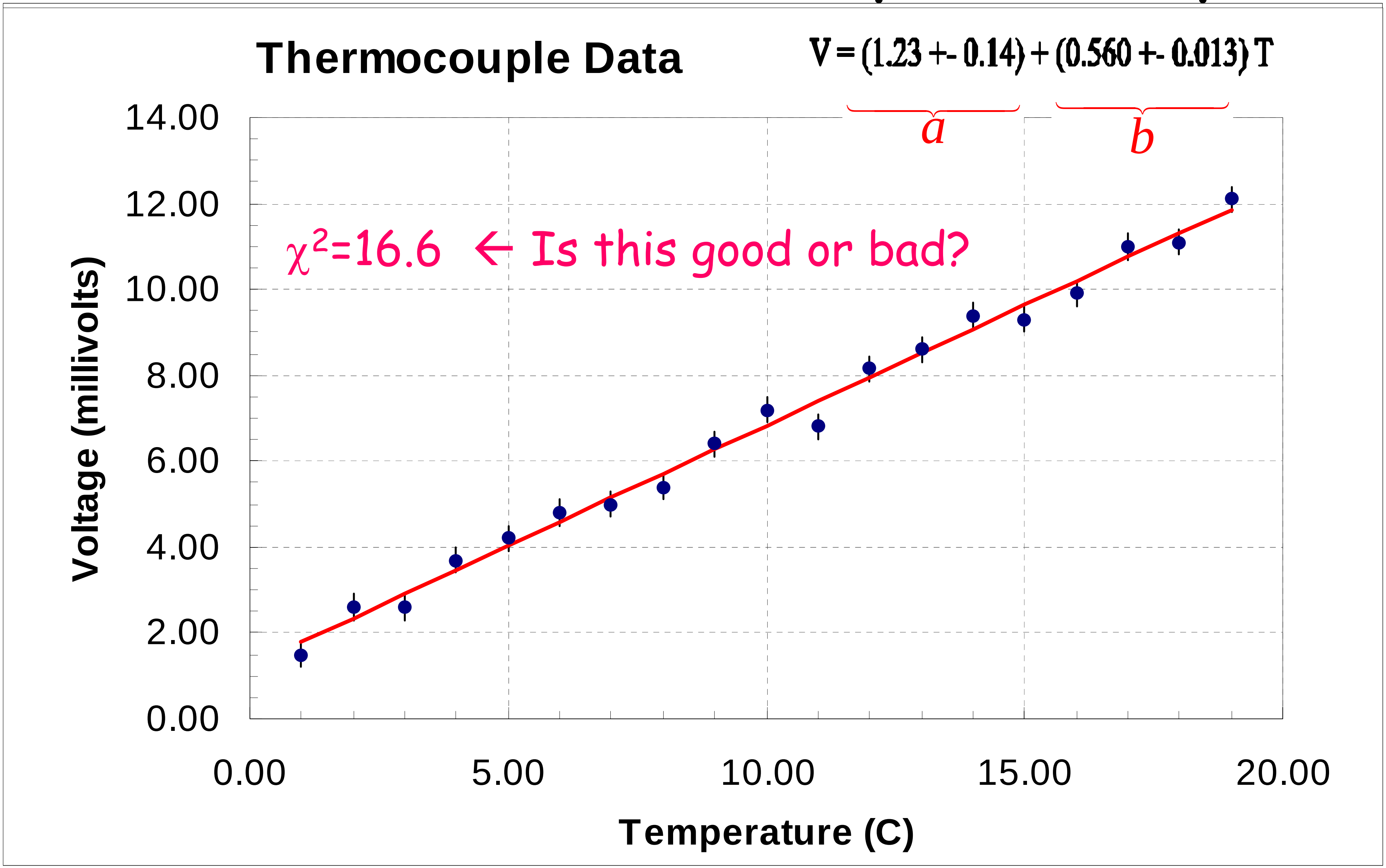
$$\sigma_b^2 = \sum_{i=1}^{N} \left(\frac{\partial b}{\partial y_i} \right)^2 \sigma_i^2$$

Can be done analytically for linear fits

See Bevington Eq 6.21 & 6.22!

The equations are straightforward but awkward to compute. For straight-line fits, they are implemented in mpl_LSF3.

Recall our Thermocouple Example



When is a fit "good"?

χ^2=16.6 ← Is this good or bad?

$$N - n = 19 - 2 = 17 \equiv \nu$$

Number of data points

Number of fit parameters or "constraints"

Number of <u>Degrees of Freedom</u> of the fit

$$\chi_\nu^2 \equiv \chi^2 / \nu = 0.98$$

Yes, this is a good fit…but why?

