Fitting data

Data fitting is the systematic variation of the parameters \( p_m (m = 1, \ldots, P) \) of a theory in order to find a parameter combination which minimises the discrepancies with experimental. This is most commonly done by minimising the \( \chi^2 \) measure of these differences, as defined below, and in section 15.5 we discuss strategies for this minimisation. The fitting parameters of a reaction model could be those which specify Woods-Saxon potentials, or they could be structural parameters such as deformations or spectroscopic factors. Sometimes the parameters linearly determine the model predictions (such as spectroscopic factors in one-step DWBA), but more often the predictions depend non-linearly on the model parameters (especially those defining the potentials).

15.1 \( \chi^2 \) measures

15.1.1 Method

We want to compare theoretical with experiment cross sections, and use this comparison to improve the theory input parameters \( \{p_j\} \) so that observed and predicted cross sections agree better. Let the experimental cross sections be \( \sigma_{\text{ex}}(i) \) from observations at angles \( \theta_i \) and (possibly identical) energies \( E_i, i = 1, 2, \ldots, N \). Suppose that these experimental results \( \sigma_{\text{ex}}(i) \) are randomly distributed, with standard deviations \( \Delta \sigma(i) \) as absolute errors. If a theoretical model predicts values \( \sigma_{\text{th}}(i) \) at the same angles and energies, then the standard measure of their discrepancy is the \( \chi^2 \) sum

\[
\chi^2 = \sum_{i=1}^{N} \frac{(\sigma_{\text{th}}(i) - \sigma_{\text{ex}}(i))^2}{\Delta \sigma(i)^2}.
\]
If the theory and experiment agreed exactly, then $\chi^2 = 0$, but that is statistically unlikely because of the random errors $\Delta \sigma(i)$ in each data point. Even if the theory is completely accurate, the most likely outcome is that the discrepancy between theory and experiment will be of the order of $\Delta \sigma(i)$, one standard deviation error in the experimental value. This implies that the best achievable $\chi^2$ value is obtained when each fraction $(\sigma_{ex}(i) - \sigma_{th}(i))^2 / \Delta \sigma(i)^2$ is of the order of unity, so that $\chi^2 \sim N$. For this reason, we usually describe a fit in terms of the value of $\chi^2 / N$: the $\chi^2$ per degree of freedom.

The best to be expected in practice is that $\chi^2 / N \sim 1$. If $\chi^2 / N \gg 1$ then the data is not yet described fully by theory. If $\chi^2 / N \ll 1$, then the statistical improbability of this ‘perfect fit’ implies that the errors $\Delta \sigma(i)$ have been over-estimated. This commonly occurs if systematic or scaling errors have been included by mistake in the value of $\Delta \sigma(i)$ for each individual error. The values of $\Delta \sigma(i)$ are supposed to be the one standard deviation errors and all statistically independent, which means that even the best curve should miss every third error bar.

If there are uncertainties in the overall scaling of the experimental cross section, then the actual scaling factor $s$ should be regarded as another parameter to be fitted. If the expected value of $s$ is $\bar{s}$ (usually unity) but the 1-σ uncertainty around $\bar{s}$ is $\Delta s$, then the $\chi^2$ sum should have an extra term $(s - \bar{s})^2 / \Delta s^2$, and appear as

$$\chi^2 = \frac{(s - \bar{s})^2}{\Delta s^2} + \sum_{i=1}^{N} \frac{(\sigma_{th}(i) - s \sigma_{ex}(i))^2}{\Delta \sigma(i)^2}.$$  \hspace{1cm} (15.1.2)

The $\chi^2$ per degree of freedom is now $\chi^2 / (N + 1)$ because of the additional ‘data point’ $\bar{s}$.

