

Evaluation and Analysis of Nuclear Resonance Data

JEFF Report 18



FOREWORD

Nuclear data are fundamental to the development and application of all nuclear sciences and technologies. Basic nuclear data, whether measured or calculated, follow a complex process of evaluation, correction and analysis before becoming directly available in applications. This report describes such a process in the case of neutron-induced reactions in the resonance range.

As there are no predictive theories for neutron-induced reactions in the resonance energy range, the basic nuclear data have to be obtained through measurements at dedicated experimental facilities, such as linear particle accelerators. The measured raw data are then corrected for the experimental conditions, such as sample impurities, background effects and detector efficiencies. However, the experimental data thus obtained are not directly suitable for application calculations. A thorough analysis of the data is necessary to produce a coherent set of applicable data.

F. Fröhner, The author of this report, describes in detail two elements necessary to perform a correct analysis of experimental data in the resonance energy range: the neutron-nucleus interaction theory in this energy range and the mathematical formalism of statistical inference. Concerning this last point, the author expresses his preference for the Bayesian approach, which he considers the most appropriate.

This report is part of an ongoing effort co-ordinated at an international level, involving national research institutions and industry. It aims at preserving nuclear data knowledge, in a field from which a large number of specialists have recently retired. The French Atomic Energy Commission at Cadarache, Électricité de France and the OECD Nuclear Energy Agency have contributed to this report. Vincent Greissier translated the report into French and Pierre Ribon reviewed the original English version and the French translation. Laurent Carraro provided comments on the mathematical part of the report.

Evaluation and Analysis of Nuclear Resonance Data

F.H. Fröhner
Forschungszentrum Karlsruhe
Institut für Neutronenphysik und Reaktortechnik
D-76021 Karlsruhe, Germany

TABLE OF CONTENTS

| | |
|--|----|
| ABSTRACT | 3 |
| 1. THE PROBABILISTIC BASIS OF DATA EVALUATION | 3 |
| 1.1. Probability, a Quantitative Measure of Rational Expectation | 4 |
| 1.2. Bayes' Theorem, the Rule for Updating Knowledge with New Data | 5 |
| 1.3. Recommended Values from Estimation Under Quadratic Loss | 7 |
| 1.4. Generalisation to More Observations and More Parameters | 9 |
| 1.5. Closer Look at Prior Probabilities, Group-Theoretic Assignment | 9 |
| 1.6. Bayesian Parameter Estimation for the Univariate Gaussian | 11 |
| 1.7. Assignment of Probabilities by Entropy Maximisation | 15 |
| 1.8. Approximation: Maximum Likelihood | 19 |
| 1.9. Approximation: Least Squares | 19 |
| 2. EVALUATION OF NUCLEAR DATA FOR APPLICATIONS | 23 |
| 2.1. Stepwise Preparation of Nuclear Data for Applications | 23 |
| 2.2. Iterative Least-Squares Fitting | 28 |
| 2.3. Statistical Errors: Poisson Statistics | 31 |
| 2.4. Systematic Errors: Correlated Uncertainties and their Propagation | 32 |
| 2.5. Goodness of Fit | 35 |
| 2.6. Inconsistent Data | 37 |
| 2.7. Estimation of Unknown Systematic Errors | 40 |
| 3. RESONANCE THEORY FOR THE RESOLVED REGION | 44 |
| 3.1. The Blatt-Biedenharn Formalism | 48 |
| 3.2. The Exact R-Matrix Expressions | 51 |
| 3.3. The Practically Important Approximations | 54 |
| 3.3.1. Kapur-Peierls Cross Section Expressions | 56 |
| 3.3.2. SLBW Cross Section Expressions | 56 |
| 3.3.3. MLBW Cross Section Expressions | 58 |
| 3.3.4. Reich-Moore Cross Section Expressions | 59 |
| 3.3.5. Adler-Adler Cross Section Expressions | 60 |
| 3.3.6. Conversion of Wigner-Eisenbud to Kapur-Peierls Parameters | 61 |
| 3.4. External Levels | 63 |
| 3.4.1. Statistical Representation of the External Levels | 63 |
| 3.4.2. Representation of the Edge Terms by Two Broad Resonances | 65 |
| 3.4.3. Narrow Bound Level to Enforce Prescribed Thermal Cross Sections | 67 |

| | |
|---|-----|
| 3.5. Doppler Broadening | 70 |
| 3.5.1. Free-Gas Approximation | 70 |
| 3.5.2. Cubic Crystal | 72 |
| 3.5.3. Gaussian Broadening with Voigt Profiles | 72 |
| 3.5.4. Gaussian Broadening with Turing's Method | 73 |
| 3.5.5. Broadening of Tabulated, Linearly Interpolable Point Data | 74 |
| 3.6. Practical Analysis of Resonance Cross Section Data | 75 |
| 3.6.1. Observables | 76 |
| 3.6.2. Experimental Complications | 78 |
| 3.6.3. Spin and Parity Assignment | 80 |
| | |
| 4. STATISTICAL RESONANCE THEORY FOR THE UNRESOLVED REGION | 82 |
| | |
| 4.1. Level Statistics | 82 |
| 4.1.1. The Porter-Thomas Hypothesis | 82 |
| 4.1.2. Wigner's Surmise and the Gaussian Orthogonal Ensemble | 84 |
| 4.1.3. Transmission Coefficients | 86 |
| 4.1.4. Nuclear Level Densities | 88 |
| 4.1.4. Information from Resolved Resonances | 95 |
| 4.2. Resonance-Averaged Cross Sections | 97 |
| 4.2.1. Average Total Cross Sections | 98 |
| 4.2.2. Average Partial Cross Sections: Heuristic Recipes | 99 |
| 4.2.3. Average Partial Cross Sections: The Exact GOE Average | 100 |
| 4.2.4. Analysis of Resonance-Averaged Data | 102 |
| 4.3. Group Constants | 106 |
| 4.3.1. Bondarenko Factors | 106 |
| 4.3.2. Analytic and Monte Carlo Methods for Group Constant Generation | 108 |
| | |
| 5. CONCLUDING REMARKS | 110 |
| | |
| ACKNOWLEDGEMENTS | 110 |
| | |
| APPENDICES | 111 |
| | |
| Appendix A: Practically Important Probability Distributions | 111 |
| A.1. Binomial and Beta Distributions | 111 |
| A.2. Poisson and Gamma Distributions | 112 |
| A.3. Univariate Gaussian | 113 |
| A.4. Multivariate Gaussian | 115 |
| Appendix B: Mathematical Properties of the Voigt Profiles ψ and χ | 118 |
| | |
| REFERENCES | 120 |

Evaluation and Analysis of Nuclear Resonance Data

F.H. Fröhner
 Forschungszentrum Karlsruhe
 Institut für Neutronenphysik und Reaktortechnik
 Postfach 3640, D-76021 Karlsruhe
 Germany
 (e-mail: fritz.froehner@inr.fzk.de).

ABSTRACT. The probabilistic foundations of data evaluation are reviewed, with special emphasis on parameter estimation based on Bayes' theorem and a quadratic loss function, and on modern methods for the assignment of prior probabilities. The data reduction process leading from raw experimental data to evaluated computer files of nuclear reaction cross sections is outlined, with a discussion of systematic and statistical errors and their propagation and of the generalised least-squares formalism including prior information and nonlinear theoretical models. It is explained how common errors induce correlations between data, what consequences they have for uncertainty propagation and sensitivity studies, and how evaluators can construct covariance matrices from the usual error information provided by experimentalists. New techniques for evaluation of inconsistent data are also presented. The general principles are then applied specifically to the analysis and evaluation of neutron resonance data in terms of theoretical models - R-matrix theory (and especially its practically used multi-level Breit-Wigner and Reich-Moore variants) in the resolved region, and resonance-averaged R-matrix theory (Hauser-Feshbach theory with width-fluctuation corrections) in the unresolved region. Complications arise because the measured transmission data, capture and fission yields, self-indication ratios and other observables are not yet the wanted cross sections. These are obtained only by means of parametrisation. The intervening effects - Doppler and resolution broadening, self-shielding, multiple scattering, backgrounds, sample impurities, energy-dependent detector efficiencies, inaccurate reference data etc. - are therefore also discussed.

1. The Probabilistic Basis of Data Evaluation

Historically data evaluation in the modern sense began with the effort of Dunnington (1939), DuMond and Cohen (1953) and collaborators to determine a set of recommended values of the fundamental physical constants (speed of light, Planck's quantum of action, fine-structure constant, etc.), and to establish their uncertainties, by a comprehensive least-squares fit to all relevant experimental data. As measurements are invariably affected by uncontrollable instrumental errors, inaccurate standards, finite counting statistics and other sources of uncertainty, data evaluation involves reasoning from incomplete information, hence probability theory. We begin therefore with a brief review of the probability-theoretical foundations of data evaluation. This will help tie together various rules for the extraction of "best" values and uncertainties from experimental data, and prescriptions for data fitting and adjustment. Most scientists learn these rules and recipes during laboratory courses and on the job, finding most books on probability theory full of intimidating statistical terminology and awkward "ad-hockeries" (Good 1965) arising from misconceived attempts to avoid Bayes' theorem with its much-maligned a-priori probabilities. The following exposition which (a) is firmly based on Bayes' theorem and (b) utilises recent progress concerning prior probabilities, will be found to lead to a concise and mathematically simple treatment of parameter estimation and data adjustment in the

general framework of inductive inference, or learning from real, always error-affected and incomplete observations.

1.1. PROBABILITY, A QUANTITATIVE MEASURE OF RATIONAL EXPECTATION

All our results in this section will be fairly direct consequences of the basic sum and product rules of probability theory,

$$P(A|C) + P(\bar{A}|C) = 1, \quad (1)$$

$$P(AB|C) = P(A|BC)P(B|C) = P(B|AC)P(A|C), \quad (2)$$

where

| | | |
|-----------|---|---|
| A, B, C | = | propositions such as “the coin shows head” or “the cross section is larger than 12 b”, |
| AB | = | both A and B are true , |
| \bar{A} | = | A is false , |
| $P(A C)$ | = | probability of A given C . |

Our notation indicates that all probability assignments are conditional, based on either empirical or theoretical information or on assumptions. Following J. Bernoulli (1713) and Laplace (1812) we interpret these probabilities as degrees of plausibility or rational expectation on a numerical scale ranging from 0 (impossibility) to 1 (certainty), intermediate values indicating intermediate degrees of plausibility. The sum rule tells us that, under all circumstances C , the more probable A is the less probable is \bar{A} , the unit sum of both probabilities representing the certainty that one of these alternatives must be true. The product rule says that, under all circumstances C , the probability that both A and B are true is equal to the probability of A given B , times the probability that, in fact, B is true. Since A and B enter symmetrically one can also take the probability of B given A and multiply it by the probability of A .

The interpretation of the P as degrees of plausibility (not the equations among them) has been criticised by statisticians who insist that by probability one must mean only “relative frequency in a random experiment” such as coin tossing, in the limit of very many repetitions, and that one can assign “direct” probabilities of effects (observations) if causes (stochastic laws and their parameters) are given, but not the “inverse” probabilities of various possible causes if observations are given. They argued that since physical constants are not random variables that assume given values with certain frequencies one ought to associate probabilities only with the observational errors, not with the physical constants. For scientists in general, and data evaluators in particular, this viewpoint is too narrow. It would not permit them to say that, according to measured data, a physical constant has such and such a probability to lie between given limits. The task to infer the values of natural constants, half-lives, reaction cross sections etc. from error-affected, uncertain and incomplete data is not a random experiment that can be repeated at will, but rather an exercise in inductive inference (reasoning in the face of uncertainty). Laplace’s probability concept seems therefore more appropriate for the evaluation of scientific data.

All doubts were dispelled by R.T. Cox (1946). Using the arithmetic of logic, Boolean algebra, he proved that any formal system of logical inference using degrees of plausibility must either be equivalent to probability theory as derived from the basic sum and product rules, or violate elementary consistency conditions. In his proof the most general consistency conditions are cast in the form of two functional equations whose solutions are just the basic rules. Criticism that Cox had assumed differentiability of his probability

functions was met by A. Rényi (1954) who gave a proof without that assumption. It is interesting that Schrödinger (1947), one of the founders of quantum mechanics, arrived independently at essentially the same conclusions as Cox: The basic rules, clearly valid for relative frequencies, are equally valid for Laplacean probabilities. In quantum theory it has always been understood that probabilities quantify incomplete knowledge but it is widely believed that classical and quantum mechanical probabilities differ somehow. Actually, it can be shown that the quantum mechanical probability formalism is perfectly consistent with the Laplacean probability concept and the basic sum and product rules (Fröhner 1998). After Cox's proof two things should be clear.

1. *Probabilities are not relative frequencies. They can equally well be applied to non-repetitive situations as to repeated trials.*
2. *Allegedly superior schemes of logical inference, such as Fuzzy Logic or Artificial Intelligence, are equivalent to probability theory at best – if not, they are bound to violate elementary consistency requirements.*

Although probabilities are not the same as frequencies, the two are surely related. In repetitive situations one finds that the probabilities are essentially expectation values of relative frequencies – see e. g. Jaynes (1968) and Fröhner (1997).

1.2. BAYES' THEOREM, THE RULE FOR UPDATING KNOWLEDGE WITH NEW DATA

Scientific experiments are usually describable by a statistical model, statistical elements being introduced by uncontrollable, seemingly random instrumental effects, by unknown errors and often by the theory itself (e. g. statistical mechanics or quantum theory or the level-statistical theory of compound-nuclear reactions). The statistical model enables one to calculate the “direct” probability of some set of observed data (“sample”), provided the physical quantities and statistical parameters of the model are given. In empirical science the situation is usually reversed: A sample of experimental data is given, and one wants to find the “inverse” probabilities for the various possible values of the physical quantities and statistical parameters of the model. Direct probabilities (of effects given the causes) and inverse probabilities (of causes given the effects) are related by Bayes' (1763) theorem. In its simplest form,

$$P(A|BC) = \frac{P(B|AC)P(A|C)}{P(B|C)}, \quad (3)$$

it is an immediate consequence of the symmetry of the product rule (2) with respect to A and B . The typical situation is that we have data B which depend on the value of an unknown physical quantity A and on other circumstances C . If we have a statistical model, represented by the so-called likelihood function $p(B|AC)$, telling us how likely observation of the data B would be under the circumstances C if the unknown quantity were in fact A , and if we have also an a priori probability (“prior” for short) $P(A|C)$, then the updated or a posteriori probability (“posterior”) $P(A|BC)$ is proportional to the product $P(B|AC)P(A|C)$. The prior summarises what we knew about A before the data became available, the likelihood function conveys the impact of the data, and the posterior contains the complete information available for further inference and prediction. Laplace (1812) gave the generalisation to several distinct, mutually exclusive alternatives A_j :

$$P(A_j|BC) = \frac{P(B|A_jC)P(A_j|C)}{\sum_j P(B|A_jC)P(A_j|C)}, \quad j = 1, 2, \dots, n, \quad (4)$$

normalised to unity as demanded by the sum rule. For continuous alternatives A and B we replace the finite discrete probabilities $P(A|C)$, $P(B|AC)$ etc. by infinitesimal probabilities $p(A|C)dA$, $p(B|AC)dB$ etc. with probability densities $p(A|C)$, $p(B|AC)$ etc., and the sum over alternatives by an integral,

$$p(A|BC) dA = \frac{p(B|AC) p(A|C) dA}{\int p(B|AC) p(A|C) dA}, \quad A_{\min} \leq A \leq A_{\max}. \quad (5)$$

These forms of Bayes' theorem can be considered as the cornerstone of data evaluation and adjustment. They show how prior knowledge (an existing data file) is to be updated with new evidence (new data). In all forms, the denominator is just a normalisation constant, so the formal rule for learning from observations can be stated briefly as

$$\text{posterior} \propto \text{likelihood} \times \text{prior}.$$

It should be understood that the expressions prior and posterior have a logical rather than temporal meaning. They simply mean without and with the new data taken into account.

As a fairly realistic illustration let us consider the determination of the decay constant λ of some short-lived radioisotope from decays registered at times t_1, t_2, \dots, t_n . Obviously we must identify λ with A , and the data t_1, \dots, t_n with B , while C consists of all other information about the situation such as applicability of the exponential decay law, purity of the sample, reliability of the recording apparatus, sufficiently long observation time for all observable decays to be recorded, etc. The statistical model for the experiment is represented by the so-called sampling distribution, i. e. by the probability with which we may "reasonably expect" the various alternatives if we sample once, given the parameters of the model. In our example this is the probability that, given λ , one particular decay, let us say the i -th one, is recorded in the time interval dt_i at t_i ,

$$p(t_i|\lambda) dt_i = \exp(-\lambda t_i) \lambda dt_i, \quad 0 < t_i < \infty. \quad (6)$$

We shall normally write continuous probability distributions in this form, with the probability density p multiplied by the relevant differential, i. e. as an infinitesimal probability, and with the range of possible values explicitly stated. This emphasises the fact that ultimately all probability distributions are used for the calculation of expectation values, so that they are parts of integrands, subject to possible change of variables. Since we generally use the letter p for probability densities irrespective of their functional form, a change of variable results in $p(x|\cdot)dx = p(x|\cdot)|dx/dy|dy \equiv p(y|\cdot)dy$. (We simplified the notation here by omitting explicit reference to the conditioning background information C).

According to the product rule the joint probability of observing the mutually independent data t_1, \dots, t_n , given λ , is

$$p(t_1, \dots, t_n|\lambda) dt_1 \dots dt_n = \exp\left(-\lambda \sum_{j=1}^n t_j\right) \lambda^n dt_1 \dots dt_n. \quad (7)$$

This corresponds to $p(A|B)dB$ above. Multiplying the likelihood function $p(t_1, \dots, t_n|\lambda)$ by the prior $p(\lambda)d\lambda$ we get

$$p(\lambda|t_1, \dots, t_n) d\lambda \propto \exp\left(-\lambda \sum_{i=1}^n t_i\right) \lambda^n p(\lambda) d\lambda, \quad 0 < \lambda < \infty. \quad (8)$$

We note that in our problem the likelihood function does not depend on all the individual sample values. They appear only in the form $\sum_i t_i \equiv n\bar{t}$, so that for given sample size n the sample average \bar{t} carries all the information contained in the data. In statistical jargon \bar{t} is a “sufficient statistic”, n an “ancillary statistic”, where statistic means any function of the sample, i. e. of the data.

If we consider all values of λ between 0 and ∞ as equally probable a priori, so that $p(\lambda)d\lambda \propto d\lambda$, we get

$$p(\lambda|n\bar{t})d\lambda \propto e^{-\lambda n\bar{t}}\lambda^n d\lambda, \quad 0 < \lambda < \infty. \quad (9)$$

Now the gamma function is defined by

$$\Gamma(n+1) \equiv \int_0^\infty e^{-x} x^n dx \quad (10)$$

(which for non-negative integers is just the factorial $n!$). It follows that the final result of our Bayesian estimation, properly normalised, can be written as

$$p(\lambda|n\bar{t})d\lambda = \Gamma(n+1)^{-1} e^{-x} x^n dx, \quad 0 < x \equiv \lambda n\bar{t} < \infty. \quad (11)$$

This posterior, a gamma distribution (also known as chi-square distribution with $\chi^2 \equiv 2x$ and $\nu \equiv 2n+2$ degrees of freedom) represents the complete information about λ which is contained in the data and the assumed prior. Fig. 1 shows chi-square distributions for various values of ν . As the sample size n increases, our posterior becomes more and more concentrated: The more data are collected the better known is λ .

1.3. RECOMMENDED VALUES FROM ESTIMATION UNDER QUADRATIC LOSS

Most users of radioactive-decay data do not want to be bothered with the details of a posterior distribution. What they usually want is a recommended decay constant and its uncertainty, and nothing else. So we calculate, with Eq. 11, the expectation value,

$$\langle \lambda \rangle = \int_0^\infty \lambda p(\lambda|n\bar{t}) d\lambda = \frac{n+1}{n\bar{t}}, \quad (12)$$

and the root-mean-square error (also called standard deviation or dispersion or one-sigma uncertainty),

$$\Delta\lambda = \left[\int_0^\infty (\lambda - \langle \lambda \rangle)^2 p(\lambda|n\bar{t}) d\lambda \right]^{1/2} = \sqrt{\frac{n+1}{n\bar{t}}}, \quad (13)$$

and state our result summarily as $\langle \lambda \rangle \pm \Delta\lambda$. This choice is justified as follows. The expectation value is that estimate λ_0 which minimises the expected squared error,

$$\int_0^\infty (\lambda_0 - \lambda)^2 p(\lambda|n\bar{t}) d\lambda = \min, \quad (14)$$

as is readily verified by differentiation with respect to λ_0 and equating to zero. With this recommended value the mean squared error (expectation value) is just the variance, $\text{var } \lambda \equiv (\Delta\lambda)^2$, which justifies also our uncertainty specification. What we just did is called “estimation under quadratic loss” in decision theory. The basic idea is that there is usually a penalty for bad estimates, the more severe the more the estimate differs from

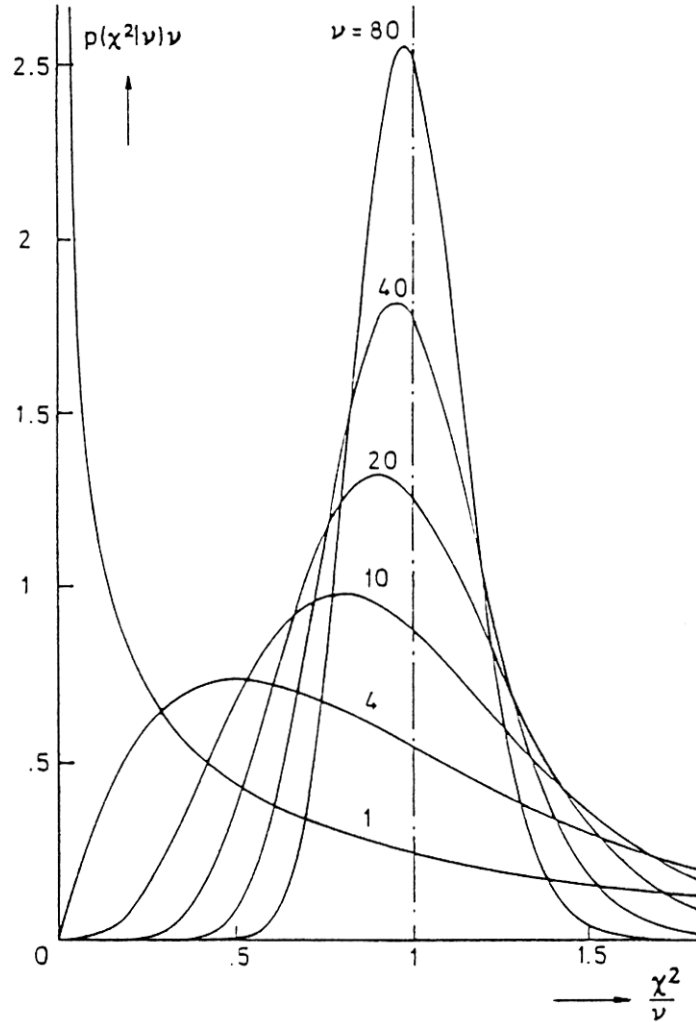


Fig. 1. Chi-square distributions for various degrees of freedom ν (cf. Eq. 105 and Korn & Korn 1968 for standard definition and main properties).

the true value, and that this penalty can be described by a “loss function” which vanishes at the true value and is positive everywhere else. In the vicinity of the true value a reasonably smooth loss function can be taken as quadratic in the error $(\lambda_0 - \lambda)$, since its Taylor expansion about zero error begins with the quadratic term (cf. e. g. DeGroot 1970, Berger 1985). Eq. 14 is therefore the condition for minimal expected penalty in this parabolic approximation.

Obviously the recommendation $\langle \lambda \rangle \pm \Delta \lambda$ tends to obscure the asymmetry of the gamma distribution which especially for small values of n is quite strong (see Fig. 1). If such details are important, one must return to the full posterior distribution. So far

our result looks reasonable enough, but as we shall see, there is a problem caused by the rather cavalier fashion in which we have assigned the prior probabilities.

1.4. GENERALISATION TO MORE OBSERVATIONS AND MORE PARAMETERS

Before we deal more carefully with priors, let us see what impact a second measurement (with a fresh radioactive sample) would have on our knowledge of the decay constant. Using the posterior distribution of the first measurement as the prior for the second one, we find as the new posterior distribution

$$p(\lambda|t_1, \dots, t_n, t'_1, \dots, t'_m) d\lambda \propto p(t'_1, \dots, t'_m|\lambda) p(t_1, \dots, t_n|\lambda) d\lambda, \quad (15)$$

where t'_1, \dots, t'_m are the new data. More generally, if there are k measurements, with associated data sets D_1, \dots, D_k and likelihood functions L_1, \dots, L_k , one gets

$$p(\lambda|D_1, \dots, D_k) d\lambda \propto \left[\prod_{j=1}^k L_j(D_j|\lambda) \right] p(\lambda) d\lambda, \quad (16)$$

which shows nicely how Bayes' theorem models the process of learning by experience: Each new experimental result can be formally incorporated into the existing body of knowledge by multiplication of the associated likelihood function into the existing probability distribution (and renormalisation). It is by no means necessary that all experiments are of the same type. In nuclear resonance analysis, for instance, one usually combines likelihood functions from transmission, capture, scattering and fission experiments involving all kinds of detector and sample geometries in order to obtain best values of resonance energies and partial widths. With each additional data set the posterior distribution becomes narrower which means the uncertainty of the estimated parameter becomes smaller. Our example shows this explicitly: For large n the relative uncertainty of λ approaches zero as $1/\sqrt{n}$.

A last generalisation concerns the estimated parameters. In data evaluation and adjustment we deal not only with large bodies of data from many different experiments, but also with many, usually correlated, parameters that must be determined simultaneously. Instead of one parameter λ one has then a parameter vector $\boldsymbol{\lambda}$ in the equations, instead of the increment $d\lambda$ one has a volume element $d^N\lambda$ in the parameter space, and the prior and posterior distributions represent joint probabilities for all N parameters (coordinates of the parameter vector), complete with correlations. Again, resonance analysis provides examples. With modern shape analysis codes one can simultaneously estimate the resonance energies and widths of dozens of resonances by fitting appropriate resonance-theoretical expressions to the combined data from several types of resonance measurements, each experiment furnishing hundreds or thousands of data points (see Fig. 4 below).

1.5. CLOSER LOOK AT PRIOR PROBABILITIES, GROUP-THEORETIC ASSIGNMENTS

We must now deal more rigorously with prior probabilities. In our decay rate example we have used the prior $p(\lambda) d\lambda \propto d\lambda$, which in terms of the mean life $\tau \equiv 1/\lambda$ can be rewritten as $p(1/\tau) d\tau/\tau^2 \equiv p(\tau) d\tau \propto d\tau/\tau^2$. Obviously we could have just as well estimated τ instead of λ , and assumed all τ equally probable, i. e. $p(\tau) d\tau \propto d\tau$. This, however, would have resulted in a different posterior distribution. It is true that the posterior depends only weakly on the prior if data are abundant, but from a fundamental viewpoint this is no consolation. There seems to be a basic arbitrariness about priors, particularly for continuous parameters.

For more than a century this seeming arbitrariness has led many statisticians to repudiate Bayesian parameter estimation and to seek alternative methods that circumvent priors. Others, comparing this to an attempt to do arithmetic without zero, defended Bayes' theorem as derivable in a few lines from the basic sum and product rules. So they used "subjective" priors or, as H. Jeffreys (1939), invoked invariance arguments to find priors which avoided ambiguities. An important step forward was taken by A. Wald (1950). He had started out to find better-founded (decision-theoretical) methods of statistical inference, without Bayes' theorem, but finished by proving that the optimal strategies for making decisions (recommending a value, for instance) in the face of uncertainty are just the Bayesian rules.

Even more important was the application of group theory and information theory to the problem of priors by E.T. Jaynes (1968, 1973). He demonstrated for a number of simple but practically important cases that, even if one is completely ignorant about the numerical values of the estimated parameters, the symmetry of the problem determines the prior unambiguously. If a so-called location parameter is to be estimated, for instance the mean μ of a Gaussian, the form of the prior must be invariant under a shift c of location, $p(\mu)d\mu = p(\mu + c)d(\mu + c)$. Otherwise not all locations of the Gaussian would be equiprobable a priori, contrary to the assumption of complete ignorance. The functional equation has the solution

$$p(\mu)d\mu \propto d\mu, \quad -\infty < \mu < \infty, \quad (17)$$

a thoroughly plausible result. Not quite so obvious is the case of a scale parameter such as the standard deviation σ of a Gaussian. If there is no preferred scale, one expects invariance under rescaling, $p(\sigma)d\sigma = p(c\sigma)d(c\sigma)$. The solution of this functional equation is

$$p(\sigma)d\sigma \propto \frac{d\sigma}{\sigma}, \quad 0 < \sigma < \infty, \quad (18)$$

as advocated already by H. Jeffreys (1939). Despite its importance and simplicity Jaynes' (1968) proof seems so little known that we quote it here almost verbatim for the case of a rate constant which multiplies (or scales) all times and time intervals in a problem (as λ in Eq. 6 does):

Suppose that two observers, Mr. X and Mr. X', wish to estimate a rate constant from a number of events. If their watches run at different rates so that their measurements of a given time interval are related by $t = ct'$, their rate or scale parameters will be related by $\lambda' = c\lambda$. They assign prior probabilities $p(\lambda)d\lambda$ and $q(\lambda')d\lambda'$, and if these are to represent the same state of ignorance, p and q must be the same function so that $p(\lambda)d\lambda = p(\lambda')d\lambda'$. From the two equations for λ and λ' one gets the functional equation $p(\lambda) = cp(c\lambda)$. Its unique solution is Jeffreys' prior,

$$p(\lambda)d\lambda \propto \frac{d\lambda}{\lambda}, \quad 0 < \lambda < \infty. \quad (19)$$

Obviously this is the appropriate prior for our decay rate example, since the decay constant is just such a scale parameter in our equations. It satisfies $p(\lambda)d\lambda \propto d\lambda/\lambda \propto d\tau/\tau$ which removes all ambiguity: Whether we estimate the scale parameter λ or the scale parameter τ , we always get the same posterior,

$$p(\lambda|n\bar{t})d\lambda = \frac{e^{-x} x^n}{\Gamma(n)} \frac{dx}{x}, \quad 0 < x \equiv \lambda n\bar{t} < \infty, \quad (20)$$

with

$$\langle \lambda \rangle = \frac{1}{\bar{t}}, \quad (21) \quad \frac{\Delta \lambda}{\langle \lambda \rangle} = \frac{1}{\sqrt{n}}, \quad (22)$$

This looks neater than our previous result, illustrating Ockham's (1349) principle of parsimony ("Ockham's razor"): The simpler result is usually the more correct one. If not the decay constant but the mean life is to be estimated, it is equally easy to find

$$\langle \tau \rangle = \langle \lambda^{-1} \rangle = \frac{n\bar{t}}{n-1}, \quad (23) \quad \frac{\Delta \tau}{\langle \tau \rangle} = \frac{1}{\sqrt{n-2}}, \quad (24)$$

applicable if $n > 2$.

Other examples of priors derived from group-theoretical invariances can be found in the works of Jaynes (1968, 1973, 1976, 1980). In the language of group theory the prior corresponding to invariance under a group of transformations is the right invariant Haar measure on this group (see Berger 1985, ch. 6.6)

The non-normalisability of such "least informative" priors is sometimes criticised, and they are called "improper" priors. Now one can employ instead a broad normalisable prior of convenient ("conjugate") mathematical form. In our example this would be a gamma distribution. The posterior would then, of course, depend on the width of this prior. If one lets the width grow indefinitely one finds always that the posterior tends toward the posterior obtained much more easily with the least informative prior. Our least informative priors can therefore be considered as limits of extremely broad, normalisable distributions on the linear ($d\mu$) and on the logarithmic ($d\sigma/\sigma = d \ln \sigma$) scale, just as Dirac's delta function is the limiting case of extremely narrow, normalised distributions. There are no conceptual or mathematical difficulties if one keeps in mind that both the least informative priors and the "most informative" delta function are, in this sense, nothing but convenient shorthand notations for extremely broad and extremely narrow distributions, meaningful only in convolution with other, less extreme distributions.

1.6. BAYESIAN PARAMETER ESTIMATION FOR THE UNIVARIATE GAUSSIAN

Let us apply least informative priors also to the principally and practically important univariate Gaussian distribution. Suppose a repeated measurement of the same physical quantity μ has produced the results x_1, \dots, x_n , with experimental errors that can be assumed as normally distributed. The sampling distribution is then

$$p(x_j|\mu, \sigma) dx_j = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x_j - \mu}{\sigma} \right)^2 \right] dx_j, \quad -\infty < x_j < \infty, \quad (25)$$

with unknown error dispersion σ . The prior expressing complete ignorance of location (mean) and scale (width) of the Gaussian is (see Jaynes 1968)

$$p(\mu, \sigma) d\mu d\sigma \propto \frac{d\mu d\sigma}{\sigma}, \quad -\infty < \mu < \infty, \quad 0 < \sigma < \infty. \quad (26)$$

The posterior is thus

$$p(\mu, \sigma|x_1, \dots, x_n) d\mu d\sigma \propto \frac{1}{\sigma^n} \exp \left[-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2 \right] \frac{d\mu d\sigma}{\sigma}. \quad (27)$$

In terms of sample mean and sample variance,

$$\bar{x} \equiv \frac{1}{n} \sum_{j=1}^n x_j, \quad (28)$$

$$s'^2 \equiv \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2, \quad (29)$$

the exponent can be written as

$$\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2 = \left[1 + \left(\frac{\mu - \bar{x}}{s'} \right)^2 \right] \frac{ns'^2}{2\sigma^2} \equiv (1 + u^2)v. \quad (30)$$

The posterior joint probability for μ and σ , properly normalised and in simplified notation, can thus be presented in the following two forms that correspond to the two factorisations of the basic product rule (2),

$$\begin{aligned} p(\mu, \sigma | \bar{x}, s', n) d\mu d\sigma &= p(u|v, n) du p(v|n) dv = \frac{e^{-vu^2}}{\sqrt{\pi}} \sqrt{v} du \frac{e^{-v} v^{(n-1)/2}}{\Gamma(\frac{n-1}{2})} \frac{dv}{v} \\ &= p(v|u, n) dv p(u|n) du = \frac{e^{-(1+u^2)v} [(1+u^2)v]^{n/2}}{\Gamma(\frac{n}{2})} \frac{dv}{v} \frac{du}{B(\frac{1}{2}, \frac{n-1}{2}) (1+u^2)^{n/2}}, \\ &\quad -\infty < u \equiv \frac{\mu - \bar{x}}{s'} < \infty, \quad 0 < v \equiv \frac{ns'^2}{2\sigma^2} < \infty, \end{aligned} \quad (31)$$

where $B(\frac{1}{2}, \frac{n-1}{2}) \equiv \Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})/\Gamma(\frac{n}{2})$ is a beta function. Note that u is essentially μ , and v is essentially σ^{-2} , and that in both factorisations the posterior depends on the sample only through the sample mean \bar{x} and the sample variance s'^2 (apart from the sample size n). These quantities are therefore “jointly sufficient statistics” in frequentist terminology. In the first factorisation the probability distribution of μ given σ is Gaussian while that of σ^{-2} is a gamma distribution. In the second one the probability distribution of σ given μ is a gamma distribution while that of μ is a Student t-distribution. The two alternative factorisations display explicitly the two marginal distributions for μ and σ :

If only μ is of interest, whatever σ may be, we integrate over all possible values of the “nuisance parameter” σ (or v) to get the marginal distribution of possible μ values,

$$p(\mu | \bar{x}, s') d\mu = \frac{du}{B(\frac{1}{2}, \frac{n-1}{2}) (1+u^2)^{n/2}}, \quad -\infty < u \equiv \frac{\mu - \bar{x}}{s'} < \infty. \quad (32)$$

This is Student’s t-distribution for $t \equiv u/\sqrt{\nu}$ with $\nu = n - 1$ degrees of freedom, see Fig. 2 and Appendix A. Its mean and variance are $\langle u \rangle = 0$ and $\langle u^2 \rangle = 1/(n - 3)$, hence

$$\langle \mu \rangle = \bar{x}, \quad (33) \quad \text{var } \mu = \frac{s'^2}{n - 3}. \quad (34)$$

We encounter here the (plausible) frequentist recipe to use the sample mean as “estimator” for the population mean. No finite and real standard error can be stated as long as $n \leq 3$. On the other hand the half width is always well defined and can be used to indicate the width of the t-distribution, as is common practice in the case $n = 2$, the Cauchy distribution (known to physicists also as Lorentzian or as symmetric Breit-Wigner profile).

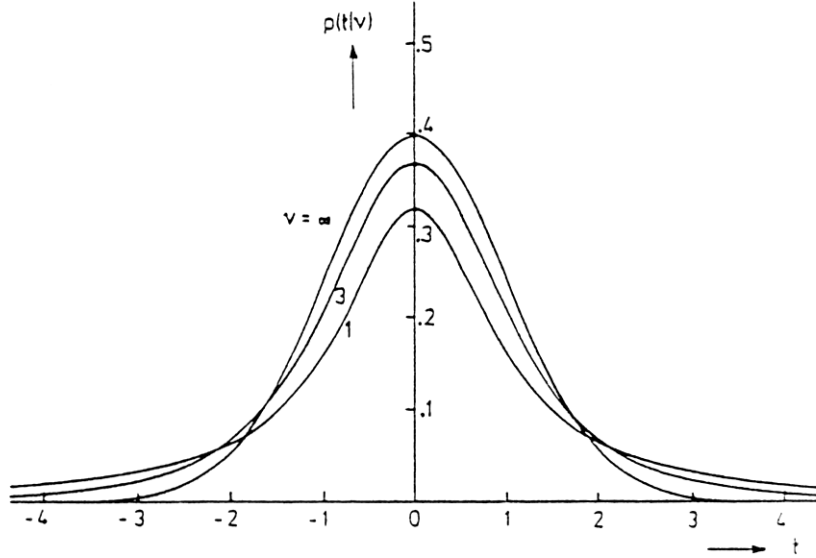


Fig. 2. Student's t -distribution for various degrees of freedom (see Korn & Korn 1986). The Cauchy distribution (Lorentzian, $\nu = 1$) and the Gaussian ($\nu = \infty$) are limiting cases.

If only σ is of interest, regardless of μ , one gets upon integration over u the gamma (or chi-square) distribution

$$p(\sigma|\bar{x}, s') d\sigma = \frac{e^{-v} v^{(n-1)/2}}{\Gamma(\frac{n-1}{2})} \frac{dv}{v}, \quad 0 < v \equiv \frac{ns'^2}{2\sigma^2} < \infty. \quad (35)$$

Its mean and variance are equal, $\langle v \rangle = \text{var } v = (n-1)/2$. The estimate under quadratic loss for σ^{-2} is therefore

$$\langle \sigma^{-2} \rangle = \frac{n-1}{n} s'^{-2} \equiv s^{-2}, \quad (36) \quad \frac{\Delta(\sigma^{-2})}{\langle \sigma^{-2} \rangle} = \sqrt{\frac{2}{n-1}}. \quad (37)$$

This supports to some extent the other (not so plausible) frequentist recipe to take not the sample variance s'^2 but s^2 as estimator for σ^2 , although this estimator is “biased” which means its expectation value, the average over all possible samples, is not equal to the estimated parameter, $\langle s^2 \rangle \neq \sigma^2$. Now $\sigma^{-2}/2$ is also called the precision (cf. e. g. DeGroot 1970) so we recognise that actually $s^{-2}/2$ is an unbiased estimator for the precision. The correct Bayesian estimate of σ^2 follows, however, from $\langle v^{-1} \rangle = 2/(n-3)$ and $\langle v^{-2} \rangle = 4/[(n-3)(n-5)]$. Without any guesswork one gets

$$\langle \sigma^2 \rangle = \frac{n}{n-3} s'^2, \quad (38) \quad \frac{\Delta(\sigma^2)}{\langle \sigma^2 \rangle} = \sqrt{\frac{2}{n-5}}. \quad (39)$$

We note that a sample of less than six values does not contain enough information for a complete estimate of σ^2 whereas σ^{-2} can be estimated already with a sample of two.

The covariance of u and v vanishes because after integration over v the remaining integrand is an odd function of u . As a consequence one has

$$\text{cov}(\mu, \sigma^{-2}) = 0. \quad (40)$$

The case with one datum only, $n = 1$, must be treated differently because $s' = 0$ precludes the definition of u , but this is easy. The posterior is simply

$$p(\mu, \sigma | x_1) d\mu d\sigma \propto \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{\mu - x_1}{\sigma} \right)^2 \right] \frac{d\mu d\sigma}{\sigma}, \quad (41)$$

from which one gets the marginals

$$p(\mu | x_1) d\mu \equiv \frac{d\mu}{|\mu - x_1|}, \quad (42)$$

$$p(\sigma | x_1) d\sigma \propto \frac{d\sigma}{\sigma}. \quad (43)$$

The marginal distribution of μ has a sharp maximum at the observed value but that of σ is seen to be still equal to the least informative prior – which makes sense, because a sample of size $n = 1$ can tell something about the location but nothing whatsoever about the spread of a distribution. This is but one example showing that the Bayesian method is consistent with common sense even in extreme cases, in particular for very small samples, where other methods tend to fail. We mention that the posterior (31) for $n > 1$ was found long before the prior (26) became available, but those who know R.A. Fisher’s (1935) “fiducial” approach will appreciate how much simpler and more straightforward the Bayesian derivation is, and how easily it is extended to the case $n = 1$ (Jeffreys 1939). Furthermore, the derivation can be extended in a straightforward way to the multivariate Gaussian distribution. With appropriate generalisations of scalar relationships to matrix form, from variance (mean square error) to covariance matrix ($\sigma^2 \rightarrow \mathbf{C}$), from univariate differential to multivariate volume element ($dx \rightarrow d(\mathbf{x}) \equiv \prod_{\nu} dx_{\nu}$, $d(\sigma^2) \rightarrow d(\mathbf{C}) \equiv \prod_{\nu \leq \kappa} C_{\nu, \kappa}$) etc., one finds matrix expressions that look very similar to the scalar expressions for the univariate Gaussian – see Appendix A (and Fröhner 1990).