When is a fit “good”?

$$\chi^2 \equiv \sum_{i=1}^{N} \underbrace{\frac{\left(y_i - y_{fit}(x_i)\right)^2}{\sigma_i^2}}_{\approx 1,\ \text{on the average}}$$

Actual spread of data

Expected spread

$$\simeq N$$

But, we “used up” n constraints in the fitting function, so

$\chi^2 \to N - n = \nu$ and hence $\chi^2 / \nu \simeq 1$

When is a fit "good"?

$$\chi^2 \equiv \sum_{i=1}^{N} \underbrace{\frac{\left(y_i - y_{fit}(x_i)\right)^2}{\sigma_i^2}}_{\color{red}{\approx 1,\ \text{on the average}}}$$

Actual spread of data

Expected spread

$$\boxed{\chi^2/\nu \simeq 1 \quad \text{for a good fit}}$$

Remember this rule

Important: if the errors on the data points are badly estimated and too big (you messed up…), this value will be too small. The fit is "too good to be true"!

How does χ^2 vary near its minimum?

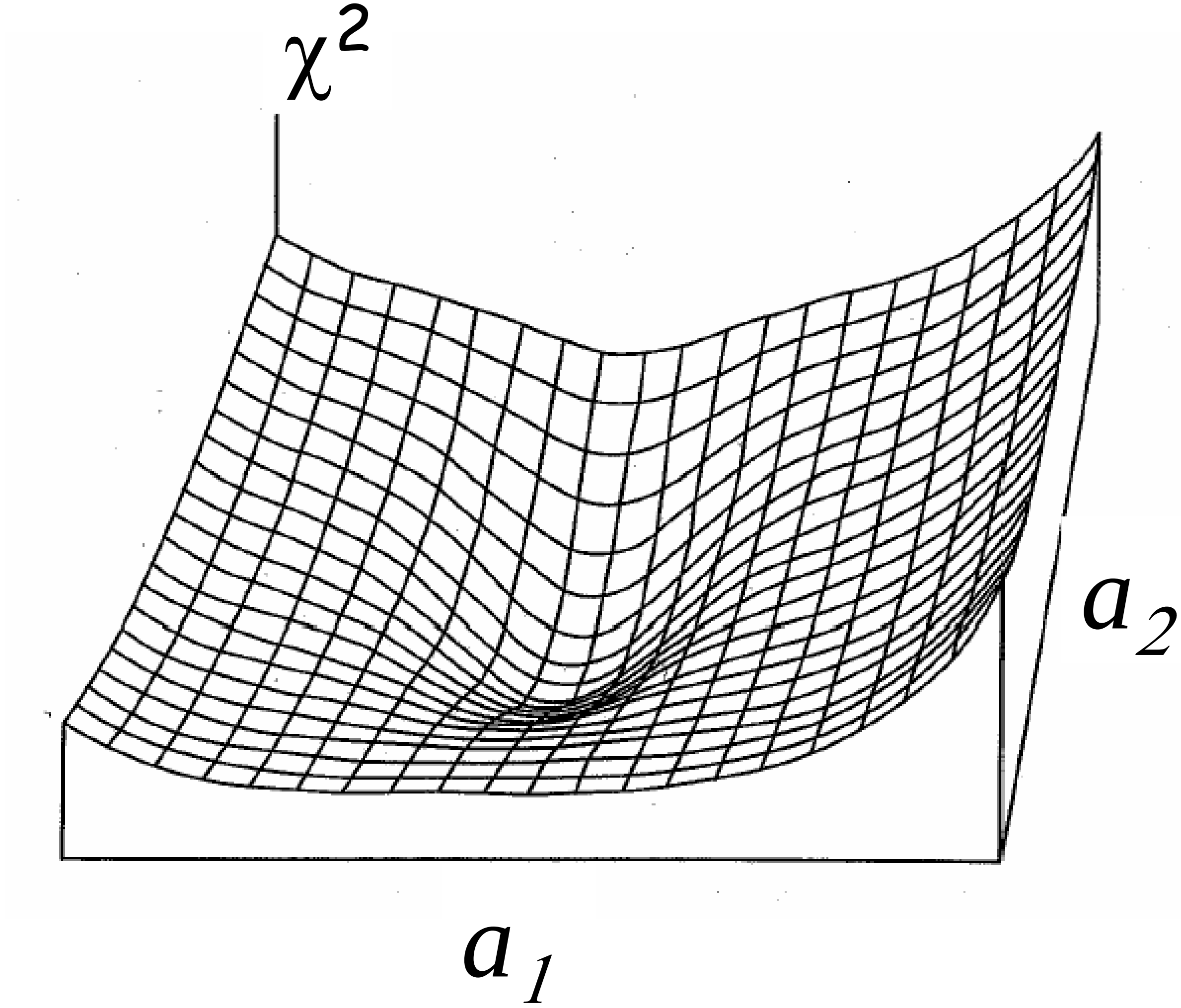
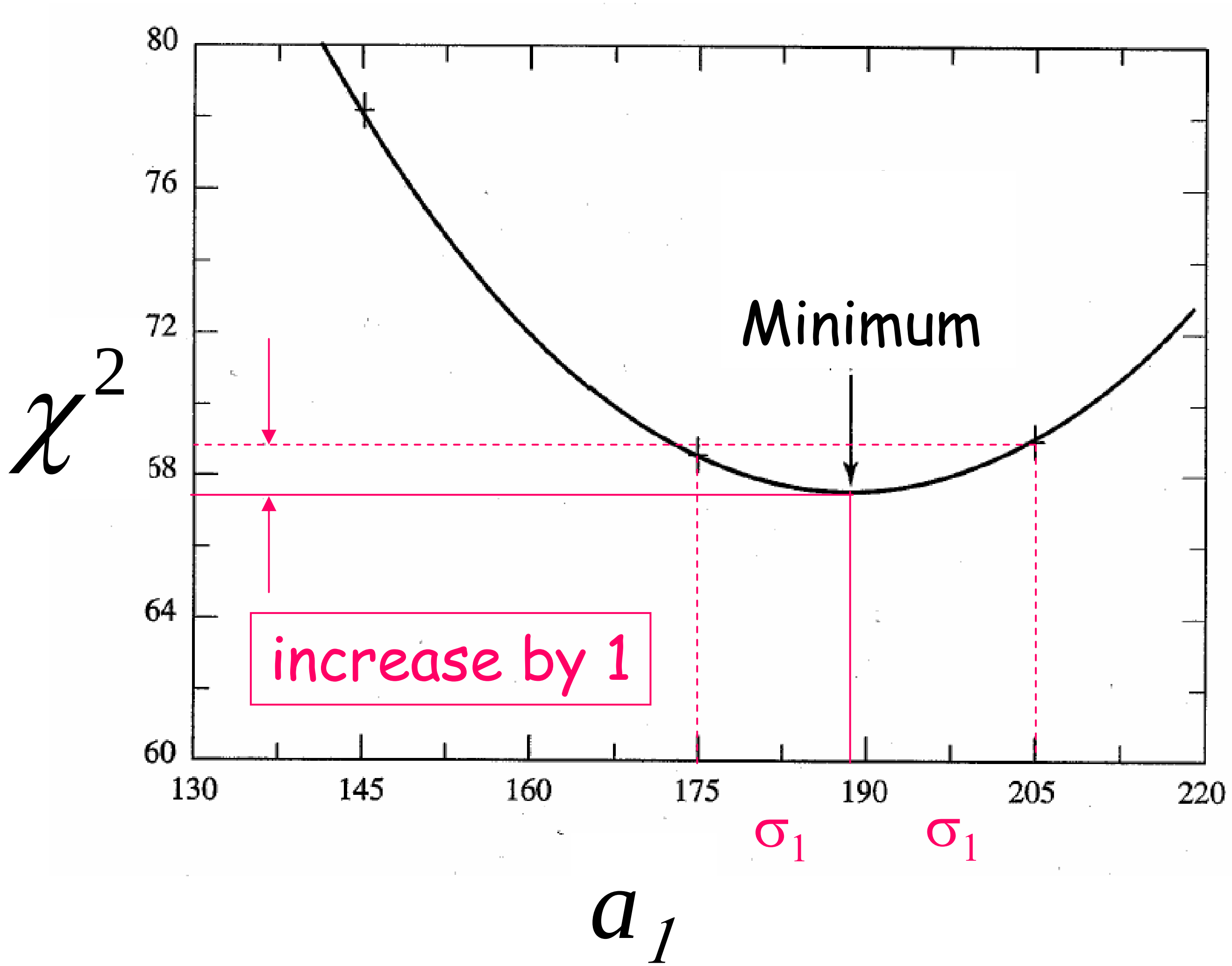


FIGURE 8.2
Chi-square hypersurface as a function of two parameters.

How does χ^2 vary near its minimum?



How does χ^2 vary near its minimum?