### 15.1.2 Multivariate theory

#### Multivariate normal distribution

In order to understand the principles underlying $\chi^2$ measures, we need first to know the general probability distribution for a set of random variables (‘variates’) $x_i$. If we have no other knowledge, we would generally assume that they follow normal distributions, for which the probability density function for a single variate is

$$f(x) = \frac{1}{\sqrt{2\pi} \Delta} \exp \left[-\frac{(x - \mu)^2}{2\Delta^2}\right]$$  \hspace{1cm} (15.1.3)
when the mean is $\mu = E[x]$ and the standard deviation $\Delta$ where $\Delta^2 = E[(x - \mu)^2]$ is the variance. We use the notation $E[X]$ for ‘expectation value’ for some expression $X$ depending on $x$, the value averaged over the probability density function for $X$:

$$E[X] \equiv \int X f(x) dx ,$$

(15.1.4) so

$$\Delta^2 = E[(x - \mu)^2] = E[x^2] - 2\mu E[x] + \mu^2 = E[x^2] - \mu^2 .$$

(15.1.5) and the standard deviation $\Delta = \sqrt{E[(x - \mu)^2]}$ is the root mean square deviation from the mean.

For many correlated variables $x = \{x_1, \ldots, x_N\}$, their joint probability distribution is the more general formula [1, p. 62]

$$f(x) = (2\pi)^{-\frac{N}{2}} |V|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} (x - \mu)^T V^{-1} (x - \mu) \right] ,$$

(15.1.6) where $V$ is the symmetric covariance matrix

$$V_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)] ,$$

(15.1.7) and $|V|$ is its determinant. The diagonal elements $V_{ii}$ of the covariance matrix are the individual variances $\Delta_i^2$, and its off-diagonal elements are

$$V_{ij} = \rho_{ij} \Delta_i \Delta_j$$

(15.1.8) where $\rho_{ij}$ is the correlation coefficient between the fluctuations of the data points $x_i$ and $x_j$ about their respective means.

### The $\chi^2$ sum

The probability that a single data point $x_i$ with variance $\Delta_i^2$ is correctly fitted by a prediction $y_i$ may be found using Eq. (15.1.3):

$$f_i(y_i) = \frac{1}{\sqrt{2\pi \Delta_i}} \exp \left[ -\frac{(x_i - y_i)^2}{2\Delta_i^2} \right] .$$

(15.1.9) For many statistically independent data points $x_i$, the joint probability of being correctly fitted by $y_i$ is the product of such single probabilities:

$$P_{\text{tot}} = (2\pi)^{-\frac{N}{2}} \Delta^{-1} \exp \left[ -\frac{1}{2} \sum_{i=1}^{N} \frac{(x_i - y_i)^2}{\Delta_i^2} \right]$$

(15.1.10) \hspace{1cm} (15.1.11)
Fitting data

once we define $\Delta = \prod_i^N \Delta_i$ and

$$\chi^2 = \sum_i^N \frac{(x_i - y_i)^2}{\Delta_i^2} \quad (15.1.12)$$

using the definition of Eq. (15.1.1). Here, we write $x_i$ as the data value for which $y_i$ is the theoretical prediction.

Comparison of Eq. (15.1.11) with Eq. (15.1.6) shows that a generalised definition of the $\chi^2$ sum, suitable for correlated data, is

$$\chi^2 = (x - y)^T V^{-1} (x - y) \quad (15.1.13)$$

$$= \sum_{ij=1}^N (x_i - y_i)[V^{-1}]_{ij}(x_j - y_j) \quad (15.1.14)$$

where $V$ is the covariance matrix of the data points.

In both cases, minimising $\chi^2$ is equivalent to maximising $P_{\text{tot}}$, the probability that the statistical means of the data points $x$ are described by the predictions $y$.

**The $\chi^2$ distribution**

If we add together $N$ squares of independent normal distributions, $z_i^2$, with zero mean and unit variance, then the sum

$$\chi^2 = \sum_{i=1}^N z_i^2 \quad (15.1.15)$$

has what is called a $\chi^2_{(N)}$ distribution. This is the well known probability distribution

$$f(\chi^2) = \frac{1}{2\Gamma(N/2)} \left(\frac{\chi^2}{2}\right)^{N/2-1} e^{-\chi^2/2} , \quad (15.1.16)$$

which has mean, variance and standard deviation

$$E[\chi^2] = N ; \quad V(\chi^2) = 2N ; \quad \sigma(\chi^2) = \sqrt{2N} . \quad (15.1.17)$$

The $\Gamma(z)$ is the Gamma function, $\Gamma(z) = (z - 1)!$.