Quite generally Bayesian parameter estimation is logically and mathematically simpler than alternative approaches, and at least as rigorous. Concepts like bias, efficiency, sufficiency, admissibility, James-Stein shrinking (see e. g. Berger 1985), essential for frequentist estimation methods, need not be introduced at all since they appear automatically as more or less incidental features of the posterior distribution and its mean and variance. There is no danger that best estimates are obtained outside the range of allowed values, as happens sometimes with other methods. Although reaction rates or cross sections are inherently positive quantities, measured data may well contain negative data points after background subtraction. It would be wrong to discard the negative points. One can rely on the fact that the prior always guarantees the correct range of the estimated quantities, e. g. $\sigma > 0$, regardless of the range of observable values admitted by the likelihood function, e. g. $-\infty < x_j < +\infty$. The basic simplicity and superiority of the Bayesian approach as compared to other estimation methods has been demonstrated quite forcefully by Jaynes (1976) with a whole series of real-life examples.

The joint posterior distribution of the mean μ and the variance σ^2 is the complete information about the parameters of a Gaussian which can be extracted from the data. From the posterior we can obtain recommended values and their uncertainties for quadratic or other loss functions as we have seen. Often, however, one is not only interested in the parameters of the statistical model but also in predictions of the outcome of further measurements. This, in fact, may be the reason why a statistical model was introduced in the first place. What can we say, after having deduced the posterior (31) for the parameters of the Gaussian model, about the outcome of one further measurement, $x \equiv x_{n+1}$? One

might think of taking the Gaussian with the a posteriori most probable, or with the average values of the parameters, or of averaging the Gaussian over the posterior distribution of its parameters. The last alternative is the correct one. This becomes clear if we write down the joint probability of x , μ and σ for given data \bar{x} and s'^2 , then use the product rule, and finally integrate out the “nuisance parameters” μ and σ , which yields

$$\begin{aligned} p(x|\bar{x}, s') dx &\propto dx \int d\sigma \int d\mu p(x|\mu, \sigma) p(\mu, \sigma|\bar{x}, s') \\ &\propto dx \int_0^\infty dv \int_{-\infty}^\infty du e^{-(u-w)^2 v/n} e^{-v} v^{(n-1)/2}, \end{aligned} \quad (44)$$

where u , v are defined as before and $w \equiv (x - \bar{x})/s'$. Integrating first over the Gaussian (all u), then over the remaining gamma distribution (all v) we get, for $n > 1$, the “predictive” distribution

$$p(x|\bar{x}, s') dx = \frac{dy}{\text{B}(\frac{1}{2}, \frac{n-1}{2}) (1+y^2)^{n/2}}, \quad -\infty < y \equiv \frac{x - \bar{x}}{s' \sqrt{n+1}} < \infty. \quad (45)$$

Although the sampling distribution is Gaussian, the predictive distribution for the outcome of an additional measurement is not a Gaussian but a t-distribution. It is true that the t-distribution approaches a Gaussian for large n , but for finite n it is always broader (see Fig. 2). The best estimate for any function $f(x)$ of the next datum is its expectation value $\langle f \rangle$ with respect to the predictive distribution.

1.7. ASSIGNMENT OF PROBABILITIES BY ENTROPY MAXIMISATION

Jaynes (1968, 1973, 1978, 1980) considered also the case that one is not completely ignorant a priori about numerical values. He showed how probabilities can be assigned in a well defined way if at least vague information is available about average quantities, for instance estimates of expectation values such as first and second moments. The key concept is that of information entropy, introduced by C.E. Shannon (1948) as the unique measure of the indeterminacy or missing information implied by a given probability distribution. The information entropy of a discrete probability distribution p_j with mutually exclusive alternatives j is (up to a constant)

$$S = - \sum_j p_j \ln p_j. \quad (46)$$

Shannon proved that this is the only measure of indeterminacy that satisfies the following requirements:

- (i) It is a smooth function of the p_j .
- (ii) If there are N alternatives, all equally probable, then the indeterminacy and hence S must grow monotonically as N increases.
- (iii) Mere grouping of alternatives cannot make any difference: If we add the entropy quantifying ignorance about the true group, and the suitably weighted entropies quantifying ignorance about the true member within each group, we ought to find the same overall entropy S as for ungrouped alternatives.

For continuous distributions with probability density $p(x)$ we take the seemingly analogous expression

$$S = - \int dx p(x) \ln p(x). \quad (47)$$

Let us now assume that we do not know $p(x)$ but that we have global information about it in the form of expectation values for several known functions $f_k(x)$,

$$\langle f_k \rangle = \int dx p(x) f_k(x), \quad k = 1, 2, \dots, K. \quad (48)$$

What is the probability density $p(x)$ which satisfies these K equations without implying any other, spurious information or hidden assumptions? The answer is provided by the *principle of maximal entropy*: If we want compatibility with the given information, yet minimal information content otherwise, we must vary $p(x)$ in such a way that its entropy is maximised, $S = \max$, subject to the K constraints (48). The well known solution to this variational problem, obtained with the technique of Lagrange multipliers, is

$$p(x) = \frac{1}{Z} \exp \left(- \sum_{k=1}^K \lambda_k f_k(x) \right). \quad (49)$$

This probability density is manifestly positive for real λ_k , and properly normalised to unity with

$$Z = \int dx \exp \left(- \sum_{k=1}^K \lambda_k f_k(x) \right). \quad (50)$$

The Lagrange multipliers λ_k must be found from the K constraints (48) or from the equivalent equations

$$\langle f_k \rangle = - \frac{\partial}{\partial \lambda_k} \ln Z. \quad (51)$$

The latter way is more convenient if Z can be expressed as an analytic function of the Lagrange parameters.

Each time a new “global” datum $\langle f_k \rangle$ becomes available one must multiply the existing distribution by a factor $\exp[-\lambda_k f_k(x)]$ (and renormalise). This shows how one can generalise to the case where a given prior distribution $m(x)dx$ is to be updated with new global data $\langle f_k \rangle$: The updated probability density must have the form

$$p(x) = \frac{m(x)}{Z} \exp \left(- \sum_{k=1}^K \lambda_k f_k(x) \right) \quad (52)$$

with

$$Z = \int dx m(x) \exp \left(- \sum_{k=1}^K \lambda_k f_k(x) \right). \quad (53)$$

The Lagrange multipliers can be found from Eqs. 48 or 51 as before. This result can be obtained by maximisation of the relative information entropy (cross entropy)

$$S = - \int dx p(x) \ln \frac{p(x)}{m(x)}, \quad (54)$$

subject to the constraints (48). The cross entropy has the required invariance under change of variable, in contrast to Eq. 47 which we now recognise as restricted to the special case of a uniform prior for the integration variable x .

The maximum entropy algorithm (48)-(51) ought to look familiar to physicists: It constitutes nothing less than the missing rationale for Gibbs' axiomatic approach to thermodynamics. There the maximised information entropy, multiplied by Boltzmann's constant, is Clausius' thermodynamic entropy, and the normalisation constant Z is the partition function from which all macroscopically observable or controllable ensemble averages can be obtained by suitable differentiation. For instance, if E is the (non-negative) energy of the particles of a thermodynamical system, about which nothing is known except their average energy (determined by the temperature of the system), one gets the canonical distribution, $p(E|\langle E \rangle) \propto \exp(-\lambda E)$, i. e. a Boltzmann factor with the inverse temperature appearing as Lagrange parameter. If also the average number of particles is known one obtains the grand-canonical ensemble, $p(E, N|\langle E \rangle, \langle N \rangle) \propto \exp(-\lambda E - \mu N)$, with the chemical potential as a second Lagrange parameter, and so forth.

Data evaluators are mostly confronted with data reported in the form $\langle x \rangle \pm \Delta x$. Our notation indicates that *we interpret these numbers as the measurer's best estimate under quadratic loss*. The first two moments, $\langle x \rangle$ and $\langle x^2 \rangle$, of an unknown distribution are thus given. (Remember that $(\Delta x)^2 = \text{var } x = \langle x^2 \rangle - \langle x \rangle^2$.) If the range of the variate is $-\infty < x < \infty$, the maximum entropy algorithm yields for this kind of information $p(x) \propto \exp(-\lambda_1 x - \lambda_2 x^2)$ as the least restrictive, least informative, hence most conservative and most objective probability density. This is obviously a Gaussian that in terms of the input data must have the form (confirmed, of course, by the maximum entropy algorithm)

$$p(x|\langle x \rangle, \Delta x) dx = \frac{1}{\sqrt{2\pi(\Delta x)^2}} \exp \left[-\frac{1}{2} \left(\frac{x - \langle x \rangle}{\Delta x} \right)^2 \right] dx, \quad -\infty < x < \infty. \quad (55)$$

The case of an inherently positive variate, $0 < x < \infty$, is reduced to the case just considered if we substitute $y = \ln x$, so that $-\infty < y < \infty$. With known first and second moment on the log (y -)scale we get a Gaussian on the log scale, i. e. a lognormal distribution on the linear (x -)scale. If we know only the mean $\langle x \rangle$, and x is inherently positive, we get a decreasing exponential. We encounter here one of the reasons for the ubiquity of these distributions in statistics and data analysis.

Traditionally the Gaussian is considered appropriate only if the variate is affected by very many independent "random" errors so that the central limit theorem is applicable, or it is invoked merely for mathematical convenience, with stern warnings about the dire consequences if the true distribution is *not* Gaussian. The maximum entropy principle can not eliminate those consequences but it prevents bad conscience or paralysis: *If nothing but the best value (mean) and the root-mean-square error (standard deviation) is given, the optimal probability distribution for further inference is the corresponding Gaussian*, whatever the unknown true distribution may happen to be. In contrast to the central limit theorem, the maximum entropy principle works also for correlated data.

Another myth is that systematic errors are to be described by rectangular probability distributions. If we do not know their sign but have at least a vague idea about their possible magnitude (from the state of the art, for instance), the maximum entropy principle tells us to use a Gaussian with zero mean and a width corresponding to that magnitude, rather than a rectangular distribution.

Generalisation to multivariate distributions is straightforward. Take, for instance, a set of experimental errors $\langle x_j \rangle - x_j \equiv \epsilon_j$, $j=1, 2, \dots, n$, about which only the expectation values $\langle \epsilon_j \epsilon_k \rangle$, i. e. their variances and covariances are known. The errors ϵ_j are the Cartesian coordinates of the vector variate ϵ , and the expectation values $\langle \epsilon_j \epsilon_k \rangle$ are the elements of the symmetric, positive definite covariance matrix $\mathbf{C} = \langle \epsilon \epsilon^\dagger \rangle$, where the dagger

denotes the transpose. For each expectation value $C_{jk} = C_{kj}$ we introduce one Lagrange multiplier $\Lambda_{jk} = \Lambda_{kj}$. The maximum entropy distribution is then

$$p(\boldsymbol{\epsilon}|\mathbf{C}) d^n \boldsymbol{\epsilon} = \frac{1}{Z} \exp \left(- \sum_j \sum_k \epsilon_j \Lambda_{jk} \epsilon_k \right) d^n \boldsymbol{\epsilon}. \quad (56)$$

Normalisation is easy in the coordinate system in which the square symmetric matrix $\boldsymbol{\Lambda}$ is diagonal. We therefore substitute $\boldsymbol{\epsilon}' = \mathbf{O}\boldsymbol{\epsilon}$, where \mathbf{O} is the orthogonal matrix that renders $\boldsymbol{\Lambda}' = \mathbf{O}\boldsymbol{\Lambda}\mathbf{O}^\dagger$ diagonal, $\det \boldsymbol{\Lambda}' = \det \boldsymbol{\Lambda}$, and $d^n \boldsymbol{\epsilon}' = d^n \boldsymbol{\epsilon}$. The n -dimensional integral Z factorises then into n elementary integrals over univariate Gaussians,

$$Z = \prod_j \int d\epsilon_j' \exp \left(- \Lambda'_{jj} \epsilon_j'^2 \right) = \sqrt{\frac{\pi^n}{\det \boldsymbol{\Lambda}}}. \quad (57)$$

The relationship between the matrix $\boldsymbol{\Lambda}$ of Lagrange parameters and the given covariance matrix \mathbf{C} can be obtained by differentiation of $\ln Z$, see Eq. 51. One finds

$$C_{jk} = - \frac{\partial}{\partial \Lambda_{jk}} \ln Z = \frac{1}{2} (\boldsymbol{\Lambda}^{-1})_{jk}, \quad (58)$$

since differentiation of the determinant with respect to an element of the matrix yields the cofactor for this element, which for a nonsingular matrix is equal to the corresponding element of the inverse matrix times the determinant. The distribution with the highest entropy among all those having the same covariance matrix \mathbf{C} is therefore

$$p(\boldsymbol{\epsilon}|\mathbf{C}) d(\boldsymbol{\epsilon}) = \frac{1}{\sqrt{\det (2\pi\mathbf{C})}} \exp \left(- \frac{1}{2} \boldsymbol{\epsilon}^\dagger \mathbf{C}^{-1} \boldsymbol{\epsilon} \right) d(\boldsymbol{\epsilon}), \quad -\infty < \epsilon_j < \infty, \quad (59)$$

where $\langle \boldsymbol{\epsilon} \rangle = 0$, $\langle \boldsymbol{\epsilon} \boldsymbol{\epsilon}^\dagger \rangle = \mathbf{C}$, $\det (2\pi\mathbf{C}) = (2\pi)^n \det \mathbf{C}$, and $d(\boldsymbol{\epsilon}) \equiv d^n \boldsymbol{\epsilon}$. Thus we find, for given second moments, an n -variate Gaussian centred at the origin. Since nothing was assumed about first moments of the errors, i. e. about the center of their distribution, there is no reason to prefer either negative or positive ϵ_j , and so the algorithm yields symmetry about zero. What if only the variances C_{jj} are known, but not the covariances C_{jk} , as often happens in practice? In this case only the Lagrange parameters Λ_{jj} appear in (56), i. e. the matrices $\boldsymbol{\Lambda}$ and \mathbf{C} are a priori diagonal. Unknown covariances can thus be taken as zero, and this simple rule can be applied also in cases where some covariances are known and others not. This is another example, after the vanishing first moments, for a general property of maximum entropy distributions: All expectation values vanish unless the constraints demand otherwise. Thus there is no difference whether we set unknown averages equal to zero after introducing Lagrange parameters for them, or whether we ignore them right from the beginning. With $\boldsymbol{\epsilon} = \langle \mathbf{x} \rangle - \mathbf{x}$ one can rewrite the error distribution (59) in the form

$$p(\mathbf{x}|\langle \mathbf{x} \rangle, \mathbf{C}) d(\mathbf{x}) = \frac{1}{\sqrt{\det (2\pi\mathbf{C})}} \exp \left(- \frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^\dagger \mathbf{C}^{-1} (\mathbf{x} - \langle \mathbf{x} \rangle) \right) d(\mathbf{x}), \quad -\infty < x_j < \infty \quad (60)$$

which is the multivariate generalisation of the univariate Gaussian distribution (55).

It should be clear by now that entropy maximisation is a powerful tool for the assignment of prior or any other probabilities. A fascinating and very informative review of the maximum entropy method, including a wide variety of applications from hypothesis testing to non-equilibrium thermodynamics and fluctuation theory, was given by Jaynes (1978).

1.8. APPROXIMATION: MAXIMUM LIKELIHOOD

The more abundant the data are, the less important is the prior distribution. Therefore it is often a reasonable approximation to use a constant prior, as we did initially in our decay constant example. This means that the posterior probability density is taken as equal to the likelihood function. The widely used maximum likelihood method consists essentially of the rule to recommend that parameter value (or vector) which maximises the likelihood function. Since the likelihood function depends only on the sample, the parameter value (or vector) maximising it can depend on nothing else: It is a statistic, and its “direct” probability distribution, i. e. the probability that its deviation from the true value (or vector) will fall between given limits if a sample is taken, can be calculated as the integral of the likelihood function over a corresponding domain of the sample space. If the distribution so obtained is narrow, the true parameter is likely to be close to the particular maximum likelihood estimate obtained from a particular sample. The distribution of the statistic must therefore be related to the Bayesian posterior. In simple cases, *when sufficient statistics exist*, one can, in fact, rigorously deduce the Bayesian posterior from the distribution of the maximum likelihood estimate. This is R.A. Fisher’s (1935) fiducial approach. In such favorable cases one finds that the maximum-likelihood result coincides with the Bayesian result obtained with the appropriate least informative prior.

We illustrate this with our decay constant example. The likelihood function in Eq. 7 becomes maximal for $\lambda = 1/\bar{t}$, the maximum likelihood estimate is therefore the same as the Bayesian “point” estimate (21). The probability for the sufficient statistic \bar{t} to be found in the infinitesimal interval $d\bar{t}$ if a sample is drawn can be obtained by integration of the likelihood function over a spherical shell with radius \bar{t} and thickness $d\bar{t}$ in the space of the t_i . With polar coordinates, $\bar{t} = r^2$, $d\bar{t} = 2rdr$, $d^n t \propto r^{2n-1} dr d\Omega$, and (trivial) integration over the angular coordinates Ω one gets

$$p(\bar{t}|n, \lambda) d\bar{t} \propto e^{-\lambda n r^2} r^{2n-1} dr, \quad 0 < r \equiv \sqrt{\bar{t}} < \infty. \quad (61)$$

After normalisation the right-hand side of this proportionality relation is the same as the Bayesian posterior (20) obtained with Jeffreys’ prior. Obviously we have here the universal probability distribution of the product $\lambda n \bar{t}$ which we may interpret either as the probability of λ given $n\bar{t}$ (one particular sample, a continuum of possible decay constants) or, equally well, as that of $n\bar{t}$ given λ (one particular decay constant, an n -dimensional continuum of possible samples for each positive integer n).

The maximum likelihood method, one of the techniques invented to circumvent priors, is thus in favorable cases equivalent to the Bayesian approach, but even then it is more cumbersome: First one must identify sufficient statistics, then one must calculate their “direct” probability distribution by integration of the likelihood function over a suitable domain in sample space, and finally one must “invert” to get the probability distribution of the parameters. In more complicated cases sufficient statistics may not exist at all, and even if they exist approximations may be required so that the maximum likelihood result only approximates the exact Bayesian result.

1.9. APPROXIMATION: LEAST SQUARES

The next approximation to be discussed, the least-squares method, is the most important one for data evaluation and adjustment. As the simplest example we consider the case that a quantity μ has been measured n times, under different experimental conditions, and that we are given the results in the form $x_j \pm \sigma_j$. These numbers are best interpreted as means and standard errors of unspecified probability distributions. Whether these are

Gaussians or not, the maximum entropy principle tells us to base all further inference on Gaussians if the unknown errors can have any value between $-\infty$ and $+\infty$. The likelihood function is the product of these Gaussian error distributions, and the appropriate prior for the location parameter μ is uniform, hence the posterior is also a Gaussian,

$$p(\mu|\{x_j, \sigma_j\}) d\mu \propto \exp \left[-\frac{1}{2} \sum_j \left(\frac{x_j - \mu}{\sigma_j} \right)^2 \right] d\mu. \quad (62)$$

Introducing σ^{-2} -weighted sample averages,

$$\bar{x} \equiv \frac{\sum_j \sigma_j^{-2} x_j}{\sum_j \sigma_j^{-2}}, \quad (63) \quad \overline{\sigma^2} \equiv \frac{\sum_j \sigma_j^{-2} \sigma_j^2}{\sum_j \sigma_j^{-2}} = \frac{n}{\sum_j \sigma_j^{-2}}, \quad (64)$$

likewise $\overline{x^2}$, and normalising properly, we get

$$p(\mu|\{x_j, \sigma_j\}) d\mu = \frac{1}{\sqrt{2\pi\overline{\sigma^2}/n}} \exp \left[-\frac{(\mu - \bar{x})^2}{2\overline{\sigma^2}/n} \right] d\mu, \quad (65)$$

with mean and variance

$$\langle \mu \rangle = \bar{x}, \quad (66) \quad \text{var } \mu \equiv (\Delta\mu)^2 = \frac{\overline{\sigma^2}}{n}. \quad (67)$$

The relative standard error of the result, $\Delta\mu/\langle\mu\rangle$, is seen to be proportional to $1/\sqrt{n}$ again. The best estimate under quadratic loss is the σ^{-2} -weighted average over all data. It minimises the sum of squares in the exponent of the posterior (62). This least-squares property will be met again in the multivariate generalisation:

Let us consider

- observables y_j , $j = 1, 2, \dots, J$ (e. g. neutron capture data),
- parameters x_μ , $\mu = 1, 2, \dots, M$ (e. g. resonance parameters),
- a theoretical model $\mathbf{y} = \mathbf{y}(\mathbf{x})$ (e. g. the R-matrix theory of resonance reactions),

where $\mathbf{x} = \{x_1, \dots, x_M\}$, $\mathbf{y} = \{y_1, \dots, y_J\}$ are vectors in parameter space and in sample space, respectively. Usually $M < J$ but we shall see that this is not necessary. Now suppose that before the data became available one had prior knowledge about the parameter vector, in the form of an estimated vector $\boldsymbol{\xi}$ and a covariance matrix $\mathbf{A} = \langle \delta\boldsymbol{\xi} \delta\boldsymbol{\xi}^\dagger \rangle$, with $\delta\boldsymbol{\xi} \equiv \boldsymbol{\xi} - \mathbf{x}$, describing the uncertainties and correlations of the estimated parameters. The prior probability distribution of \mathbf{x} , given $\boldsymbol{\xi}$ and \mathbf{A} , is then to be taken as

$$p(\mathbf{x}|\boldsymbol{\xi}, \mathbf{A}) d(\mathbf{x}) \propto \exp \left[-\frac{1}{2} (\boldsymbol{\xi} - \mathbf{x})^\dagger \mathbf{A}^{-1} (\boldsymbol{\xi} - \mathbf{x}) \right] d(\mathbf{x}), \quad (68)$$

where $d(\mathbf{x}) \equiv d^M x$ is the volume element of the M-dimensional parameter space (not to be mistaken for the infinitesimal vector $d\mathbf{x}$).

Suppose further that measurements yielded a data vector $\boldsymbol{\eta}$, affected by experimental errors whose uncertainties and correlations are specified by the covariance matrix $\mathbf{B} =$

$\langle \delta \boldsymbol{\eta} \delta \boldsymbol{\eta}^\dagger \rangle$, with $\delta \boldsymbol{\eta} \equiv \boldsymbol{\eta} - \mathbf{y}$, so that the likelihood to obtain these values, given the true vector \mathbf{y} of observables, is

$$p(\boldsymbol{\eta}|\mathbf{y}, \mathbf{B}) d(\boldsymbol{\eta}) \propto \exp \left[-\frac{1}{2}(\boldsymbol{\eta} - \mathbf{y})^\dagger \mathbf{B}^{-1}(\boldsymbol{\eta} - \mathbf{y}) \right] d(\boldsymbol{\eta}), \quad (69)$$

where $d(\boldsymbol{\eta}) \equiv d^J \boldsymbol{\eta}$ is the volume element in the J -dimensional sample space.

These probability assignments are the multivariate Gaussians dictated by the maximum entropy principle (compare Eq. 59). Multiplying prior distribution and likelihood function we get the posterior distribution,

$$p(\mathbf{x}|\boldsymbol{\xi}, \mathbf{A}, \boldsymbol{\eta}, \mathbf{B}) d(\mathbf{x}) \propto \exp \left[-\frac{1}{2}(\boldsymbol{\xi} - \mathbf{x})^\dagger \mathbf{A}^{-1}(\boldsymbol{\xi} - \mathbf{x}) - \frac{1}{2}(\boldsymbol{\eta} - \mathbf{y}(\mathbf{x}))^\dagger \mathbf{B}^{-1}(\boldsymbol{\eta} - \mathbf{y}(\mathbf{x})) \right] d(\mathbf{x}). \quad (70)$$

So far we neglected correlations between prior information and new data. This is, however, not always possible, for instance if the prior information stems from older measurements in which the same techniques and standards were employed as in the new measurements. Generalisation to this situation is not difficult. Eq. 70 shows that we can consider the old parameter estimates and the new data on the same footing. Prior estimates and their uncertainties have exactly the same impact as if they were data obtained in a measurement of the special observables $y_\mu = x_\mu$. So we can combine the vectors \mathbf{x} and $\mathbf{y}(\mathbf{x})$ in a hypervector $\mathbf{z}(\mathbf{x}) \equiv \{\mathbf{x}, \mathbf{y}(\mathbf{x})\}$. If the corresponding data vector $\boldsymbol{\zeta} \equiv \{\boldsymbol{\xi}, \boldsymbol{\eta}\}$ and the associated covariance matrix $\mathbf{C} \equiv \langle \delta \boldsymbol{\zeta} \delta \boldsymbol{\zeta}^\dagger \rangle$ are given, the maximum entropy distribution is

$$p(\mathbf{x}|\boldsymbol{\zeta}, \mathbf{C}) d(\mathbf{x}) \propto \exp \left[-\frac{1}{2}(\boldsymbol{\zeta} - \mathbf{z}(\mathbf{x}))^\dagger \mathbf{C}^{-1}(\boldsymbol{\zeta} - \mathbf{z}(\mathbf{x})) \right] d(\mathbf{x}), \quad (71)$$

where \mathbf{C} now contains also covariances between prior estimates and new data, $\langle \delta \xi_\mu \delta \eta_i \rangle$. This posterior distribution is the most general, most detailed result of our Bayesian parameter estimation. It contains in analytic form all information about the parameter vector \mathbf{x} that the given prior information and the new data contain.

The remaining task is condensation into a recommended parameter vector and its uncertainty. Decision theory tells us what to do if a loss function is given. If none is given we assume quadratic loss which entails that we must recommend the updated posterior mean vector $\langle \mathbf{x} \rangle$ and specify the uncertainties and correlations by the posterior covariance matrix $\langle \delta \mathbf{x} \delta \mathbf{x}^\dagger \rangle$, with $\delta \mathbf{x} \equiv \mathbf{x} - \langle \mathbf{x} \rangle$. For a linear model $\mathbf{y}(\mathbf{x})$ the necessary integrations are easy as we shall see in Sect. 2.2. For a nonlinear model one must either integrate numerically, which is not practical if many parameters are to be estimated (except perhaps with Monte Carlo techniques), or one must resort to the method of steepest descent (Laplace approximation). The latter means essentially that the exact posterior is replaced by a multivariate Gaussian with the same maximum and the same curvature tensor at the maximum so that the integrand is well approximated at least in the domain that contributes most to the integral. This is accomplished by Taylor expansion of the exponent of Eq. 71 about its minimum and truncation after the bilinear terms. (see e. g. Bernardo and Smith 1994, Lange 1999). The maximum, i. e. the a posteriori most probable parameter vector $\hat{\mathbf{x}}$, is specified by

$$(\boldsymbol{\zeta} - \mathbf{z}(\hat{\mathbf{x}}))^\dagger \mathbf{C}^{-1}(\boldsymbol{\zeta} - \mathbf{z}(\hat{\mathbf{x}})) = \min. \quad (72)$$

This is the formal statement of the principle of least squares in its most general form. (In the principal-axes system the quadratic form appears as a sum of squares, whence the name).

In frequentist statistics the least-squares principle is introduced more or less ad hoc, or it is derived from the maximum-likelihood principle which in turn is introduced ad hoc. In both cases only the likelihood function is utilised. Here we recognise that the least-squares condition is a natural consequence of more fundamental premises, and that it demands maximisation not only of the likelihood function but of the full posterior distribution. Without simplifying too much one may say that *generalised least squares is nothing but Bayesian parameter estimation under quadratic loss in Laplace approximation, in problems where only data and data uncertainties are given* so that the maximum entropy principle demands Gaussians whether the unknown true distributions are Gaussian or not. An ad hoc least-squares principle is not needed.

We conclude this chapter by summarising the essential difference between the frequentist (also called orthodox or classical or sampling-theoretical) and the Bayesian approach to inductive inference as follows. *Frequentists average over all imaginable outcomes of a measurement, conditional on given causes, whereas Bayesians average over all possible causes, conditional on the one observed outcome and prior knowledge.* There can be no doubt that the Bayesian approach is the appropriate one for a physicist who must infer cross sections or cross section parameters from uncertainty-affected observations.

2. Evaluation of Nuclear Data for Applications

In the following sections we shall discuss some of the more practical aspects of data evaluation, with special attention to the least-squares formalism, to statistical versus systematic errors and how the latter cause correlations. Historically data evaluation in the modern sense began with the effort of Dunnington (1939), Cohen, DuMond and collaborators (1957, 1992) to determine a recommended set of fundamental constants (Planck's constant, fine-structure constant, electron mass, etc.), and to assess the uncertainties, by a comprehensive least-squares fit to all relevant experimental data. At about the same time the rapidly growing nuclear industry began to develop a voracious appetite for accurate nuclear data, especially neutron and photon cross sections, but also nuclear structure and decay data. The data for nuclear reactions induced by neutrons having "thermal" energies around 25.3 meV were evaluated with first priority (see Westcott et al. 1965, Lemmel 1975), but the expanding scope of nuclear technology, from thermal fission reactors to fast fission reactors and eventually to fusion reactors, brought a corresponding expansion of the energy range of interest. Modern neutron data files contain millions of cross section values covering the whole range from 10 μeV to at least 20 MeV for hundreds of isotopes, and computers are indispensable for their maintenance and utilisation. Nuclear data provide a highly developed example of the process leading from experimental raw data to evaluated data files.

2.1. STEPWISE PREPARATION OF NUCLEAR DATA FOR APPLICATIONS

Nuclear (and other scientific) data for technological applications are usually prepared in several steps, by different groups of specialists. Let us take neutron cross section data to illustrate these steps.

1. **Measurement:** Experimenters take data, typically at steady-state or pulsed accelerators, the latter permitting use of the time-of-flight method. This method produces, in one experimental run and therefore under exactly identical conditions, large numbers of data points covering wide energy ranges with high resolution. The simplest type of measurement is that of the total cross section σ . One measures that fraction of a beam of particles of given energy (given flight time) which traverses, without interacting, a sample of given thickness n (atoms/barn). This fraction, the transmission, is $1 - \sigma\Delta n$ for a very thin layer of material. For the whole sample it is

$$T = \lim_{\Delta n \rightarrow 0} (1 - \sigma\Delta n)^{n/\Delta n} = e^{-n\sigma}. \quad (73)$$

In practice the transmission is obtained as the ratio of the count rates from a "sample-in" and a "sample-out" run. The incoming flux and the detector efficiency cancel out, so there is no calibration uncertainty. Background noise, however, requires corrections. If the cross section has resonance structure one employs both "thin" and "thick" samples in order to obtain good accuracy of the extracted cross sections, at the resonance peaks as well as in the valleys between resonances.

Partial cross sections are more difficult to measure. Experimentally one obtains a reaction yield, for instance by recording the fission products or capture gamma rays emitted from a thin sample under neutron bombardment. The reaction yield is defined as that fraction of beam particles which undergoes a reaction of the measured type in the sample. It is a sum of contributions from multiple-collision events where the beam particle undergoes zero, one, two etc. scattering collisions before it finally induces the recorded reaction,

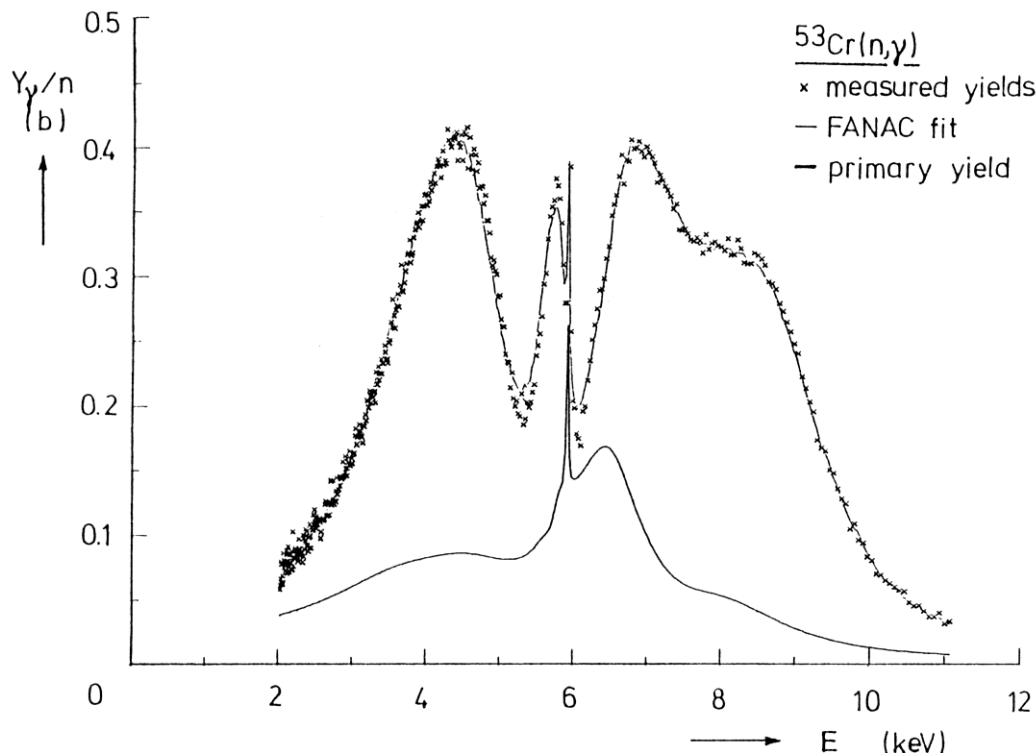


Fig. 3. Neutron capture yields, measured with the time-of-flight method at the Geel linear accelerator (Gelina) across four overlapping s-wave resonances and one narrow p-wave resonance, including instrumental and Doppler broadening and heavy multiple-collision capture. Thin curve: least-squares fit with the R-matrix code FANAC (Fröhner 1977), thick curve: corresponding calculated first-collision yield (Brusegan et al. 1992).

$$Y_x = Y_{x0} + Y_{x1} + Y_{x2} + \dots \quad (74)$$

The first-collision yield is the product of the interaction probability of incident neutrons times the ratio σ_x/σ of the number of (n,x) events to the total number of interactions,

$$Y_{x0} = (1 - e^{-n\sigma}) \frac{\sigma_x}{\sigma}, \quad (75)$$

where σ_x is the partial cross section for the (n,x) reaction of interest. For very thin samples, $n\sigma \ll 1$, one can neglect multiple collisions and take $Y_x \simeq Y_{x0} \simeq n\sigma_x$, but for thicker samples the full first-collision term (75) must be used. Especially for resonance reactions the multiple-collision terms are increasingly complicated functionals of the cross sections, so that they must be calculated by Monte Carlo simulation (see Fröhner 1989). The observed count rate is the product of yield, incoming flux and detector efficiency. The latter two quantities must be measured separately and introduce correlated normalisation uncertainties as will be discussed below. Fig. 3 shows neutron capture data measured across overlapping resonances of $^{53}\text{Cr}+n$ with heavy contribution from multiple-collision capture (Brusegan et al. 1992). Fig. 4 demonstrates the quality of modern time-of-flight data: there are three sets of neutron transmission data and two sets of neutron-induced fission yields, all five showing the same resonances of the compound system $^{239}\text{Pu}+n$ (Derrien et al. 1988).

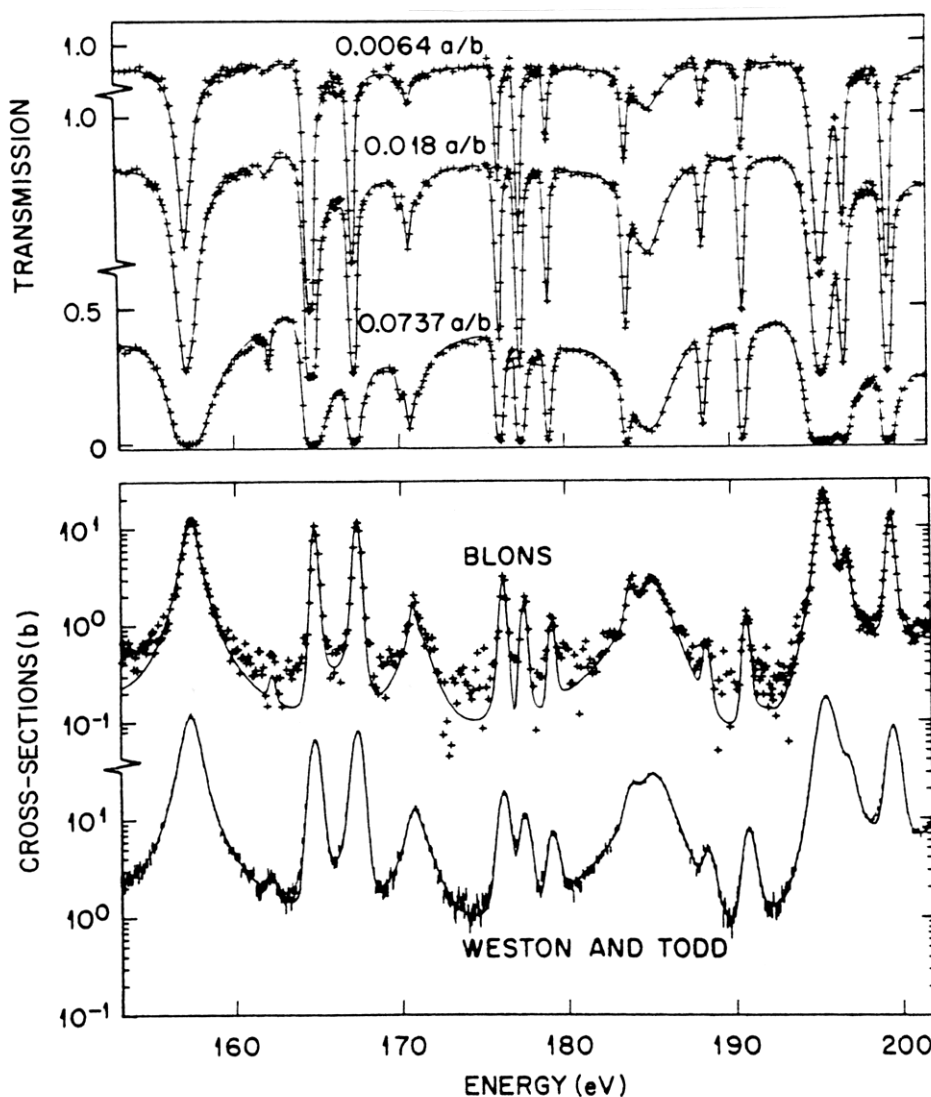


Fig. 4. Transmission data for three samples and two sets of fission cross sections, measured with the time-of-flight method at the electron linear accelerators at Saclay (Blons) and at Oak Ridge (all others), showing resonances of $^{239}\text{Pu}+n$. The curves correspond to a simultaneous R-matrix fit to the Oak Ridge data with the generalised least-squares code SAMMY (Larson and Perey, 1980). From Derrien et al. (1988).

2. **Reduction of raw data:** Constant and time-dependent backgrounds are subtracted, sample impurities are corrected for, and, in the case of partial cross section data (yields), flux and detector efficiency are factorised out. Corrections for multiple scattering and instrumental resolution are usually deferred to the next step if they require e. g. resonance theory. Due to the complicated relationship between cross sections and observables (Eqs. 73-75) cross sections are still unavailable at this stage, except perhaps from thin-sample measurements.