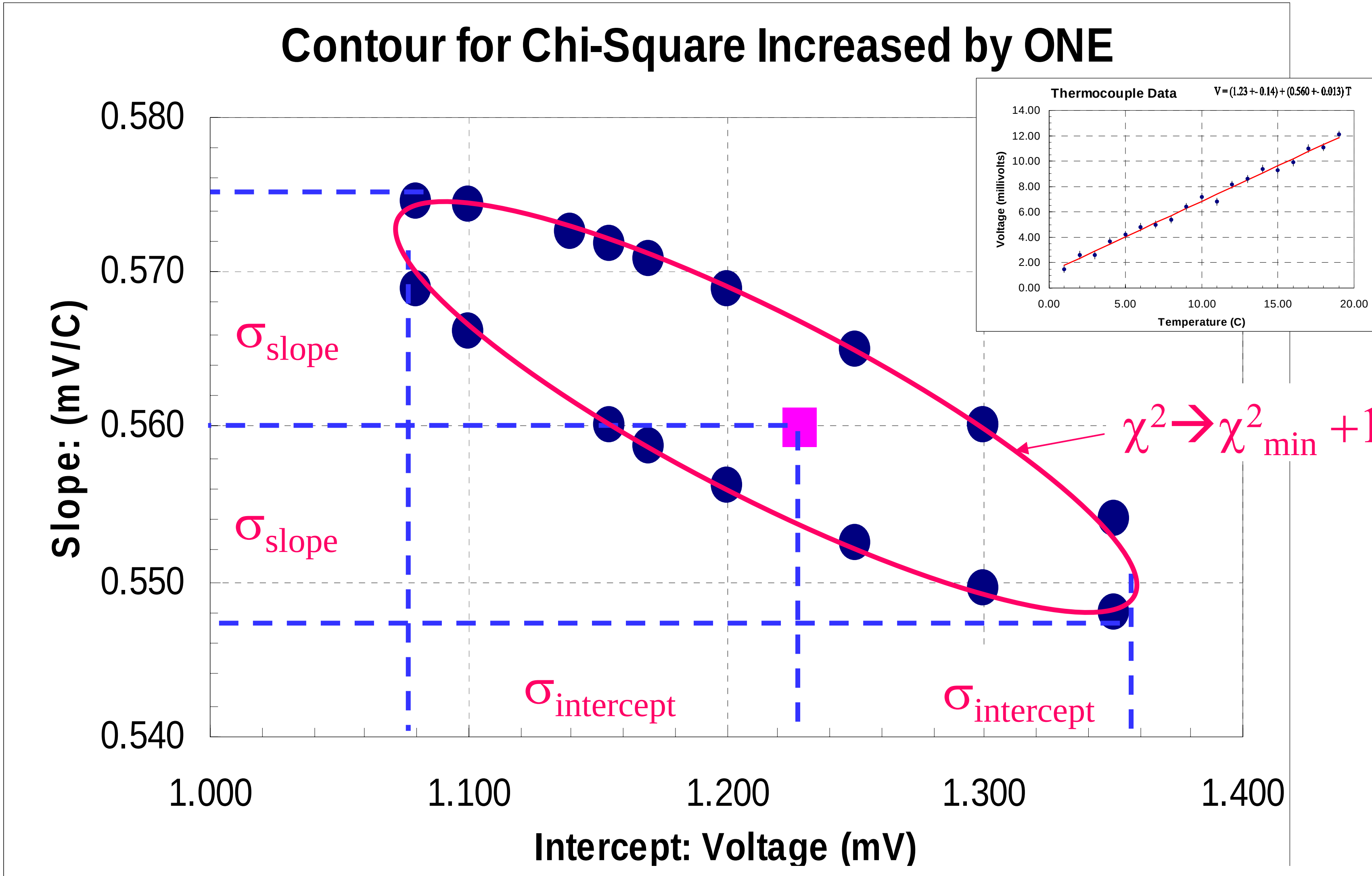
Near the minimum, for large enough data samples, varying parameter a_j varies χ^2 quadratically.
See Bevington Section 8.1

$$\chi^2 = \frac{\left(a_j - a_j^{\min}\right)^2}{\sigma_j^2} + C$$

C depends on all the data points and all the other fit parameters

We see that changing a_j by σ_j in either direction causes χ^2 to increase by 1.

How does χ^2 vary near its minimum?



Tilting of the ellipse means the errors are correlated

Summary

- Data errors must be estimated before fitting
- $\chi^2/\nu \sim 1$ is the benchmark for a "good" fit
- Errors on fit parameters can be estimated by letting χ^2 increase by 1
- χ^2 quadratic near a well-behaved minimum
- Errors on fit parameters tend to be correlated (for good or ill…)

Backup Slides…