For large $N \gtrsim 20$, the $\chi^2_{(N)}$ distribution is approximately normal with the means and variances of Eq. (15.1.17). This implies that $\chi^2/N$, with mean 1 and standard deviation $\sqrt{2/N}$, is also approximately normal.
Perfect and non-perfect fits

The optimal fit of a set of data occurs when the theory predicts exactly the statistical means of all the experimental values. This is not to give the actual values on any particular occasion, but the average if the same experiment were repeated many times at the same angle and the same energy. In this ‘perfect’ case we have \( y_i = E[x_i] = \mu_i \), and the value of \( \chi^2 \) is a sum of terms

\[
z_i^2 = \frac{(x_i - y_i)^2}{\Delta_i^2} = \frac{(x_i - \mu_i)^2}{\Delta_i^2},
\]

(15.1.18)

which therefore have zero means and unit variances. If these were normally distributed, then the sum \( \chi^2 = \sum_i^N z_i^2 \) would follow the \( \chi^2(N) \) distribution, which we saw above has mean \( N \) and variance \( 2N \). This is the reason we stated earlier that the best to be expected in practice is that \( \chi^2/N \sim 1 \), and that (for a good fit) deviations both above and below unity are statistically improbable and must thus have other reasons.

Non-perfect fits will therefore yield larger \( \chi^2/N \gg 1 \) values, and will have been constructed from \( z_i^2 \) with different statistical properties. If the \( z_i \) means are non-zero but their variances are still unity, then \( \chi^2 \) will follow what is called [1, p. 66] a ‘non-central \( \chi^2 \) distribution with \( N \) degrees of freedom’.

More generally again, when the parameters giving the predictions \( y_i \) have been estimated by those values which minimise \( \chi^2 \), the distribution \( \chi^2 \) is no longer \( \chi^2(N) \). The minimum value of \( \chi^2 \) is then the sum of squares of correlated, non-central random variables of non-zero mean, which are not, in general, normally distributed, and hence difficult to calculate. However, we can still numerically search for a minimum of \( \chi^2 \) as a function of the theory parameters \( \mathbf{p} = \{p_1, ..., p_P\} \).

Properties of the \( \chi^2 \) minimum

At a smooth local \( \chi^2 \) minimum for parameter values \( \{p_0\} \), the \( \chi^2 \) function will have zero first derivatives, and hence be well described by the multivariate Taylor series

\[
\chi^2(p_1, ..., p_P) \approx \chi^2(p_0^0, ..., p_0^0) + \frac{1}{2} \sum_{m,n=1}^{P} H_{mn}(p_m - p_0^m)(p_n - p_0^n) \] \[ \equiv \chi^2(\mathbf{p}^0) + \frac{1}{2}(\mathbf{p} - \mathbf{p}^0)^T \mathbf{H} (\mathbf{p} - \mathbf{p}^0),
\]

(15.1.19)

(15.1.20)

where the matrix of second derivatives, the Hesse matrix \( \mathbf{H} \), is

\[
H_{mn} = \frac{\partial^2}{\partial p_m \partial p_n} \chi^2(p_1, ..., p_P).
\]

(15.1.21)
The fitted parameters $p_m^0$ will in general have correlated errors, even if the original data had uncorrelated errors. We may therefore define a parameter covariance matrix by

$$V_{mn}^p = E[(p_m - p_m^0)(p_n - p_n^0)] \tag{15.1.22}$$

for $m, n = 1, ..., P$, the number of parameters. If we now substitute Eq. (15.1.19) into Eq. (15.1.11), we find that, as the parameters are varied around the minimum point, the fitting probability $P_{tot}$ changes to

$$P_{tot} = \frac{1}{(2\pi)^{\frac{N}{2}} \Delta} e^{-\frac{1}{2} \sum_{mn} (p_m - p_m^0)H_{mn}(p_n - p_n^0)} \tag{15.1.23}$$

This probability of the data being fit by the theory can also be used for fixed data points, to show how the fit varies when the theory parameters are slightly changed. If we compare this equation with Eq. (15.1.6), we see that it describes a multivariate normal distribution for the parameters $p$ if we identify

$$(V^p)^{-1} = \frac{1}{2}H \quad \text{or} \quad V^p = 2H^{-1} \tag{15.1.24}$$

This important equation allows us to extract the correlated variances of the extracted parameters from the properties of the $\chi^2$ surface around its minimum point.