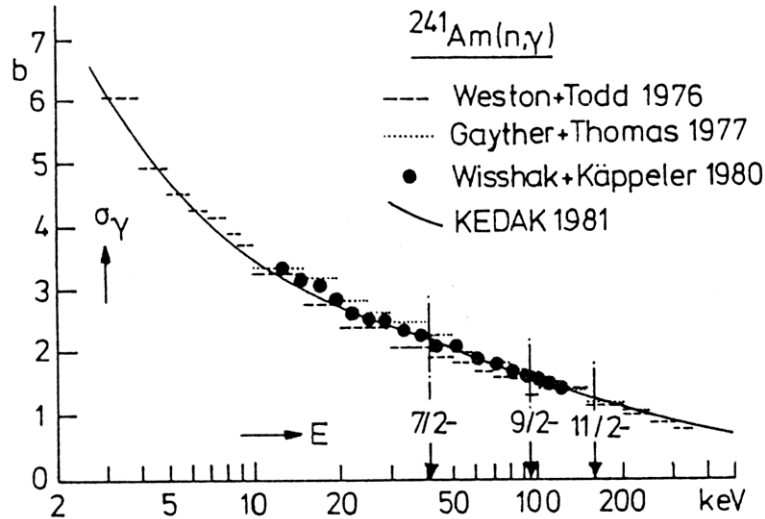


Fig. 5. Three sets of capture cross section data measured at Oak Ridge (1976), Harwell (1977) and Karlsruhe (1980) for ^{241}Am in the unresolved resonance region. The (piecewise) smooth curve is a generalised least-squares fit obtained with the Hauser-Feshbach code FITACS. From Fröhner et al. (1982)

3. **Analysis of clean data:** Whenever possible, measured cross sections are parametrised by means of nuclear reaction theory. This ensures consistency between partial and total cross sections, permits inter- and extrapolation into regions where no data were measured, and prevents unphysical values from being recommended. In order of increasing energy the following theories and models are employed:

- R-matrix theory in the thermal and resolved resonance region (parameters: level energies, level spins, partial widths), see Figs. 3, 4;
- level-statistical (Hauser-Feshbach) theory in the unresolved resonance region (parameters: level density, strength functions and average partial widths, or the equivalent channel transmission coefficients), see Fig. 5;
- the optical model at higher energies, where levels overlap strongly but compound reactions still dominate (parameters: radius, depth, diffuseness and deformation of the real and imaginary potential well), see Fig. 6;
- precompound, direct and multistep theories at still higher energies where direct and preequilibrium processes are important.

This is complemented by the giant dipole resonance model for photon reactions, fission barrier models for fission reactions, etc. At this stage parameter estimation (curve fitting) techniques are used extensively. From the estimated parameters one generates the cross sections and uncertainties which the user finds in files of evaluated nuclear data. Where suitable models are not available, polynomial or similar fits are used to interpolate between evaluated data points. Gaps in the measured data are filled by model calculations or with the help of systematics. Figs. 3-6 show examples of theoretical curves fitted to measured data.

4. **Generation of group constants:** Doppler-broadened point cross sections for all open reaction channels and for various temperatures can now be calculated, and averaged over finite energy intervals as required for reactor or shielding calculations. The result is a set of group constants, usually including group-averaged (“infinite-

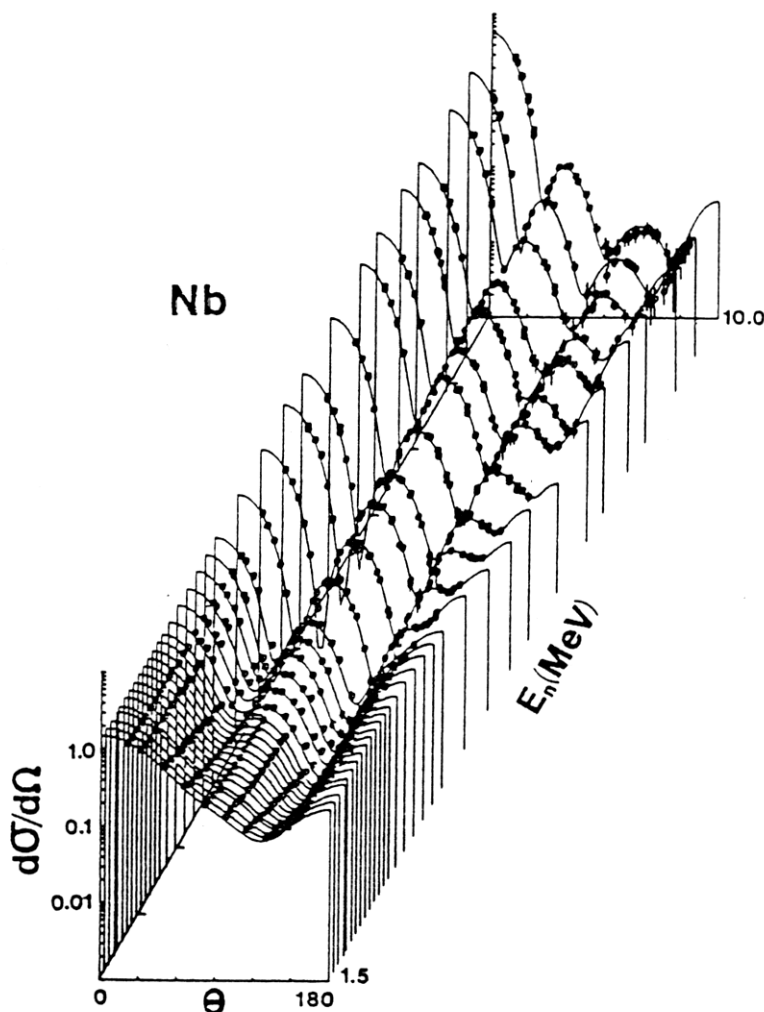


Fig. 6. Angle-dependent elastic-scattering cross sections of ^{93}Nb measured at Argonne National Laboratory. Differential cross sections are given in b/sr and scattering angles (θ , laboratory system) in degrees. Point symbols represent data and their errors, curves are calculated from an adjusted optical model potential. From Smith et al. (1985).

dilution”) cross sections for room temperature and temperature- and dilution-dependent self-shielding factors on a specified grid of temperatures and “dilutions” (describing admixtures of other nuclides), see e. g. Fröhner (1989). These group constant sets are usually special-purpose files for particular applications, in contrast to the general-purpose “microscopic” files from which they are generated. In order to enhance their special-purpose reliability they are often adjusted (least-squares fitted) to data obtained in “integral” experiments with special test reactors or similar facilities.

All these steps require time and several years may pass before measured nuclear data needed for technological or scientific applications become available in the form of evaluated computer files. This has motivated efforts to coordinate the work on an international scale.

Requests for Nuclear Data, known as the WRENDA list.

Measured data are collected by a network of data centers, each operating within its agreed service area:

| | |
|-------|--|
| NNDC | (National Nuclear Data Center) at Brookhaven, USA, servicing the USA and Canada; |
| NEADB | (NEA Data Bank, OECD) at Saclay, France, servicing the non-American OECD countries; |
| CJD | (Centr po Jadernym Dannym) at Obninsk, Russia, servicing the territory of the former Soviet Union; |
| NDS | (Nuclear Data Section, IAEA), at Vienna, Austria, servicing all other countries. |

Regular data exchange ensures that the data base is essentially the same at all four centers. Evaluated data are also collected, notably the files ENDF (USA), JEF (NEADB member countries), JENDL (Japan), and BROND (Comecon countries). The centers produce also the widely used bibliographic Computer Index to Neutron Data (CINDA). The well known “barn book” (Mughabghab et al. 1984) containing comprehensive resonance parameter tables and cross section plots, is a product of NNDC. Similar networks of data centers compile and distribute charged-particle data and nuclear structure and decay data. The ENSDF file contains evaluated data of the latter type, it is the machine-readable offspring of the well known Nuclear Data Sheets and the Table of Isotopes (Lederer et al. 1979).

Comparable international cooperation involving data bank networks exists in meteorology, high-energy physics, materials research, aerospace research, and many other scientific and technological areas.

2.2. ITERATIVE LEAST-SQUARES FITTING

Most of the parameter estimation work in the analysis of clean data (step 3. in the last section) is based on the least-squares method. Let us therefore return to the least squares condition,

$$(\zeta - \mathbf{z}(\hat{\mathbf{x}}))^{\dagger} \mathbf{C}^{-1} (\zeta - \mathbf{z}(\hat{\mathbf{x}})) = \min. \quad (76)$$

We recall that the vector ζ is the combined set of prior parameter estimates and of measured data, $\zeta = \{\xi, \eta\} = \{\xi_1, \dots, \xi_M, \eta_1, \dots, \eta_J\}$. The measured data may come from quite different types of measurements which, of course, must be mathematically described by the corresponding coordinates of the model vector $\mathbf{y}(\mathbf{x})$.

Without prior information about the parameters one has simply

$$(\eta - \mathbf{y}(\mathbf{x}))^{\dagger} \mathbf{B}^{-1} (\eta - \mathbf{y}(\mathbf{x})) = \min. \quad (77)$$

If one neglects also the correlations between the data η_j , so that the matrix \mathbf{B} becomes diagonal, one gets the equations for “primitive” least-squares fitting which is still employed in many fitting codes. It utilises only the new data and their uncertainties, ignoring all prior information that may be available. The resulting parameter estimates and their uncertainties must then be combined with the prior information, derived for instance from previous measurements, by some kind of weighted averaging after the fit. Now a fundamental tenet of inductive logic is that the *probabilities should encode all the available information*, prior knowledge as well as new evidence. It is therefore more correct – and also more convenient – to include the prior estimates (or first guesses) and their uncertainties (variances) right from the beginning in the form of a prior distribution as in

Eq. 70. Moreover, the convergence of nonlinear least-squares fitting is always improved – often dramatically – after inclusion of prior information by means of Gaussian or similar priors. One need not worry too much about unknown input correlations. According to the maximum entropy principle we may simply neglect those that we do not know, and set the corresponding elements of the input covariance matrix equal to zero. On the other hand it is not difficult to construct complete input covariance matrices if the uncertainties of the input data are well documented, with clear specification of the various error components including their root-mean-square errors. This will be explained below but for now we shall assume that the input covariance matrix \mathbf{C} is given, at least in diagonal form.

Let us try to calculate the parameter estimate under quadratic loss, i.e. the mean and the (output) covariance matrix of the posterior distribution (71). Numerical integration is always possible but in order to get analytic expressions we must employ the Laplace approximation (method of steepest descent, saddle point integration) which is strictly exact only in the case of a linear model $\mathbf{y}(\mathbf{x})$. For nonlinear models it is adequate for most practical purposes but one must keep in mind that it may fail if nonlinearities are severe across the peak of the posterior. Taylor expansion of the exponent of Eq. 71 around its minimum at $\mathbf{x} = \hat{\mathbf{x}}$ yields

$$\begin{aligned} Q(\mathbf{x}) &\equiv (\zeta - \mathbf{z}(\mathbf{x}))^\dagger \mathbf{C}^{-1} (\zeta - \mathbf{z}(\mathbf{x})) \\ &= Q(\hat{\mathbf{x}}) + (\mathbf{x} - \hat{\mathbf{x}})^\dagger \mathbf{A}'^{-1} (\mathbf{x} - \hat{\mathbf{x}}) + \dots \end{aligned} \quad (78)$$

The most probable parameter vector $\hat{\mathbf{x}}$ is defined by $\nabla Q = 0$ (with $\nabla_\mu \equiv \partial/\partial x_\mu$), i. e. by the “normal” equations

$$\mathbf{S}(\hat{\mathbf{x}})^\dagger \mathbf{C}^{-1} (\zeta - \mathbf{z}(\hat{\mathbf{x}})) = 0, \quad (79)$$

and the matrix \mathbf{A}' by

$$\mathbf{A}'^{-1} \equiv \frac{1}{2} [\nabla \nabla^\dagger Q]_{\mathbf{x}=\hat{\mathbf{x}}} = \mathbf{S}(\hat{\mathbf{x}})^\dagger \mathbf{C}^{-1} \mathbf{S}(\hat{\mathbf{x}}) + \dots, \quad (80)$$

where \mathbf{S} is the rectangular matrix of sensitivity coefficients

$$S_{j\mu} = \frac{\partial z_j}{\partial x_\mu}. \quad (81)$$

The vector $\hat{\mathbf{x}}$ can be found from the normal equations (79) by Newton-Raphson iteration: If one has, after n steps, an approximate solution \mathbf{x}_n , one can insert the linear approximations

$$\mathbf{z}(\hat{\mathbf{x}}) \simeq \mathbf{z}(\mathbf{x}_n) + \mathbf{S}(\mathbf{x}_n)(\hat{\mathbf{x}} - \mathbf{x}_n) \quad (82)$$

and $\mathbf{S}(\hat{\mathbf{x}}) \simeq \mathbf{S}(\mathbf{x}_n)$ in the normal equations and solve for $\hat{\mathbf{x}}$. The improved solution found in this way is

$$\mathbf{x}_{n+1} = \mathbf{x}_n + [\mathbf{S}(\mathbf{x}_n)^\dagger \mathbf{C}^{-1} \mathbf{S}(\mathbf{x}_n)]^{-1} \mathbf{S}(\mathbf{x}_n)^\dagger \mathbf{C}^{-1} [\zeta - \mathbf{z}(\mathbf{x}_n)]. \quad (83)$$

So one can start with the a priori most probable value, $\mathbf{x}_0 = \boldsymbol{\xi}$, on the right-hand side, calculate an improved value \mathbf{x}_1 , reinsert it, and so on, until stationarity is achieved within single precision of the computer (or until some other reasonable convergence criterion is satisfied). In each step we must recalculate $\mathbf{z}(\mathbf{x})$ and the sensitivity matrix $\mathbf{S}(\mathbf{x})$. For a linear model \mathbf{z} is a linear function of \mathbf{x} , i.e. Eq. 82 is exact for $n = 0$, with S not depending

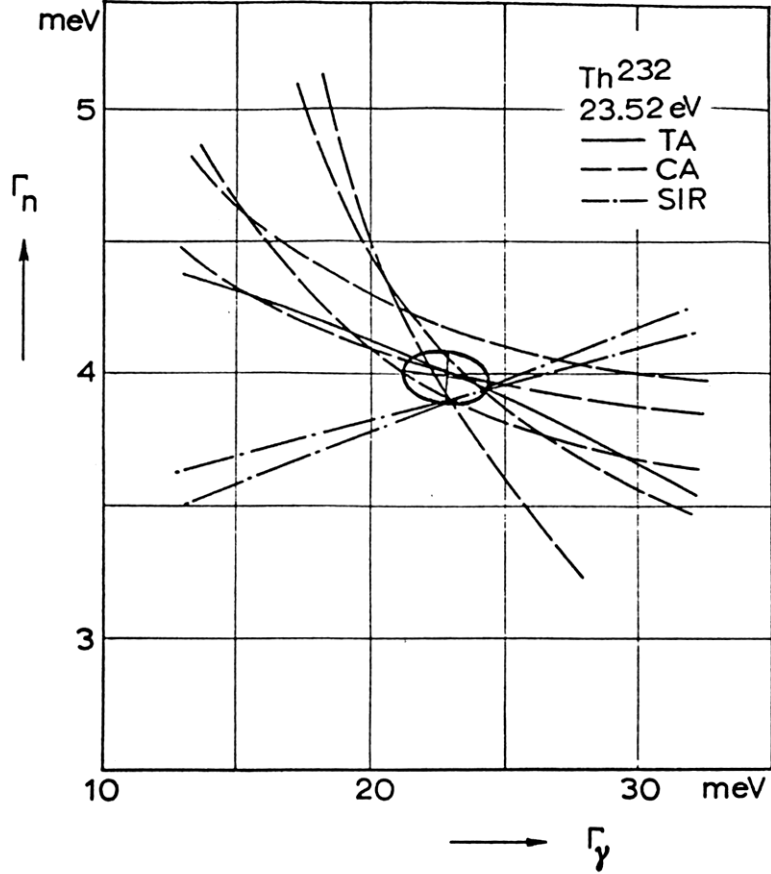


Fig. 7. Illustration of nonlinearities and correlations in least-squares fitting: neutron width Γ_n and radiation width Γ_γ of the 23.52 eV resonance of $^{232}\text{Th}+n$ were estimated by simultaneous adjustment to two measured transmission dip areas (TA), five capture peak areas (CA) and two self-indication ratios (SIR). The curves represent the eight components of the theoretical-model vector $\mathbf{y}(\mathbf{x})$, the parameter vector \mathbf{x} having two coordinates, Γ_n and Γ_γ . The “error ellipse” shows the best estimate and the correlated uncertainties (TACASI code). From Fröhner (1966).

on \mathbf{x} . As a consequence the final result is already obtained in the first step, $\hat{\mathbf{x}} = \mathbf{x}_1$, as expected.

Once we have established the final result $\hat{\mathbf{x}}$ of the iteration we can calculate the recommendation (in saddle point approximation)

$$\langle \mathbf{x} \rangle = [\mathbf{S}(\hat{\mathbf{x}})^\dagger \mathbf{C}^{-1} \mathbf{S}(\hat{\mathbf{x}})]^{-1} \mathbf{S}(\hat{\mathbf{x}})^\dagger \mathbf{C}^{-1} [\zeta - \mathbf{z}(\hat{\mathbf{x}})], \quad (84)$$

$$\langle \delta \mathbf{x} \delta \mathbf{x}^\dagger \rangle = [\mathbf{S}(\hat{\mathbf{x}})^\dagger \mathbf{C}^{-1} \mathbf{S}(\hat{\mathbf{x}})]^{-1}. \quad (85)$$

For a linear model we simply put $\hat{\mathbf{x}} = \boldsymbol{\xi}$, the mean values thus obtained are exact.

Note that in the process of iteration one has to calculate improved vectors $\mathbf{z}(\mathbf{x})$ containing both improved parameters \mathbf{x}_n and improved data $\mathbf{y}(\mathbf{x}_n)$. The changes are small if the sensitivity coefficients are small and if the calculated values are already close to

the prior estimates of parameters and data. In neutron data evaluation and cross section parametrisation the number M of parameters is usually much smaller than the number J of experimental data points. The opposite is true for group constant adjustment, since the number M of adjusted group constants in a data library is usually much larger than the number J of integral data (measured reactor responses). The equations are the same in both cases. The iterative Bayesian least-squares formalism is employed in the resonance analysis code SAMMY (Larson and Perey 1980) and in the Hauser-Feshbach program FITACS (Fröhner et al. 1982) that is used to analyse resonance-averaged neutron cross section data in the unresolved resonance region. Experience with these codes has clearly shown the advantage of explicit inclusion of a-priori information via a Gaussian prior. Because the prior constrains the parameter search smoothly to a reasonable domain, the (linear programming) problems encountered with sharp boundaries are avoided, and convergence is dramatically improved relative to earlier “primitive” least-squares versions of these codes which did not utilise prior uncertainties. A typical problem with “primitive” least-squares fitting, viz. lack of convergence if too many parameters are adjusted simultaneously, is practically eliminated with the generalised least-squares scheme which tolerates large numbers of adjusted parameters (e. g. several dozens in resonance analysis), even if the prior uncertainties are quite large.

It must be kept in mind, however, that the covariance formalism based on first-order error propagation, and hence the least-squares approach to data adjustment and model fitting, is exact only for linear models but not in general. Different estimates can result in nonlinear least-squares data combination depending on whether an evaluator has the raw data available or whether he can only use reduced data with covariance information. A much discussed example of this was “Peelle’s Pertinent Problem” (see Fröhner 1997).

In reactor physics one often works with relative errors, $(\zeta_j - z_j)/\zeta_j$, and with the corresponding relative covariance matrix elements $\langle \delta\zeta_j \delta\zeta_k \rangle / (\zeta_j \zeta_k)$, and relative sensitivity coefficients, $S_{j\mu}/\zeta_j$. This amounts to the replacements

$$\zeta - \mathbf{z} \rightarrow \mathbf{D}(\zeta - \mathbf{z}), \quad \mathbf{C} \rightarrow \mathbf{D}\mathbf{C}\mathbf{D}, \quad \mathbf{S} \rightarrow \mathbf{D}\mathbf{S} \quad \text{with} \quad D_{jk} = \delta_{jk}/\zeta_j. \quad (86)$$

Obviously the inserted diagonal matrices \mathbf{D} are canceled by their inverses in the least-squares equations, so the use of relative quantities changes only the form of the equations but not the content. The advantage is, of course, that relative (e.g. percentage) errors and sensitivities are easier to grasp and to remember which makes it easier to compare them and to assess their relative importance in problems with many physically different parameters. Difficulties arise, however, if the values ζ_j are much smaller than their uncertainties so that one must divide essentially by zero. In this case it is better to use absolute rather than relative values.

Occasionally the question is raised whether it is correct to use the (measured or a priori estimated) data ζ_j as reference or whether it is not better to use the true values z_j . The answer is simple: True values are principally unknown, so it is not possible to use them as reference for instance in computer files of relative covariance matrices and sensitivity coefficients. All one can do is to *use the best estimates that are available at a given stage*. Before least-squares adjustment these are the prior parameter estimates ξ_μ and the experimental data η_j , hence the quantities ζ_j . After the adjustment one has improved parameter estimates $\langle x_\mu \rangle$ and improved calculated data $y_j(\langle \mathbf{x} \rangle)$, and those are then the correct reference values for the relative output covariance matrices. True values are never available, they appear only in the equations as arguments of probability distributions, i. e. as *possible* values, and are integrated out whenever expectation values or other recommended quantities are computed for practical applications.

2.3. STATISTICAL ERRORS: POISSON STATISTICS

We must now discuss the error information that is needed for the construction of the covariance matrix \mathbf{C}_η that describes the uncertainties of the data and their correlations. In practically all nuclear data measurements one detects and counts particles of a particular type, for instance fission fragments signalling nuclear fission events, or gamma quanta signalling radiative capture events. The count rates are a measure of the corresponding fission or capture probabilities (conventionally expressed as nuclear fission or capture cross sections). In the limit of infinite counting time, and in the absence of other errors, one would measure the probabilities (in the frequentist sense) directly, but in practice there is always some statistical uncertainty about the limiting count rate (or cross section) because of the finite counting time. What can we say about the true rate λ , and especially about its uncertainty, if n events have been registered during a time t ?

Constant count rate means that there is a well-defined average time interval $\langle t \rangle$ between counts. With this global information the maximum entropy principle yields immediately the familiar exponential interval distribution of Poisson statistics,

$$p(t|\lambda) dt = e^{-\lambda t} \lambda dt, \quad 0 < t < \infty. \quad (87)$$

The Lagrange multiplier, the rate constant, is the reciprocal of the mean interval, $\lambda = 1/\langle t \rangle$. Knowing the interval distribution one can write down the joint probability that counts are registered in infinitesimal time intervals dt_1, dt_2, \dots, dt_n within some time span t , and integrate over all possible locations of these intervals. The result is the probability of n counts registered at arbitrary times within the interval t , given the rate constant λ ,

$$P(n|\lambda, t) d\lambda = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (88)$$

This is the Poisson distribution. Bayes' theorem with Jeffreys' prior for the scale parameter λ yields immediately the inverse probability (Sect. 1.5, Eq. 20)

$$p(\lambda|t, n) d\lambda = \frac{e^{-x} x^n}{\Gamma(n)} \frac{dx}{x}, \quad 0 < x \equiv \lambda t < \infty, \quad (89)$$

with

$$\langle \lambda \rangle = \frac{n}{t}, \quad (90) \quad \frac{\Delta \lambda}{\langle \lambda \rangle} = \frac{1}{\sqrt{n}}. \quad (91)$$

The relative uncertainty is nothing but the familiar $1/\sqrt{n}$ rule for the estimation of statistical uncertainties, widely used for all kinds of tallies in statistics generally and in Monte Carlo work in particular. More general problems, for instance the determination of a count rate in the presence of a background, are also readily solved with the Bayesian method (see Loredano 1990, Fröhner 1997).

2.4. SYSTEMATIC ERRORS: CORRELATED UNCERTAINTIES AND THEIR PROPAGATION

We must now discuss briefly the basic types of systematic errors and how they induce correlations between estimates. As before we denote the unknown errors of the data by $\delta\eta_i = \eta_i - y_i$. If they were of purely statistical origin they would be uncorrelated, and the elements of the covariance matrix \mathbf{B} would be

$$\langle \delta\eta_j \delta\eta_k \rangle = \delta_{jk} \text{var } \eta_j \equiv \delta_{jk} (\Delta\eta_j)^2, \quad (92)$$

i. e. the matrix would be diagonal. This is assumed in many “primitive” least-squares codes: each squared error is weighted by the reciprocal variance. Besides the statistical errors there are, however, always experimental errors from flux determination, detector calibration, timing uncertainty, etc. In contrast to the statistical errors these so-called systematic errors are common to a whole set of data, for instance to all the data accumulated in the time channels during a time-of-flight experiment. Quite generally, *correlations between data are usually caused by common errors*. To see this we write the unknown total errors in the form

$$\delta\eta_j = \delta\eta'_j + \delta\eta', \quad (93)$$

where $\delta\eta'_j$ is the statistical, and $\delta\eta'$ the systematic error. The latter is the same for the whole data set so it does not carry a subscript. The elements of the covariance matrix \mathbf{B} are

$$\langle \delta\eta_j \delta\eta_k \rangle = \langle (\delta\eta'_j)^2 \rangle \delta_{jk} + \langle (\delta\eta')^2 \rangle, \quad (94)$$

since the statistical errors of different data points are uncorrelated, $\langle \delta\eta'_j \delta\eta'_k \rangle = \langle (\delta\eta'_j)^2 \rangle \delta_{jk}$, with zero mean, $\langle \delta\eta'_j \rangle = 0$, and since there is no correlation between statistical and systematic errors, $\langle \delta\eta'_j \delta\eta' \rangle = 0$. We conclude that common, i. e. systematic, errors always produce correlations between the elements of an experimental data set.

The most frequent systematic errors are those that shift the observed value (e.g. background errors) and those that multiply them (e.g. errors in flux normalisation or detector calibration). If we vary a subtracted background b for one of the data points, we must do the same for all others: all the reported values η_j must be varied together. This is, of course, the literal meaning of covariance. More generally, if the reported data were obtained from raw count rates a_j by application of common corrections b, c, \dots , the total errors are $\delta\eta_j \simeq (\partial\eta_j/\partial a_j)\delta a_j + (\partial\eta_j/\partial b)\delta b + (\partial\eta_j/\partial c)\delta c + \dots$. Now the statistical errors and usually the common errors, too, are uncorrelated, so that all their covariances vanish, $\langle \delta a_j \delta a_k \rangle = \langle \delta a_j \delta b \rangle = \langle \delta b \delta c \rangle = 0$. The overall covariance elements depend then only on mean square errors (variances) of the components and on sensitivity coefficients (derivatives),

$$\langle \delta\eta_j \delta\eta_k \rangle \simeq \left(\frac{\partial\eta_j}{\partial a_j} \right)^2 (\Delta a_j)^2 \delta_{jk} + \frac{\partial\eta_j}{\partial b} (\Delta b)^2 \frac{\partial\eta_k}{\partial b} + \frac{\partial\eta_j}{\partial c} (\Delta c)^2 \frac{\partial\eta_k}{\partial c} + \dots \quad (95)$$

This shows that *for the construction of covariance matrices the evaluator needs* from the measurers

- root-mean-square errors for all error components,
- information about the data reduction so that he can calculate the sensitivity coefficients.

If he is told that the data reduction consisted in subtraction of a common background b and multiplication by a calibration factor c , i. e. $\eta_j = (a_j - b) \cdot c$, he has no difficulty to calculate $\langle \delta\eta_j \delta\eta_k \rangle \simeq \delta_{jk}(c\Delta a_j)^2 + (c\Delta b)^2 + \eta_j\eta_k(\Delta c)^2/c^2$. Obviously it is essential that experimentalists state clearly, and in sufficient detail, the various statistical and systematic error components when they report data, whereas they need not worry about correlations since these can be constructed readily from the error components, even for very large data sets (by computer, if necessary, see N. Larson 1986, 1992). For an instructive example of correlated data uncertainties and their impact on estimated parameters see the discussion of resonance energy standards by F. Perey (1978).

The merging of statistical and systematic errors into a total error is often considered as incorrect. Our equations show, however, that it is perfectly straightforward to *calculate*

the mean square total error as the sum of the mean square errors (variances) of all error components, statistical as well as systematic. The dividing line between the two error types is debatable anyway: One person's statistical error may well be another person's systematic error. The only problem with a sum of mean square errors is that by itself it does not reveal how much of it is statistical and how much systematic, i. e. how much correlation there may be. This question can only be answered if the physical meaning and the standard errors of all the error components are known.

A somewhat related criticism has been raised against least-squares combination of data obtained from different sources: Suppose a constant count rate, for example background noise, is measured repeatedly, with counts n_j registered during time intervals t_j . The σ_j^{-2} -weighted sample average, calculated with relative uncertainties equal to $1/\sqrt{n_j}$ as is appropriate for counts, turns out to be $\sum_j t_j / \sum_j (t_j^2/n_j)$. This does not look right and has been mistaken as evidence that there is something fundamentally wrong with the least-squares formalism which needs to be cleared up and remedied. The cause of the difficulty is, however, not the least-squares formalism but its employment in a situation where a non-Gaussian sample distribution – the Poisson distribution – is known to apply. There is no need here to replace unknown distributions by Gaussians via entropy maximisation or the central limit theorem and then use least squares. The correct posterior is a product of Poisson probabilities, Eq. 88, and of Jeffreys' prior for the rate constant. The resulting estimates, $\langle \lambda \rangle = \sum_j n_j / \sum_j t_j$ and $\Delta \lambda / \langle \lambda \rangle = 1/\sqrt{\sum_j n_j}$, involve the total number of counts and the total counting time in precisely the way common sense expects from Eqs. 90 and 91. The lesson is that *the least-squares method should not be used blindly if the available information admits a more rigorous treatment*. This is another illustration of the fundamental truth that the correct solution of a probabilistic problem requires *all* given information to be utilised.

The statement that experimentalists need not worry much about correlations should not be misunderstood. It does not mean that correlations are unimportant. It means only that they are not needed for the construction of covariance matrices *provided the data reduction procedure and all relevant root-mean-square errors are adequately documented*. The correlated output uncertainties of cross sections or cross section parameters, contained in the posterior covariance matrix $\langle \delta \mathbf{x} \delta \mathbf{x}^\dagger \rangle$, constitute highly relevant information for users of the data. The uncertainty of any function f of the cross sections x_μ , for example the calculated criticality of a nuclear reactor, is given in linear approximation by the square root of

$$\langle (\delta f)^2 \rangle = \sum_\mu \sum_\nu \frac{\partial f}{\partial x_\mu} \langle \delta x_\mu \delta x_\nu \rangle \frac{\partial f}{\partial x_\nu}, \quad (96)$$

where $\langle \delta x_\mu \delta x_\nu \rangle$ is an element of the covariance matrix, and where the derivatives or sensitivity coefficients are to be calculated with the best estimates $\langle x_\mu \rangle$. It is obvious that a good sensitivity study is not feasible without the covariance matrix. In the past it appeared repeatedly as if nuclear data were not accurate enough for certain applications when covariance information was ignored. When it was properly taken into account the accuracy turned out to be quite acceptable due to negative terms in the double sum from anticorrelated (compensating) errors. Hence those who extract cross section parameters from experimental data should not just state the parameters and their uncertainties, but also at least the more important elements of the covariance matrix.

We close the subsection by noting that the covariance matrix elements $C_{\mu\nu}$ are related to the standard errors $\Delta x_\mu = \sqrt{\text{var } x_\mu}$ and the correlation coefficients $\rho_{\mu\nu}$ as follows,

$$C_{\mu\nu} \equiv \langle \delta x_\mu \delta x_\nu \rangle = \Delta x_\mu \rho_{\mu\nu} \Delta x_\nu, \quad (97)$$

with $\rho_{\mu\mu} = 1$. The Schwarz inequality constrains the range of the correlation coefficients,

$$-1 \leq \rho_{\mu\nu} \equiv \frac{\text{cov}(x_\mu, x_\nu)}{\sqrt{\text{var } x_\mu} \sqrt{\text{var } x_\nu}} \leq +1. \quad (98)$$

Uncertainties can thus be stated in several equivalent ways:

1. as variances and covariances,
2. as relative variances and relative covariances,
3. as standard errors and correlation coefficients,
4. as relative standard errors and correlation coefficients.

By far the most economic and mnemotechnically safest way is the last one (at least if the data are not vanishingly small compared to the uncertainties). First, relative (percentage) standard errors are more easily grasped, remembered and compared than absolute ones, especially if the data vectors contain physically distinct quantities (e. g. resonance parameters and cross sections, or group cross sections and self-shielding factors). Second, relative standard errors and correlation coefficients have a clearcut intuitive meaning in contrast to variances and covariances. If I am told that the variances of x and y are 0.04 and 0.000009 and that their covariance is -0.000012 I am helpless. Only with the further information that the recommended values are $x = 10$ and $y = 0.1$ can I find out that the relative standard errors are 2% for x and 3% for y and that the correlation coefficient is -0.02. Had I been told the last three figures right away I would have understood immediately that both quantities are known with good accuracy and that their anticorrelation is negligible for most purposes.

It is hard to understand why the present ENDF-6 format (Rose and Dunford 1990) admits only variances and covariances. *A format extension to standard errors and correlation coefficients would enable evaluators and users to work with uncertainty files that are much easier to construct, read, and update, hence considerably less error-prone than the existing ones, without being bulkier.*

2.5. GOODNESS OF FIT

The quality and consistency of a least-squares fit is indicated by the posterior value of

$$\chi^2 \equiv (\boldsymbol{\zeta} - \mathbf{z})^\dagger \mathbf{C}^{-1} (\boldsymbol{\zeta} - \mathbf{z}). \quad (99)$$

The smaller the residues $|\zeta_j - z_j|$ are, the better is the fit, and the smaller is χ^2 . Inconsistencies, such as a wrong theoretical model (an overlooked resonance, for instance) or underestimated errors, tend to make χ^2 too large, while overestimated errors make it too small. In order to see what too large or too small means we consider the probability distribution of the variate χ^2 (in sample space, for given parameters). For a linear model $\mathbf{y}(\mathbf{x})$ it is easy to see that the distribution of χ^2 is a gamma distribution: The maximum entropy distribution of the data vector $\boldsymbol{\zeta}$, given the true vector \mathbf{z} , is

$$p(\boldsymbol{\zeta}|\mathbf{z}, \mathbf{C}) d(\boldsymbol{\zeta}) \propto \exp \left[-\frac{1}{2} (\boldsymbol{\zeta} - \mathbf{z})^\dagger \mathbf{C}^{-1} (\boldsymbol{\zeta} - \mathbf{z}) \right] d(\boldsymbol{\zeta}) = \exp \left(-\frac{\chi^2}{2} \right) d(\boldsymbol{\zeta}). \quad (100)$$

We simplify by an orthogonal transformation to the principal-axes system in which χ^2 is a sum of $N = M + J$ squares,

$$\chi^2 = \sum_{j=1}^N \left(\frac{\zeta'_j - z'_j}{\sigma'_j} \right)^2 \equiv \sum_j \chi_j^2. \quad (101)$$

The primes indicate the principal-axes system and the $\sigma_j'^2$ are the elements of the diagonal matrix $\mathbf{C}' = \mathbf{O}^\dagger \mathbf{C} \mathbf{O}$, with $\mathbf{O}^\dagger \mathbf{O} = \mathbf{1}$. The relative deviations χ_j can be considered as Cartesian coordinates in the space of the ζ_i . The corresponding volume element is invariant under the orthogonal transformation (rotation) \mathbf{O} ,

$$d(\zeta) = d(\zeta') \propto d(\chi) \propto \chi^{N-1} d\chi d\Omega. \quad (102)$$

In the last term we introduced polar coordinates with the radial coordinate χ , as suggested by Eq. 101. With this volume element it is trivial to integrate Eq. 100 over the angular coordinates Ω . The resulting radial distribution can be written as a χ^2 distribution with N degrees of freedom,

$$p(\chi^2 | \mathbf{z}, \mathbf{C}) d\chi^2 = \Gamma\left(\frac{N}{2}\right)^{-1} \exp\left(-\frac{\chi^2}{2}\right) \left(\frac{\chi^2}{2}\right)^{N/2-1} \frac{d\chi^2}{2},$$

$$0 < \chi^2 \equiv (\zeta - \mathbf{z})^\dagger \mathbf{C}^{-1} (\zeta - \mathbf{z}) < \infty. \quad (103)$$

Unfortunately we do not know the true vector observable $\mathbf{z} = \mathbf{z}(\mathbf{x})$. All we know is the estimate $\hat{\mathbf{z}} \equiv \mathbf{z}(\hat{\mathbf{x}})$, where $\hat{\mathbf{x}}$ satisfies the least-squares condition (76) or the normal equations (79) which we now write in more abbreviated form as

$$\hat{\mathbf{S}}^\dagger \mathbf{C}^{-1} (\zeta - \hat{\mathbf{z}}) = 0. \quad (104)$$

Because of this system of M equations not all ζ_j are mutually independent but only J of them. Taylor expansion about the estimate $\hat{\mathbf{x}}$ yields (compare Eqs. 78 and 80)

$$\chi^2 = \hat{\chi}^2 + (\mathbf{x} - \hat{\mathbf{x}})^\dagger (\hat{\mathbf{S}}^\dagger \mathbf{C}^{-1} \hat{\mathbf{S}}) (\mathbf{x} - \hat{\mathbf{x}}), \quad (105)$$

where

$$\hat{\chi}^2 = \sum_{j=1}^N \left(\frac{\zeta_j' - \hat{z}_j'}{\sigma_j'} \right)^2 \equiv \sum_j \hat{\chi}_j^2. \quad (106)$$

For a linear theory one has $\mathbf{z}' = \hat{\mathbf{z}}' + \mathbf{O}^\dagger \mathbf{S} (\mathbf{x} - \hat{\mathbf{x}})$ and therefore

$$\chi_j = \hat{\chi}_j + \frac{1}{\sigma_j'} \sum_{k,\mu} O_{jk}^\dagger S_{k\mu} (x_\mu - \hat{x}_\mu). \quad (107)$$

One can now replace the integration variables χ_j by $\hat{\chi}_j$ for $j = 1, 2, \dots, J$ and by x_μ for $j = J+1, J+2, \dots, N$. The Jacobian for this substitution is constant because of the linear relationship (107) between old and new variables, so that

$$d(\zeta) \propto d(\chi) \propto \hat{\chi}^{J-1} d\hat{\chi} d\hat{\Omega} d^M x. \quad (108)$$

Integrating (100) over all angular coordinates and all x_μ one finds finally the distribution of the sum of square deviations $\hat{\chi}^2$ that can be calculated from known quantities,

$$p(\hat{\chi}^2 | \hat{\mathbf{z}}, \mathbf{C}) d\hat{\chi}^2 = \Gamma\left(\frac{J}{2}\right)^{-1} \exp\left(-\frac{\hat{\chi}^2}{2}\right) \left(\frac{\hat{\chi}^2}{2}\right)^{J/2-1} \frac{d\hat{\chi}^2}{2},$$

$$0 < \hat{\chi}^2 \equiv (\zeta - \hat{\mathbf{z}})^\dagger \mathbf{C}^{-1} (\zeta - \hat{\mathbf{z}}) < \infty. \quad (109)$$

This chi-square distribution with J degrees of freedom is broader than the chi-square distribution (103) with $J + M$ degrees of freedom. Its mean and variance are

$$\langle \hat{\chi}^2 \rangle = J, \quad (110) \quad \text{var } \hat{\chi}^2 = 2J. \quad (111)$$

This result is exact for linear theoretical models $\mathbf{z}(\mathbf{x})$, for nonlinear models it is valid in saddle point (Laplace) approximation. It certainly applies whenever iterative least-squares fitting (see Sect. 2.2) works. The last two equations show that a good fit is characterised by a chi-square value close to the number J of data points, roughly between $J - \sqrt{2J}$ and $J + \sqrt{2J}$. The statement that a fit “yielded a chi-square of 1031 for 1024 data points” would therefore indicate good consistency between input data and theoretical model. A much higher chi-square could point to underestimated uncertainties, and it is fairly common practice in this case to multiply all input uncertainties (i. e. the matrix elements of \mathbf{C}) by a common scale factor so as to enforce $\chi^2 = J$ (and to adjust the posterior parameter covariance matrix accordingly). This is dangerous, however, because a high value of chi-square could also be due to an inadequate theoretical model (incorrectly assigned level spins or overlooked weak resonances, for instance, in resonance analysis). A suspiciously small chi-square, on the other hand, could indicate too conservative uncertainty estimates, but it could also be due to tampering with the data such as elimination of data (“outliers”) which seem to contradict a favourite theory. *Rescaling of the uncertainties should therefore be considered only after one has made sure that there is nothing obviously wrong with the model or the input.*

In “primitive” least-squares fitting, where prior information is neglected, the same argumentation leads to the widely used result that $\hat{\chi}^2 = (\boldsymbol{\eta} - \hat{\mathbf{y}})^\dagger \mathbf{B}^{-1} (\boldsymbol{\eta} - \hat{\mathbf{y}})$ obeys a chi-square distribution with $J - M$ degrees of freedom so that $\hat{\chi}^2$ is expected roughly between $(J - M) - \sqrt{2(J - M)}$ and $(J - M) + \sqrt{2(J - M)}$. The statement of the goodness of fit must now include both the number of measured data points and of parameters. A “chi-square of 1031 for 1024 data points and 20 adjusted parameters (1004 degrees of freedom)” would indicate reasonable consistency.

2.6. INCONSISTENT DATA

One of the thorniest problems in data evaluation is that of inconsistent data. Suppose we are given the results of n completely independent and experimentally different measurements of the same physical quantity, μ , in the form $x_j \pm \sigma_j$, $j = 1, 2, \dots, n$. If the separation of any two values, $|x_j - x_k|$, is smaller or at least not much larger than the sum of the corresponding uncertainties, $\sigma_j + \sigma_k$, the data are said to be consistent or to agree “within error bars”. (The probability that two equally precise measurements yield a separation greater than $\sigma_j + \sigma_k = 2\sigma$ is only $\text{erfc } 1 \simeq 0.157$ for two Gaussian sampling distributions with the same standard deviation σ). If some or all separations are much larger, the data are not consistent with the stated uncertainties. Inconsistencies are caused by unrecognised or malcorrected experimental effects such as backgrounds, dead time of the counting electronics, instrumental resolution, sample impurities, calibration errors, etc. As already mentioned, a popular quick fix consists in increasing all input errors by a common factor until chi-square has the expected value (and all error bars overlap). This, however, does not change the relative weights, hence there is the same penalty for overoptimistic as for conservative uncertainty assignments. The following Bayesian treatment provides more justice. It considers not only the claimed uncertainties but also how far a given datum is from the average.

What can we say about unrecognised errors? If we know only the data but not how they were measured, positive and negative errors are equally probable, hence the probability distribution for the unrecognised error ϵ_j of the j -th experiment should be symmetric

about zero, and the same distribution should apply to all measurements. Let us therefore assume, in accordance with the maximum entropy principle, Gaussian distributions for all ϵ_j ,

$$p(\epsilon_j|\tau_j) d\epsilon_j = \frac{1}{\sqrt{2\pi\tau_j^2}} \exp\left[-\frac{1}{2}\left(\frac{\epsilon_j}{\tau_j}\right)^2\right] d\epsilon_j, \quad -\infty < \epsilon_j < \infty. \quad (112)$$

The probability to measure the value x_j , given the true value μ , the unrecognised error ϵ_j , and the uncertainty σ_j due to all recognised error sources, is then given by

$$p(x_j|\mu, \sigma_j, \epsilon_j) dx_j = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{1}{2}\left(\frac{x_j - \mu - \epsilon_j}{\sigma_j}\right)^2\right] dx_j, \quad -\infty < x_j < \infty. \quad (113)$$

If the dispersions τ_j of unrecognised errors are known, the joint posterior distribution for μ and the ϵ_j is

$$p(\mu, \epsilon|\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) d\mu d(\epsilon) \propto d\mu \prod_{j=1}^n d\epsilon_j \exp\left[-\frac{(x_j - \mu - \epsilon_j)^2}{2\sigma_j^2} - \frac{\epsilon_j^2}{2\tau_j^2}\right]. \quad (114)$$

(ϵ , $\boldsymbol{\sigma}$, $\boldsymbol{\tau}$ on the left-hand side are to be understood as vectors in sample space, with coordinates ϵ_j , σ_j , τ_j .) Completing squares in the exponent we can easily integrate over the ϵ_j . The resulting posterior distribution for μ is a Gaussian,

$$p(\mu|\bar{x}, s) d\mu = \frac{1}{\sqrt{2\pi(\sigma^2 + \tau^2)/n}} \exp\left[-\frac{(\mu - \bar{x})^2}{2(\sigma^2 + \tau^2)/n}\right] d\mu, \quad -\infty < \mu < \infty, \quad (115)$$

with

$$\langle \mu \rangle = \bar{x}, \quad (116) \quad \text{var } \mu = \frac{1}{n(\overline{\sigma^2 + \tau^2})} \quad (117)$$

where the overbars denote averages over j (over measurements) with weights $1/(\sigma_j^2 + \tau_j^2)$. Integrating (114) over μ we find the joint distribution of the unrecognised errors,

$$p(\epsilon|\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) d(\epsilon) \propto \exp\left[-\frac{1}{2}(\epsilon - \mathbf{x})^\dagger \mathbf{A}^{-1}(\epsilon - \mathbf{x}) - \frac{1}{2}\epsilon^\dagger \mathbf{B}^{-1}\epsilon\right] d(\epsilon), \quad (118)$$

where \mathbf{A}^{-1} and \mathbf{B}^{-1} are positive definite, symmetric matrices defined by

$$(\mathbf{A}^{-1})_{jk} \equiv \sigma_j^{-2}\delta_{jk} - \frac{\sigma_j^{-2}\sigma_k^{-2}}{\sum_l \sigma_l^{-2}}, \quad (119) \quad (\mathbf{B}^{-1})_{jk} \equiv \tau_j^{-2}\delta_{jk}. \quad (120)$$

This product of two multivariate Gaussians is a multivariate Gaussian again, with mean vector $\langle \epsilon \rangle = \mathbf{C}\mathbf{A}^{-1}\mathbf{x}$ and inverse covariance matrix $\mathbf{C}^{-1} = \mathbf{A}^{-1} + \mathbf{B}^{-1}$, so that

$$(\mathbf{A}^{-1} + \mathbf{B}^{-1})\langle \epsilon \rangle = \mathbf{A}^{-1}\mathbf{x}. \quad (121)$$

Solving the last equation for $\langle \epsilon_j \rangle$ one obtains

$$\langle \epsilon_j \rangle = \frac{\tau_j^2}{\sigma_j^2 + \tau_j^2}(x_j - \bar{x}). \quad (122)$$

The best estimate of ϵ_j is thus the deviation $x_j - \bar{x}$ of the j -th datum from the (weighted) mean \bar{x} , multiplied by a “shrinking factor” $\tau_j^2/(\sigma_j^2 + \tau_j^2)$ that is close to zero if the

expected unrecognised error is much smaller than the known uncertainty σ_j , and close to unity if it is much larger. This, however, is trivial: If we know the variances τ_j^2 of the unrecognised errors then these are not really unrecognised. We know as much about them as about the other errors. We can therefore add variances as usual to get the total mean square errors $\sigma_j^2 + \tau_j^2$. Their reciprocals are expected to appear as weights in all j -averages, and this is in fact what we have just found.

The simplest nontrivial case is obtained if the τ_j can be considered not as the actual root-mean-square unrecognised errors but merely as their estimates, based for instance on the general quality of the various measurements, on the accuracy of the techniques employed, perhaps even on the credibility of the experimentalists as judged from their past record. (Note that it is perfectly all right to put $\tau_j = 0$ for those experiments that can be considered as unaffected by unrecognised errors). The unknown true variance can then be taken as τ_j^2/c , where c is an adjustable common scale parameter with prior $p(c)dc$, and the joint probability for μ and the vector ϵ as

$$p(\mu, \epsilon | \mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) d\mu d(\epsilon) \propto d\mu \int_0^\infty dc p(c) \prod_{j=1}^n c^{1/2} d\epsilon_j \exp \left[-\frac{(x_j - \mu - \epsilon_j)^2}{2\sigma_j^2} - \frac{c\epsilon_j^2}{2\tau_j^2} \right]. \quad (123)$$

Integrating over all ϵ_j one gets the posterior distribution of μ ,

$$p(\mu | \mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) d\mu \propto d\mu \int_0^\infty dc p(c) \prod_{j=1}^n (\sigma_j^2 + \tau_j^2/c)^{-1/2} \exp \left[-\frac{1}{2} \frac{(\mu - \bar{x})^2}{(\sigma_j^2 + \tau_j^2/c)} \right]. \quad (124)$$

If we have no numerical information at all about the scale parameter c Jeffreys' prior dc/c appears appropriate. The integration over c is then easy if the known uncertainties are unimportant. With $\sigma_j = 0$ for all j the integrand becomes essentially a gamma distribution of c , integration over which yields Student's t-distribution,

$$p(\mu | \bar{x}, s') d\mu = \frac{du}{B(\frac{1}{2}, \frac{n-1}{2}) (1+u^2)^{n/2}}, \quad -\infty < u \equiv \frac{\mu - \bar{x}}{s'} < \infty, \quad (125)$$

with

$$\langle \mu \rangle = \bar{x}, \quad (126) \quad \text{var } \mu = \frac{s'^2}{n-3}, \quad (127)$$

where $s'^2 \equiv \overline{x^2} - \bar{x}^2$, the overbars denoting sample averages weighted by τ_j^{-2} . Thus the uncertainty of μ in this extreme case is equal to the sample variance s'^2 , determined by the scatter of the data x_j (sometimes called the "external error"). This is, of course, just what we found when we discussed estimation of μ from a sample drawn from a Gaussian with unknown standard deviation (Eq. 34). For large n the distribution of μ is practically Gaussian.

In the general case, $\sigma_j > 0$, $\tau_j > 0$, the integral (124) with Jeffreys' least informative prior $p(c)dc \propto dc/c$ diverges logarithmically because the integrand becomes proportional to $1/c$ for $c \rightarrow 0$. The Bayesian formalism signals in this way that the prior information is not sufficient for definite predictions. Is there anything we know in addition to the fact that c is a scale parameter? Actually, if the τ_j are our best estimates of the uncertainties caused by unrecognised errors, we expect c to be close to unity. The maximum-entropy prior constrained by $\langle c \rangle = 1$ is

$$p(c) dc = e^{-c} dc, \quad 0 < c < \infty. \quad (128)$$

This is almost as noncommittal as Jeffreys' prior, decreasing also monotonically as c increases, but normalisable and giving less weight to the extrema. With this prior both the c -integral and the normalisation constant of the posterior μ -distribution (124) are finite and can be calculated numerically without difficulty. Fig. 8 shows a real-life example, the posterior distribution of the ^{239}Pu fission cross section for 14.7 MeV neutrons, together with the Gaussian distributions representing the six experimental results listed in Table I. Prior uncertainties of $\tau_j = 0.1$ b were assigned to all experiments indiscriminately, based on the state of the art. The posterior mean and the root-mean-square uncertainty, computed numerically, are also given in Table I.

2.7. ESTIMATION OF UNKNOWN SYSTEMATIC ERRORS

What can we learn about the unrecognised systematic errors ϵ_j from the set of inconsistent data, $x_j \pm \sigma_j$, $j = 1, \dots, n$? With the prior (128) it is easy to integrate the posterior probability distribution (124) first over the gamma distribution of c , then over the Gaussian distribution of μ . The result can be written in the form

$$p(\epsilon|\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) d(\epsilon) \propto \exp\left[-\frac{1}{2}(\epsilon - \mathbf{x})^\dagger \mathbf{A}^{-1}(\epsilon - \mathbf{x})\right] \left[1 + \frac{1}{2}\epsilon^\dagger \mathbf{B}^{-1}\epsilon\right]^{-n/2-1} d(\epsilon), \quad (129)$$

with the matrices \mathbf{A} and \mathbf{B} defined as before, eqs. 119 and 120. In order to get the mean vector and the covariance matrix of this distribution in analytic form we employ saddle point integration, replacing the exact distribution by a multivariate Gaussian with the same maximum and the same curvature at the maximum,

$$p(\epsilon|\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\tau}) \equiv \exp[-F(\epsilon)] \simeq \exp[-F(\hat{\epsilon}) - \frac{1}{2}(\epsilon - \hat{\epsilon})^\dagger (\nabla\nabla^\dagger F)_{\epsilon=\hat{\epsilon}}(\epsilon - \hat{\epsilon})], \quad (130)$$

where the nabla operator has coordinates $\nabla_i \equiv \partial/\partial\epsilon_j$. Thus we have

$$\langle \epsilon \rangle \simeq \hat{\epsilon}, \quad (131) \quad \langle \delta\epsilon \delta\epsilon^\dagger \rangle \simeq [(\nabla\nabla^\dagger F)_{\epsilon=\hat{\epsilon}}]^{-1}. \quad (132)$$

The most probable vector $\hat{\epsilon}$ must be found by solving the equation

$$\nabla F = \mathbf{A}^{-1}(\epsilon - \mathbf{x}) + \frac{n+2}{2} \frac{\mathbf{B}^{-1}\epsilon}{1 + \epsilon^\dagger \mathbf{B}^{-1}\epsilon/2} = 0 \quad (133)$$

and the approximate covariance matrix as the inverse of

$$\nabla\nabla^\dagger F = \mathbf{A}^{-1} + \frac{n+2}{2} \frac{\mathbf{B}^{-1}(1 + \epsilon^\dagger \mathbf{B}^{-1}\epsilon/2) - \mathbf{B}^{-1}\epsilon \epsilon^\dagger \mathbf{B}^{-1}}{(1 + \epsilon^\dagger \mathbf{B}^{-1}\epsilon/2)^2}, \quad (134)$$

evaluated at $\epsilon = \hat{\epsilon}$. With the definitions (119) and (120) of \mathbf{A} and \mathbf{B} we obtain

$$\hat{\epsilon}_j = x_j - \bar{x} + \bar{\epsilon} - \frac{n+2}{2} \frac{\tau_j^{-2}\sigma_j^2\hat{\epsilon}_j}{1 + \sum_k \tau_k^{-2}\hat{\epsilon}_k^2/2}, \quad (135)$$

where \bar{x} and $\bar{\epsilon}$ are σ_j^{-2} -weighted averages. This is suitable for iteration. Inserting $\hat{\epsilon}_j \simeq x_j - \bar{x}$ for all j as first estimate on the right-hand side one finds the second approximation,

$$\hat{\epsilon}_j \simeq \left[1 - \frac{n+2}{2} \frac{\tau_j^{-2}\sigma_j^2}{1 + \sum_k \tau_k^{-2}(x_k - \bar{x})^2}\right] (x_j - \bar{x}), \quad (136)$$

and by repetition better and better approximations. Our treatment of unrecognised systematic errors is an example of the “hierarchical” (here: two-stage) Bayesian method. It involves repeated application of Bayes’ theorem: The sampling distribution (113) depends on parameters ϵ_j with the prior (112) that, in turn, depends on the “hyperparameter” c with the “hyperprior” (128) if we replace τ_j^2 by τ_j^2/c .

The final estimates and their uncertainties (square roots of the diagonal elements of the matrix $\langle \delta\epsilon \delta\epsilon^\dagger \rangle$) for our ^{239}Pu problem, obtained in this way, are listed in the last column of Table I. No significant unrecognised errors are found for the measurements 1, 5 and 6, whereas 2, 4 and perhaps 3 seem to be affected by unrecognised errors that are of the same order of magnitude as the uncertainties stated by the authors. Of course, these conclusions could have been drawn already from the experimental data (and especially from the pictorial representation by Gaussians in Fig. 8), but our formalism provides quantitative support for our common sense also under less obvious circumstances. Furthermore, it sheds some light on a much advertised recent estimation method:

Our second approximation (136) resembles the James-Stein estimators (Stein 1956) that, since their introduction, have caused a great deal of excitement, confusion and a flood of papers. C. Stein showed, using the frequentist definition of risk (square error averaged over all possible samples, for a given set of parameters) that estimators similar to (136) have sometimes lower risk than the estimates resulting from Bayesian estimation under quadratic loss (that minimise the square error averaged over all possible parameters, for the sample at hand). Many improvements have been suggested to Stein’s original estimators, based on frequentist distribution theory and more or less educated guesswork. For instance, the “plus rule” estimator

$$\epsilon_j^* \simeq \left[1 - \frac{n-2}{n} \frac{\sigma^2}{s'^2} \right]_+ (x_j - \bar{x}), \quad (137)$$

was proposed for the special case of equal uncertainties ($\sigma_j = \sigma$ for all j). The subscript + means that only positive values of the “shrinking factor” are accepted. For negative values one must put $\epsilon_j^* = 0$. Moreover, the “plus rule” estimator is to be used only for $n \geq 3$. Wild discussions arose about the “paradox” that the estimator for ϵ_j depends not only on x_j but on all the other, independently sampled x_k , $k \neq j$. The question was raised whether inclusion of other unrelated data would not improve the estimate. (“If I want to estimate tea consumption in Taiwan will I do better to estimate simultaneously the speed of light and the weight of hogs in Montana?” (Efron and Morris 1973). So-called parametric empirical Bayes recipes seemed to offer some insight, e. g. replacement of τ^2 in Eq. 121 (in the case $\tau_j = \tau$ for all j) by the sample variance s'^2 . However plausible such recipes may seem (see e.g. Berger 1985), without rigorous justification they remain ad-hockeries.

Under the same circumstances ($\sigma_j = \sigma$, $\tau_j = \tau$) our Eq. 136 yields

$$\langle \epsilon_j \rangle \simeq \left[1 - \frac{(n+2)\sigma^2}{ns'^2 + 2\tau^2} \right] (x_j - \bar{x}), \quad (138)$$

valid for all $n \geq 2$, without discontinuities or questions of interpretation. A paradox exists only for frequentists who deny themselves the use of priors. For a Bayesian there is no paradox. He or she encodes the information that the data are all of the same type, measured in related and hence comparable experiments, in a second-stage prior which induces correlations and “shrinking” factors in a natural way. Speed of light data are unrelated, hence excluded. From the Bayesian viewpoint, on the other hand, it seems odd

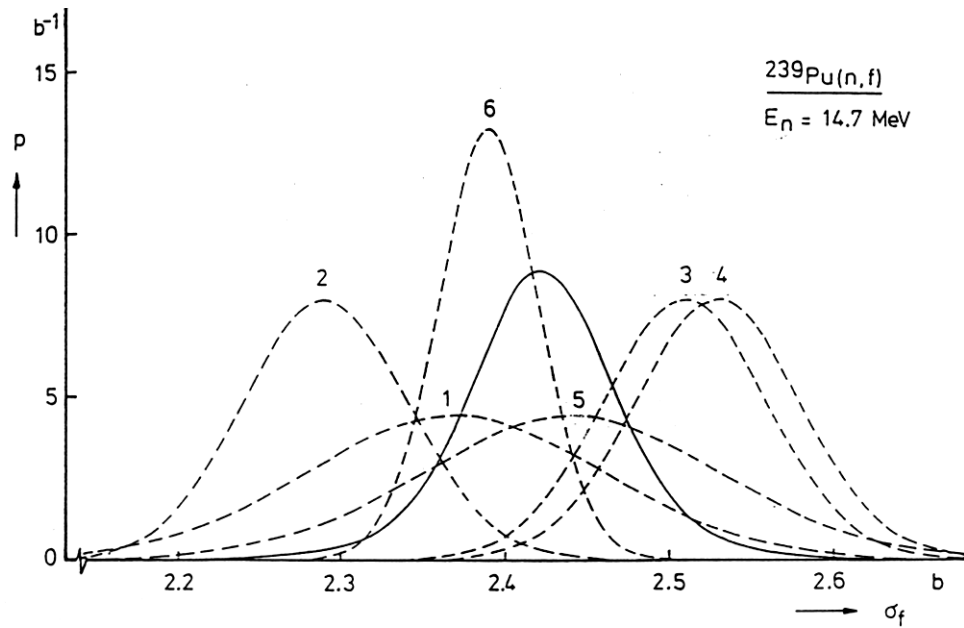


Fig. 8. Probability densities representing the experimental results in Table I (dashed Gaussians), and posterior density of the true value (solid curve) estimated with the two-stage Bayesian model, Eq. 124, with hyperprior Eq. 128. Inconsistencies are evident between the experiments 2, 3 and 4 (curve labels correspond to the first column of the table).

Table 1. Experimental input and estimation results for $^{239}\text{Pu}(n,f)$ at 14.7 MeV incident neutron energy

| No. | Authors | Year | measured fission cross section (barn) | estimated unrecognised error (barn) |
|-----|------------------|------|--|--|
| 1 | Kari | 1978 | $2.37 \pm .09$ | $-.019 \pm .056$ |
| 2 | Cancé et al. | 1978 | $2.29 \pm .05$ | $-.086 \pm .050$ |
| 3 | Adamov et al. | 1979 | $2.51 \pm .05$ | $.056 \pm .048$ |
| 4 | Li et al. | 1982 | $2.53 \pm .05$ | $.069 \pm .049$ |
| 5 | Mahdawi et al. | 1982 | $2.44 \pm .09$ | $.006 \pm .056$ |
| 6 | Arlt et al. | 1983 | $2.39 \pm .03$ | $-.027 \pm .041$ |
| | best estimate: | | $2.42 \pm .05$ | |
| | prior estimates: | | $\tau_j = 0.1 \text{ b}$ | for all systematic errors |

to base the risk criterion not on the given sample and on prior knowledge, but on the given sample and all other unobserved samples that can be imagined, without utilisation of prior knowledge. Frequentist claims about the superiority of estimators “in the long

run” (averaged over many samples) are not very relevant for data evaluation, where one must infer best values (for quadratic or any other loss) from one available sample. It may be true that an estimator with low risk is closer to the true value for a larger fraction of all possible samples than an estimator which ensures minimal quadratic loss, but for the remaining fraction of samples the errors tend to be so much larger that the apparent advantage turns into a net disadvantage (see Jaynes 1976). In any case, the Bayesian two-stage method yields, in second saddle point approximation, estimators which are similar to, but especially for small samples better than, James-Stein estimators. Moreover, iteration yields all the possible improvement and also the uncertainties in a systematic and unambiguous way, without the bizarre discontinuities (Efron and Morris 1973) of many improved James-Stein estimators. *The Bayesian method leaves no room for guesswork once loss function, statistical model and priors are specified.*

Obviously, the correct interpretation of data inconsistencies is never easy, and the old saying that “data evaluation is more an art than a science” remains true to a certain extent, in spite of much progress toward formalisation and quantification during the last three decades.

3. Resonance Theory for the Resolved Region

If nuclei of a given species, e. g. ^{235}U , are bombarded by neutrons, one observes nuclear reactions such as elastic scattering, radiative capture, or fission. The probabilities for those (n,n), (n, γ), or (n,f) processes, customarily expressed as cross sections in units of barn ($1 \text{ b} = 10^{-24} \text{ cm}^2$), depend sensitively on the energies of the incident neutrons. The scattering cross section, for instance, is mostly close to the geometrical cross section of the nuclei (several barns) but at certain energies it rises suddenly by several orders of magnitude. Similar resonance behaviour, at the same energies, is exhibited by the capture and fission cross sections. Fig. 9 (top and middle) shows this behaviour for the nucleus ^{238}U for which elastic scattering and radiative capture are the only energetically allowed neutron reactions at low energies. (Subthreshold fission is negligibly small.) For ^{235}U one would see resonances also in fission in this energy range. Each of these resonances is due to excitation of a relatively long-lived (quasi-stationary) state of the compound nucleus that is formed temporarily if a neutron interacts with a target nucleus. Note the different peak shapes: The resonances in the capture cross section are symmetric whereas those in the scattering cross section are asymmetric with pronounced minima and sizable "potential" scattering between resonances.

The impact of the resonances on the neutron spectrum in a power reactor is shown in Fig. 9 (bottom). The conspicuous dips in the neutron flux coincide with the resonance peaks in the cross sections. The explanation is simple: Neutrons cannot survive long at energies where ^{238}U , the main fuel constituent, has high cross sections, because there they are soon captured (removed completely) or scattered (transferred to some other, usually lower) energy. As a result the flux is depleted at the ^{238}U resonances. Smaller dips in the flux are due to other, less abundant fuel components such as the fissile ^{235}U .

At low energies the resonances appear fairly well separated but as the energy increases their spacings decrease while their widths increase. Eventually they overlap so much that the compound resonance structure is averaged out and only much broader structures survive, such as the size (or single-particle) resonances described by the optical model, or the giant dipole resonances observed in photonuclear reactions. As a rule only the resonances at relatively low energies can be observed directly. At intermediate energies they are not fully resolved because of limited instrumental resolution, although the real disappearance of compound resonance structure due to excessive level overlap occurs only at still higher energies. Thus one distinguishes the resolved resonance region from the unresolved (or at best partially resolved) resonance region.

The more nucleons belong to the compound system the finer is the resonance structure. Typical level spacings observed in neutron reactions are of the order

| | |
|-----|--------------------|
| MeV | for light, |
| keV | for medium-weight, |
| eV | for heavy nuclei. |

The level spacings of target nuclei with even nucleon number are generally larger than those of nuclei with odd nucleon number. Magic or near-magic nuclei have untypically large level spacings. The heavy, doubly-magic nucleus ^{208}Pb , for instance, has level spacings resembling those of light nuclei.

Thermal motion of the target nuclei causes Doppler broadening of the resonance peaks observed in the laboratory system: As the target temperature increases, the peaks become broader while their areas remain practically constant. This changes the average scattering, capture and fission rates and the whole neutron balance in a fission reactor. As a

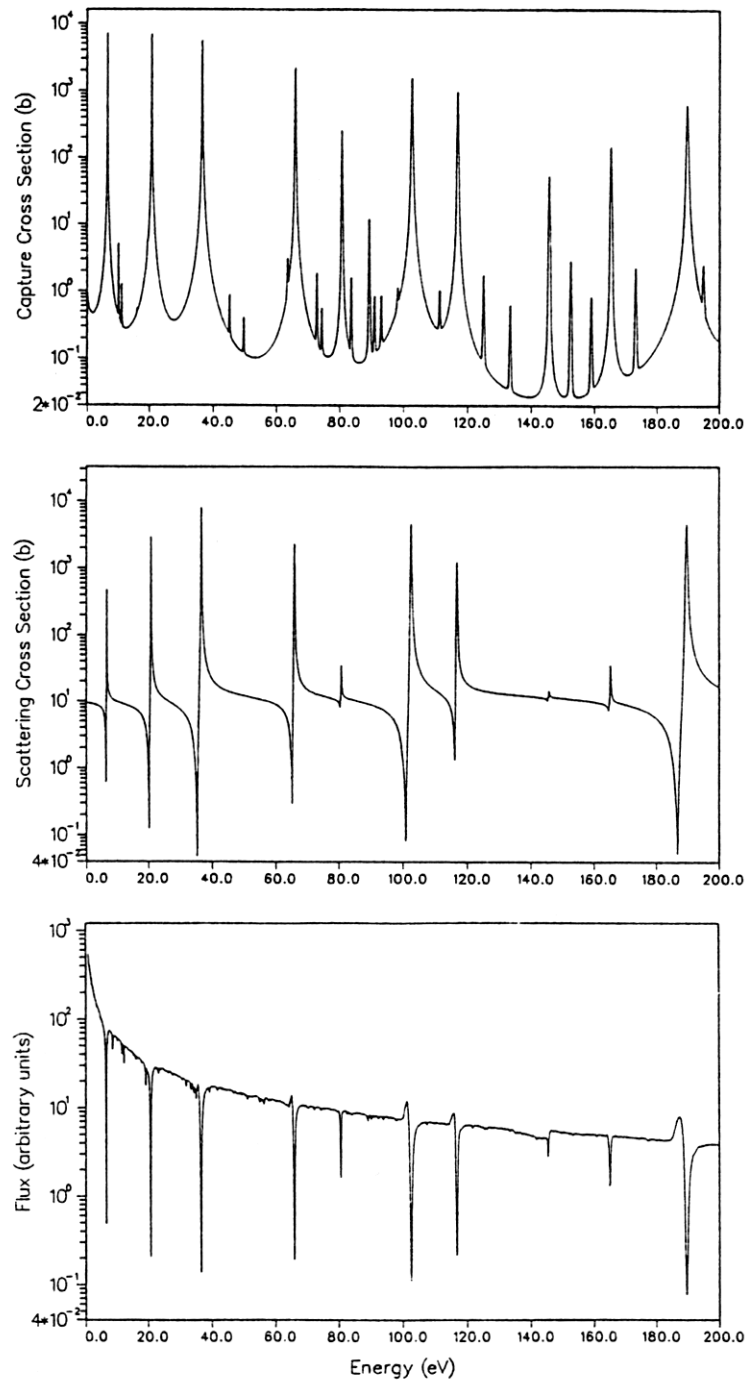


Fig. 9. Top: ^{238}U neutron capture cross section below 200 eV, Doppler broadened to 300 K. Middle: ^{238}U neutron scattering cross section, broadened to 300 K. Bottom: Neutron spectrum in an advanced pressurised-water reactor (C. Broeders, FZK, private communication). Note the logarithmic ordinate scales. Energies are given in the laboratory system.

consequence the safety characteristics of the various fission reactor designs depend crucially on the cross section resonances of the main fuel constituents and in particular on their Doppler broadening. The Doppler effect is the only natural phenomenon that promptly counteracts a sudden power excursion in a fission reactor. Thermal expansion has the same tendency but is much slower. Quite generally one demands that the temperature rise accompanying a power excursion must result in less neutrons produced per neutron absorbed so that the fission chain reaction does not get out of hand. In more technical terms the prompt Doppler reactivity coefficient must be negative.

In shielding applications the minima displayed by the scattering and hence also by the total cross section (the sum of all partial cross sections for scattering, capture, fission etc.) provide dangerous energy "windows" for neutrons. In fusion reactor designs such as ITER (International Thermonuclear Experimental Reactor), at present on the drawing boards, steel shielding is foreseen for the superconducting magnet coils. The windows in the total cross sections of the main steel components limit the efficiency of the shielding significantly (see Fig. 10 for ^{56}Fe).

These examples should suffice to illustrate the importance of cross section resonances in nuclear science and technology. Resolved resonances are described most conveniently by R-matrix theory. It attained its standard form with the comprehensive review article of Lane and Thomas (1958). This article is required reading for each specialist in the field. Briefly the principles of R-matrix theory are as follows. All collisions are considered as binary, an ingoing wave function describing the two incident particles, an outgoing wave function the two emerging reaction products. The incident particles could be, for instance, a neutron and a ^{235}U nucleus, the reaction products could be two fission fragments, or an excited ^{236}U nucleus and a photon, or a ^{235}U nucleus in its ground state and an elastically scattered neutron. Since the nuclear forces have short range but are not well understood otherwise, one divides configuration space into (i) an external region, where nuclear forces are negligible so that the well known wave functions for free or at most electromagnetically interacting particles can be used, and (ii) an internal region, where nuclear forces predominate. Although the internal wave function is unknown, it can at least be written as a formal expansion in terms of the eigenfunctions of an eigenvalue problem. This eigenvalue problem is defined by the (nonrelativistic) Schrödinger equation with prescribed logarithmic derivatives of the eigenfunctions at the boundary between the two regions. Matching external and internal wave functions at the boundary, and demanding finite probabilities everywhere, one finds that for a given ingoing wave all outgoing waves, and hence all cross sections, are parametrised by the eigenvalues and eigenvector components of the problem. These can be identified with the energies and decay amplitudes of the quasi-stationary compound states. All this will be discussed in detail below.

Although the principles of resonance theory are quite simple, the general expressions can look rather formidable to the beginner. We cannot give full derivations (those can be found in the review paper of Lane and Thomas 1958), and the basic quantum-mechanical collision theory (see e. g. Lynn 1968) will be assumed to be known, but we shall try to present the formalism in all the detail that is needed for applications in science and technology. The practically important variants are

- the Blatt-Biedenharn formalism
- " single-level Breit-Wigner approximation (SLBW)
- " multi-level " - " " (MLBW)
- " (multi-level) Adler-Adler "
- " (multi-level) Reich-Moore "

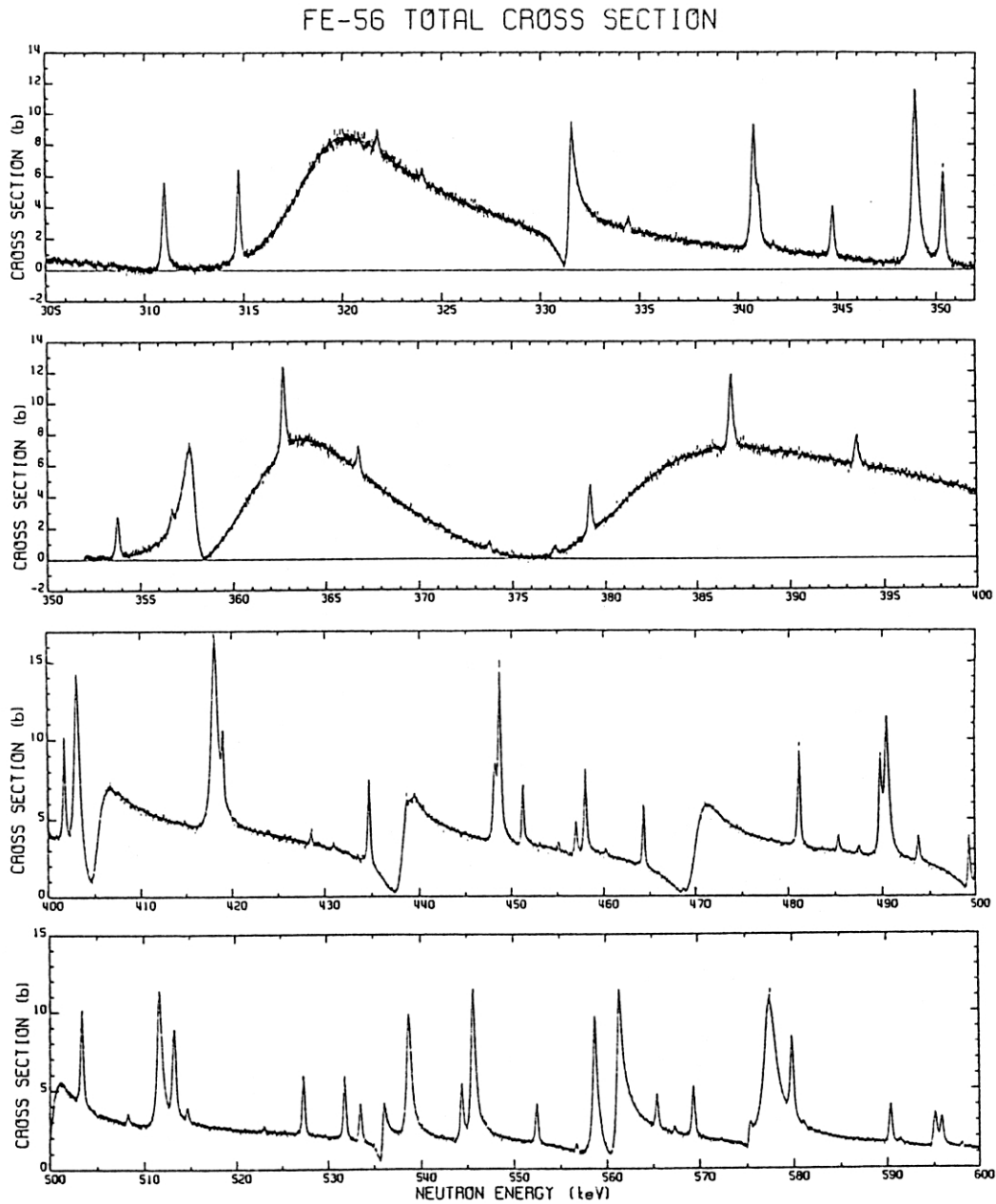


Fig. 10. Total cross section data for $^{56}\text{Fe}+n$ from 305 to 600 keV (bars) and resonance-theoretical fit (solid line, barely visible) with the one-channel Reich-Moore formalism of R-matrix theory. The broad asymmetric peaks are s-wave resonances, the narrow symmetric peaks are p- and d-wave resonances. The data were obtained in a time-of-flight transmission measurement with 201.6 m flight path at the Oak Ridge Electron Linear Accelerator. (From Perey et al. 1990)

The first one is quite general. It shows how cross sections can be expressed in terms of the unitary, symmetric collision matrix (S matrix) with special emphasis on angular distributions and on particle spins. It can be combined with any of the other four which provide different approximations to the collision matrix.

3.1. THE BLATT-BIEDENHARN FORMALISM

Our notation in this and the following subsections will be basically that of Lane and Thomas (1958). We recall that in nuclear reaction theory one talks about reaction channels. Each channel is fully specified by

| | |
|------------|---|
| α , | the partition of the compound system into reaction partners, |
| | e. g. $^{235}\text{U} + \text{n}$ or $^{236}\text{U} + \gamma$ (both involving the same compound nucleus) |
| J , | the total angular momentum in units of \hbar , |
| ℓ , | " orbital " " " " " \hbar , |
| s , | " channel spin " " " " \hbar . |

The angular momenta satisfy the quantum-mechanical triangle relations

$$\vec{J} = \vec{\ell} + \vec{s}, \quad \text{i.e.} \quad |\ell - s| \leq J \leq \ell + s, \quad (139)$$

$$\vec{s} = \vec{I} + \vec{i}, \quad \text{i.e.} \quad |I - i| \leq s \leq I + i, \quad (140)$$

where \vec{I} and \vec{i} are the spins (in units of \hbar) of the two collision partners. Total energy, total angular momentum and (for all practical purposes) parity are conserved in nuclear reactions.

We further remember that for spinless, neutral particles one can solve the nonrelativistic Schrödinger equation for the boundary condition "stationary ingoing plane wave + stationary outgoing spherical wave" with the result

$$d\sigma_{\alpha\alpha} = \pi\lambda_{\alpha}^2 \left| \sum_{\ell=0}^{\infty} (2\ell + 1)(1 - U_{\ell})P_{\ell}(\cos\vartheta) \right|^2 \frac{d\Omega}{4\pi} \quad (141)$$

for the differential elastic-scattering cross section. The de Broglie wave length $2\pi\lambda_{\alpha} = \hbar/(\mu_{\alpha}v_{rel})$ corresponds to the relative motion of the collision partners, with reduced mass μ_{α} and relative speed v_{rel} . The angular-momentum eigenfunctions P_{ℓ} are Legendre polynomials of order ℓ . The sum terms with $\ell = 0, 1, 2, \dots$ are said to belong to the s-, p-, d-, f-... wave, a historical nomenclature taken over from atomic spectroscopy (where it refers to the so-called sharp, principal, diffuse, fundamental series of spectral lines). The collision function U_{ℓ} describes the modification of the ℓ -th outgoing partial wave relative to the case without interaction. Its absolute value gives the reduction in amplitude, its argument the phase shift caused by the interaction. With $P_{\ell}P_{\ell'} = (\ell\ell'00, L0)^2 P_L$, where $(\ell\ell'00, L0)$ is a Clebsch-Gordan coefficient vanishing unless $|\ell - \ell'| \leq L \leq \ell + \ell'$ and $(-)^{\ell+\ell'} = (-)^L$, one can write the differential cross section as a linear expansion in Legendre polynomials,

$$d\sigma_{\alpha\alpha} = \lambda^2 \sum_{L=0}^{\infty} B_L P_L(\cos\vartheta) d\Omega, \quad (142)$$

with coefficients

$$B_L = \frac{1}{4} \sum_{\ell, \ell'} (2\ell + 1)(2\ell' + 1)(\ell\ell'00, L0)^2 (1 - U_{\ell}^*)(1 - U_{\ell'}). \quad (143)$$

Blatt and Biedenharn (1952) worked out the generalisation for particles with spin and for partition-changing (rearrangement) collisions. For zero Coulomb interaction they obtained

$$d\sigma_{\alpha\alpha'} = \frac{\lambda^2}{(2i+1)(2I+1)} \sum_{s,s'} \sum_{L=0}^{\infty} B_L(\alpha s, \alpha' s') P_L(\cos \vartheta) d\Omega, \quad (144)$$

with coefficients

$$B_L(\alpha s, \alpha' s') = \frac{(-)^{s-s'}}{4} \sum_{J_1, J_2} \sum_{\ell_1, \ell_2} \sum_{\ell'_1, \ell'_2} \bar{Z}(\ell_1 J_1 \ell_2 J_2, sL) \bar{Z}(\ell'_1 J_1 \ell'_2 J_2, s'L) \\ \times (\delta_{\alpha\alpha'} \delta_{\ell_1 \ell'_1} \delta_{s s'} - U_{\alpha \ell_1 s, \alpha' \ell'_1 s'}^{J_1})^* (\delta_{\alpha\alpha'} \delta_{\ell_2 \ell'_2} \delta_{s s'} - U_{\alpha \ell_2 s, \alpha' \ell'_2 s'}^{J_2}), \quad (145)$$

$$\bar{Z}(\ell_1 J_1 \ell_2 J_2, sL) = \sqrt{(2\ell_1 + 1)(2\ell_2 + 1)(2J_1 + 1)(2J_2 + 1)} \\ \times (\ell_1 \ell_2 00, L0) W(\ell_1 J_1 \ell_2 J_2, sL), \quad (146)$$

where $W(\ell_1 J_1 \ell_2 J_2, sL)$ is a Racah coefficient, see e. g. Fano and Racah (1959) or de-Shalit and Talmi (1963). Our phase convention is that of Lane and Thomas (1958); a slightly different convention is used in the Z tables of Biedenharn (1953). The Z coefficients vanish unless the quantum-mechanical triangle relations for the vector sums,

$$\vec{\ell}_1 + \vec{\ell}_2 = \vec{L} = \vec{\ell}'_1 + \vec{\ell}'_2 \quad (147)$$

$$\vec{\ell}_i + \vec{s} = \vec{J}_i = \vec{\ell}'_i + \vec{s}' \quad (i = 1, 2) \quad (148)$$

are fulfilled. Parity conservation demands $(-)^{\ell} \Pi_{\alpha} = \Pi_i = (-)^{\ell'} \Pi_{\alpha'}$, where $\Pi_{\alpha}, \Pi_{\alpha'}$ are the eigenparities of the ingoing and outgoing particles (positive for neutrons, protons, α -particles and photons) and Π_i is the parity of the compound system with total angular momentum J_i ($i = 1, 2$). If there is Coulomb interaction between the collision partners additional terms must be included, see Lane and Thomas (1958).

Let us integrate Eq. 144 over angles. All terms with $L < 0$ vanish because of the orthogonality of the P_L and because of

$$\bar{Z}(\ell_1 J_1 \ell_2 J_2, s0) = (-)^{J_1+s} \sqrt{2J_1+1} \delta_{J_1 J_2} \delta_{\ell_1 \ell_2}, \quad (149)$$

(see de-Shalit and Talmi 1963) and one finds

$$\sigma_{\alpha\alpha'} = \pi \lambda_{\alpha}^2 \sum_J g_J \sum_{\ell, \ell'} \sum_{s, s'} |\delta_{\alpha\alpha'} \delta_{\ell \ell'} \delta_{s s'} - U_{\alpha \ell s, \alpha' \ell' s'}^J|^2, \quad (150)$$

where the so-called spin factors

$$g_J \equiv \frac{2J+1}{(2i+1)(2I+1)} \quad (151)$$

are weights for the various possible total angular momenta.