Moreover, from Eq. (15.1.3), we see that the $1\sigma$ deviations away from a mean are when the exponential in the normal formula is $\frac{1}{2}$. This implies that the set of parameter combinations that are $1\sigma$ away from the minimum may be determined from

$$\frac{1}{2} \sum_{mn} (p_m - p_m^0)H_{mn}(p_n - p_n^0) = 1 \tag{15.1.25}$$

and hence, even more simply by using Eq. (15.1.19), when

$$\chi^2(p_1, ..., p_P) = \chi^2(p_1^0, ..., p_P^0) + 1 \tag{15.1.26}$$

The sets of parameters satisfying this condition will in general be multi-dimensional ellipses around the minimum point. These results for parameter correlations will be used in section 15.5.2.
15.2 Legendre polynomial fits

The simplest description of an angular cross section is as a linear combination of Legendre polynomials

\[ \sigma(\theta) = \sum_{\lambda \geq 0} a_\lambda P_\lambda(\cos \theta) \]  (15.2.1)

for some real coefficients \( a_\lambda \) which could be determined from experimental data by e.g. \( \chi^2 \) minimisation. The comparison of theory with experiment is most often done by comparing cross sections directly, but is sometimes performed by comparing the fitted and predicted values for these coefficients, a procedure which is most practical at low energies where not too many \( \lambda \) values are needed.

To find the theoretical predictions for these coefficients, we note that in chapter 3 the formulae (3.1.46) and (3.2.24) give the scattering amplitude \( f(\theta) \) as linear combinations of Legendre polynomials: of \( P_L(\cos \theta) \) in the (3.1.46) case, and of associated Legendre polynomials \( P^M_L(\cos \theta) \) in the (3.2.24) case. If the square modulus \( \sigma = |f(\theta)|^2 \) for the cross section is expanded into a sum over two partial wave indices, then each term will contain a product of two Legendre polynomials. We will see below how the sum of all these products can be rewritten as a sum exactly of the form (15.2.1).

15.2.1 Cross section in a spin zero channel

Legendre expansion does not work with the point-Coulomb scattering amplitude, since the partial wave series of Eq. (3.1.80) with Coulomb phase shifts does not converge. We therefore restrict ourselves to neutron elastic scattering, or else a non-elastic cross section. From the general form of Eq. (3.1.46), we have

\[ \sigma(\theta) = \frac{1}{k^2} \sum_{L=0}^{\infty} (2L + 1) P_L(\cos \theta) T_L \]
\[ = \frac{1}{k^2} \sum_{L L'} (2L+1)(2L'+1) P_L(\cos \theta) P_{L'}(\cos \theta) T_L T_{L'} . \]  (15.2.2)

The product of two Legendre polynomials for the same argument is

\[ P_L(z) P_{L'}(z) = \sum_{\lambda = \mid L - L' \mid}^{L+L'} \langle L0, L'0 | \lambda0 \rangle^2 P_\lambda(z) , \]  (15.2.3)

where \( \langle L0, L'0 | \lambda0 \rangle \) is a Clebsch-Gordan coefficient. Non-zero contributions require even values of \( L+L'+\lambda \).
We thus derive the general form of Eq. (15.2.1) with coefficients

\[ a_\lambda = \frac{1}{k^2} \sum_{LL'} (2L+1)(2L'+1) \langle L0, L'0|\lambda0 \rangle^2 \mathbf{T}_L^* \mathbf{T}_{L'} . \]  

(15.2.4)

Note that \( a_0 \) is proportional to the angle-integrated total cross section of Eq. (3.1.50) by \( \sigma_T = 4\pi a_0 \), so often it is convenient to write

\[ \sigma(\theta) = \frac{\sigma_T}{4\pi} \left[ 1 + \sum_{\lambda>0} \bar{a}_\lambda P_\lambda(\cos \theta) \right] \]  

(15.2.5)

where \( \bar{a}_\lambda = a_\lambda/a_0 = 4\pi a_\lambda/\sigma_T \).

15.2.2 General cross sections

When some of the initial and final participating nuclei have non-zero spin, the above procedure may be repeated starting from Eq. (3.2.24). Because the cross section (3.2.20) in this case is a sum over all the magnetic substates of the nuclei, the Legendre expansion of the cross section is still in terms of simply the \( P_L \) as in Eq. (15.2.1).

The Legendre expansions for tensor analysing powers \( T_{kq}(\theta) \) of Eq. (3.2.31), by contrast, require associated Legendre polynomials \( P^M_L(z) \) for \( 0 \leq M \leq q \).

15.3 Fitting optical potentials

15.3.1 Sensitivities

Because the optical potentials go to zero as \(-V_{ws} \exp(-(R - R_{ws})/a)\) outside the sum the radii of the nuclei, they have diminishing effects on the forward angle cross sections. The Coulomb repulsion at radius \( R \) scatters to forward angles \( \theta \approx 2\eta/Rk \) by Eq. (3.1.76), so cross section at small angles should become equal to Rutherford \( \sigma \sim \sigma_R \) of Eq. (3.1.82). This fact can often be used to check the normalisation the absolute magnitude of experimental cross sections, by comparison to the Rutherford cross section at the smallest angles.

The spin-orbit potentials described in section 4.3.2 have their biggest effects on the vector analysing powers \( iT_{11}(\theta) \), but still have a small influence on the elastic cross section \( \sigma(\theta) \). If a polarised projectile or target enters into the reaction, then it is essential to include the spin-orbit forces for that nucleus.
15.3.2 Ambiguities in optical potentials

There are three principal kinds of ambiguities in fitting optical potentials, one kind associated with phase shifts at low energies, another with potential volume integrals at low energies, and the third concerning surface properties in heavy ion scattering.

The scattering at the lowest energies is determined by the phase shifts, and the pattern of phase shifts is largely unchanged if the potential is changed in depth so that the nearest bound or resonant state in a given partial wave remains the same distance from threshold. If the potential is made deeper, then a bound state would have an extra node, and guidance from shell-model filling orders may then be used to resolve these ambiguities.

At medium energies, it is often possible to make small changes in $V_{ws}$ and $R_{ws}$, while keeping $V_{ws}R_{ws}^2$ approximately constant, and not change the quality of the scattering fit. In these cases, the volume integral of the optical potential

$$J = \int V(r)dr = 4\pi \int_0^{\infty} V(r)r^2dr$$

(15.3.1)

is determined more precisely than the detailed parameters. Nucleon scattering potentials for targets of size $A$ generally scale as $J \sim AJ_0$ for $J_0 \approx 450$ MeV at zero energy.

For collisions of heavier nuclei, ‘grazing’ collisions are those in which the semiclassical impact parameter $b = L/k$ are comparable to the sum of the radii of the interacting nuclei. Any collisions closer than grazing will almost certainly lead to strong imaginary parts, because of the great probability of reactions leaving the elastic channel. In such reactions, the diffraction that occurs for grazing impact parameters will be governed by just the tail of the optical potentials at $R \gtrsim R_{ws} + a$. In this region, the real part is

$$V(R) \approx -V_{ws}e^{-(R-R_{ws})/a_{ws}} = -V_{ws}e^{R_{ws}/a_{ws}}e^{-R/a_{ws}}$$

(15.3.2)

and hence $R_{ws}$ and $V_{ws}$ may be varied together as long as $V_{ws}e^{R_{ws}/a_{ws}}$ is constant, because only the exponential tail is important.

15.4 Multi-channel fitting

15.4.1 Elastic fits

If a fitted optical potential is used as a diagonal potential in a coupled-channels system, then the predicted elastic cross sections will be different from the one-channel case, because of the extra coupling terms.
Fitting data

If we take a schematic pair of coupled channels

\[
[T_1 + U_1 - E_1] \psi_1(R) + V_{12} \psi_2(R) = 0
\]

\[
[T_2 + U_2 - E_2] \psi_2(R) + V_{21} \psi_1(R) = 0 ,
\]

(15.4.1)

then we can formally solve these equations using Green’s function methods to find the effective equation for \( \psi_1 \) alone. We find

\[
[T_1 + U_1 + V_{12} \hat{G}_2^+ V_{21} - E_1] \psi_1(R) = 0 .
\]