We cannot go into the details of angular distributions but we do point out that they show interference between different partial waves, e. g. s and p wave, whereas angle-integrated cross sections do not. The latter are simple sums over terms with given ℓ and

| Table 2. Possible combinations of target spin I , orbital angular momentum ℓ and channel spin s resulting in total spin J , parity Π and spin factor g , for positive target parity Π_0 and projectile spin $i = 1/2$. (For negative target parity all signs must be reversed.) | | | | | | |
|--|--------|------------------------|--------------------|---------------|----------|------|
| $I\Pi_0$ | ℓ | s | $J\Pi$ | g | $\sum g$ | wave |
| 0+ | 0 | 1/2 | 1/2+ | 1 | 1 | s |
| | 1 | 1/2 | 1/2-, 3/2- | 1, 2 | 3 | p |
| | 2 | 1/2 | 3/2+, 5/2+ | 2, 3 | 5 | d |
| | etc. | | | | | |
| 1/2+ | 0 | 0 | 0+ | 1/4 | 1 | s |
| | | 1 | 1+ | 3/4 | | |
| | 1 | 0 | 1- | 3/4 | 3 | p |
| | | 1 | 0-, 1-, 2- | 1/4, 3/4, 5/4 | | |
| | 2 | 0 | 2+ | 5/4 | 5 | d |
| | 1 | 1+, 2+, 3+ | 3/4, 5/4, 7/4 | | | |
| etc. | | | | | | |
| 1+ | 0 | 1/2 | 1/2+ | 1/3 | 1 | s |
| | | 3/2 | 3/2+ | 2/3 | | |
| | 1 | 1/2 | 1/2-, 3/2- | 1/3, 2/3 | 3 | p |
| | | 3/2 | 1/2-, 3/2-, 5/2- | 1/3, 2/3, 3/3 | | |
| | 2 | 1/2 | 3/2+, 5/2+ | 2/3, 3/3 | 5 | d |
| | 3/2 | 1/2+, 3/2+, 5/2+, 7/2+ | 1/3, 2/3, 3/3, 4/3 | | | |
| etc. | | | | | | |

s , without cross terms, see Eq. 150. Nevertheless, a certain connexion exists between different partial waves. As already mentioned, the compound system and its quasi-stationary states are characterised, apart from energy, by total angular momentum, J , and parity, Π . Table 2 shows, for given target spin and positive parity, the possible combinations of ℓ , s and J for incident particles with spin $i = 1/2$. Certain $J\Pi$ combinations can be formed through more than one channel if $\ell > 0$ and $I > 0$. If $I\Pi_0 = 1/2+$, for instance, resonances with $J\Pi = 1-$ can be excited by the two p waves with $s = 0$ and $s = 1$, and the $2+$ levels can be excited by the d waves with $s = 0$ and $s = 1$. The neutron widths (that give the strength of the excitation, see below) for $1-$ and $2+$ levels are therefore sums of two partial widths for the two channel spins. For $I\Pi_0 = 1+$ the $1/2+$ levels can be excited even by partial waves with different ℓ , an s wave with $s = 1/2$ and a d wave with $s = 3/2$, while the $3/2+$ levels are excitable by three partial waves, an s wave with $s = 3/2$ and two d waves with $s = 1/2$ and $s = 3/2$, and so on.

So we find that each quasi-stationary compound state shows up as a resonance in all open channels that are not excluded by the spin and parity selection rules. The intensities (peak areas) may differ, but the resonance width must be the same in all those channels, being proportional to the reciprocal half life of the compound state. In this context it should be understood that the customary terms s- or p-wave resonance actually mean that the level can be excited at least by the s or p wave but possibly also by higher-order partial waves with the same parity. To give an example: The $3/2+$ s-wave resonances observed in neutron reactions with target nuclei having $I\Pi_0 = 1+$ contain also a d-wave component. It is true that at low incident energies the s-wave component is much larger

because of the higher centrifugal barrier for d-wave neutrons (see below) but it must be realised that certain d, f, ... resonance sequences are masked by coinciding s, p, ... sequences. This is important e. g. for the statistical interpretation of observed level densities or for the simulation of resonance effects in the unresolved resonance region with resonance “ladders” obtained by Monte Carlo sampling (see Section 4 below).

3.2. THE EXACT R-MATRIX EXPRESSIONS

The angle-integrated cross section $\sigma_{\alpha\alpha'}$, Eq. 150, can be written as a sum over partial cross sections, $\sigma_{cc'}$, obtained by summing over all entrance channels $c \equiv \{\alpha J \ell s\}$ and exit channels $c' \equiv \{\alpha' J' \ell' s'\}$ that lead from partition α to partition α' . In slightly simplified notation we write

$$\sigma_{cc'} = \pi \lambda_c^2 g_c |\delta_{cc'} - U_{cc'}|^2. \quad (152)$$

Note that for $c \neq c'$ the partial cross section is proportional to the quantum-mechanical probability $|U_{cc'}|^2$ of a transition from channel c to channel c' , and to the probability g_c of getting the correct angular momentum J from the spins of the collision partners. The Kronecker symbol $\delta_{cc'}$ arises since ingoing and outgoing particles cannot be distinguished if $c = c'$. The kinematical factor $\pi \lambda_c^2$ relates probability and cross section. The collision matrix \mathbf{U} , often called scattering or S matrix, is symmetric because for all practical purposes we can consider nuclear (and Coulomb) interactions as invariant under time reversal. Moreover, \mathbf{U} is unitary since the probabilities for transitions into the various channels must add up to unity, $\sum_{c'} |U_{cc'}|^2 = 1$. From the unitarity of \mathbf{U} and Eq. 152 it follows that the total cross section for entrance channel c is a linear function of U_{cc} ,

$$\sigma_c \equiv \sum_{c'} \sigma_{cc'} = 2\pi \lambda_c^2 g_c (1 - \text{Re } U_{cc}), \quad (153)$$

in contrast to the partial cross sections which depend quadratically on the $U_{cc'}$. The expressions obtained are thus simplest for the total cross section, most complicated for elastic scattering (because of the Kronecker symbol). It is therefore more convenient to calculate σ_{cc} as the difference between σ_c and the other partial cross sections rather than directly from Eq. 152. The reciprocity relation between the cross sections for a reaction $c \rightarrow c'$ and the inverse reaction $c' \rightarrow c$,

$$\frac{\sigma_{c'c}}{g_{c'} \lambda_{c'}^2} = \frac{\sigma_{cc'}}{g_c \lambda_c^2} \quad (154)$$

follows immediately from the symmetry of \mathbf{U} .

These equations are quite general. In order to introduce resonances we invoke R-matrix theory which allows us to express \mathbf{U} in terms of the channel matrix \mathbf{R} (see Lane and Thomas 1958, Lynn 1968),

$$\begin{aligned} U_{cc'} &= e^{-i(\varphi_c + \varphi_{c'})} P_c^{1/2} \{ [1 - \mathbf{R}(\mathbf{L} - \mathbf{B})]^{-1} [1 - \mathbf{R}(\mathbf{L}^* - \mathbf{B})] \}_{cc'} P_{c'}^{-1/2} \\ &= e^{-i(\varphi_c + \varphi_{c'})} \{ \delta_{cc'} + 2i P_c^{1/2} [(1 - \mathbf{R}\mathbf{L}^\circ)^{-1} \mathbf{R}]_{cc'} P_{c'}^{1/2} \}, \end{aligned} \quad (155)$$

$$R_{cc'} = \sum_{\lambda} \frac{\gamma_{\lambda c} \gamma_{\lambda c'}}{E_{\lambda} - E}, \quad (156)$$

$$L_{cc'}^0 \equiv L_{cc'} - B_{cc'} = (L_c - B_c) \delta_{cc'} \equiv (S_c + iP_c - B_c) \delta_{cc'}. \quad (157)$$

Alternatively the collision matrix can be expressed in terms of the level matrix \mathbf{A} ,

$$U_{cc'} = e^{-i(\varphi_c + \varphi_{c'})} \left(\delta_{cc'} + i \sum_{\lambda, \mu} \Gamma_{\lambda c}^{1/2} A_{\lambda \mu} \Gamma_{\mu c'}^{1/2} \right), \quad (158)$$

$$\Gamma_{\lambda c}^{1/2} \equiv \gamma_{\lambda c} \sqrt{2P_c}, \quad (159)$$

$$(\mathbf{A}^{-1})_{\lambda \mu} = (E_\lambda - E) \delta_{\lambda \mu} - \sum_c \gamma_{\lambda c} L_c^o \gamma_{\mu c}. \quad (160)$$

Note: Roman subscripts refer to reaction channels, Greek subscripts to compound levels, and $\mathbf{1}$ is the unit matrix. Three groups of physical quantities appear in these equations:

First, there are the resonance parameters, viz. formal level energies E_λ and probability amplitudes $\gamma_{\lambda c}$ for decay (or formation) of compound states λ via exit (or entrance) channels c , all neatly wrapped up in the R matrix (156), each level contributing one sum term (a hyperbola in terms of energies E). The $\gamma_{\lambda c}$ can be positive or negative, with practically random signs except near the ground state. Cross section formulae are usually written in terms of partial widths $\Gamma_{\lambda c}$ and total widths $\Gamma_\lambda \equiv \sum_c \Gamma_{\lambda c}$ rather than decay amplitudes.

The second group, hard-sphere phases φ_c and logarithmic derivatives L_c , depend only on the (known) in- and outgoing radial wave functions I_c and O_c at the channel radius a_c ,

$$\varphi_c \equiv \arg O_c(a_c) = \arctan \frac{\text{Im } O_c(a_c)}{\text{Re } O_c(a_c)}, \quad (161)$$

$$L_c \equiv a_c \frac{O'_c(a_c)}{O_c(a_c)} = \left[r_c \frac{\partial \ln O_c}{\partial r_c} \right]_{r_c=a_c}. \quad (162)$$

The $S_c \equiv \text{Re } L_c$ are called shift factors for reasons that will become clear later on, the $P_c \equiv \text{Im } L_c$ are centrifugal-barrier penetrabilities. The quantities B_c and a_c form the third group. They define the eigenvalue problem with eigenvalues E_λ . Their choice is largely a matter of convenience. The B_c are logarithmic derivatives of the radial eigenfunctions at the channel radii a_c . These radii define the boundary between the internal and the external region. They must be chosen so large that the nuclear interaction can be safely neglected if the distance r_c between the collision partners is larger, otherwise they are arbitrary. It is best to choose a_c just slightly larger than the radius of the compound nucleus (see Lynn 1968). A reasonable choice for neutron channels is $a_c = (1.23A^{1/3} + 0.80)$ fm, where A is the number of nucleons in the target nucleus. We mention here that in applied work all energies, resonance widths etc. are given in the laboratory system, as for instance in the widely used resonance parameter compilation of Mughabghab et al. (1981, 1984) known as the ‘‘barn book’’, or in that of Sukhoruchkin et al. (1998), or in computer files of evaluated nuclear data.

For neutral projectiles the outgoing radial wave functions are proportional to the spherical Hankel functions of the first kind, $h_\ell^{(1)}$,

$$O_c = I_c^* = ik_c r_c h_\ell^{(1)}(k_c r_c) \left(\simeq i^\ell e^{ik_c r_c} \quad \text{if } k_c r_c \gg \sqrt{\ell(\ell+1)} \right), \quad (163)$$

where $k_c = 1/\lambda_c$. The properties of the Hankel functions yield the recursion relations

$$L_0 = ik_c a_c = iP_0, \quad L_\ell = -\ell - \frac{(k_c a_c)^2}{L_{\ell-1} - \ell}, \quad (164)$$

$$\varphi_0 = k_c a_c, \quad \varphi_\ell = \varphi_{\ell-1} + \arg(\ell - L_{\ell-1}), \quad (165)$$

Table 3. Channel wave functions and related quantities for neutral projectiles ($\rho \equiv k_c r_c$, $\alpha \equiv k_c a_c$)

| ℓ | O_c | φ_c | S_c | P_c |
|----------|---|---|---|---|
| 0 | $e^{i\rho}$ | α | 0 | α |
| 1 | $e^{i\rho} \left(\frac{1}{\rho} - i \right)$ | $\alpha - \arctan \alpha$ | $\frac{-1}{\alpha^2 + 1}$ | $\frac{\alpha^3}{\alpha^2 + 1}$ |
| 2 | $e^{i\rho} \left(\frac{3}{\rho^2} - \frac{3i}{\rho} - 1 \right)$ | $\alpha - \arctan \frac{3\alpha}{3 - \alpha^2}$ | $\frac{-3(\alpha^2 + 6)}{\alpha^4 + 3\alpha^2 + 9}$ | $\frac{\alpha^5}{\alpha^4 + 3\alpha^2 + 9}$ |
| \vdots | \vdots | \vdots | \vdots | \vdots |

with which Table 3 is constructed. Note that $S_c = 0$ for $\ell = 0$, and that $S_c \rightarrow -\ell$ for $k_c a_c \rightarrow 0$ (at low energies). Therefore, $B_c = -\ell$ is a good choice for the resolved resonance region: Quite generally it simplifies all R-matrix expressions, and in particular it eliminates shift factors rigorously for s waves and, as we shall see below, approximately also for higher-order partial waves. This means that the cross section peaks occur at the formal resonance energies E_λ as they should, instead of being shifted. S_c and P_c for photon and fission channels are usually taken as constant.

The basic resonance parameters E_λ , $\gamma_{\lambda c}$ depend on the unknown nuclear interaction. They can therefore not be calculated from first principles (except for simple models like a square well potential, see below). In typical applications of R-matrix theory they are just fit parameters, adjustable to experimental data. Depending on the choice of B_c they can be either real and constant or complex and energy-dependent.

The Wigner-Eisenbud version of R-matrix theory is obtained if the boundary parameters B_c are chosen as real constants. (Wigner and Eisenbud 1947). The resonance parameters E_λ and $\gamma_{\lambda c}$ are then also real and constant, and the energy dependence of the collision matrix \mathbf{U} is solely due to the φ_c and L_c , both known functions of $k_c a_c$, i. e. of energy. This makes the Wigner-Eisenbud version the most convenient formalism for most purposes, especially with the choice $B_c = -\ell$. It is easily checked that the real R matrix yields a unitary collision matrix, which means the partial cross sections, Eq. 152, add up exactly to the correct total cross section, Eq. 153. A certain problem is, however, the need to invert either the channel matrix $\mathbf{1} - \mathbf{RL}^o$ of Eq. 155, or the inverse level matrix \mathbf{A}^{-1} of Eq. 160. Both matrices have very high rank. In practice the difficulty is overcome by various approximations to the inverse level matrix as will be shown below.

The Kapur-Peierls version of R-matrix theory is obtained with the choice $B_c = L_c$, i. e. $L_c^o = 0$ (Kapur and Peierls 1938). This removes the need for matrix inversion completely, since $\mathbf{1} - \mathbf{RL}^o = \mathbf{1}$, but leads to complex resonance parameters which depend implicitly on energy in a rather obscure way as now the very definition of the eigenvalue problem varies with energy, and thus the eigenvalues and the whole system of eigenfunctions, too. Moreover, the unitarity of the collision matrix is not manifest because the R

matrix is complex. In spite of these handicaps, formulae of the Kapur-Peierls type are useful in narrow energy ranges, in particular for the description of Doppler broadening. We shall write the complex and energy-dependent Kapur-Peierls parameters as \mathcal{E}_λ , $g_{\lambda c}$ in order to distinguish them from the real and constant Wigner-Eisenbud parameters E_λ , $\gamma_{\lambda c}$. Thus

$$U_{cc'} = e^{-i(\varphi_c + \varphi_{c'})} \left(\delta_{cc'} + i \sum_\lambda \frac{G_{\lambda c}^{1/2} G_{\lambda c'}^{1/2}}{\mathcal{E}_\lambda - E} \right), \quad (166)$$

$$G_{\lambda c}^{1/2} = g_{\lambda c} \sqrt{2P_c}. \quad (167)$$

Note that the Kapur-Peierls form of the collision matrix (and hence the corresponding total cross section expression) involves a simple sum over levels, whereas the Wigner-Eisenbud expression (158) involves a double sum.

The R-matrix equations reviewed so far are practically all that is needed in applied work from the whole apparatus of resonance theory. They ought to be thoroughly understood, however, and experience shows that this is not easy for the beginner. He might therefore wish to look at a simple illustration that shows the essential steps in the development of the theory and exhibits the meaning of the various quantities without the complications of spin algebra and matrix notation. Such an illustration is offered by the spherical optical model, especially with a square-well complex potential, for which everything can be calculated explicitly (see Fröhner 1998a), the results being of practical relevance for the unresolved resonance region.

3.3. THE PRACTICALLY IMPORTANT APPROXIMATIONS

A convenient starting point for the practically important versions of R-matrix theory is the inverse level matrix. We shall consider the following representations and approximations.

Wigner-Eisenbud representation (exact)

with B_c real and constant:

$$(\mathbf{A}^{-1})_{\lambda\mu} = (E_\lambda - E) \delta_{\lambda\mu} - \sum_c \gamma_{\lambda c} L_c^o \gamma_{\mu c} \quad (168)$$

(eigenvalues E_λ and decay amplitudes $\gamma_{\lambda c}$ real, constant, energy dependence of L_c^o known)

Kapur-Peierls representation (exact)

with $B_c = L_c$:

$$(\mathbf{A}^{-1})_{\lambda\mu} = (\mathcal{E}_\lambda - E) \delta_{\lambda\mu} \quad (169)$$

(eigenvalues \mathcal{E}_λ and decay amplitudes $g_{\lambda c}$ complex, energy dependences implicit, obscure)

Single-level Breit-Wigner approximation (SLBW)

Only one level retained, all others neglected:

$$(\mathbf{A}^{-1})_{\lambda\mu} \rightarrow E_0 - E - \sum_c L_c^o \gamma_c^2 \equiv E_0 + \Delta - E - i\Gamma/2 \quad (170)$$

(level shift Δ and total width $\Gamma = \sum_c \Gamma_c$ real, energy dependences explicit, well known)

Multi-level Breit-Wigner approximation (MLBW)

Off-diagonal elements of \mathbf{A}^{-1} neglected:

$$(\mathbf{A}^{-1})_{\lambda\mu} = (E_\lambda - E - \sum_c L_c^o \gamma_{\lambda c}^2) \delta_{\lambda\mu} \equiv (E_\lambda + \Delta_\lambda - E - i\Gamma_\lambda/2) \delta_{\lambda\mu} \quad (171)$$

(level shift Δ_λ and total width $\Gamma_\lambda = \sum_c \Gamma_{\lambda c}$ real, energy dependences explicit, well known)

Reich-Moore approximation

Off-diagonal contributions from photon channels, $c \in \gamma$, are neglected:

$$(\mathbf{A}^{-1})_{\lambda\mu} = (E_\lambda + \Delta_{\lambda\gamma} - E - i\Gamma_{\lambda\gamma}/2) \delta_{\lambda\mu} - \sum_{c \notin \gamma} \gamma_{\lambda c} L_c^o \gamma_{\mu c} \quad (172)$$

(real level shift $\Delta_{\lambda\gamma}$ from photon channels usually absorbed in real, constant E_λ , radiation width $\Gamma_{\lambda\gamma} = \sum_{c \in \gamma} \Gamma_{\lambda c}$ real, usually taken as constant; other energy dependences explicit)

Adler-Adler approximation

Energy dependence of L_c^o neglected:

$$(\mathbf{A}^{-1})_{\lambda\mu} = (E_\lambda - E) \delta_{\lambda\mu} - \sum_c \gamma_{\lambda c} \sqrt{L_c^o(E_\lambda) L_c^o(E_\mu)} \gamma_{\mu c} \quad (173)$$

The Reich-Moore approximation is most, SLBW least accurate among these approximations. With a suitable choice of the boundary parameters the level shifts Δ_λ vanish at least locally. At low energies this is accomplished for neutron channels with $B_c = -\ell$ (see Table 3) as mentioned before. Table 3 shows also that the centrifugal-barrier penetrabilities $P_c = P_\ell$ for neutrons, and hence all neutron widths,

$$\Gamma_{\lambda c}(E) \equiv 2P_\ell(E) \gamma_{\lambda c}^2 = \Gamma_{\lambda c}(|E_\lambda|) \frac{P_\ell(E)}{P_\ell(|E_\lambda|)} \quad (c \equiv \{\alpha J \ell s\} \in n), \quad (174)$$

contain (at least) a factor \sqrt{E} . Additional factors in the p-, d-, ... penetrabilities behave for low energies as E, E^2, \dots . As a consequence s-wave levels dominate at low energies while p-wave levels show up only at higher energies, d-wave levels at still higher energies, etc.

The absolute values in the conventional definition (170) of the neutron width make it applicable not only to compound states with $E_\lambda > 0$ but also to subthreshold (bound, "negative") states with $E_\lambda < 0$ although, strictly speaking, the centrifugal-barrier penetrabilities $P_\ell(E)$ and thus the widths vanish below the reaction threshold ($E < 0$). Neutron widths given in tables and computer files are to be understood as $\Gamma_{\lambda c}(|E_\lambda|)$. Another convention concerns the signs of the width amplitudes. They are important in the multi-channel case but get lost when widths are calculated. It is therefore customary to tabulate partial widths with the sign (relative to the neutron width amplitude) of the corresponding width amplitude. From a principal viewpoint it would be more appropriate (and less confusing) to list the width amplitudes rather than the partial widths – neither do they depend on energy, nor do they vanish below the threshold, and the signs must not be explained either – but it is too late to change ingrained habits.

The shifts and penetrabilities for photon and fission channels can usually be taken as constant. Hence these shifts vanish if we choose $B_c = S_c$, and the fission and radiation widths do not depend on energy. Let us now look at the cross section expressions resulting from the various representations and approximations.

3.3.1. Kapur-Peierls Cross Section Expressions

In anticipation of Doppler broadening we write the Kapur-Peierls collision matrix (166) in the form

$$U_{cc'} = e^{-i(\varphi_c + \varphi_{c'})} \left[\delta_{cc'} - \sum_{\lambda} \frac{G_{\lambda c}^{1/2} G_{\lambda c'}^{1/2}}{G_{\lambda}/2} (\psi_{\lambda} + i\chi_{\lambda}) \right], \quad (175)$$

where the symmetric and asymmetric resonance profiles or line shape functions ψ_{λ} and χ_{λ} are defined by

$$\psi_{\lambda} + i\chi_{\lambda} \equiv \frac{iG_{\lambda}/2}{E - \mathcal{E}_{\lambda}} = \frac{G_{\lambda}^2/4}{(E - \tilde{E}_{\lambda})^2 + G_{\lambda}^2/4} + i \frac{(E - \tilde{E}_{\lambda})G_{\lambda}/2}{(E - \tilde{E}_{\lambda})^2 + G_{\lambda}^2/4} \quad (176)$$

and the real Kapur-Peierls resonance energies \tilde{E}_{λ} and widths G_{λ} by

$$\mathcal{E}_{\lambda} \equiv \tilde{E}_{\lambda} - iG_{\lambda}/2. \quad (177)$$

The symmetric resonance profile is essentially (if we disregard the weak energy dependences of \mathcal{E}_{λ} and of G_{λ}) a Lorentzian, and the asymmetric profile is its energy derivative. The resulting cross section expressions are

$$\sigma_c = 4\pi\lambda_c^2 g_c \left\{ \sin^2 \varphi_c + \text{Re} \left[e^{-2i\varphi_c} \sum_{\lambda} \frac{G_{\lambda c}}{G_{\lambda}} (\psi_{\lambda} + i\chi_{\lambda}) \right] \right\}, \quad (178)$$

$$\sigma_{cc'} = \sigma_c \delta_{cc'} - 4\pi\lambda_c^2 g_c \text{Re} \left[\sum_{\lambda} \frac{G_{\lambda c}^{1/2} G_{\lambda c'}^{1/2}}{G_{\lambda}} W_{cc'}(\mathcal{E}_{\lambda}^*)^* (\psi_{\lambda} + i\chi_{\lambda}) \right], \quad (179)$$

$$W_{cc'}(\mathcal{E}_{\lambda}^*) \equiv \delta_{cc'} + i \sum_{\mu} \frac{G_{\mu c}^{1/2} G_{\mu c'}^{1/2}}{\mathcal{E}_{\mu} - \mathcal{E}_{\lambda}^*}. \quad (180)$$

The resonance profiles contain the rapid, resonance-related energy variations that are sensitive to Doppler broadening, while the other quantities vary slowly with energy. We stress that although the weak energy dependences of the Kapur-Peierls parameters are not known explicitly, the Kapur-Peierls formalism is formally exact.

3.3.2. SLBW Cross Section Expressions

The collision matrix for a single level,

$$U_{cc'} = e^{-i(\varphi_c + \varphi_{c'})} \left(\delta_{cc'} + \frac{i\Gamma_c^{1/2} \Gamma_{c'}^{1/2}}{E_0 + \Delta - E - i\Gamma/2} \right), \quad (181)$$

is unitary. The resulting cross section expressions are

$$\sigma_c = 4\pi\lambda_c^2 g_c \left\{ \sin^2 \varphi_c + \frac{\Gamma_c}{\Gamma} (\psi \cos 2\varphi_c + \chi \sin 2\varphi_c) \right\}, \quad (182)$$

$$\sigma_{cc'} = 4\pi\lambda_c^2 g_c \frac{\Gamma_c \Gamma_{c'}}{\Gamma^2} \psi \quad (c \neq c'), \quad (183)$$

$$\sigma_{cc} = \sigma_c - \sum_{c' \neq c} \sigma_{cc'}, \quad (184)$$

with the resonance profiles given by

$$\psi + i\chi \equiv \frac{i\Gamma/2}{E - E_0 + i\Gamma/2} = \frac{\Gamma^2/4}{(E - E_0)^2 + \Gamma^2/4} + i \frac{(E - E_0)\Gamma/2}{(E - E_0)^2 + \Gamma^2/4}. \quad (185)$$

The basic resonance shape displayed by partial cross sections with $c \neq c'$ is essentially symmetric, whereas that of the total and of the scattering cross section is a sum of three terms: the nearly constant potential-scattering cross section, a symmetric resonance term, and an asymmetric term arising from interference between potential (hard-sphere) and resonance scattering. It is easy to deduce from Eqs. 182 and 185 that the total cross section reaches its peak value,

$$\sigma_c(E_+) = 4\pi\lambda_c^2 g_c \left(1 - \frac{\Gamma_a}{\Gamma} \cos^2 \varphi_c\right), \quad (186)$$

at the energy

$$E_+ = E_0 + \frac{\Gamma}{2} \tan \varphi_c, \quad (187)$$

while the minimum value in the interference dip is

$$\sigma_c(E_-) = 4\pi\lambda_c^2 g_c \frac{\Gamma_a}{\Gamma} \sin^2 \varphi_c = \frac{\Gamma_a}{\Gamma} \sigma_{p,c} \quad (188)$$

at

$$E_- = E_0 - \frac{\Gamma}{2} \cot \varphi_c, \quad (189)$$

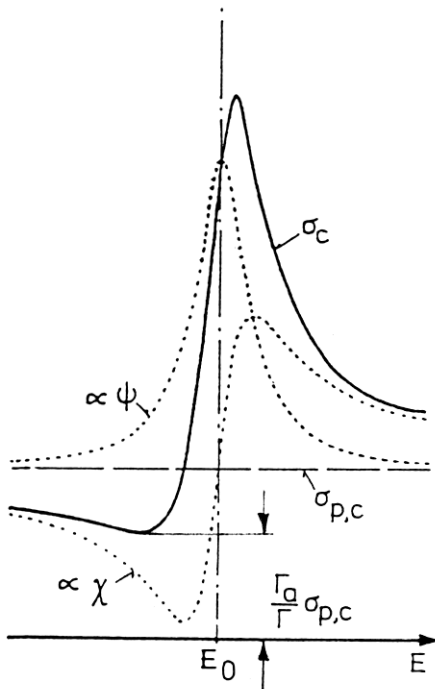


Fig. 11.

where $\Gamma_a \equiv \Gamma - \Gamma_n$ is the absorption width and $\sigma_{p,c}$ the potential-scattering cross section for entrance channel c . These last expressions are valid if the slow energy dependences of λ_c^2 , φ_c , Γ and Γ_n are negligible. For pure elastic scattering ($\Gamma_a = 0$) the minimum cross section is zero, the peak cross section is equal to the unitarity limit $4\pi\lambda_c^2 g_c$ (compare Eq. 153). The spin factor, hence the level spin J , can thus often be obtained by just measuring the resonance height. This works best for light nuclei or structural materials such as ^{56}Fe and other iron, nickel and chromium isotopes that are almost pure scatterers, and for broad, "isolated" resonances that are virtually unaffected by Doppler broadening and multi-level interference, and observable with good instrumental resolution. The depth of the interference minimum, most important for shielding calculations, is essentially given by the ratio Γ_a/Γ times the potential scattering cross section. At low energies s-wave reactions dominate, the potential scattering cross section is practically constant, whereas the capture and fission cross sections behave like $1/v$. This low-energy behaviour is quite generally true, not just in SLBW approximation.

Because of the slow variation of the sines and cosines with energy the total cross section resonances look differently at different energies: At low energies they look as in Fig. 11, with the interference minimum ("window") on the low-energy side. This shape is typical for the resolved region. At higher energies the symmetric term becomes less and less important until the asymmetric term dominates. At still higher energies, when $\varphi_c \approx \pi$, resonances appear as dips rather than peaks (a famous example is the 2.35 MeV resonance of $^{16}\text{O}+\text{n}$) and eventually the interference windows reappear on the high-energy side of the peaks.

In practice one must, however, describe cross sections with many resonances. One can simply add SLBW resonance terms (and add potential scattering for σ_c and σ_{cc}). This is the SLBW definition of the ENDF format (cf. Rose and Dunford 1990) that is used world-wide for applications-oriented, computer-readable libraries of evaluated neutron data. Since this ad-hoc recipe does not originate from a unitary collision matrix the unitarity constraint $0 < \sigma_c < 4\pi\lambda_c^2 g_c$ is not guaranteed. In fact, this "many-level" SLBW approximation is notorious for the occurrence of nonphysical negative total and scattering cross sections. The reason is easy to understand: At low energies negative contributions can only come from the asymmetric profiles of resonances above. On average they are compensated by positive contributions from resonances below, but if the resonances above are unusually strong or those below unusually weak, scattering cross sections can become negative in the interference minima. Less noticeable but often equally bad is the opposite effect: SLBW peak cross sections can exceed the unitarity limit if resonances above are weak or those below strong.

3.3.3. MLBW Cross Section Expressions

The MLBW approximation is better than the many-level SLBW approximation. The collision matrix following from Eq. 171,

$$U_{cc'} = e^{-i(\varphi_c + \varphi_{c'})} \left(\delta_{cc'} + i \sum_{\lambda} \frac{\Gamma_{\lambda c}^{1/2} \Gamma_{\lambda c'}^{1/2}}{E_{\lambda} + \Delta_{\lambda} - E - i\Gamma_{\lambda}/2} \right), \quad (190)$$

involves a simple sum over resonances, as the Kapur-Peierls collision matrix does. It follows that we can take over the Kapur-Peierls expressions with the replacements $E'_{\lambda} \rightarrow E_{\lambda} + \Delta_{\lambda}$, $G_{\lambda} \rightarrow \Gamma_{\lambda} = \sum_c \Gamma_{\lambda c}$, $G_{\lambda c}^{1/2} \rightarrow \Gamma_{\lambda c}^{1/2}$, whence

$$\sigma_c = 4\pi\lambda_c^2 g_c \left[\sin^2 \varphi_c + \sum_c \frac{\Gamma_{\lambda c}}{\Gamma_{\lambda}} (\psi_{\lambda} \cos 2\varphi_c + \chi_{\lambda} \sin 2\varphi_c) \right], \quad (191)$$

$$\sigma_{cc'} = \sigma_c \delta_{cc'} - 4\pi\lambda_c^2 g_c \operatorname{Re} \left[\sum_{\lambda} \frac{\Gamma_{\lambda c}^{1/2} \Gamma_{\lambda c'}^{1/2}}{\Gamma_{\lambda}} W_{cc'}(\mathcal{E}_{\lambda}^*)^* (\psi_{\lambda} + i\chi_{\lambda}) \right], \quad (192)$$

$$W_{cc'}(\mathcal{E}_{\lambda}^*) = \delta_{cc'} + i \sum_{\mu} \frac{\Gamma_{\mu c}^{1/2} \Gamma_{\mu c'}^{1/2}}{E_{\mu} - E_{\lambda} - i(\Gamma_{\mu} + \Gamma_{\lambda})/2}. \quad (193)$$

Since the partial cross sections (192) were derived from the collision matrix as absolute squares (see Eq. 152), they are guaranteed to be positive, and they are again linear functions of the line profiles ψ_{λ} and χ_{λ} defined exactly as in the SLBW case, Eq. 185. We recognise further that σ_c , Eq. 191, is just the "many-level" SLBW approximation. As

the MLBW collision matrix is not unitary, however, σ_c is not the sum of the partial cross sections, Eq. 192. The MLBW approximation as defined in the ENDF format (cf. Rose and Dunford 1990) is even cruder, in fact it is an SLBW/MLBW hybrid: Only elastic scattering is actually calculated in MLBW approximation. All other partial cross sections are calculated in (many-level) SLBW approximation, and the total cross section as the sum over all partials. This avoids negative cross sections yet prevents neither unphysical peak cross sections nor badly described interference minima for strongly overlapping levels. For light and medium-mass nuclei and for fissile actinides the MLBW approximation is therefore often inadequate, although it works quite well for compound systems with widely spaced, narrow levels like $^{232}\text{Th}+n$ or $^{238}\text{U}+n$.

Note that the calculation of MLBW partial cross sections according to Eqs. 192 and 193 involves double sums over levels. Even with modern computers this can be time-consuming if hundreds of levels are to be included, as is not unusual with modern evaluated files. It is then better to calculate the partial cross section directly from the collision matrix (i. e. from Eqs. 152 and 190) which involves only a single sum over levels. For Doppler broadening, however, the representation (192), (193) in terms of line shape profiles has advantages as will be seen below.

3.3.4. Reich-Moore Cross Section Expressions

Usually very many photon channels contribute to the sum $\sum_c \gamma_{\lambda c} L_c^o \gamma_{\mu c}$ in the inverse level matrix \mathbf{A}^{-1} , Eq. 160. While their contributions all add up with the same sign in the diagonal elements, they tend to cancel in the off-diagonal elements because the decay amplitudes have practically random signs but comparable magnitudes. Therefore the error is quite small if one simply neglects all photon channel contributions to the off-diagonal elements, as proposed independently by Thomas (1955) and by Reich and Moore (1958). The resulting inverse level matrix, Eq. 172, belongs evidently to an eigenvalue problem with E_λ replaced by $E_\lambda - i\Gamma_{\lambda\gamma}/2$, with a "reduced" R matrix

$$R_{cc'} = \sum_{\lambda} \frac{\gamma_{\lambda c} \gamma_{\lambda c'}}{E_\lambda - E - i\Gamma_{\lambda\gamma}/2} \quad (c, c' \notin \gamma), \quad (194)$$

reduced in the sense that it is defined in the subspace of nonphotonic channels only. The only traces of the eliminated photon channels are the total radiation widths, $\Gamma_{\lambda\gamma}$, in the denominators. A similar complex R function is encountered in the R-matrix treatment of the optical model (see Fröhner 1998a) which suggests that the imaginary part of the denominators in the reduced R matrix and the imaginary part of the complex potential are different consequences of the same phenomenon: absorption into compound states and subsequent decay into eliminated channels.

From the reduced R matrix one gets a reduced collision matrix and therefrom the cross sections for all retained nonphotonic channels in the usual way, Eqs. 152-153. The reduced R matrix is of low rank, hence inversion of $\mathbf{1} - \mathbf{R}\mathbf{L}^o$ is easy. In fact, the highest rank employed in neutron resonance analyses up to now is 3 (1 elastic, 2 fission channels). Cases with rank 2 involve 1 elastic plus 1 fission or 1 inelastic channel. For the overwhelming majority of neutron resonance data the only energetically allowed processes are merely elastic scattering and radiative capture, for which 1-channel Reich-Moore expressions are sufficient, with R functions instead of R matrices. (An example is the 1-channel Reich-Moore fit to ^{56}Fe total cross section data displayed in Fig 10.) The capture cross section can be found from

$$\sigma_{c\gamma} = \pi \lambda_c^2 g_c \sum_{\lambda} \Gamma_{\lambda\gamma} \left| \sum_{c' \notin \gamma} \frac{P_c^{1/2} [(\mathbf{1} - \mathbf{R}\mathbf{L}^o)^{-1}]_{cc'} P_{c'}^{-1/2} \Gamma_{\lambda c'}^{1/2}}{E_\lambda - E - i\Gamma_{\lambda\gamma}/2} \right|^2 \quad (195)$$

(cf. Reich and Moore 1958). We point out that this approximation is exact in the limit of vanishing radiation widths (more generally: vanishing widths for eliminated channels) in which it reduces to the general Wigner-Eisenbud formalism. It is also exact in the limit of one single level since in this case the Reich-Moore level matrix \mathbf{A} reduces to the corresponding SLBW level matrix. Otherwise it is so exact that although the reduced collision matrix cannot be unitary - because of transitions into eliminated channels - the overall collision matrix can still be considered as unitary, i. e. as conserving probability flux, so that the capture cross section may alternatively be obtained as the difference

$$\sigma_{c\gamma} = \sigma_c - \sum_{c' \notin \gamma} \sigma_{cc'}, \quad (196)$$

with σ_c calculated from the reduced collision matrix element U_{cc} according to Eq. 153. Experience has shown that with this approximation all resonance cross section data can be described in detail, in the windows as well as in the peaks, even the weirdest multilevel interference patterns (see Fig. 10). It works equally well for light, medium-mass and heavy nuclei, fissile and nonfissile. It is often believed that the Reich-Moore approximation can only be applied to fissile nuclei, but actually the retained channels can be of any type - elastic, inelastic, fission, even individual photon channels such as those for transitions to the ground state or to specific metastable states. Furthermore, computer programs written for the Reich-Moore formalism can be used for general Wigner-Eisenbud R-matrix calculations - one must simply set all radiation widths (eliminated-channel widths) equal to zero.

It might be expected that with all these advantages the Reich-Moore formalism is the most widely used one, but this is not true. The main reason is that Reich-Moore cross sections cannot be expressed as sums over Breit-Wigner resonance profiles, at least not without some preparatory work. This is often considered as a disadvantage for Doppler broadening computations. We shall see below, however, that the problem is not as serious as some believe. Among resonance analysts there is no question that the Reich-Moore approximation is superior to the other single- and multi-level variants of R-matrix theory.

3.3.5. Adler-Adler Cross Section Expressions

The approximation (173) for the matrix \mathbf{A}^{-1} is a generalisation for the s-wave expression used by Adler and Adler (1970), a generalisation that preserves symmetry with respect to the level indices λ and μ . Diagonalisation of the level matrix \mathbf{A} yields the collision matrix in Kapur-Peierls form, Eqs. 172-173, but with parameters \mathcal{E}_λ and $g_{\lambda c}$ that do not depend on energy, in contrast to genuine Kapur-Peierls parameters. The corresponding cross section expressions are often not written for specific channels (c, c', \dots) but for specific reaction types ($x = f, \gamma, \dots$, total), restricted to $\ell = 0$:

$$\sigma \equiv \sum_{c \in n} \sigma_c = \sigma_p + \frac{1}{\sqrt{E}} \sum_{\lambda} \frac{1}{\nu_{\lambda}} \left(G_{\lambda}^{(T)} \psi_{\lambda} - H_{\lambda}^{(T)} \chi_{\lambda} \right), \quad (197)$$

$$\sigma_x \equiv \sum_{c \in n} \sum_{c' \in x} \sigma_{cc'} = \frac{1}{\sqrt{E}} \sum_{\lambda} \frac{1}{\nu_{\lambda}} \left(G_{\lambda}^{(x)} \psi_{\lambda} - H_{\lambda}^{(x)} \chi_{\lambda} \right), \quad (x = \gamma, f, \dots), \quad (198)$$

where σ_p is the potential-scattering cross section, $G_{\lambda}^{(x)}/(\nu_{\lambda}\sqrt{E})$ and $H_{\lambda}^{(x)}/(\nu_{\lambda}\sqrt{E})$ are sums over all coefficients of ψ_{λ} and χ_{λ} in Eqs. 178-180, with $\nu_{\lambda} \equiv \Gamma_{\lambda}/2$ and \sqrt{E} coming from $P_c(E)$. The sums over λ are over levels irrespective of $J\Pi$, with spin factors absorbed in

the coefficients $G_\lambda^{(x)}$ and $H_\lambda^{(x)}$. These coefficients, together with the level energies $\mu_\lambda \equiv E'_\lambda$, (half) widths ν_λ and the potential-scattering cross section σ_p (or an effective radius) are the Adler-Adler parameters. In principle one could define them even for isotopic mixtures, by similarly absorbing the relative abundances in the coefficients. The approximation (173) means essentially that the energy dependence of level shifts and total widths is neglected in the resonance denominators. Therefore the Adler-Adler approximation works well for fissile nuclei, for which $\Gamma_\lambda \approx \Gamma_{\lambda\gamma} + \Gamma_{\lambda f} = \text{const}$, but not so well for light or medium-mass nuclei, for which $\Gamma_\lambda \approx \Gamma_{\lambda n} = 2P_c(E)\gamma_{\lambda n}^2$.