(15.4.2)

Here we use the distorted-wave Green’s function for channel 2, \( \hat{G}_2^+ = [E_2 - T_1 - U_1]^{-1} \), with the notation of Eq. (3.3.10). This implies that the scattering in channel 1 is not governed purely by \( U_1 \), but by an additional effective potential \( V_{12} \hat{G}_2^+ V_{21} \) that depends on energy \( E_2 \) and is non-local. This additional part is called a dynamic polarisation potential (DPP):

\[
V_{\text{DPP}} = V_{12} \hat{G}_2^+ V_{21} ,
\]

(15.4.3)

and its effect is to change the elastic scattering from that caused by \( U_1 \) alone. The dynamic polarisation potential will have an imaginary part that reflects the loss of flux from channel 1 into channel 2, and will also have a real part that may modify potential barriers and hence any low-energy tunnelling probabilities.

This means that we must distinguish from optical potentials, which fit some experimental data, from the bare potentials \( U_1, U_2 \) that are the diagonal potentials to be used in a coupled set of equations such as Eq. (15.4.1). Because of the dynamic polarisation effects arising from the couplings, the bare potentials will be different from any optical potentials. The bare potentials will typically have weaker imaginary parts than an optical potential, since the imaginary part of a bare potential is used to describe the loss of flux out of the model space, which is less than the flux leaving the elastic channel since the channel 2 is still part of the coupled-channels model space.

Different strategies may be used to deal with this difference, in order to solve coupled channels equations while fitting elastic data. One is to initially ignore the DPP, start with a bare potential equal to an optical one, and then later make small corrections to the parameters of the bare potential to improve the fit. Typically, the strength of its imaginary part will be reduced by small fractions in order to restore the elastic fit. A second strategy is redo the entire fitting of the optical potential, now including all of the equations (15.4.1) within the search process. This is probably the most accurate strategy overall. A third strategy, however, may still prove useful: this is to artificially set to zero the reverse coupling \( V_{12} \) that feeds back to the elastic channel. This means that the elastic channel still reproduces
15.4 Multi-channel fitting

the result of the optical potential, so this strategy is another form of the Distorted Wave Born approximation introduced on page 95. This step does not require the couplings to be small, as we only argue that the DPP is already included within the optical potential, since it fits the experimental scattering, and therefore the elastic fit should not be disturbed by couplings since this would count the coupling effects twice.

15.4.2 Fitting inelastic scattering

To fit the inelastic scattering of a nucleus to states in its rotational band by the collective mechanisms discussed in section 4.4.1, we normally proceed in two steps. First we use a simple first-order rotational model with quadrupole couplings to construct transition potentials between states of the rotational band, the only unknown parameter being the deformation $\beta_2$ (or deformation length $\delta_2$, or $B(E2)$) which linearly scales the potential. This transition potential will have both nuclear and Coulomb contributions, but these scale in the same way. Then, comparison with the overall magnitudes of the observed inelastic cross sections will give a good estimate of the deformation. Then afterwards, the fit can be improved by detailed adjustments of the optical potential, inclusion of the deformations non-linearly as in Eq. (4.4.25), possible $\beta_4$ effects, along with the coupled-channels effects from multiple couplings. If necessary, the Coulomb and nuclear deformations may be made different: the Coulomb matrix elements will particularly affect the forward-angle cross sections in a way that strongly depends also on the excitation energies of the levels in the rotational band.

15.4.3 Transfer fits

Transfer reactions are individually rather weak processes, and so most commonly are treated in first order in some kind of Born approximation. As mentioned above, that approximation is equivalent to neglecting some set of reverse couplings.