3.3.6. Conversion of Wigner-Eisenbud to Kapur-Peierls Parameters

Wigner-Eisenbud parameters can be converted to Kapur-Peierls parameters as follows (Fröhner 1978). The collision matrix must be invariant under a change of boundary parameters, e. g. from $B_c = -\ell$ to $\tilde{B}_c = L_c^o$. (We shall use the tilde for Kapur-Peierls quantities.) Eq. 155 shows that this implies $(\mathbf{1} - \mathbf{R}\mathbf{L}^o)^{-1}\mathbf{R} = \tilde{\mathbf{R}}$, which with the abbreviations

$$\mathbf{K} \equiv \mathbf{L}^{o1/2}\mathbf{R}\mathbf{L}^{o1/2}, \quad \tilde{\mathbf{K}} \equiv \mathbf{L}^{o1/2}\tilde{\mathbf{R}}\mathbf{L}^{o1/2} \quad (199)$$

yields

$$(\mathbf{1} - \mathbf{K})^{-1} = \mathbf{1} + \tilde{\mathbf{K}}. \quad (200)$$

The Kapur-Peierls resonance energies \mathcal{E}_λ are the complex poles of $\tilde{\mathbf{K}}$, i. e. the solutions of

$$\det[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)] = 0 \quad (201)$$

because $\mathbf{A}^{-1} = \mathbf{C}[\mathbf{A}]/\det \mathbf{A}$ for any nonsingular matrix \mathbf{A} , where we use the notation $\det(\mathbf{A})$ for the determinant and $\mathbf{C}[\mathbf{A}]$ for the matrix of cofactors. The residues are obtained from Eq. 200. In the limit $E \rightarrow \mathcal{E}_\lambda$ one gets $[\mathbf{1} + \tilde{\mathbf{K}}(E)]_{cc'} \simeq L_c^{o1/2}g_{\lambda c}g_{\lambda c'}L_{c'}^{o1/2}/(E - \mathcal{E}_\lambda)$ on the right hand side, while on the left one has $\{\mathbf{C}[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)]\}_{cc'}/\det[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)]$, where Taylor expansion of the determinant gives $\det[\mathbf{1} - \mathbf{K}(E)] \simeq (E - \mathcal{E}_\lambda)\text{tr}\{\mathbf{C}[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)]\mathbf{K}'(\mathcal{E}_\lambda)\}$. Hence the residues at the pole \mathcal{E}_λ are

$$g_{\lambda c}g_{\lambda c'} = \frac{1}{\sqrt{L_c^o(\mathcal{E}_\lambda)L_{c'}^o(\mathcal{E}_\lambda)}} \frac{\{\mathbf{C}[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)]\}_{cc'}}{\text{tr}\{\mathbf{C}[\mathbf{1} - \mathbf{K}(\mathcal{E}_\lambda)]\mathbf{K}'(\mathcal{E}_\lambda)\}}, \quad (202)$$

where tr denotes the trace and \mathbf{K}' is the derivative of \mathbf{K} ,

$$K'_{cc'}(E) = \frac{\partial}{\partial E}L_c^{o1/2}R_{cc'}L_{c'}^{o1/2} \simeq \sqrt{L_c^o(E)L_{c'}^o(E)} \sum_{\mu} \frac{\gamma_{\mu c}\gamma_{\mu c'}}{(E_\mu - E)^2}. \quad (203)$$

So we know how to calculate residuals from given poles, but how do we find the poles corresponding to given Wigner-Eisenbud parameters, i. e. how can we solve the deceptively simple-looking Eq. 201? Fortunately we know already the MLBW approximation $\mathcal{E}_\lambda \approx E_\lambda + \Delta_\lambda - i\Gamma_\lambda/2$, see Eq. 190. We may take it as an initial guess to be improved by iteration. In order to find an iteration scheme we write the determinant (201) in the form

$$\det(\mathbf{1} - \mathbf{K}) = 1 - \text{tr} \mathbf{K} + F(\mathbf{K}), \quad (204)$$

where $-\text{tr} \mathbf{K} + F(\mathbf{K})$ is the sum of $\det(-\mathbf{K})$ and all its principal minors (cf. e. g. Korn and Korn 1968), in particular

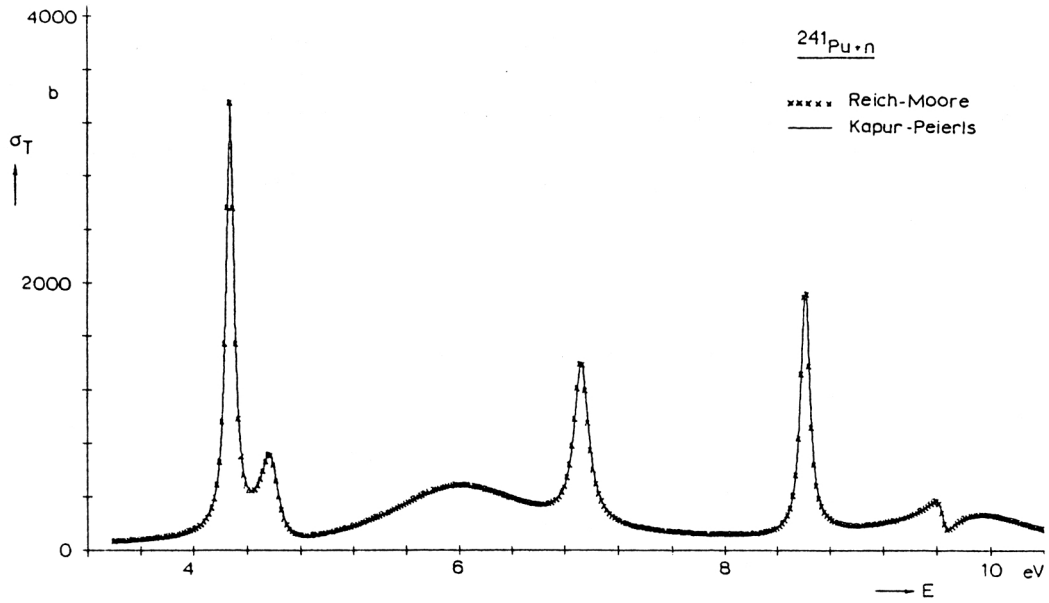


Fig. 12. Verification of the resonance parameter conversion technique explained in Subject. 3.3.6: 3-channel Reich-Moore cross sections (symbols) and Kapur-Peierls cross sections calculated from converted resonance parameters (solid line) agree. (From Fröhner 1978)

$$\begin{aligned}
 F &= 0 && \text{for 1 (elastic) channel,} \\
 F &= \det(-\mathbf{K}) && \text{for 2 channels,} \\
 F &= \det(-\mathbf{K}) + \text{tr } \mathbf{C}[-\mathbf{K}] && \text{for 3 channels.}
 \end{aligned}$$

Next we pull out the λ -th term of $\text{tr } \mathbf{K}$,

$$\text{tr } \mathbf{K}(\mathcal{E}_\lambda) = \sum_{\mu} \frac{1}{E_{\mu} - \mathcal{E}_\lambda} \sum_c L_c^o \gamma_{\mu c}^2 = \frac{\Delta_\lambda - i\Gamma_\lambda/2}{E_\lambda - \mathcal{E}_\lambda} + \sum_{\mu \neq \lambda} \frac{\Delta_\mu - i\Gamma_\mu/2}{E_\mu - \mathcal{E}_\mu}, \quad (205)$$

which together with Eq. 204 permits us to rewrite Eq. 201 in the form

$$\mathcal{E}_\lambda = E_\lambda + \frac{\Delta_\lambda - i\Gamma_\lambda/2}{1 + \sum_{\mu \neq \lambda} \frac{\Delta_\mu - i\Gamma_\mu/2}{E_\mu - \mathcal{E}_\lambda} + F(\mathcal{E}_\lambda)}. \quad (206)$$

This equation is convenient for iteration: Inserting the initial MLBW approximation on the right-hand side one gets an improved value that can be reinserted on the right, and so on. After a few iterations the result becomes stable within some reasonable accuracy and can be inserted in Eq. 202 to yield the residues. Once all Kapur-Peierls parameters \mathcal{E}_λ and $g_{\lambda c} g_{\lambda c'}$ are known one can insert them in the Kapur-Peierls cross section expressions that involve the resonance profiles.

Conversion of Reich-Moore to Kapur-Peierls parameters works in the same way, the only change being that E_μ must be replaced by $E_\mu - i\Gamma_{\mu\gamma}/2$, and Γ_μ by $\Gamma_\mu - \Gamma_{\mu\gamma}$ everywhere. Fig. 12 shows cross sections calculated from Reich-Moore parameters directly and from Kapur-Peierls parameters after conversion. Conversion of Wigner-Eisenbud to Adler-Adler parameters by matrix inversion is possible for instance with the POLLA code (de Saussure and Perez 1969).

3.4. EXTERNAL LEVELS

R-matrix theory shows that the cross sections in a limited energy range depend not only on the “internal” levels in that range but also on the “external” levels below and above. Problems arise in practical resonance fitting and parametrisation work because below the neutron or proton threshold ($E < 0$) the compound levels are unobservable and therefore unknown. Above the analysed range, resonances may still be observable but less and less well resolved as energy increases, because instrumental resolution worsens while level density grows and resonance widths increase – all of which makes the distinction between single resonances and unresolved multiplets increasingly difficult and eventually impossible. If the unknown external levels are simply omitted one cannot fit the experimental data satisfactorily. In particular with elastic scattering and total cross section data one gets troublesome edge effects and problems with the potential scattering cross section between resonances. Various ad hoc methods have been developed in the past to cope with the unknown external levels, from simulating them by “picket fence” or Monte Carlo sampled fictitious resonance sequences (“ladders”) to repeating the internal levels periodically below and above the internal region. The following sections present better founded, well tested and more convenient methods that have been available since decades but are not as widely used as they deserve.

3.4.1. Statistical Representation of External Levels

Modern evaluated nuclear-data libraries contain parameters for hundreds of resonances per isotope. Such large numbers suggest a statistical treatment of the more distant levels if a cross section is to be calculated at a given energy. Moreover, there are always enormous numbers of unknown levels both below and above the resolved resonance region contributing noticeably to the R matrix, in particular near the edges of this region. In order to include them at least statistically we split the (Reich-Moore) R matrix for a given level sequence (given $J\Pi$) into a sum over the unknown (“distant” or “external” levels) and another one over the known (“local” or “internal”) levels,

$$R_{cc'} = R_{cc'}^0 + \sum_{\lambda=1}^{\Lambda} \frac{\gamma_{\lambda c} \gamma_{\lambda c'}}{E_\lambda - E - i\Gamma_{\lambda\gamma}/2}, \quad (207)$$

and replace the sums in the distant-level term by integrals,

$$\begin{aligned} R_{cc'}^0 &= \left(\sum_{\lambda} - \sum_{\lambda=1}^{\Lambda} \right) \frac{\gamma_{\lambda c} \gamma_{\lambda c'}}{E_\lambda - E - i\Gamma_{\lambda\gamma}/2} \\ &\simeq \left(\int_{-\infty}^{\infty} - \int_{\bar{E}-I/2}^{\bar{E}+I/2} \right) \frac{dE'}{D_c} \langle \gamma_c \gamma_{c'} \rangle \frac{E' - E - i\bar{\Gamma}_\gamma/2}{(E' - E)^2 + \bar{\Gamma}_\gamma^2/4}, \end{aligned} \quad (208)$$

where \bar{E} and I are midpoint and length of the interval containing the local levels, $1/D_c = 1/D_J$ is the density of levels with spin J (and given parity) that is needed if the sums

are to be approximated by integrals, and $\bar{\Gamma}_\gamma$ is the average radiation width. Especially for heavy nuclei the radiation width, as a sum over very many partial radiation widths, does not vary much from level to level so that $\Gamma_{\lambda\gamma} \simeq \bar{\Gamma}_\gamma$. Since $(E' - E)^2 \gg \bar{\Gamma}_\gamma^2/4$ for the distant levels we can neglect $\bar{\Gamma}_\gamma^2/4$ in the last expression. Moreover we can neglect the off-diagonal elements of the average matrix $\langle \gamma_c \gamma_{c'} \rangle$ because of the practically random signs of the $\gamma_{\lambda c}$. With the usual definition of the pole strength s_c and its Hilbert transform, the distant-level parameter R_c^∞ ,

$$s_c \equiv \frac{\langle \gamma_c^2 \rangle}{D_c}, \quad (209) \quad R_c^\infty(E) \equiv \oint_{-\infty}^{\infty} dE' \frac{s_c(E')}{E' - E}, \quad (210)$$

where \oint denotes a (Cauchy) principal-value integral, and neglecting the (weak) energy variation of both these quantities in the internal region, we find in Reich-Moore approximation

$$R_{cc'}^0(E) = \left[R_c^\infty + 2s_c \left(\arctanh \frac{E - \bar{E}}{I/2} + \frac{i\bar{\Gamma}_\gamma I/4}{I^2/4 - (E - \bar{E})^2} \right) \right] \delta_{cc'} \quad (211)$$

The cyclometric or area function $\arctanh x = (1/2) \ln[(1+x)/(1-x)]$ (where ar stands for area, not arcus) is the inverse hyperbolic tangent, also written, in somewhat misleading fashion, $\tanh^{-1} x$ or $\text{arc tanh } x$. The analogous distant-level contribution to the general Wigner-Eisenbud R matrix is obtained if one simply puts $\Gamma_{\lambda\gamma} = 0$ and $\bar{\Gamma}_\gamma = 0$ everywhere:

$$R_{cc'}^0(E) = \left[R_c^\infty + 2s_c \arctanh \frac{E - \bar{E}}{I/2} \right] \delta_{cc'}. \quad (212)$$

If the pole strength is not taken as constant in the internal region but as varying linearly the only modification is that s_c is to be interpreted as $s_c(\bar{E})$ and that an additional term $-s'_c(\bar{E})I$ appears which can, however, be absorbed in R_c^∞ . Experience has shown that it is usually quite adequate for fitting purposes to adjust merely two constants, R_c^∞ and s_c .

The pole strength s_c is related to the strength function S_ℓ commonly employed in applied neutron resonance work by

$$S_\ell \equiv 2k_c a_c s_c \sqrt{1\text{eV}/E}. \quad (213)$$

This follows from the (historical) definition of reduced neutron widths as the neutron widths taken at the arbitrary reference energy $E_r = 1$ eV. For s-wave resonances one has $\Gamma_n^0 = 2P_0(E_r)\gamma_n^2$. The same convention has been used later on also for p-, d-, ... resonances so that reduced neutron widths for single channels $c = \{nJ\ell s\}$ are quite generally defined as $\Gamma_n^c = 2P_0(E_r)\gamma_c^2$. Averaging and dividing by the mean level spacing D_c one obtains the strength function $S_c = 2P_0(E_r)s_c$ with $P_0(E_r) = k_c a_c \sqrt{1\text{eV}/E}$, that is the right-hand side of Eq. (213). The optical model suggests, and experiment confirms, that one can usually take $s_c = s_\ell$, hence $S_c = S_\ell$, which we used on the left-hand side.

The distant-level parameter $R_c^\infty(E)$ is essentially the difference between the contributions to the R matrix from the resonances below and above E . It is negative if the levels below (including the bound ones) have more strength than those above, and positive if the levels above are stronger. Near E the integrand is practically an odd function of $E' - E$, so that the contributions from nearby levels tend to cancel. As a consequence mostly distant levels contribute, whence the name, and typical values are small, $|R_c^\infty| \ll 1$. In applied neutron resonance work the effective nuclear radius,

$$R'_c = a_c(1 - R_c^\infty) \quad (\text{for s-wave channels}), \quad (214)$$

is often employed instead of the distant-level parameter. The reason is that at low energies the potential scattering cross section appears modified by a smooth contribution from the distant levels with the result

$$\sigma_c^{\text{pot}} \rightarrow 4\pi a_c^2 (1 - R_c^\infty)^2 = 4\pi R'^2 \quad \text{for } k_c \rightarrow 0, \quad (215)$$

It has been concluded that the hard-sphere phases ought to be computed as functions of $k_c R'_c$ rather than $k_c a_c$ but that is wrong, notwithstanding fairly common practice and misleading ENDF formats. The effective radius is well defined and applicable only in the low-energy limit, and only for the s-wave. For accurate scattering or total cross section calculations beyond the thermal range one needs the distant-level parameter as the more fundamental, generally valid concept. It modifies the R matrix, not the hard-sphere phase shift.

We conclude that input from optical-model calculations, e. g. from the plots of s- and p-wave strength functions and of effective nuclear radii given in the barn book (Mughabghab et al. 1981, 1984), can be used to estimate the contribution of distant levels. If it is neglected, one gets troublesome edge effects near the boundaries of the internal range (with explicitly given resonances).

3.4.2. Representation of the Edge Terms by Two Broad Resonances

The statistical representation of external levels is quite convenient for cross section parametrisation but an even simpler one is obtained if we approximate the energy-dependent “edge” (area function) terms in Eqs. 211 and 212 by the tails of two very broad resonances of equal strength, located symmetrically below and above the internal range,

$$2s_c \left\{ \arctan \frac{E - \bar{E}}{I/2} + i \frac{\bar{\Gamma}_\gamma}{I} \left[1 - \left(\frac{E - \bar{E}}{I/2} \right)^2 \right]^{-1} \right\} \simeq \frac{\gamma_n^2}{E_- - E - i\Gamma_\gamma/2} + \frac{\gamma_n^2}{E_+ - E - i\Gamma_\gamma/2}. \quad (216)$$

We want to fix the parameters $E_+ - \bar{E} = \bar{E} - E_-$, γ_n^2 , and Γ_γ in such a way that the right-hand side becomes similar to the left-hand side. A suitable degree of similarity is attained, for example, if both sides have equal values, slopes (first derivatives) and curvatures (second derivatives) at the mid-energy \bar{E} . The resulting three equations for three unknowns can be solved rigorously. The solution can be further simplified with $\bar{\Gamma}_\gamma \ll I$ which yields the final approximations

$$E_\pm \simeq \bar{E} \pm \frac{\sqrt{3}}{2} I, \quad (217)$$

$$\gamma_n^2 \simeq \frac{3}{2} I s_c \quad (218)$$

$$\Gamma_\gamma \simeq \bar{\Gamma}_\gamma. \quad (219)$$

(Eq. 218 can be rewritten with Eqs. 174 and 213 as $\Gamma_{n\pm} \simeq (3/2) I S_\ell \sqrt{|E_\pm|/1 \text{ eV}} v_\ell(|E_\pm|)$ with $v_\ell \equiv P_\ell/P_0$.) Insertion on the right-hand side of Eq. 216 reveals that they are tantamount to the approximations $\arctan x \simeq 3x/(3-x^2)$ and $1/(1-x^2) \simeq 3(3+x^2)/(3-x^2)^2$. Fig. 13 shows that the differences between the original functions and the approximations are small over most of the range. Towards the edges they become larger but since the approximations stay finite there, in contrast to the original functions, this is not necessarily

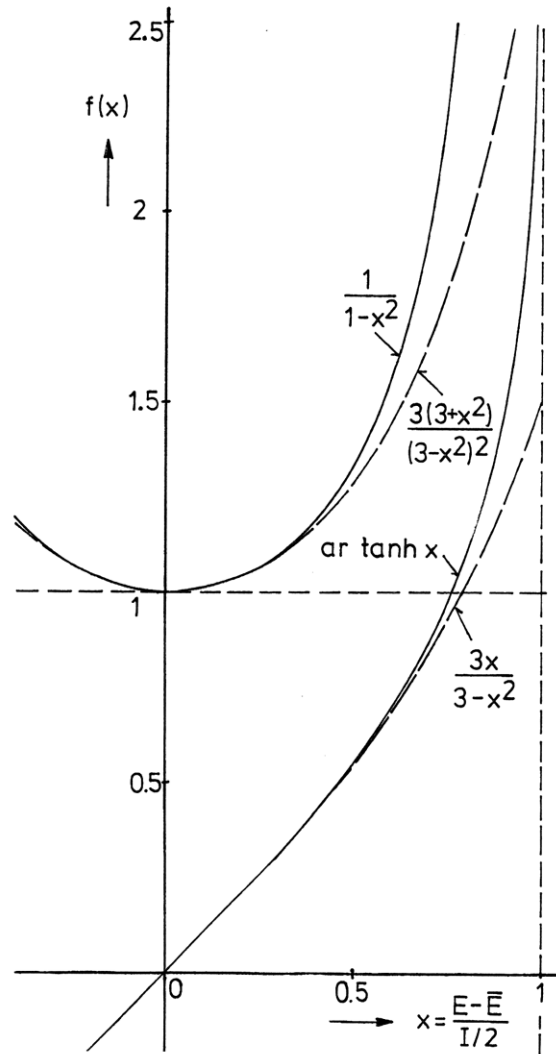


Fig. 13. Comparison of functions describing edge effects. Solid lines: statistically averaged external levels (left-hand side of Eq. 216); dashed lines: approximation by two fictitious broad levels (right-hand side of Eq. 216).

distant-level parameter plus two symmetrically located very broad external levels for the energy-dependent edge effects is convenient and suitable as input for resonance fitting, readily adjusted (if necessary at all) along with the internal resonances. In the current ENDF format it is trivial to store two more resonances but for the distant-level parameter there is no place. It is therefore best to put it equal to zero in ENDF-oriented resonance fitting (which means $R'_c = a_c$ at low energies, see Eq. 214) and adjust the two broad levels independently of each other (together with the internal resonances). Then one of them can become stronger than the other one which produces the imbalance originally described by the distant-level parameter. Fig. 14 shows how well this worked in a recent

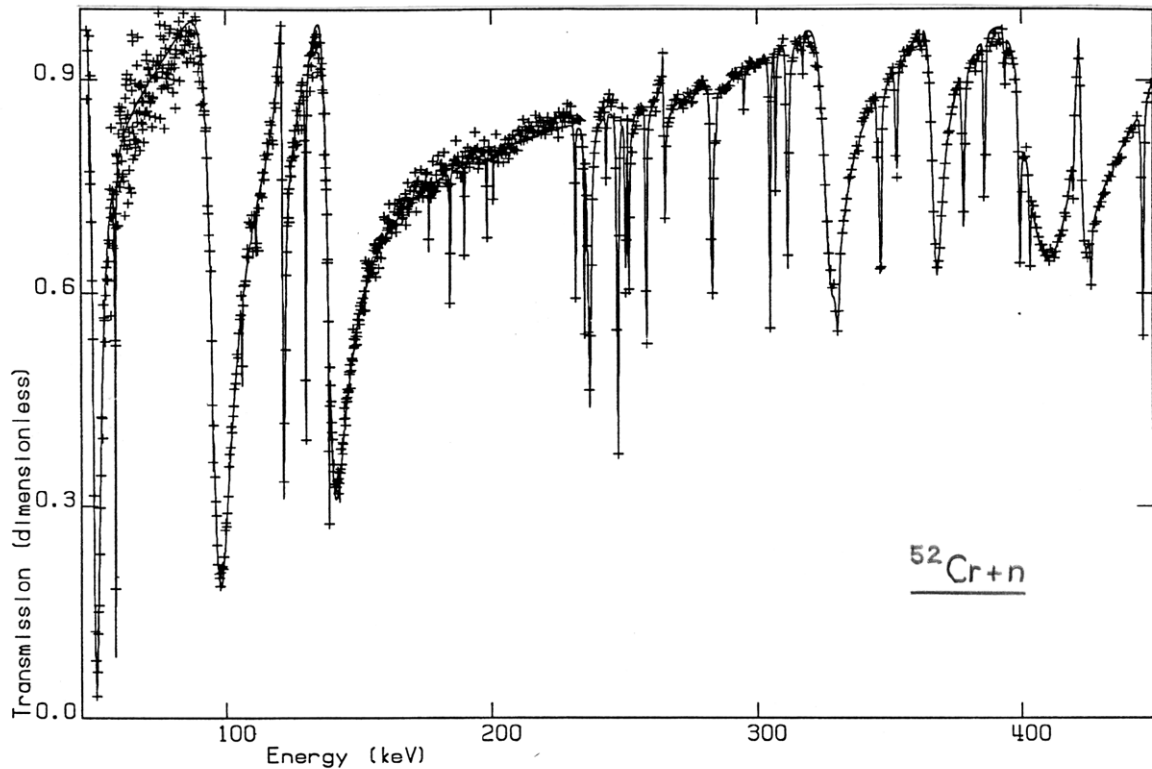


Fig. 14. Transmission data from ORNL (Harvey 1995) fitted with the SAMMY code (Larson and Perey 1980) by O. Bouland (1999). Account of external levels by two broad fictitious levels (with $R_c^\infty = 0$) is seen to be sufficient for a good cross section description all the way down to the lower edge (and in fact also up to the upper edge) of the internal range, 45-1400 keV.

fit to ^{52}Cr transmission data.

In order to speed up point cross section calculations if very many resonances are given explicitly, one can retain only those in a restricted range around the energy of interest and let the others be described summarily by the external-level part (211) of the R matrix. Explicit inclusion of something like fifty levels below and fifty levels above the energy of interest ($I \approx 100 D_c$) is enough for most purposes.

3.4.3. Narrow Bound Level to Enforce Prescribed Thermal Cross Sections Bound Level

The simulation of external levels by the constant distant-level parameter and either the level-statistical (area function) edge term in Eqs. 211 and 212 or a fictitious pair of broad levels with parameters given by Eqs. 217-219 is usually not quite enough to guarantee, together with the known internal levels, the thermal cross sections that for many nuclei are known very accurately. Given thermal cross sections can, however, be reproduced exactly with just one more fictitious bound level (Fröhner 1978, 1981).

Let us consider a nonfissile nucleus ($\sigma_c = \sigma_{cc} + \sigma_{c\gamma}$, $\Gamma = \Gamma_n + \Gamma_\gamma$) for which the parameters of all internal levels are given, and also the level-statistical approximation R_{cc}^0

parameters of all internal levels are given, and also the level-statistical approximation R_{cc}^0 to the external part of the R matrix. These parameters, usually determined by fitting internal resonance data, will not reproduce given thermal cross sections exactly but we can fine-tune by adding one bound (“negative”) level with appropriate parameters. At thermal energies only s-wave resonances must be considered, all other resonances being negligible due to the small centrifugal-barrier penetration factors P_ℓ . With the natural choice $B_0 = 0$ one has $L_c^0 = i\varphi_c = ik_c a_c$, so the Reich-Moore collision function for each s-wave channel (one for zero, two for nonzero target spin) is

$$U_{cc} = e^{-2ik_c a_c} \frac{1 + i \sum_{\lambda} \frac{\Gamma_{\lambda n}/2}{E_{\lambda} - E - i\Gamma_{\lambda\gamma}/2}}{1 - i \sum_{\lambda} \frac{\Gamma_{\lambda n}/2}{E_{\lambda} - E - i\Gamma_{\lambda\gamma}/2}} \quad (c \in n). \quad (220)$$

The summation is over all s-wave levels, internal as well as external ones, that have the spin and parity implied by c . We now split the sum into three parts, namely the internal-level part ($\lambda = 1, 2, \dots, \Lambda$), the external-level part R_{cc}^0 calculated either from Eq. 211 or with two broad levels, Eqs. 217-219, and a third part from the additional bound level ($\lambda = 0$). Solving for the third part ($\lambda = 0$) we get

$$\Delta_{cc} \equiv \frac{i\Gamma_n/2}{E_0 - E - i\Gamma_{\gamma}/2} = - \sum_{\lambda=1}^{\Lambda} \frac{i\Gamma_{\lambda n}/2}{E_{\lambda} - E - i\Gamma_{\lambda\gamma}/2} - ik_c a_c R_{cc}^0 + \frac{U_{cc} - e^{-2ik_c a_c}}{U_{cc} + e^{-2ik_c a_c}}. \quad (221)$$

The right-hand side, denoted by Δ_{cc} , can be calculated from the given resonance parameters and the prescribed thermal cross sections with

$$U_{cc} = \left(1 - \frac{\sigma_c}{2\pi\lambda_c^2 g_c}\right) \pm i \sqrt{\frac{\sigma_{cc}}{\pi\lambda_c^2 g_c} - \left(\frac{\sigma_c}{2\pi\lambda_c^2 g_c}\right)^2}, \quad (222)$$

as follows from the basic equations (152) and (153). Separating real and imaginary part of Eq. (225) one finds

$$\text{Re } \Delta_{cc} = \frac{-\Gamma_n \Gamma_{\gamma}/4}{(E - E_0)^2 + \Gamma_{\gamma}^2/4} < 0, \quad (223)$$

$$\text{Im } \Delta_{cc} = \frac{-(E - E_0)\Gamma_n/2}{(E - E_0)^2 + \Gamma_{\gamma}^2/4} < 0 \quad \text{for } E_0 < 0, \quad (224)$$

and finally

$$E_0 = E - \frac{\text{Im } \Delta_{cc}}{\text{Re } \Delta_{cc}} \frac{\Gamma_{\gamma}}{2}, \quad (225)$$

$$\frac{\Gamma_n}{2} = - \frac{|\Delta_{cc}|^2}{\text{Re } \Delta_{cc}} \frac{\Gamma_{\gamma}}{2}. \quad (226)$$

With only two equations for the three unknowns E_0 , Γ_n , Γ_{γ} we can choose one of them arbitrarily. The weak variation of radiation widths from level to level suggests to set Γ_{γ} equal to the average radiation width of the internal levels,

$$\Gamma_{\gamma} = \bar{\Gamma}_{\gamma}, \quad (227)$$

but exact reproduction of the thermal cross sections is ensured also with any other choice. The sign ambiguity in Eq. 222 is due to the fact that the cross sections depend only on $\text{Re } U_{cc}$ and $|U_{cc}|^2$. Usually the plus sign can be discarded immediately because it yields $E_0 > 0$ contrary to the assumption of a bound level.

For thermally fissile nuclei one finds that Eqs. 225 and 226, although no longer rigorous, are valid in good approximation at least in the usual situation where there is no resonance very close to the thermal energy, so that $\sigma_c \ll 4\pi\lambda_c^2 g_c$. There is now an additional equation for the fission width,

$$\frac{\Gamma_f}{2} = -\frac{|\Delta_{cf}|^2}{\text{Re } \Delta_{cc}} \frac{\Gamma_\gamma}{2}, \quad (228)$$

with

$$|\Delta_{cf}|^2 \equiv \frac{\sigma_{cf}}{\pi\lambda_c^2 g_c} - \sum_{\lambda=1}^{\Lambda} \frac{\Gamma_{\lambda n} \Gamma_{nf}/4}{(E - E_\lambda)^2 + \Gamma_\lambda^2/4} - S_0 \sqrt{\frac{E}{1 \text{ eV}}} \frac{\bar{\Gamma}_f}{I} \left[1 - \left(\frac{E - \bar{E}}{I/2} \right)^{-1} \right], \quad (229)$$

where σ_{cf} is the thermal fission cross section for the entrance channel c .

If, for target nuclei with nonzero spin, the level spins are unknown and only $g\Gamma_n$ is known for the unbound levels but not g and Γ_n separately, one finds the equations

$$E_0 = E - \frac{\text{Im } \Delta_{nn}}{\text{Re } \Delta_{nn}} \frac{g\Gamma_\gamma}{2}, \quad (230)$$

$$\frac{g\Gamma_n}{2} = -\frac{|\Delta_{nn}|^2}{\text{Re } \Delta_{nn}} \frac{g\Gamma_\gamma}{2}, \quad (231)$$

$$\frac{g\Gamma_f}{2} = -\frac{|\Delta_{nf}|^2}{\text{Re } \Delta_{nn}} \frac{g\Gamma_\gamma}{2}, \quad (232)$$

with

$$\Delta_{nn} = -i \sum_{\lambda=1}^{\Lambda} \frac{(g\Gamma_n)_\lambda/2}{E_\lambda - E - i\Gamma_{\lambda\gamma}/2} - ikaR_{nn}^0 + \frac{U_{nn} - e^{-2ika}}{U_{nn} + e^{-2ika}}, \quad (233)$$

$$\Delta_{nf} = -\sum_{\lambda=1}^{\Lambda} \frac{(g\Gamma_n)_\lambda \Gamma_{\lambda f}/4}{(E - E_\lambda)^2 + \Gamma_\lambda^2/4} - S_0 \sqrt{\frac{E}{1 \text{ eV}}} - \frac{\sigma_f}{\pi\lambda^2}, \quad (234)$$

and

$$U_{nn} = \left(1 - \frac{\sigma}{2\pi\lambda^2} \right) \pm i \sqrt{\left(2ka_{\text{coh}} \right)^2 - \left(\frac{\sigma}{2\pi\lambda^2} \right)^2}. \quad (235)$$

The directly observable total and fission cross sections at the thermal energy are sums over the two s-wave channels c (with spins $I + 1/2$ and $I - 1/2$), $\sigma = \sum_c \sigma_c$ and $\sigma_f = \sum_c \sum_{c' \in f} \sigma_{cc'}$. The total widths are to be approximated by $\Gamma_\lambda \simeq (2g\Gamma_n)_\lambda + \Gamma_{\lambda\gamma} + \Gamma_{\lambda f}$. The same external-level term R_{nn}^0 , i. e. the same distant-level parameters, strength functions and average radiation widths were assumed for both spin states, and the channel subscript of λ_c and k_c was dropped. Furthermore we used the relationship between the coherent scattering length a_{coh} and the elastic scattering cross sections for the two spin states,

$$a_{\text{coh}} = \sum_c \sqrt{\frac{g_c \sigma_{cc}}{4\pi}}. \quad (236)$$

Specialisation to one spin state (target spin zero, $g_\lambda = g_c$) or to nonfissile nuclei ($\sigma_f = 0$, $\Gamma_{\lambda f} = \bar{\Gamma}_f = 0$) leads to the equations given above.

With the bound-level parameters calculated analytically in this way the cross sections are usually well reproduced not only at the thermal energy, $E = 0.0253$ eV, but in the entire thermal region below the first resonance. Sometimes it happens, however, that the calculated fictitious bound level is much closer to the neutron threshold than the first unbound level (i. e. $|E_0| < E_1$). Although the calculated capture or fission cross section curve goes through the prescribed point it shows normal $1/v$ behaviour only below the “mirror energy” $|E_0|$. Around that energy the curve changes towards $1/v^5$ behaviour (before the resonance at E_1 makes it rise again). The single-level Breit-Wigner formulae (183) and (185) explain this: The asymptotic behaviour at low and high energies due to a *bound level at $E_0 < 0$* is the same as that of an *unbound level at the mirror energy $|E_0| > 0$* (see Fig. 15). It is easy, however, to restore an actually observed $1/v$ shape all the way up to the first unbound resonance, without changing the correctly calculated thermal cross sections, simply by increasing the (arbitrarily chosen) radiation width that scales the other resonance parameters, see Eqs. 225-226 and 230-232. The computed changeover from $1/v$ to $1/v^5$ can thereby be shifted toward energies above E_1 where it is harmless because other resonances dominate.

3.5. DOPPLER BROADENING

In practical applications resonance cross sections are mostly needed in Doppler-broadened form. It is sometimes claimed that for light nuclei Doppler broadening can be neglected. This may be true for the broad s-wave levels but certainly not for the narrow p-, d-, ... wave levels that in the case of the so-called structural materials (iron, nickel, chromium, cobalt, manganese etc.) contribute significantly to resonance absorption and activation.

3.5.1. Free-Gas Approximation

Doppler broadening in nuclear reactions is caused by the thermal motion of target nuclei. Consider a parallel beam of monoenergetic particles with laboratory velocity \mathbf{v} , colliding with target nuclei whose velocities \mathbf{u} are distributed in such a way that $p(\mathbf{u})d^3u$ is the fraction with velocities in a small three-dimensional region d^3u around \mathbf{u} in velocity space. If ρ_1 and ρ_2 are the densities of beam and target particles, respectively, the number of reactions occurring per unit time and unit volume is

$$\rho_1 \rho_2 \int d^3u p(\mathbf{u}) |\mathbf{v} - \mathbf{u}| \sigma(|\mathbf{v} - \mathbf{u}|) \equiv \rho_1 \rho_2 v \bar{\sigma}(v), \quad (237)$$

where $\sigma(|\mathbf{v} - \mathbf{u}|)$ is the unbroadened cross section for a relative speed $|\mathbf{v} - \mathbf{u}|$ between the collision partners, and $\bar{\sigma}(v)$ the effective or Doppler-broadened cross section for incident particles with speed v . It is obvious from this definition that a $1/v$ cross section is not affected by Doppler broadening. Let us now assume that the target nuclei have the same velocity distribution as the atoms of an ideal gas, viz. the Maxwell-Boltzmann distribution

$$p(\mathbf{u})d^3u = \frac{1}{\pi^{3/2}} \exp\left(-\frac{u^2}{u_T^2}\right) \frac{d^3u}{u_T^3} \quad \left(\frac{Mu_T^2}{2} \equiv kT\right), \quad (238)$$

where M is the mass of the target nucleus and kT the gas temperature in energy units. Integrating over all possible relative velocities $\mathbf{w} \equiv \mathbf{v} - \mathbf{u}$ and employing polar coordinates

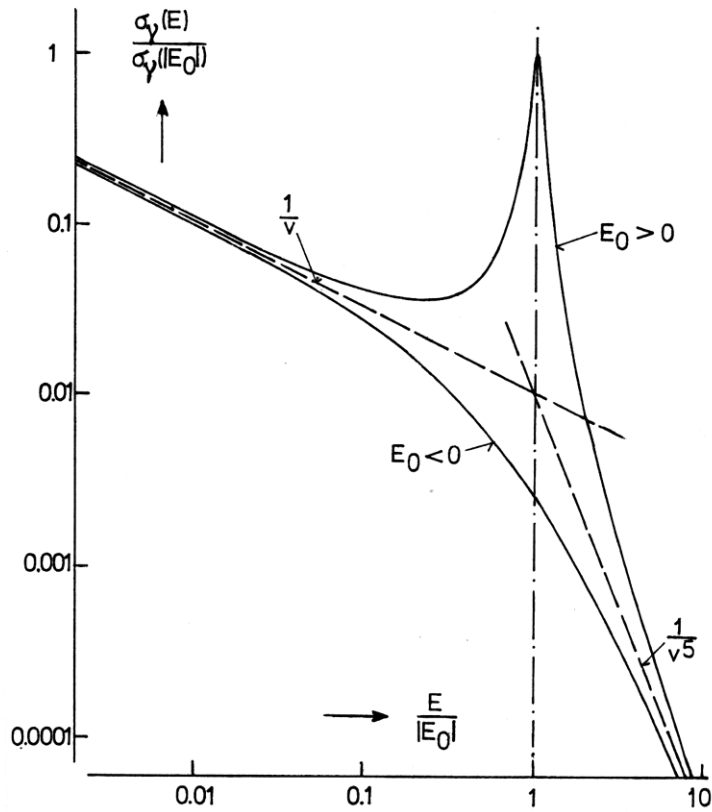


Fig. 15. SLBW capture cross section of an unbound and a bound s-wave level, as given by Eqs. 183, 185 with $\Gamma_n(E) = \Gamma_n(|E_0|)\sqrt{E/|E_0|}$. Both show the same asymptotic behaviour at low and high energies.

with the polar axis parallel to the beam, $d^3u = d^3w = w^2 dw d\mu d\varphi$ with $\mu = \cos\theta$, one finds easily the exact free-gas expression for the Doppler-broadened cross section (Solbrig 1961)

$$\begin{aligned}\bar{\sigma}(v) &= \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{dw}{u_T} \left\{ \exp\left[-\left(\frac{w-v}{u_T}\right)^2\right] - \exp\left[-\left(\frac{w+v}{u_T}\right)^2\right] \right\} \frac{w^2}{v^2} \sigma(w) \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^\infty \frac{dw}{u_T} \exp\left[-\left(\frac{w-v}{u_T}\right)^2\right] \frac{w|w|}{v^2} \sigma(|w|).\end{aligned}\quad (239)$$

This means Gaussian broadening of the odd function $v|v|\sigma(|v|)$ on a speed scale ranging from $-\infty$ to $+\infty$, with a broadening width u_T . In terms of laboratory energies, $E = mv^2/2$, one has

$$\bar{\sigma}(E) = \frac{1}{\Delta\sqrt{\pi}} \int_0^\infty dE' \left\{ \exp\left[-\left(\frac{E - \sqrt{EE'}}{\Delta/2}\right)^2\right] - \exp\left[-\left(\frac{E + \sqrt{EE'}}{\Delta/2}\right)^2\right] \right\} \sqrt{\frac{E'}{E}} \sigma(E'),\quad (240)$$

where

$$\Delta \equiv \sqrt{\frac{4EkT}{M/m}} \quad (241)$$

is called the Doppler width. For $E \gg \Delta$, which is usually satisfied above a few eV, one can simplify by retaining only the first two terms of the expansion $\sqrt{EE'} = E + (E' - E)/2 + \dots$ in the exponent, by neglecting the second exponential, and by shifting the lower limit of the integral to $-\infty$. The result is

$$\sqrt{E}\bar{\sigma}(E) = \frac{1}{\Delta\sqrt{\pi}} \int_{-\infty}^{\infty} dE' \exp\left[-\left(\frac{E' - E}{\Delta}\right)^2\right] \sqrt{E'} \sigma(E'), \quad (242)$$

which means Gaussian broadening of the reaction rate on the energy scale with a width Δ .