The simplest transfer model, therefore, is the one-step DWBA, in which optical potentials are used for both the entrance and exit channels and the transfer coupling is taken to first order, to couple just the entrance to the exit channel (and not also the reverse). This then becomes exactly the DWBA considered on page 95, where the entrance optical potential reproduces some elastic data of these nuclei at the needed or a nearby energy. In first-order DWBA, the cross section is exactly proportional to the spectroscopic factor $S$ of Eq. (5.3.11). We might therefore attempt to find the empirical spec-
The second-simplest model is one in which other inelastic couplings are permitted to act to higher or all orders, but the transfer coupling still only to first order. This scheme is called the CCBA, or coupled channels Born approximation. One of the participating nuclei, for example, might have a rotational band that can be easily excited. Note first that if all the couplings were only from and to the ground state, then these rotational effects would have already been included in the entrance optical potential. A more detailed treatment is necessary, therefore, if there are transfers to or from excited states of the band. In this case, at least the CCBA scheme is essential, as multiple routes to a transfer state are possible. There could be transfers from ground-state to ground-state, or first an excitation and then a transfer, or an excitation followed by transfer followed by a de-excitation. These higher-order routes with increasing numbers of steps become individually less significant, but may be essential if for some reason the direct route is suppressed on structural grounds, or if there are many such intermediate steps energetically allowed. One complication to note when multiple routes may contribute, is that then the relative phases become important, as the amplitudes for the different processes may interfere constructively or destructively.

The detailed shape of a transfer cross section is affected by these higher-order effects, and also by the features of the bound state wave function that comes, for example, from the geometry of its binding potential. Transfers at sub-barrier energies are expected to probe only the remote tail of the bound state wave function, where its amplitude is governed mostly by the asymptotic normalisation coefficient (ANC) defined in Eq. (4.5.24). We may therefore expect low-energy reactions to determine these ANC values to good accuracy. At higher energies, above the Coulomb barrier, there will be more penetration to the surface and perhaps into the interior, so more global properties of the transfer wave function govern the cross section. It may well be that a selection of experiments at different energies is needed to probe the entire wave function.

15.4.4 A Progressive Improvement Policy
We may summarise the strategies for the detailed fitting of experimental data as arising from a ‘progressive improvement policy’. Fitting data, by this policy, should start with the simplest data and the simplest reaction
model. In most cases this should be elastic data and some optical potential, preferably over a range of energies so that local peculiarities will not upset the modelling. Then for inelastic excitation, transfer or breakup, one-step DWBA should be used for the specific final states that can be reached from the initial bound state. Only after that has been done and the results studied, should more complicated models be employed, such as those with inelastic excitations in the entrance or exit channels, or those with two-step or coupled-channels transfers.

Only in special cases should coupled-channels be started at the beginning, such as those where the entrance optical potential is unknown and has to be calculated dynamically including the effects of many inelastic or breakup couplings.

The progressive improvement policy is also particularly useful when performing $\chi^2$ searches, as we see next.

### 15.5 Searching

The tasking of searching is to take a defined set of theoretical parameters which may be varied (perhaps with experimental absolute normalisation factors), and numerically vary these in a systematic way while monitoring the total $\chi^2$ value, until $\chi^2$ becomes as small as possible. For linear variations this can be achieved by solving simultaneous equations, but for non-linear problems as here, iterative searches are necessary. The most widely used tool to achieve this is the program Minuit [2]. Within this suite the most efficient minimizing technique is called Migrad, and this is a gradient search method using Fletcher’s switching variation [3] to the Davidson-Fletcher-Powell variable metric algorithm [4]. It approaches a local minimum closely and generates parabolic error estimates, which will be true errors when the $\chi^2$ function around the minimum is accurately quadratic with respect to each parameter according to Eq. (15.1.19).

#### 15.5.1 Strategies

It is useful to remember the following strategies for $\chi^2$ searching:

(i) The process can be restarted from any intermediate stage, if subsequently either strange or no results were found. Intermediate sets of parameter values should therefore be saved during the search.

(ii) Ambiguities in searching can be dealt with by means of grid searches. That is, if for example there is a mutual ambiguity between $R_{ws}$ and
Fitting data

$V_{ws}$, then searches should be conducted for finite increments in $R_{ws}$. In this case, the potentials for each $R_{ws}$ value will, when plotted, often cross at a ‘sensitive region’ for the scattering. See the examples for $^{16}$O [5] and $^6$He scattering [6] scattering on lead around the Coulomb barrier.

Without grid searches, searching ambiguities will give highly correlated errors between the related variables. These correlations will be discussed below.