3.5.2. Cubic Crystal

Lamb (1939) found the expression (241) also for radiative capture of neutrons by the nuclei of a Debye crystal, in the practically most important case $\Gamma + \Delta > 4kT_D$, where T_D is the Debye temperature that is a measure of the binding force holding the atoms at their positions in the lattice, high for tightly bound and low for weakly bound atoms. The only difference between an ideal gas and a Debye crystal is that one must calculate the Doppler width not with the true temperature T but with an effective temperature T_L given by

$$T_L = T \left(\frac{T}{T_D}\right)^3 \frac{3}{2} \int_0^{T_D/T} dx x^3 \coth \frac{x}{2} = T \left(1 + \frac{1}{20} \frac{T_D^2}{T^2} + \dots\right) \quad (243)$$

that is usually - at room temperature - a few percent higher than T . In the approximation of quasi-free scattering one finds the same result for scattering, and for cubic crystals in general (Fröhner 1970). The correction as a function of T_D/T is given in curve form by Lamb (1939). Problems with the Debye temperature of crystals containing both light and heavy nuclei - example: $^{238}\text{UO}_2$ - are discussed by Lynn (1968).

3.5.3. Gaussian Broadening with Voigt Profiles

In Kapur-Peierls representation, Eqs. 178-180, all resonance cross sections appear as linear superpositions of symmetric and asymmetric line shape profiles (plus a slowly varying potential scattering cross section in case of σ_c and σ_{cc}). Since the shape profiles contain the rapid, resonance-type variations while everything else varies slowly we get Doppler-broadened cross sections in good approximation if we simply replace the unbroadened ("natural") line shapes of the Kapur-Peierls expressions by the Gaussian-broadened profiles introduced by Voigt (1912)

$$\psi_\lambda = \frac{1}{\Delta\sqrt{\pi}} \int_{-\infty}^{\infty} dE' e^{-(E' - E)^2/\Delta^2} \frac{G_\lambda^2/4}{(E' - \tilde{E}_\lambda)^2 + G_\lambda^2/4}, \quad (244)$$

$$\chi_\lambda = \frac{1}{\Delta\sqrt{\pi}} \int_{-\infty}^{\infty} dE' e^{-(E' - E)^2/\Delta^2} \frac{(E' - \tilde{E}_\lambda)G_\lambda/2}{(E' - \tilde{E}_\lambda)^2 + G_\lambda^2/4}, \quad (245)$$

where Δ , \tilde{E}_λ and G_λ are to be taken at $E' = E$. This means that all weak energy dependences are neglected locally, over the range (few Doppler widths) of the Gaussian weight

function, but that their long-range effect is fully taken into account. Doppler broadening by means of the Voigt profiles is popular because fast subroutines are available for their computation (see e. g. Bhat and Lee-Whiting 1967). In Adler-Adler approximation their utilisation is straightforward. In other representations one must first convert from Wigner-Eisenbud to Kapur-Peierls parameters. In SLBW and MLBW approximation this is trivial, one has simply $\tilde{E}_\lambda = E_\lambda + \Delta_\lambda$, $G_{\lambda c}^{1/2} = \Gamma_{\lambda c}^{1/2}$, $G_\lambda = \Gamma_\lambda$ (cf. Eqs. 170-171). In Reich-Moore approximation one must first convert iteratively as explained in Subsect. 3.3.6. This is easy to program and does not add significantly to computing time, especially if used together with a fast algorithm for Gaussian broadening.

3.5.4. Gaussian Broadening with Turing's Method

A fast algorithm for Gaussian broadening of functions having poles in the complex plane (meromorphic functions) was proposed by Turing (1943). The simplest meromorphic function, with a single pole, is the combination $\psi + i\chi$ of natural resonance profiles that we encountered in the resonance formulae of Sect. 3.3. Turing's method is therefore widely used for the calculation of Voigt profiles. He introduced artificial, equidistant poles along the real axis and applied contour integration (see e.g. Bhat and Lee-Whiting 1967) to get

$$\begin{aligned} \psi + i\chi = & \frac{1}{\Delta\sqrt{\pi}} \sum_{n=-\infty}^{\infty} \delta E e^{-(E_n - E)^2/\Delta^2} \frac{i\Gamma/2}{E_n - E_0 + i\Gamma/2} \\ & + \sqrt{\pi} \frac{\Gamma}{\Delta} \frac{e^{-(E - E_0 + i\Gamma/2)^2/\Delta^2}}{1 - e^{-2\pi i(E - E_0 + i\Gamma/2)/\delta E}} P + F, \end{aligned} \quad (246)$$

where δE is the (arbitrary) spacing of the artificial poles, $E_n = E + n\delta E$ is a grid point (artificial pole), and

$$P = \begin{Bmatrix} 0 \\ 1/2 \\ 1 \end{Bmatrix} \quad \text{for} \quad \frac{\Gamma/2}{\Delta} \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \frac{\pi\Delta}{\delta E}, \quad (247)$$

$$|F| \geq \frac{2}{\sqrt{\pi}} \left[1 + \left(\frac{E - E_0}{\Gamma/2} \right)^2 \right]^{1/2} \left| 1 - \left(\frac{2\pi\Delta^2}{\Gamma\delta E} \right)^2 \right|^{-1} \frac{e^{-(\pi\Delta/\delta E)^2}}{1 - e^{-2(\pi\Delta/\delta E)^2}}. \quad (248)$$

We recognise that Turing's approximation consists of (i) a simple sum approximation to the integral with bin width δE , (ii) a term involving the pole energy $E_0 + i\Gamma/2$ and a discontinuous factor P, and (iii) an error term F which becomes small for $\delta E < \Delta$ because of the factor $\exp[-(\pi\Delta/\delta E)^2]$. The pole term is a correction to the sum, needed only in the neighbourhood of narrow peaks (poles close to the real axis) for which the bin width of the sum approximation is too coarse, but negligible elsewhere as specified by the factor P. With the choice $\delta E \simeq 0.7\Delta$ one can neglect the error term completely and still obtain relative accuracies of 10^{-7} or better (Bhat and Lee-Whiting 1967). Applying Turing's method to each term of the Kapur-Peierls cross section expressions (178) or (179) one finds

$$\begin{aligned} \sqrt{E}\sigma(E) \simeq & \frac{1}{\Delta\sqrt{\pi}} \sum_{n=-N}^N \delta E e^{-(E_n - E)^2/\Delta^2} \sqrt{E_n}\sigma(E_n) \\ & + \pi\sqrt{E} \text{Re} \sum_{\lambda} C_{\lambda} G_{\lambda} \frac{e^{-(E - \mathcal{E}_{\lambda})^2/\Delta^2}}{1 - e^{-2\pi i(E - \mathcal{E}_{\lambda})/\delta E}} P_{\lambda}, \end{aligned} \quad (249)$$

where C_λ is the coefficient of $\psi_\lambda + i\chi_\lambda$ in Eq. 178 (for total cross sections) or in Eq. 179 (for partial cross sections), and the factors P_λ are analogous to P , Eq. 247.

The first term on the right-hand side is again the sum approximation to the integral. Due to the rapidly decreasing weight in the wings of the Gaussian one needs only the sum terms with $-5 \leq n \leq +5$ for the usual accuracy of about 0.1% required in typical applications. Moreover, the natural (unbroadened) cross section $\sigma(E_n)$ can be calculated directly from the unconverted Wigner-Eisenbud or Adler-Adler parameters given in the evaluated files. Double sums are not needed: Natural MLBW cross sections are directly obtained from the collision matrix (190), Reich-Moore cross sections from the reduced R matrix (194). In both cases one needs only single sums over levels. The computer time needed for the histogram approximation (first sum in Eq. 250) is therefore practically the same in all four approximations: SLBW, MLBW, Reich-Moore and Adler-Adler.

The pole term in Eq. 249, on the other hand, requires Kapur-Peierls parameters, but only for narrow resonances (nonvanishing P_λ) and only near their peaks where weak energy dependences can be neglected. Adler-Adler parameters need not be converted at all, for SLBW and MLBW the conversion is trivial. Only in Reich-Moore approximation must one convert by iteration as explained in Subsect. 3.3.6, but merely at few energies, namely at the formal resonance energies of the narrow resonances. The extra time needed for this preparation is only a small fraction of the total time required for comprehensive point cross section calculations for which time savings are important.

Turing's method can be applied, of course, not only to Gaussian broadening on the energy scale, Eq. 242, but also to Gaussian broadening on the speed (or momentum) scale with the free-gas kernel, Eq. 239. In the latter case there is even an extra bonus: The width of the Gaussian weight function does not depend any more on energy (or momentum), so the Gaussian weights needed (for $-5 \leq n \leq +5$, say) can be computed once and for all before the calculation begins. Another bonus of Turing's method is the introduction of a natural grid depending only on the effective temperature, which is convenient for fast point cross section calculation, producing automatically less points at higher temperatures where broadened cross sections are smoother. The method is convenient not only for cross section fitting, as is sometimes thought, but quite generally whenever Doppler-broadened multi-level point cross sections are needed. The program DOBRO is written along these lines (Fröhner 1980). Employing the exact free-gas kernel it generates Doppler-broadened MLBW and Reich-Moore cross sections about equally fast as SLBW cross sections from given resonance parameters. The key idea is not to insist on Voigt profiles but to apply the best technique for their computation – Turing's method – directly to the multi-level cross section expressions.

3.5.5. Broadening of Tabulated, Linearly Interpolable Point Data

A widely used method for the generation of Doppler-broadened resonance cross sections starts from natural cross sections σ_k given at energies E_k such that for any intermediate energy E linear interpolation is possible,

$$\sigma(E) = \frac{(E - E_k)\sigma_{k+1} + (E_{k+1} - E)\sigma_k}{E_{k+1} - E_k} \quad (E_k \leq E \leq E_{k+1}) \quad (250)$$

with some specified accuracy. The linear variation with energy translates into a quadratic variation with speed,

$$\sigma(v) = a_k + b_k v^2, \quad (251)$$

where a_k and b_k are constant coefficients. Linearly interpolable point cross section tables are typical for evaluated nuclear data files. Insertion in Eq. 239 yields

$$\bar{\sigma}(v) = \sum_k \int_{w_k}^{w_{k+1}} \frac{dw}{u_T} \frac{w^2}{v^2} \left[e^{-(w-v)^2/u_T^2} - e^{-(w+v)^2/u_T^2} \right] (a_k + b_k w^2). \quad (252)$$

Each sum term corresponds to a linear piece of the cross section representation. Substituting $x = (w-v)/u_T$ we find that for each sum term we need the integrals

$$\frac{2}{\sqrt{\pi}} \int_{x_k}^{x_{k+1}} dt e^{-t^2} t^n = I_n(x_k) - I_n(x_{k+1}) \quad \text{for} \quad n = 0, 1, 2, 3, 4 \quad (253)$$

with

$$I_n(x) \equiv \frac{2}{\sqrt{\pi}} \int_x^\infty dt e^{-t^2} t^n = \frac{1}{\sqrt{\pi}} e^{-x^2} x^{n-1} + \frac{n-1}{2} I_{n-2}(x). \quad (254)$$

I_0 and I_1 are easily calculated whereupon the others can be obtained with the last recursion relation (that results from partial integration):

$$\begin{aligned} I_0(x) &= \operatorname{erfc} x, \\ I_1(x) &= \frac{1}{\sqrt{\pi}}, \\ I_2(x) &= \frac{1}{2} \operatorname{erfc} x + \frac{1}{\sqrt{\pi}} e^{-x^2} x \\ I_3(x) &= \frac{1}{\sqrt{\pi}} e^{-x^2} (x^2 + 1), \\ I_4(x) &= \frac{3}{4} \operatorname{erfc} x + \frac{1}{\sqrt{\pi}} e^{-x^2} \left(x^3 + \frac{3}{4} x \right), \end{aligned} \quad (255)$$

This is the basis of the SIGMA1 code (Cullen and Weisbin 1976). It should be noted that in spite of the title of the paper the method is not exact since the linear interpolation between the tabulated cross sections is an approximation that introduces some error. (In modern evaluated files relative deviations of up to 0.5 % or at best 0.1 % are admitted within each linear piece.) It should also be realised that exponentials and error functions must be calculated for each linear piece of the cross section representation. If the cross sections σ_k are not given but must be calculated first, the SIGMA method is definitely slower and in any case less accurate than the Turing approach, and the choice of an irregular grid permitting interpolation with a specified accuracy, with a minimum of grid points, may be problematic, whereas the Turing method provides a suitable grid automatically.

3.6. PRACTICAL ANALYSIS OF RESONANCE CROSS SECTION DATA

We mentioned in Section 2 that the determination of cross sections from experimental data is best accomplished via extraction of resonance parameters. In fact all resolved resonance cross sections that go into reactor calculations and similar applications are generated from resonance parameters. It might be asked whether one cannot use the best measured high-resolution cross sections directly and thus eliminate the need for resonance parameter extraction. There are several reasons why the determination of resonance parameters cannot be avoided if resonance reactions are to be described and predicted accurately.

- (1) Resonance parameters along with consequent utilisation of resonance theory enable us to represent the often staggering detail of cross section structure by comparably few numbers.

Example: The presently analysed number of resonances of the compound system $^{238}\text{U}+n$ is of the order of 1000. If subthreshold fission is neglected they are specified by about 4000 parameters ($E_0, \Gamma_n, \Gamma_\gamma, J\Pi$) whereas a reasonably accurate pointwise representation of the scattering and capture cross section requires about $5 \cdot 10^4$ data points or 10^5 numbers. If one considers also angular distributions and different temperatures one gets easily several million data points that would be needed to describe the behaviour of ^{238}U in a reactor.

- (2) Doppler broadening of resonances for arbitrary temperatures can be calculated reliably only from resonance parameters but not from point data.
- (3) Resonance parameters and the R-matrix formalism guarantee consistency with physical constraints such as the unitary limits for the total cross section in each reaction channel ($0 \leq \sigma_c \leq 4\pi\lambda_c^2 g_c$) or Wick's limit for scattering in the forward direction ($d\sigma_{cc}(0)/d\Omega \geq \sigma_c^2/(4\pi\lambda_c)^2$).

Another consistency is more subtle but practically at least equally important, especially for the calculation of self-shielding. Theory tells us that there is a rigid relationship between the line shape in one reaction channel and the line shape corresponding to the same compound level in other channels. This relationship is guaranteed if cross sections are generated coherently from resonance parameters, whereas for experimental data sets a common energy scale is always problematic.

- (4) At least equally important is the fact that even the best measured resonance data are affected by resolution and Doppler broadening and (except transmission data) by self-shielding and multiple scattering. The only reliable way to correct for these effects is full-scale parametrisation by fitting resonance-theoretical curves to the data. The fitted quantities should not be some sort of reduced data resembling cross sections, such as logarithms of transmission values, but the observables themselves, e. g. transmissions and capture, fission, or scattering yields. It is then straightforward to include resolution and temperature broadening, self-shielding and multiple scattering, sample impurities and other effects in the theoretical model.
- (5) Extrapolation into the region of unmeasured or unresolved resonances by level-statistical (Hauser-Feshbach) cross section calculations requires statistical parameters such as level densities and strength functions. These in turn must be estimated from resolved resonance parameters.

In order to understand the more practical problems of resonance fitting let us review in some detail the principal types of experimental resonance data that must be modelled by the fitting algorithm. The observables are more or less complicated functions or functionals of the cross sections, rather than cross sections themselves.

3.6.1. Observables

As already mentioned in Sect. 2.1, the simplest measurement is that of the total cross section σ . One measures the fraction of a beam of particles of given energy that traverses without interaction a sample of given areal thickness n (nuclei/b),

$$T = e^{-n\sigma}. \quad (256)$$

The total cross section is thus proportional to the logarithm of the observable.

The (n,x) reaction yield Y_x ($x = f, \gamma, n', p, \alpha, \dots$), i. e. the fraction of beam particles inducing an (n,x) reaction in the sample, is a sum of contributions from events where the (n,x) reaction is preceded by 0, 1, 2, ... scattering collisions,

$$Y_x = Y_{x0} + Y_{x1} + Y_{x2} + \dots, \quad (257)$$

with

$$\begin{aligned} Y_{x0} &= (1 - T) \frac{\sigma_x}{\sigma}, \\ Y_{x1} &= (1 - T) \frac{\sigma_n}{\sigma} \left\langle (1 - T_1) \frac{\sigma_{x1}}{\sigma_1} \right\rangle_1, \\ Y_{x2} &= (1 - T) \frac{\sigma_n}{\sigma} \left\langle (1 - T_1) \frac{\sigma_{n1}}{\sigma_1} \left\langle (1 - T_2) \frac{\sigma_{x2}}{\sigma_2} \right\rangle_2 \right\rangle_1 \\ &\text{etc.} \end{aligned} \quad (258)$$

The numerical subscripts indicate the number of preceding collisions so that $1 - T_1$, for example, is the probability that after the first collision the scattered neutron interacts again somewhere in the sample. The brackets $\langle \rangle_1, \langle \rangle_2, \dots$ denote spatial and angular averages over all possible 1st, 2nd, ... collisions. In each elastic collision the energy of the projectile changes from E to

$$E' = E \frac{A^2 + 2A\mu_c + 1}{(A + 1)^2} \quad (259)$$

if the target particle is initially at rest. Here μ_c is the cosine of the centre-of-mass scattering angle and A the projectile-to-target mass ratio. Note that in the resonance region small energy changes can cause dramatic cross section changes. The multiple-collision yields Y_{x1}, Y_{x2}, \dots are therefore increasingly complicated functionals of the cross sections σ_x, σ_n and σ . If inelastic scattering is energetically allowed the brackets $\langle \rangle_1$ etc. include also averages over all possible scattering modes (residual reactions). The thin-sample approximation,

$$Y_x = n\sigma_x \quad \text{if} \quad n\sigma \ll 1, \quad (260)$$

is often accurate enough for fission yields since fissile samples must be extremely thin so that the fission fragments signalling (n,f) events can get out. In capture data analysis, on the other hand, one must usually include the self-shielding factor $(1 - T)/(n\sigma)$ and multiple-collision contributions because the weak self-absorption of the photons signalling (n, γ) events enables measurers to improve count rates by employing thick samples.

Sample thickness effects, i. e. self-shielding and multiple scattering, are also important in scattering measurements. In analogy to Eqs. 257-258 one has

$$dY_n = dY_{n1} + dY_{n2} + dY_{n3} + \dots \quad (261)$$

with

$$\begin{aligned} dY_{n1} &= \frac{1 - T}{\sigma} \frac{d\sigma_n}{d\Omega} \left\langle T_1 \right\rangle_1 d\Omega \\ dY_{n2} &= \frac{1 - T}{\sigma} \sigma_n \left\langle \frac{1 - T_1}{\sigma_1} \frac{d\sigma_{n1}}{d\Omega_1} \left\langle T_2 \right\rangle_2 \right\rangle_1 d\Omega \\ dY_{n3} &= \frac{1 - T}{\sigma} \sigma_n \left\langle \frac{1 - T_1}{\sigma_1} \sigma_{n1} \left\langle \frac{1 - T_2}{\sigma_2} \frac{d\sigma_{n2}}{d\Omega_2} \left\langle T_3 \right\rangle_3 \right\rangle_3 \right\rangle_1 d\Omega \\ &\text{etc.} \end{aligned} \quad (262)$$

where $d\Omega$ is a solid-angle element covered by the detector.

From our discussion of reaction yields it should be clear that, unless samples are very thin, extraction of (n,x) cross sections from (n,x) yields involves also the total cross section. Quite generally one can say that total cross section data are a prerequisite to good partial cross section analysis. Another data type, valuable in particular for level-statistical tests in the unresolved resonance region is obtained in self-indication measurements. One places two samples in the beam, a filter sample (thickness n_1) and a detector sample (thickness n_2), both consisting of the same material. The probability for a beam particle to induce an (n,x) reaction in the second sample is

$$S_x(n_1, n_2) = T(n_1)Y_x(n_2) . \quad (263)$$

In this way one measures essentially the transmission of the filter sample with a detector system that has enhanced efficiency across the resonance peaks (across the transmission dips).

Ideally the resonance parameter analysis is based on data measured with isotopically pure samples and proceeds more or less as follows.

- (1) From transmission data one determines basically

$$\begin{array}{ll} E_0, \Gamma_n, \Gamma, g & \text{for } \ell = 0 , \\ E_0, g\Gamma_n, & \text{for } \ell \geq 0 . \end{array}$$

- (2) The transmission results permit calculation of sample-thickness corrections for (n,x) yield data from which one obtains basically

$$\begin{array}{ll} E_0, \Gamma_x & \text{if } \Gamma_n, g \text{ are known} \\ E_0, g\Gamma_x & \text{if only } g\Gamma_n \text{ is known} \end{array}$$

- (3) If transmission results are not available (p-, d-, ... levels are not easily seen in transmission measurements) one gets only

$$E_0, g\Gamma_n\Gamma_x/\Gamma \quad \text{if } g\Gamma_n \text{ is not known}$$

In less ideal cases there are complications because of sample impurities – mostly other isotopes of the same element in enriched materials, oxygen in oxide samples or from corrosion, but also hydrogen from adsorbed water vapour. Other unavoidable experimental complications are briefly described in the next subsection.

3.6.2. Experimental Complications

Backgrounds are a main source of uncertainties in resonance analysis. In time-of-flight measurements there are always two types of background: constant and time-dependent. Constant backgrounds may be due to radioactivity of the sample and its environment or to cosmic radiation. Time-dependent backgrounds are induced by the accelerator pulses or by sample effects. An example is the background caused by resonance-scattered neutrons in time-of-flight measurements of neutron transmission or capture. It reflects the resonance structure of the scattering cross section, hence fluctuates violently with time of flight (or energy). This influence of the sample makes “sample-out” background determinations often questionable. Therefore one uses “notch” filters, special samples placed in front of the sample under study. The ideal notch filter has a few widely spaced resonances and is so thick that at the corresponding dips (notches) all beam particles are removed and only the background is observed during the actual run. Of course, no “true” data can be measured across the notches, so one uses several complementary filters. Notch filters

provide an improvement over sample-out background determination but do not completely remove the problems caused by the presence of the sample.

Resolution broadening is another source of complications. All experimental data are resolution broadened. The true observables are

$$\bar{T}(E) = \int dE' r(E', E) T(E), \quad (264)$$

$$\bar{Y}_x(E) = \int dE' r(E', E) Y_x(E), \quad (265)$$

etc., where $dE' r(E', E)$ is the probability that an event observed at the energy E (or the corresponding flight time) was actually due to a beam particle with an energy E' in dE' . The main causes for the deviations $E - E'$ in time-of-flight data are

- finite accelerator burst width (t_b),
- finite time channel width (t_c),
- electronic drift, jitter (t_d),
- uncertain starting point of flight path
(e. g. in moderator slab or booster)
and end point
(e. g. in sample or Li glass detector) (ΔL),
- finite angular resolution ($\Delta\theta$).

The resolution function $r(E', E)$ is often taken as a Gaussian,

$$r(E', E) = \frac{1}{W\sqrt{\pi}} e^{-(E' - E)^2/W^2} \quad (266)$$

with, for instance (Fröhner and Haddad 1965),

$$W = 2E \left[2 \left(\frac{\Delta L}{L} \right)^2 + \frac{E}{3mL^2} (t_b^2 + t_c^2 + t_d^2) \right]^{1/2} = E \sqrt{c_1 + c_2 E}. \quad (267)$$

Slight adjustment of c_1 and c_2 may improve the fit but frequently the true resolution functions have tails and the Gaussians must be replaced by other, asymmetric resolution functions such as χ^2 functions (Fröhner 1978) or Gaussians with tails.

Detector efficiency and flux are a third important source of uncertainties for partial cross section measurements where the observables are count rates,

$$c = \varphi Y_x \epsilon \quad (\simeq \varphi n \sigma_x \epsilon \quad \text{if } n\sigma \ll 1). \quad (268)$$

Absolute determination of the flux φ and the efficiency ϵ is difficult and is therefore avoided where possible. Often one measures relative to a reference sample (subscript r) in the same flux to get

$$\frac{c}{c_r} = \frac{Y_x \epsilon}{Y_r \epsilon_r} \quad (\simeq \frac{n \sigma_x \epsilon}{n_r \sigma_r \epsilon_r} \quad n\sigma \ll 1, \quad n_r \sigma_r \ll 1), \quad (269)$$

where Y_r is known with good accuracy. This eliminates the need to know the flux but one may still have problems with n/n_r and ϵ/ϵ_r as the thin-sample approximation shows. If the energy dependence of ϵ/ϵ_r is known one can calibrate by normalising to an accurately known cross section value, for example the thermal cross section. If no suitable known value is available one can often use the saturated-resonance (“black sample”) technique.

One uses a special sample which is so thick that at a well known resonance the transmission is negligibly small. Quite generally one has

$$(1 - T) \frac{\sigma_x}{\sigma} < Y_x < 1 - T . \quad (270)$$

Because of $c = \varphi Y_x \epsilon$ this yields at the resonance peak, $E = E_0$, where the sample is black, $T \simeq 0$,

$$c < \epsilon \varphi < c \frac{\sigma(E_0)}{\sigma_x(E_0)} . \quad (271)$$

If $\sigma(E_0) \simeq \sigma_x(E_0)$ (i. e. $\Gamma \simeq \Gamma_x$) this defines, without further calculation, a quite accurate value of $\epsilon \varphi$. The 4.91 eV resonance of $^{197}\text{Au} + n$, for example, has frequently been used in this way for black-sample normalisation of capture data. Serious problems are encountered if the detector efficiency varies from isotope to isotope or, even worse, from resonance to resonance. This has been a persistent source of difficulties with capture measurements. Here the detector response depends on the gamma spectrum (binding energy, transition strength to low-lying levels etc.) which fluctuates from level to level in an unpredictable way, especially for light and medium-mass nuclei. The problem could be overcome only with massive liquid-scintillator or crystal scintillator detectors that surround the sample in 4π geometry and absorb most of the capture gamma rays.

Self-shielding and multiple scattering affect mostly neutron capture and scattering data. As the Eqs. 257-258 and 261-262 show the two effects are intertwined and cannot be treated separately. Both together are referred to as sample-thickness effects. An analytical treatment is not possible in the resolved resonance region because of the violently fluctuating scattering and capture cross sections and the need to describe the data in detail, not just in an average sense. The only reliable way is Monte Carlo simulation of multiple-collision neutron histories based on the detailed resonance cross sections, on the appropriate probability distributions for free paths and scattering angles, and on the exact sample geometry (cf. Fröhner 1989 for details). The feasibility of straightforward Monte Carlo simulation of sample-thickness effects in capture resonance fitting was demonstrated with the FANAC code (Fröhner 1978).

3.6.3. Spin and Parity Assignment

The conventional least-squares fitting algorithm employs derivatives (sensitivities), hence it is directly applicable only to continuous probability distributions. Therefore resonance parameter determination by nonlinear least-squares fitting is straightforward only for resonance energies and widths, for which a continuum of possible values exists so that derivatives for Newton-Raphson iteration can be calculated. For spins and parities, having discrete values, derivatives are not available. One could imagine a combinatorial generalisation of the least-squares method including also discrete probability distributions, but in resonance analyses involving dozens or even hundreds of levels the number of possible spin-parity combinations for which least squares must be found is forbidding. Therefore a first assignment of spins and parities is usually based on inspection of transmission data. Most s-wave resonances are easily recognised because pronounced interference between resonant and potential scattering makes them quite asymmetric (see Fig. 11), whereas p- and d-wave resonances tend to be narrower and symmetric because of small potential-scattering cross sections and small centrifugal-barrier penetrabilities at low energies.

A first rough categorisation of those narrow levels for which only $g\Gamma_n$ is known from transmission analysis can be based on the expected average neutron widths. If the expectation values $\langle g\Gamma_n \rangle_{\ell J}$ for the possible (ℓ, J) combinations are given, one can calculate the

corresponding probabilities for (ℓ, J) by means of Bayes' theorem (Bollinger and Thomas 1968). The prior is proportional to the density of levels with spin J and parity $\Pi = (-)^\ell$ which can be taken as independent of parity, $\rho_{\ell J} = \rho_J$. With the rough approximation $\rho_J \propto 2J + 1$ one gets, for example, s-, p- and d-wave level densities in the ratio 1:3:5 if the target spin is zero (see $\sum g_J$ column in Table 2). The likelihood function is given by the Porter-Thomas distribution (see Chapter 4 below)

$$p(g\Gamma_n|\ell, J) d(g\Gamma_n) = \left[\frac{e^{-x} x^{\nu/2}}{\Gamma(\frac{\nu}{2})} \frac{dx}{x} \right]_{\ell J}, \quad 0 < x_{\ell J} \equiv \frac{\nu_{\ell J}}{2} \frac{g\Gamma_n}{\langle g\Gamma_n \rangle_{\ell J}} < \infty, \quad (272)$$

where $\Gamma(\frac{\nu}{2})$ is a gamma function and $\nu_{\ell J}$ the number of channel spins (1 or 2) that can be combined with ℓ to give J (cf. Eqs. 139-140 and Table 2). Since dx/x does not depend on ℓ, J the resulting posterior probability is

$$P(\ell, J|g\Gamma_n) = \frac{\rho_J [e^{-x} x^{\nu/2}]_{\ell J}}{\sum_{\ell, J} \rho_J [e^{-x} x^{\nu/2}]_{\ell J}}, \quad \ell = 0, 1, 2, \dots, \quad |\ell - |I - 1/2|| \leq J \leq \ell + I + 1/2, \quad (273)$$

where I is the target spin. The average widths involve products of level spacings $D_J = 1/\rho_J$ and strength functions S_ℓ ,

$$\langle g\Gamma_n \rangle_{\ell J} = g_J \nu_{\ell J} D_J S_\ell \sqrt{E/1 \text{ eV}} v_\ell(E). \quad (274)$$

Estimates of the average widths can therefore be based on observed mean level spacings and on observed or optical-model strength functions S_ℓ . The functions $v_\ell \equiv P_\ell/P_0$ (with $P_0 = ka$) are relative centrifugal-barrier penetrabilities, equal to unity for the s wave. For the other partial waves one has $v_\ell \simeq (ka)^{2\ell}/[(2\ell - 1)!!]^2$ if $ka \ll \sqrt{\ell(\ell + 1)}$ (i. e. at low energies). This means that in the resonance region only partial waves with $\ell = 0, 1, 2$ and (at most) 3 need be considered. The others are effectively suppressed by high centrifugal barriers. A spin-parity assignment based merely on comparison of the observed width with estimated average widths for the various possible ℓJ combinations is, of course, purely probabilistic, subject to revision if new evidence becomes available such as a characteristic s-wave interference minimum of the total cross section.

The ultimate spin-parity information is provided by scattering data since angular distributions differ markedly for s-, p- and d-levels. Usually one compares precalculated angular distributions for isolated resonances with the observed ones. In practice the resonances are, however, rarely isolated but interfere with other levels having the same spin and parity. Moreover, angular distributions exhibit interference even between different partial waves, for instance between s- and p-wave amplitudes. A certain amount of trial and error concerning spins and parities is therefore unavoidable during the initial phase of resonance fitting, even in the ideal case where high-resolution double-differential data are available. With increasing energy the interference effects become more troublesome and eventually analysis of resolved resonances must be replaced by analysis of average cross section data as is discussed in the following chapter.

4. Statistical Resonance Theory for the Unresolved Region

We have already begun to use resonance statistics when we estimated the contribution of external (“distant”) levels to the R-matrix in Subsect. 3.4.7. Now we shall apply it more systematically to the so-called unresolved resonance region, where limited instrumental resolution permits only observation of resonance-averaged, seemingly smooth cross sections, although resonances exist and make themselves felt in phenomena such as temperature-dependent absorption and self-shielding. Cross section averages, mean square fluctuations (variances) and other cross section functionals such as beam attenuation or temperature-dependent self-shielding can be calculated and predicted at least probabilistically if one knows the statistics of the unseen resonances, in particular the probability distributions of their spacings and partial widths. The Statistical Model of nuclear (and atomic) resonance reactions emerged in the nineteen-fifties (see Porter 1965 for key publications). It is based on the probability theory of Hamiltonian matrices, i. e. on joint probability distributions (“ensembles”) of their matrix elements constrained by their symmetries and other global characteristics. The Gaussian Orthogonal Ensemble of real symmetric matrices was recognised as implying theoretical distributions of partial widths and level spacings that agree with the observed ones. Analytic cross section expressions, in terms of mean level spacings and average partial widths, could be deduced at first only for resonance-averaged total sections, whereas the expressions found for partial cross sections were approximate, valid only in the limit of well separated, weakly overlapping resonances. For strong level overlap this so-called Hauser-Feshbach problem remained unsolved until 1985.

4.1. LEVEL STATISTICS

We begin with basic level statistics, in particular with the (local) distributions of the R-matrix resonance parameters, level energies E_λ and decay amplitudes $\gamma_{\lambda c}$.

4.1.1. The Porter-Thomas Hypothesis

The decay amplitudes $\gamma_{\lambda c}$ of R-matrix theory are essentially values of the internal radial eigenfunctions at the channel entrance, representing the overlap of the λ -th eigenfunction and the external (“channel”) wave function at the matching radius $r_c = a_c$ (see Lane and Thomas 1958, Lynn 1968). For a compound system with $A + 1$ nucleons they are $(3A + 2)$ -dimensional integrals over the surface of the internal region in configuration space. The integrands oscillate rapidly so that positive and negative contributions nearly cancel. The integrals are therefore expected to be close to zero, and positive or negative with equal probability, depending on the unknown particulars of the λ -th eigenstate. Under these circumstances a Gaussian distribution of the $\gamma_{\lambda c}$ with mean zero seems to be a reasonable guess. In fact, the maximum entropy principle of probability theory (see Sect. 1.7) tells us that, if we know only that the distribution has zero mean and finite variance $\langle \gamma_c^2 \rangle$, the most conservative and objective probability distribution for all further inference is indeed the Gaussian,

$$p(\gamma_c | \langle \gamma_c^2 \rangle) d\gamma_c = \frac{1}{\sqrt{\pi}} e^{-x^2} dx, \quad -\infty < x \equiv \frac{\gamma_c}{\sqrt{2\langle \gamma_c^2 \rangle}} < \infty. \quad (275)$$

With $d\gamma_c^2 = 2\gamma_c d\gamma_c$ and $p(\gamma_c | \langle \gamma_c^2 \rangle) d\gamma_c \equiv p(\gamma_c^2 | \langle \gamma_c^2 \rangle) d\gamma_c^2$ this becomes the distribution hypothesised by Porter and Thomas (1956),

$$p(\gamma_c^2 | \langle \gamma_c^2 \rangle) d\gamma_c^2 = \frac{e^{-y}}{\sqrt{\pi y}} dy \quad 0 < y \equiv \frac{\gamma_c^2}{2\langle \gamma_c^2 \rangle} < \infty, \quad (276)$$

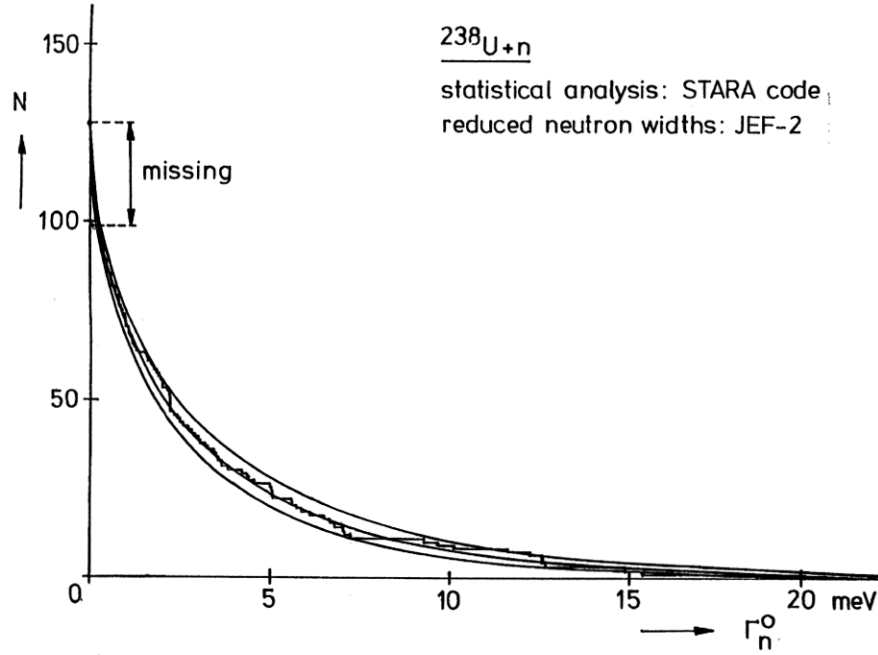


Fig. 16. Porter-Thomas distribution. Histogram: number of s-wave resonances of $^{238}\text{U}+n$ below 3 keV with reduced neutron width exceeding abscissa value. Smooth curves: cumulative Porter-Thomas distribution with confidence band, maximum-likelihood estimate. (From Fröhner 1989).

for partial widths $\Gamma_{\lambda c} = 2P_c\gamma_{\lambda c}^2$ of single channels (and at a given energy). Examples for single-channel widths are neutron widths for $I = 0$ or $\ell = 0$ (see Table 2) or partial radiation widths for single radiative transitions, not only in nuclear but also in atomic and molecular spectroscopy. The single-channel Porter-Thomas distribution agrees well with observed distributions of single-channel reduced neutron widths and single-transition photon widths. It must be realised, however, that the weakest levels – those with the smallest entrance channel widths – tend to get lost in the experimental background. According to the Porter-Thomas hypothesis they are the most frequent ones. Empirical width distributions and, in particular, empirical level densities must therefore invariably be corrected for missing levels. Fig. 16 shows an example of the good agreement between observed and theoretical width distribution above the detection threshold for weak levels. On the other hand it is seen that more than 20% of the levels are missing. (The sample contained resonances of $^{238}\text{U}+n$ from the neutron energy range 0 to 3 keV that belong to the s-wave sequence with an estimated probability of 99% or more.)

Many observable widths are, however, sums of single-channel widths, for instance neutron widths for $I > 0$ and $\ell > 0$ (see Table 2), or total radiation widths, or fission widths. If the averages $\langle\gamma_c^2\rangle$ were the same for all ν contributing channels, such an observable width would obey the generalised Porter-Thomas distribution, i. e. a χ^2 -distribution with ν degrees of freedom,

$$p(\gamma_x^2|\langle\gamma_x^2\rangle)d\gamma_x^2 = \frac{e^{-y}y^{\nu/2}}{\Gamma(\frac{\nu}{2})} \frac{dy}{y}, \quad 0 < y \equiv \frac{\nu}{2} \frac{\gamma_x^2}{\langle\gamma_x^2\rangle} < \infty, \quad (277)$$

where $\Gamma(\frac{\nu}{2})$ is a Gamma function (not a width), and

$$\gamma_x^2 \equiv \sum_{c \in x} \gamma_c^2, \quad (278) \quad \langle \gamma_x^2 \rangle = \nu \langle \gamma_c^2 \rangle. \quad (279)$$

The generalised Porter-Thomas distribution applies to two-channel neutron widths ($\nu = 2$, exponential distribution) and, with an effective (not necessarily integer) number ν of fission channels, to fission widths (ν small) and to total radiation widths (ν large, distribution delta-like: radiation widths fluctuate little from level to level). Large effective ν for total radiation widths are not unexpected because of the usually large number of allowed radiative transitions to lower-lying compound states. That ν is small for total fission widths, however, was a surprise. The hundreds of possible pairs of fission fragments, each with many possible excited states, would seem to imply equally many partial fission widths, and a correspondingly large effective ν .

The puzzle was solved by Å. Bohr (1955). He pointed out that before scission can occur the compound system must pass the saddle point of the potential-energy surface (in the space of deformation parameters) beyond which Coulomb repulsion prevails over nuclear cohesion. At the saddle point most of the excitation energy is tied up as deformation energy, so only little remains for other modes of excitation whose spectrum resembles that of the low-lying states observed at the ground state deformation. Energy, angular momentum and parity conservation allow access to only few of these transition states, regardless of the huge number of final partitions. Therefore the fission channels are correlated in such a way that the fission width can be approximated as a sum over a small number of terms, one for each transition state ("saddle point channel"). For fission, therefore, ν is the effective number of open saddle point channels rather than the number of reaction channels in the usual sense.

This illustrates that the level-statistical "laws" are not as rigid as the resonance formalism discussed in Chapter 3. They hold mainly for highly excited compound states for which all single-particle, collective, or other simplicity is lost. Reflecting more our ignorance than truly random phenomena they are not really applicable where the states considered are simple and well understood. Recognition of the role of collective transition states of a fissioning nucleus, for example, enables us to modify and, in fact, to simplify the statistical description of fission resonances. In the model case of neutron interaction with a complex square-well potential, where everything can be calculated explicitly, nothing at all is random or unspecified, and the reduced neutron widths (essentially squared decay amplitudes) turn out to be all equal instead of exhibiting a Porter-Thomas distribution (see Fröhner 1998a).