(iii) Discrete ambiguities, such as those which change the number of nodes in the wave functions for low-energy scattering, require specific restarting of a gradient search with increased or decreased potentials. The required potential depths may be estimated by finding what is needed to change the number of nodes in a one-channel bound state.

(iv) If some features of the data are never fitted properly, then the error bars for these data points can be temporarily reduced by a factor of say 10, to increase their importance in the $\chi^2$ sum. You can then learn what in the model is sufficient to fit those features, and perhaps reconsider the model to make this easier when the error bars are restored to their realistic values.

(v) If a parameter is given a limited range and then ends up near one end of this extent, then the MIGRAD procedure often gives a spurious convergence with respect to its variation because of the way it internally maps the limited range. Apparently converged results with one variable near its limit should always be repeated with a wider allowed range.

(vi) Results from other experiments may already give values and error bars for some of the parameters currently being searched for. For example, lifetime measurements may give deformations, and other transfer reactions may have given spectroscopic factors. As long as the $1\sigma$ errors for those experiments are known, they may be used to constrain your present search by including an extra contribution to the $\chi^2$ sum of the form shown in Eq. (15.1.2).

(vii) Sometimes, instead of searching for two correlated variables, they should be combined into one variable. For example, in one-step DWBA the two projectile and target spectroscopic factors must be combined (or else one of them fixed). Similarly, if both a bound state and scattering data are to be fitted with the same potential, it may be possible to find the bound state first and reset the potential depth to that necessary for this exact eigenstate.
During the search process also, we may apply the Progressive Improvement Policy outlined above. The initial stages of a $\chi^2$ search should use the simplest reaction theory that gives plausible results for the measurements of interest. There is no harm in making approximations to the scattering theory at early stages (such as zero-range rather than finite-range transfers, or omitting Coulomb excitations at large radii, or using the DWBA rather than finding coupled channels solutions), and a lot of time can be saved by making these approximations initially. Then, after an apparent convergence, more details of the reaction theory can be inserted, and/or the existing theory calculated more accurately. After these model improvements have been made, the search procedure can be restarted for progressively finer readjustments.

15.5.2 Error estimates from fitting

Using the $\chi^2$ expression of Eq. (15.1.1), the Eq. (15.1.26) tell us that the simplest estimate of the error in each the fitted parameters is that variation which increases $\chi^2$ from its converged minimum by unity (so $\chi^2/N$ is increased by $1/N$ for $N$ data points). These parameter errors give a very useful estimate how the parameters are fixed by the given data. Note that if the errors $\Delta\sigma(i)$ are all overestimated by a factor $A$, then the resulting parameter errors from the fit will be overestimated by the same factor $A$.

To examine correlations between the fitted parameters, we need to look at the error matrix: the parameter covariance matrix $V_p$, such as that calculated by Eq. (15.1.24). This is twice the inverse of the Hesse matrix $H_{mn}$ of partial second derivatives of $\chi^2$ defined in Eq. (15.1.21). The errors based on the error matrix take account of all the parameter correlations, but not their non-linearities (the $2\sigma$ error will be exactly twice the $1\sigma$ error estimate). When the covariance error matrix has been calculated, the parameter errors are the square roots of the diagonal elements of this matrix, which include the effects of correlations with the other parameters.

These correlation effects are included because the inverse of the error matrix, the second derivative Hesse matrix $H_{mn}$ of Eq. (15.1.21), has as diagonal elements the second partial derivatives $\partial^2(\chi^2)/\partial p_m^2$ with respect to one parameter at a time. These diagonal elements are not therefore coupled to any other parameters, but when the matrix is inverted, the diagonal elements of the inverse contain contributions from all the elements of the second derivative matrix, and thus includes the correlation effects on the individual errors.

It is also instructive to look at the matrix of parameter correlation coeffi-
Fitting data

\[ \rho_{mn}^p = \frac{V_{mn}^p}{\sqrt{V_{mm}^p V_{nn}^p}} \]  \hspace{1cm} (15.5.1)

found from Eq. (15.1.8). Correlation coefficients very close to one (greater than 0.99) indicate a difficult or false fit, and occur especially if there are severe ambiguities between two or more search parameters. These would occur in an ill-posed problem with more free parameters than can be determined by the model and the data.

References