4.1.2. Wigner's Surmise and the Gaussian Orthogonal Ensemble

It proved to be much more difficult to find the distribution of level spacings in a given $J\Pi$ level sequence than to find the partial-width distributions. Early in the game Wigner (1957) tried a bold guess. He took issue with the Poisson distribution tried by others according to which the probability of finding a level spacing $E_{\lambda+1} - E_\lambda$ in a small interval dD at D is just proportional to dD , independent of the distance to the preceding level. He pointed out that level energies are eigenvalues of Hamiltonian matrices, and that matrix ensembles always exhibit eigenvalue repulsion (vanishing probability for zero level spacing) so that at least for small spacings the probability should be proportional to DdD . Assuming proportionality also for large D he got immediately what is now known

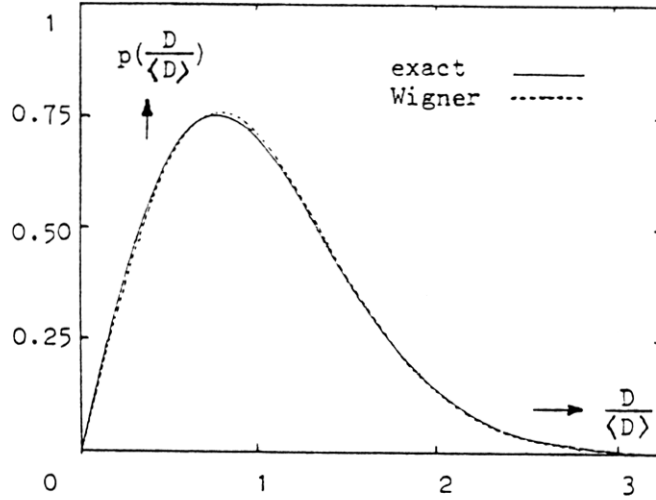


Fig. 17. Next-nearest neighbour spacing distribution for the Gaussian Orthogonal Ensemble of real, symmetric matrices. Solid line: large $N \times N$ matrices, limit $N \rightarrow \infty$ (Gaudin 1961). Dotted line: 2×2 matrices (Wigner distribution, Eq. 280)

as Wigner's surmise,

$$p(D|\langle D \rangle)dD = \exp\left(-c \int_0^D D' dD'\right) c D dD = c e^{-cD^2/2} D dD, \quad 0 < D < \infty, \quad (280)$$

where the proportionality constant is related to the mean level spacing $\langle D \rangle$ by $c = \pi/(2\langle D \rangle^2)$.

The theory of Hamiltonian matrix ensembles (that is, of probability distributions for Hamiltonian matrices) was subsequently developed by Wigner, Porter, Dyson, Mehta and others (see Porter 1965, Brody et al. 1981, Mehta 1991). Nuclear Hamiltonian matrices are Hermitean, of course, but also, due to the practical invariance of nuclear interactions under time reversal, symmetric and thus real. If we knew a probability distribution of such matrices we could derive the corresponding probability distribution of eigenvalues. The simplest ensemble is obtained if we assume nothing but a finite spread of the eigenvalue spectrum, which is well supported by the Gaussian-like eigenvalue distributions resulting from shell model calculations (see Brody et al. 1981). Maximising the entropy of the distribution under the constraint of finite spread (see e. g. Fröhner 1990, 1991a) one finds as the most objective choice

$$p(\mathbf{H}|\sigma^2)d(\mathbf{H}) \propto \prod_{\mu} \exp\left(-\lambda H_{\mu\mu}^2\right) dH_{\mu\mu} \prod_{\mu < \nu} \exp\left(-2\lambda H_{\mu\nu}^2\right) dH_{\mu\nu}, \quad \lambda = \frac{N+1}{4\sigma^2}, \quad (281)$$

where N is the rank of the matrix \mathbf{H} (the number of eigenvalues), $d(\mathbf{H}) = \prod_{\mu < \nu} dH_{\mu\nu}$ the volume element in the space of independent matrix elements, and σ the spread of the eigenvalue spectrum (around its centre at $E = 0$). Having maximal entropy (minimal information content) among all distributions of real, symmetric matrices with given standard

deviation σ , this ensemble plays a similar role for those matrices as the Gaussian distribution does for scalar distributions with given spread. It is called the Gaussian orthogonal ensemble (GOE) because it is invariant under orthogonal transformations and because the matrix elements have independent Gaussian distributions. Actually Wigner derived it from the requirements of rotational invariance (all orthogonal bases must be equivalent in quantum mechanics) and of independently distributed matrix elements, but the independence requirement was criticised as unphysical, in conflict with the predominant two-body character of nuclear forces. In the maximum entropy approach independence is a natural consequence of the limited input information. In any case Wigner's suggestion that the GOE provides a mathematically simple model of level statistics has been fully confirmed. Porter and Rosenzweig (1960) demonstrated that for very large matrices (very many compound states) the GOE yields the Porter-Thomas distribution of partial widths. The GOE level spacing distribution for 2×2 matrices is exactly Wigner's surmise, while for larger matrices it is very close as shown by Mehta (1960) and Gaudin (1961), see Fig. 17.

The level spacings are correlated in such a way that a relatively large spacing is followed by a short one more often than not, and vice versa. The resulting correlation coefficient is

$$\rho(D_\lambda, D_{\lambda+1}) \equiv \frac{\text{cov}(D_\lambda, D_{\lambda+1})}{\sqrt{\text{var}(D_\lambda) \text{var}(D_{\lambda+1})}} \simeq -0.27 \quad (282)$$

for large matrices. The eigenvalue sequence is therefore remarkably regular ("stiff"), with almost equidistant level positions, differing noticeably from a "random" sequence with Poisson interval distribution. All this is in excellent agreement with observed nuclear (and atomic) level statistics, at least in limited energy ranges where slow variations of the mean level spacing and of average partial widths can be neglected. Seeming deviations from GOE predictions usually vanish if the long-range ("secular") variations of the level-statistical parameters are properly taken into account.

The Gaussian orthogonal ensemble, constrained only by the finite spread of the eigenvalue spectrum, cannot be expected to reproduce more specific nuclear features such as fermion gas level densities, shell effects, giant dipole resonances or fission barriers. In fact, the semicircular GOE level density obtained by Wigner (1957) differs from the Gaussian-like densities found in more realistic shell model calculations (see Brody et al. 1981). Although the distributions of level energies and partial widths can locally be taken as those of the GOE, their parameters (level density, average widths) vary slowly with energy. These secular variations are described by macroscopic models of the nucleus – level densities, for instance, by the fermion gas model or, at higher energies, by the shell model with residual interaction; neutron, proton and alpha particle strength functions by the optical model; photon strength functions by the giant-dipole resonance model; fission strength functions by fission barrier models.

4.1.3. Transmission Coefficients

The appropriate theory for the unresolved resonance region is Hauser-Feshbach theory with width fluctuation corrections. It is obtained if one averages R-matrix cross section expressions over the GOE. The essential parameters are strength functions or the closely related transmission coefficients. For particle (neutron, proton, alpha) channels the latter are defined by

$$T_c \equiv 1 - |\bar{U}_{cc}|^2 = \frac{4\pi s_c P_c}{|1 - \bar{R}_{cc} L_c^o|^2}. \quad (283)$$

The denominator, with

$$\bar{R}_{cc} \equiv R_c^\infty + i\pi s_c, \quad (284)$$

accounts for overlap and interference effects due to nearby and distant levels, s_c and R_c^∞ being the pole strength and the distant-level parameter already encountered in the context of external levels, Eqs. 209 and 210. The transmission coefficient is thus essentially 2π times the ratio of average effective particle (e. g. neutron) width to mean level spacing. For photon and fission channels one uses analogously

$$T_\gamma = 2\pi \frac{\bar{\Gamma}_\gamma}{D_c}, \quad (285) \quad T_f = 2\pi \frac{\bar{\Gamma}_f}{D_c}. \quad (286)$$

Transmission coefficients for particle channels can be obtained from the optical model which describes the interaction between an incident particle and the target nucleus by a complex potential, in analogy to the diffraction and absorption of light by the proverbial cloudy crystal ball. The complex potential is adjusted so that the resonance-averaged total cross section is well reproduced over the whole unresolved resonance range and far above. For nuclei of similar size, e. g. actinides like ^{235}U , ^{238}U and ^{239}Pu , one expects similar potential wells and thus similar average total cross sections. This is in fact what one observes – see for instance the neutron transmission measurements and optical-model fits performed by Poenitz et al. (1981) on a whole series of heavy nuclides. The optical model has therefore considerable predictive and systematising power.

It must be kept in mind, however, that it is essentially a single-particle model where scattering means direct (potential) scattering only, while absorption means compound nucleus formation and includes not only final radiative capture and fission but also compound scattering, i. e. reemission of the incident particle (or one indistinguishable from it) from the compound nucleus. Moreover, the angular distributions calculated from a complex potential that reproduces the total cross section correctly are not exactly equal to the resonance-averaged double-differential scattering derived from the R matrix formalism. The only directly observable cross section type obtained from the optical model is thus the (average) total cross section. Fits to scattering or other partial cross section data require due account of compound scattering except far above the unresolved resonance region where compound processes are unimportant.

The main insights gained from optical-model studies of neutron induced reactions are that the strength functions and distant-level parameters vary but little over the relatively narrow resolved resonance range, and little from nucleus to nucleus. Furthermore, they can usually be taken as depending only on orbital angular momentum as was mentioned already in the context of Eq. 213. Transmission coefficients for inelastic, for example (n,n') or (p,p'), channels can be calculated with the same expressions as for elastic (n,n) or (p,p) channels but with the particle energy E replaced by $E - E_c$, where E_c is the excitation energy transferred to the residual nucleus.

The total photon transmission coefficient for the $J\Pi$ resonance sequence is dominated by electric and magnetic dipole transitions,

$$T_\gamma^{J\Pi} = \sum_{c' \in \gamma, J\Pi} [T_{c'}^{E1} + T_{c'}^{M1}], \quad (287)$$

where the summation is over all accessible exit channels c' , i. e. all allowed dipole transitions from the compound state with spin J and parity Π to lower lying levels. The electric dipole contributions are commonly taken as having the classical Lorentz form of giant dipole resonances,

$$T_{c'}^{E1} \propto \frac{E_\gamma^4}{(E_\gamma^2 - E_0^2)^2 + (\Gamma E_\gamma)^2}, \quad (288)$$

where E_γ is the photon energy of the transition, E_0 and Γ are energy and width of the giant dipole resonance where all protons vibrate against all neutrons. Empirically it is found that for spherical compound nuclei with A nucleons one has $E_0 \simeq \text{MeV}/A^{1/6}$, $\Gamma \simeq 33 \text{ MeV}/A^{1/3}$. For magic nuclei Γ is smaller by a factor of 0.6, for near-magic nuclei with Z or N differing from a magic number by 1 or 2 the factor is about 0.8, and for deformed nuclei it is about 1.2 (see Holmes, Woosley, Fowler and Zimmerman 1976). The magnetic dipole contributions are smaller. They are often approximated by the simple Weisskopf estimate

$$T_c^{M1} \propto E_\gamma^3 \quad (289)$$

or neglected altogether. The sum (287) over final levels, for arbitrary excitation energy U , can be calculated as an integral $\int_0^U dE_\gamma \rho_J(U - E_\gamma) \dots$, with e. g. a Gilbert-Cameron level density ρ_J , and normalised to the photon strength function $2\pi\rho_J \bar{\Gamma}_\gamma^{J\Pi}$ in the resolved resonance region (empirically obtained at least for s-wave resonances from their radiation widths and spacings, inferred for others via the theoretical spin distributions discussed below).

Fission transmission coefficients were derived by Hill and Wheeler (1939). For single-hump fission barriers of identical parabolic shapes for all transition states they have the form

$$T_f^{J\Pi} = \sum_{c \in f, J\Pi} \frac{1}{1 + \exp[2\pi(E_c - E)/\hbar\omega]}, \quad (290)$$

where E is the energy of the incident particle, E_c the height of the barrier on the same energy scale, $\hbar\omega$ is proportional to the inverse curvature of the parabola, and the sum is over all fission transition states (saddle point channels) consistent with J and Π . For double-hump barriers it is often sufficient to combine the transmission coefficients T_A for the inner barrier and T_B for the outer one by adding reciprocals,

$$\frac{1}{T_f^{J\Pi}} = \frac{1}{T_A^{J\Pi}} + \frac{1}{T_B^{J\Pi}}, \quad (291)$$

in analogy to resistors in series. More general expressions are given for example by Vandenbosch and Huizenga (1973). Again the sum (290) must be calculated as an integral over a suitable density of transition states, see Lynn (1974).

Mean level spacings or their reciprocals, nuclear level densities, are seen to play the role of scale factors in the theory. Their spin and energy dependence has a strong influence on the behaviour of resonance-averaged cross sections and will be discussed next.

4.1.4. Nuclear Level Densities

Compound nuclear levels can be observed in two energy regions – near the ground state up to few MeV (e. g. by neutron capture gamma ray spectroscopy or Coulomb excitation), and at the neutron separation energy of about 7 MeV (by observation of resonances in neutron and proton induced reactions). At those higher excitation energies the level density is found to be several orders of magnitude larger than near the ground state. An explanation of such a rapid increase of level densities with excitation energy must start from the basic features of nuclei that are incorporated in the nuclear shell model: The nucleons, obeying Fermi-Dirac statistics and therefore Pauli's exclusion principle, move almost independently in the potential well created by their mutual interaction. Let us denote the ν -th energy eigenvalue of the well by ϵ_ν , and the occupation number of the

ν -th level in the i -th nuclear state by $n_{i\nu}$ (0 or 1 for fermions). For independent nucleons the total nucleon number and the total energy of the i -th nuclear state are then

$$N_i = \sum_{\nu} n_{i\nu} , \quad n_{i\nu} = 0, 1 . \quad (292)$$

$$E_i = \sum_{\nu} n_{i\nu} \epsilon_{\nu} , \quad \epsilon_{\nu} > 0 . \quad (293)$$

The actual two-dimensional density of compound nuclear states,

$$\rho(N, E) = \sum_i \delta(N - N_i) \delta(E - E_i) , \quad (294)$$

admits only discrete possibilities. A smooth density can be obtained if we prescribe arbitrary non-negative values of N and E as weighted averages

$$N = \sum_i p_i N_i , \quad (295) \quad E = \sum_i p_i E_i . \quad (296)$$

The maximum entropy principle tells us how to choose the weights p_i under these two constraints. The most conservative choice, ensuring a minimum of spurious information, is the grand-canonical ensemble (first introduced by Gibbs in thermodynamics),

$$p_i = \frac{1}{Z} e^{\alpha N_i - \beta E_i} , \quad (297) \quad Z = \sum_i e^{\alpha N_i - \beta E_i} , \quad (298)$$

with Lagrange multipliers α and β . Noting that the partition function Z is the Laplace transform of the level density ρ , Eq. 294, we conclude that the level density must be related to the partition function by inverse Laplace transformation,

$$Z(\alpha, \beta) = \int_0^{\infty} dN \int_0^{\infty} dE \rho(N, E) e^{\alpha N - \beta E} , \quad (299)$$

$$\begin{aligned} \rho(N, E) &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\alpha \int_{-i\infty}^{i\infty} d\beta Z(\alpha, \beta) e^{-\alpha N + \beta E} \\ &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} d\alpha \int_{-i\infty}^{i\infty} d\beta e^S , \end{aligned} \quad (300)$$

where S is the information entropy for arbitrary Lagrange parameters α and β ,

$$S = - \sum_i p_i \ln p_i = \ln Z - \alpha N + \beta E . \quad (301)$$

Saddle point integration, i. e. expansion of S about its maximum at $\alpha = \hat{\alpha}$, $\beta = \hat{\beta}$ and truncation after the quadratic terms, yields the following remarkable relationship between level density and entropy

$$\rho(N, E) \simeq \frac{e^{\hat{S}}}{\sqrt{\det(2\pi \nabla \nabla^{\dagger} \hat{S})}} , \quad (302)$$

where we introduced the differential vector operator $\nabla \equiv (\partial/\partial\hat{\alpha} \quad \partial/\partial\hat{\beta})^{\dagger}$. The Lagrange parameters $\hat{\alpha}$, $\hat{\beta}$ at the maximum are just those following from the maximum entropy

algorithm, and the maximised information entropy $\hat{S} \equiv S(\hat{\alpha}, \hat{\beta})$ is the physicists' thermodynamic entropy divided by Boltzmann's constant.

Let us consider the partition function. We note that summation over all possible compound nuclear states is the same as summation over all possible sets of fermion occupation numbers,

$$Z = \sum_i e^{\alpha N_i - \beta E_i} = \prod_{\nu} (1 + e^{\alpha - \beta \epsilon_{\nu}}). \quad (303)$$

Expanding the last product one recognises in fact that each state is represented by one sum term, each sum term being a product of exponentials for all the occupied levels and of unit factors for the empty ones, as demanded by Eqs. 292 and 293. Taking the logarithm and approximating the sum by an integral one obtains

$$\ln Z = \sum_{\nu} \ln(1 + e^{\alpha - \beta \epsilon_{\nu}}) \simeq \int_0^{\infty} d\epsilon g(\epsilon) \ln(1 + e^{\alpha - \beta \epsilon}), \quad (304)$$

where $g(\epsilon)$ is the density of single-particle levels. In the ground state, with total energy E_0 , all levels are occupied up to the so-called Fermi edge ϵ_F , so that

$$N = \int_0^{\epsilon_F} d\epsilon g(\epsilon), \quad (305) \quad E_0 = \int_0^{\epsilon_F} d\epsilon g(\epsilon) \epsilon. \quad (306)$$

The nucleus is thus described as a condensed (“degenerate”) fermion gas. The condensation is weakened as excitation increases and more and more empty levels are created below the Fermi edge while levels above are filled. As long as only a relatively narrow energy band around the Fermi edge is affected, where the energy variation of $g(\epsilon)$ is negligible, one can use the approximation

$$\ln Z \simeq \alpha N - \beta E_0 + \frac{g(\epsilon_F)}{\beta} \left(\frac{(\alpha - \beta \epsilon_F)^2}{2} + \frac{\pi^2}{6} \right) \quad (307)$$

(see e. g. Bohr and Mottelson 1969). Entropy maximisation with this partition function yields two coupled equations for $\hat{\alpha}$ and $\hat{\beta}$,

$$\hat{\alpha} = \hat{\beta} \epsilon_F, \quad (308) \quad E - E_0 = \frac{\pi^2}{6} \frac{g(\epsilon_F)}{\hat{\beta}^2}, \quad (309)$$

and finally the fermion gas level density formula

$$\rho(N, E) \simeq \frac{\exp \sqrt{4aU}}{\sqrt{48} U}, \quad (310)$$

where $U \equiv E - E_0$ is the excitation energy, and $a \equiv (\pi^2/6)g(\epsilon_F)$, called the fermion gas level density parameter, depends on N because of Eq. 305.

Van Lier and Uhlenbeck pointed out, following a hint by Goudsmit, that in the special case of equidistant single-particle levels, i. e. for a harmonic-oscillator potential, the fermion gas level density can be calculated exactly (see Ericson 1960). The possible excitation energies are integer multiples of the spacing $d \equiv 1/g$. For $U/d = 1, 2, 3, 4 \dots$ one has 1, 2, 3, 5 \dots different states (occupation patterns). As can be seen from Fig. 18 (top) the number of states is equal to the number of different partitions of the integer U/d , partition being defined here in the number-theoretic sense as a decomposition in positive

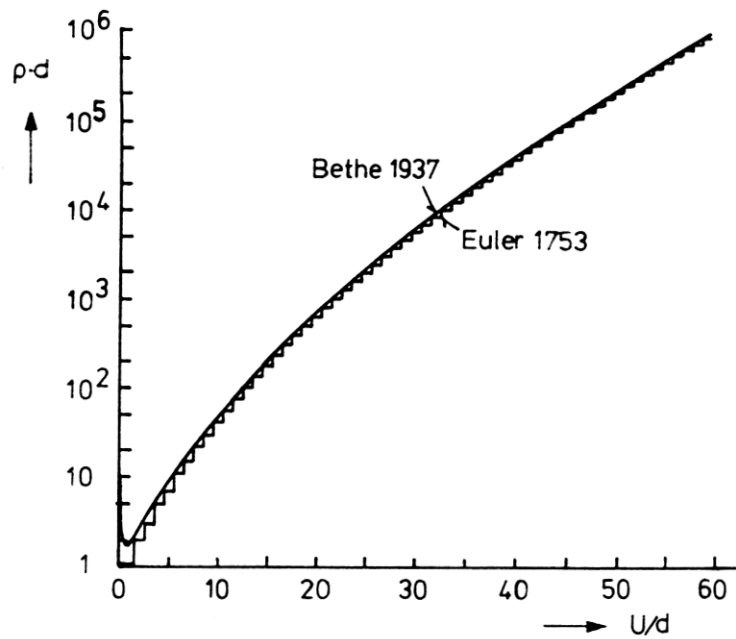
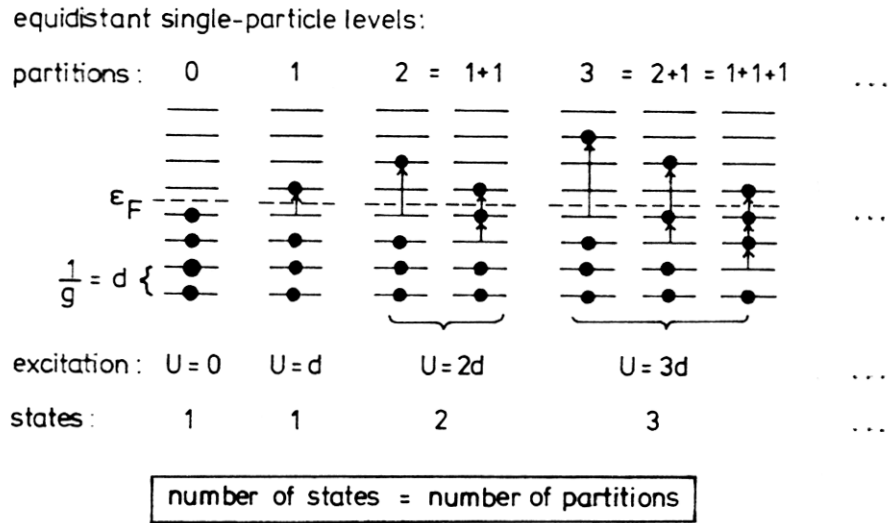


Fig. 18. Fermions occupying equidistant single-particle levels. Top: ground and first excited states. Bottom: Exact level densities from number theory (steps) and approximation from entropy maximisation (smooth curve).

integer summands. The number of partitions can be calculated with a recursion formula due to Euler (1753). The resulting rigorous level density histogram is plotted together with the approximate fermion gas curve in Fig. 18 (bottom). The agreement is good except at the lowest excitation energies. The rapid, almost exponential rise of the level density with increasing energy is evident.

So far we neglected the difference between protons and neutrons and their spins. It is straightforward to generalise to nuclei with Z protons, N neutrons, and spin orientation quantum number M . The result is Bethe's (1937) famous level density formula (see also Gilbert and Cameron 1965)

$$\rho(Z, N, E, M) \simeq \frac{\exp \sqrt{4a[U - M^2/(2g\langle m^2 \rangle)]}}{12\sqrt{2g\langle m^2 \rangle}[U - M^2/(2g\langle m^2 \rangle)]^{3/2}} \quad (311)$$

with

$$g \equiv g_p + g_n, \quad (312) \quad g\langle m^2 \rangle \equiv g_p\langle m_p^2 \rangle + g_n\langle m_n^2 \rangle, \quad (313)$$

where g_p and g_n are the single-particle level densities for protons and neutrons, m_p and m_n their spin orientation quantum numbers. (The potential well and hence the single-particle levels for protons differ from those for neutrons because of the Coulomb interaction.) Usually $M^2/(2g\langle m^2 \rangle)$ is much smaller than U . This leads to the approximate factorisation

$$\rho(Z, N, E, M) \equiv \omega_M(U) \simeq \omega(U) \frac{e^{-M^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \quad (314)$$

into the total state density

$$\omega(U) = \sum_{M=-\infty}^{\infty} \omega_M(U) = \frac{\sqrt{2\pi}}{3} \frac{e\sqrt{4aU}}{(4aU)^{5/4}} a, \quad (315)$$

and a Gaussian distribution over the various orientation quantum numbers M , with variance

$$\sigma^2 = g\langle m^2 \rangle \sqrt{\frac{U}{a}}. \quad (316)$$

The Gaussian is correctly normalised to unity since the Euler-MacLaurin summation formula yields $\sum_{M=-\infty}^{\infty} e^{-M^2/2\sigma^2} = \int_{-\infty}^{\infty} dM e^{-M^2/2\sigma^2} = \sqrt{2\pi\sigma^2}$ first for integer but then also for half-integer M . The standard deviation σ is often called the spin cut-off. Typical values are $\sigma \sim 3$ for medium weight nuclides such as the structural materials Fe, Ni, Cr and $\sigma \sim 4.5$ for actinides like Th, U, Pu.

It must be realised, however, that in the absence of external fields one can distinguish, and count as separate resonances, only nuclear states $|J, M\rangle$ with different total angular momentum (level spin) J . States differing merely in spin orientation, $M = -J, -J+1, \dots, J$, are degenerate, hence indistinguishable. This means that we ought to count only one of these alternatives, for example only the states with $M = J$, $|J, J\rangle$, if we want the density ρ_J of levels with given J . Now the states contributing to ω_J and ω_{J+1} can be arranged in two columns as follows,

$$\begin{array}{ll} |J, J\rangle & \\ |J+1, J\rangle & |J+1, J+1\rangle \\ |J+2, J\rangle & |J+2, J+1\rangle \\ \vdots & \vdots \end{array}$$

$$f_J = e^{-J^2/2\sigma^2} - e^{-(J+1)^2/2\sigma^2}$$

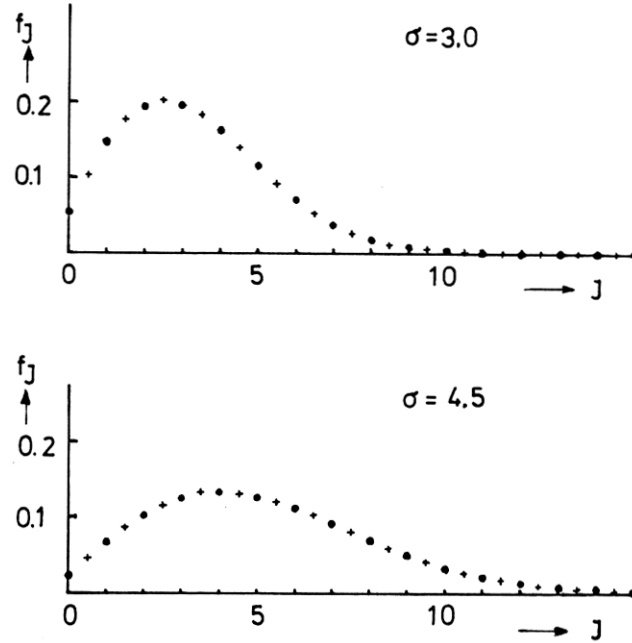


Fig. 19. Level spin distributions for $\sigma = 3$ (roughly corresponding to medium-weight nuclei such as Fe, Ni, Cr) and $\sigma = 4.5$ (roughly corresponding to actinides such as Th, U, Pu). Dots: integer spin, crosses: half-integer spin.

with states in the same row, e. g. $|J+1, J\rangle$ and $|J+1, J+1\rangle$, being degenerate. If we form the difference $\omega_J - \omega_{J+1}$ all contributions from the first column are therefore canceled by those from the second one – except for the contribution from the states $|J, J\rangle$ that is equal to the wanted level density ρ_J (Bethe 1937). Thus one has, with the approximate factorisation (314),

$$\begin{aligned} \rho_J = \omega_J - \omega_{J+1} &\simeq \frac{\omega}{\sqrt{2\pi\sigma^2}} \left[\exp\left(-\frac{J^2}{2\sigma^2}\right) - \exp\left(-\frac{(J+1)^2}{2\sigma^2}\right) \right] \\ &= \frac{2\omega}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{8\sigma^2}\right) \sinh\left(\frac{J+1/2}{2\sigma^2}\right) \exp\left(-\frac{(J+1/2)^2}{2\sigma^2}\right). \end{aligned} \quad (317)$$

This is the density of all levels with spin J , regardless of parity. The level density for either parity (+ or -) can normally be taken as just half as big (see Ericson 1960, Gilbert and Cameron 1965). Widely used approximations are obtained by replacing the hyperbolic sine by its argument or, more drastically, by putting $\rho_J \propto 2J+1 \propto g_e$ which, however, is good enough only if $J \ll \sigma$, i. e. for the two or three smallest possible spins. Fig. 19 shows the exact distribution $f_J = \exp[-J^2/2\sigma^2] - \exp[-(J+1)^2/2\sigma^2]$ for $\sigma = 3.0$ and $\sigma = 4.5$, values that roughly apply to the resolved neutron resonance region of structural materials like Fe, Ni, Cr, and to actinides like Th, U, Pu, respectively. We note that the distribution is normalised to unity – for integer spins exactly, for half-integer spins almost exactly (to $e^{-1/8\sigma^2} \simeq 1$).

Although the number-theoretic approach to nuclear level densities looks promising also for realistic (non-equidistant) single-particle level schemes (Anzaldo 1995), most cur-

rent level density theories that include explicit shell model levels, residual interaction, deformation, collective rotations and vibrations, and the superfluidity of nuclear matter at low excitation energies, are based on Bethe's thermodynamical (maximum entropy) approach.

An example is the widely used composite level density formula of Gilbert and Cameron (1965). It accounts in a heuristic way for nucleon pairing effects and for the empirically observed behaviour of level densities near the ground state where collective modes preclude a purely statistical treatment of compound states. It is composed of two parts, a constant-temperature part valid at low excitation energies and a smoothly joining fermion gas (Bethe) part valid at high energies, with parameter systematics derived from a large body of nuclear structure data near the ground state and of resonance data above the neutron and proton separation energies. The density of levels with spin J and arbitrary parity is described by Eq. 317 with

$$\frac{\omega}{\sqrt{2\pi\sigma^2}} = \begin{cases} \frac{\sqrt{2}}{3} \frac{\exp \sqrt{4aU}}{(4aU)^{3/2}} \frac{a}{\sqrt{c} A^{1/3}} & \text{if } U \geq U_x, \\ \frac{\sqrt{2}}{3} \frac{\exp \sqrt{4aU_x}}{(4aU_x)^{3/2}} \frac{a}{\sqrt{c} A^{1/3}} \exp\left(\frac{U - U_x}{T}\right) & \text{if } U \leq U_x, \end{cases} \quad (318)$$

the high-energy part of which is obtained if one puts, following Jensen and Luttinger (1952), $\langle m^2 \rangle = cA^{2/3}$ with $A = Z + N$ and $c = 0.146$. The effective excitation energy,

$$U = B + E - P(Z) - P(N), \quad (319)$$

is taken as the sum of the neutron binding energy B and the kinetic neutron energy E , corrected for the energies $P(Z)$ and $P(N)$ that are needed for pair breaking if all protons or all neutrons are paired, i. e. if the proton number Z or the neutron number N is even. Below the matching energy the spin cut-off is often taken as a linear function of U , vanishing at the ground state, so that

$$2\sigma^2 = \begin{cases} cA^{2/3} \sqrt{4aU}, & U \geq U_x, \\ cA^{2/3} \sqrt{4aU_x} \frac{U + P(Z) + P(N)}{U_x + P(Z) + P(N)} & U \leq U_x \end{cases} \quad (320)$$

Bethe's fermion gas form for high energies and the constant-temperature form for low energies are required to join smoothly at the matching energy U_x . The temperature T (in energy units) following from the matching condition is given by

$$T = \left(1 - \frac{3}{\sqrt{4aU_x}} + \frac{J_{\min}}{4aU_x cA^{2/3}}\right)^{-1} \sqrt{\frac{U_x}{a}}, \quad J_{\min} = \begin{cases} 0 & \text{for even } A, \\ \frac{1}{2} & \text{for odd } A, \end{cases} \quad (321)$$

where the third term is usually negligible. Typical values in the resolved resonance region are $T \sim 1.4$ MeV for structural materials like Fe, Ni, Cr and $T \sim 0.4$ MeV for actinides like U and Pu. Gilbert and Cameron give empirical parameters a , U_x , $P(Z)$, $P(N)$ for many compound nuclei, as well as analytical formulae for their systematics, e. g.

$$U_x = \left(2.5 + \frac{150}{A}\right) \text{ MeV}, \quad (322)$$

so that level densities can be estimated even in the absence of nuclear structure data (level schemes) for low-lying levels or resonance data (cross sections) above the neutron binding

energy. This is valuable, for instance, if fission cross sections are to be calculated, for which one needs the transition state densities at the saddle point deformation. Those are not directly observable but are expected to be similar to the state densities near the ground state deformation and therefore at least roughly describable by the constant-temperature part of the composite level density formula. Other examples are nuclei in metastable states or short-lived radionuclides, data for which are difficult to measure but are needed for burnup and transmutation calculations in nuclear technology or for nucleosynthesis studies in astrophysics.

As already mentioned the Gaussian Orthogonal Ensemble has a semicircular eigenvalue distribution, and more physical ensembles such as the shell model with statistically treated residual interaction have eigenvalue spectra resembling Gaussians. Is this not in conflict with the nearly exponentially increasing level density of the Bethe formula? The answer is that the Bethe formula is valid for modest excitation, where only few single-particle levels around the Fermi edge are affected and neither the finite depth of the single-particle potential nor unbound continuum states have any effect yet. As excitation energy grows, however, more and more nucleons are lifted from levels below the Fermi edge to levels above and even to continuum states that no longer involve the intact compound nucleus. The level density *of a given compound nucleus* rises therefore to a maximum but then it decreases again, in Gaussian-like fashion, due to growing competition by unbound states representing nuclear transmutation and destruction. The Bethe formula appears thus as an approximation to the low-energy tail of a nearly Gaussian density function (see Grimes 1980), certainly applicable in the unresolved resonance region but not at GeV excitation energies.

4.1.5. Information from Resolved Resonances

Global information about strength functions or transmission coefficients can be obtained from optical-model systematics. More specific prior information for a given compound system comes, however, from resolved resonances. The transmission coefficient for a particle channel $c = \{\alpha J \ell s\}$ is related to the corresponding average partial width by $T_c = 2\pi\bar{\Gamma}_c/D_c$ (we absorb the multi-level correction denominator of Eq. 283 in the partial width) but the *observable* (“lumped”) neutron widths are sums of partial widths for *all* channels compatible with the level characteristics J and Π . Neglecting the weak channel spin dependence of the transmission coefficients predicted by the optical model one gets for the average neutron width

$$\langle \Gamma_n \rangle_{\ell J} = \nu_{\ell J} D_J S_\ell \sqrt{E/1\text{eV}} v_\ell(E), \quad (323)$$

where $\nu_{\ell J}$ is the number of channel spins (1 or 2) that can be combined with ℓ to give J . This is essentially Eq. 274 that we employed to find the Bayesian probability of the characteristics J and ℓ (or rather Π) of a resonance with given value of $g\Gamma_n$. One can now introduce the usual definition of the reduced neutron width,

$$\Gamma_n^\ell \equiv \frac{\Gamma_n}{\sqrt{E/1\text{eV}} v_\ell(E)}, \quad (324)$$

replace the ensemble average by the sample average,

$$\langle g_J \Gamma_n^\ell \rangle \simeq \frac{1}{N} \sum_{\lambda=1}^N (g\Gamma_n^\ell)_\lambda, \quad (325)$$

multiply both sides by $g_J \rho_J$ and sum over all J compatible with ℓ . With $\sum_J g_J \nu_{\ell J} = 2\ell + 1$ and $N / \sum_J \rho_J \simeq \Delta E$ one finds eventually the widely used recipe for estimating neutron strength functions from all the $g\Gamma_n^\ell$ values found in an energy interval ΔE ,

$$S_\ell \simeq \frac{\sum_{\lambda=1}^N (g\Gamma_n^\ell)_\lambda}{(2\ell + 1)\Delta E}. \quad (326)$$

It has the advantage that neither the resonance spins nor the level densities must be known. Only the products $g\Gamma_n^\ell$ are needed which is often all that is known for the weaker resonances. Moreover, the estimator is fairly insensitive to missing levels as these have small reduced widths and therefore contribute little to the sum, and it is similarly insensitive to wrong ℓ assignments that again affect mainly the weak levels.

The problem of missing levels is encountered head-on if the level density is to be estimated. Weak resonances are often not seen at all in transmission measurements whereas at least some of them show up in capture and fission yield data. A mere counting of observed peaks in some energy interval ΔE is then not enough. Nor is a fit to the Wigner distribution very helpful because missing levels tend to distort the whole observed level spacing distribution. A better way to estimate the fraction of missing levels is to look at the neutron width distribution where only the portion below some detection threshold is affected. Without threshold we have the complete Porter-Thomas distribution (277) which we now write, with the abbreviations $G \equiv g\Gamma_n^\ell$, $\Theta \equiv \langle g\Gamma_n^\ell \rangle$, and $n \equiv \nu_{\ell J}/2$, as

$$p(G|\Theta, n) dG = \frac{e^{-x} x^n}{\Gamma(n)} \frac{dx}{x}, \quad 0 < x \equiv n \frac{G}{\Theta} < \infty. \quad (327)$$

With the likelihood function $\prod_j p(G_j|\Theta, n)$ for a sample G_1, G_2, \dots, G_N reduced widths and with Jeffreys' prior $d\Theta/\Theta$ for the scale parameter Θ one gets the posterior

$$p(\Theta|G_1, \dots, G_N, n) d\Theta = \frac{e^{-y} y^{nN}}{\Gamma(nN)} \frac{dy}{y}, \quad 0 < y \equiv nN \frac{\bar{G}}{\Theta}, \quad (328)$$

where $\bar{G} \equiv (G_1 + \dots + G_N)/N$ is the sample average. With this gamma distribution it is easy to calculate the expectation values $\langle y^{-1} \rangle = 1/(nN - 1)$ and $\langle y^{-2} \rangle = 1/[(nN - 1)(nN - 2)]$ and then the estimate under quadratic loss,

$$\langle \Theta \rangle = \frac{nN}{nN - 1} \bar{G}, \quad (329) \quad \frac{\Delta \Theta}{\langle \Theta \rangle} = \frac{1}{\sqrt{nN - 2}}, \quad (330)$$

with $n = 1/2$ for single-channel neutron widths ($\ell = 0$ or $I = 0$) and $n = 1$ for two-channel neutron widths ($\ell > 1$ and $I > 0$). Note how big the uncertainty is even for large samples – e. g. 10 % for a sample of 204 s-wave resonances.

The case of a given detection threshold $G_c > 0$ can be treated in complete analogy. The sampling distribution is now a truncated Porter-Thomas distribution of G (gamma distribution of x),

$$p(G|\Theta, n, G_c) dG = \frac{e^{-x} x^n}{\Gamma(n, x_c)} \frac{dx}{x}, \quad x_c \equiv n \frac{G_c}{\Theta} < x \equiv n \frac{G}{\Theta} < \infty, \quad (331)$$

normalised by the incomplete gamma function,

$$\Gamma(n, x_c) \equiv \int_{x_c}^{\infty} e^{-x} x^{n-1} dx, \quad (332)$$

which we recognise as the probability of a level to be observable. It depends on the estimated parameter, in contrast to the gamma function $\Gamma(n)$ that we had before. The posterior for the sample G_1, G_2, \dots, G_N is thus

$$p(\Theta|G_1, \dots, G_N, n) d\Theta = \frac{1}{Z} \frac{e^{-y} y^{nN}}{\Gamma(n, cy)^N} \frac{dy}{y}, \quad 0 < y \equiv nN \frac{\bar{G}}{\Theta}, \quad (333)$$

where $c \equiv G_c/(N\bar{G})$, so that $cy = x_c$. The normalisation is given by

$$Z = \int_0^\infty \frac{e^{-y} y^{nN}}{\Gamma(n, cy)^N} \frac{dy}{y}. \quad (334)$$

The needed expectation values $\langle y^{-1} \rangle$ and $\langle y^{-2} \rangle$ involve similar integrals,

$$\langle y^{-k} \rangle = \frac{1}{Z} \int_0^\infty \frac{e^{-y} y^{nN-k}}{\Gamma(n, cy)^N} \frac{dy}{y}, \quad k = 1, 2, \quad (335)$$

and the same is true for the expected fraction of unobservable (missing) levels, $1 - \langle \Gamma(n, cy) \rangle$ with

$$\langle \Gamma(n, cy) \rangle = \frac{1}{Z} \int_0^\infty \frac{e^{-y} y^{nN}}{\Gamma(n, cy)^{N-1}} \frac{dy}{y}. \quad (336)$$

The particular incomplete gamma functions needed for neutron (and proton) widths are

$$\Gamma\left(\frac{1}{2}, x_c\right) = \sqrt{\pi} \operatorname{erfc} \sqrt{x_c}, \quad (337)$$

$$\Gamma(1, x_c) = \exp(-x_c). \quad (338)$$

This shows that at least for $n = 1$, i. e. two possible channel spins, the integrals can be calculated analytically: The width estimate under quadratic loss is

$$\langle \Theta \rangle = N\bar{G} \langle y^{-1} \rangle = \frac{N}{N-1} (\bar{G} + G_c), \quad (339) \quad \frac{\Delta\Theta}{\Theta} = \frac{1}{\sqrt{N-2}}, \quad (340)$$

while the observed fraction of levels is expected to be

$$\langle \Gamma(1, cy) \rangle = \left(1 + \frac{G_c}{N(\bar{G} - G_c)}\right)^{-N}. \quad (341)$$

For $n = 1/2$ one must integrate numerically, or use the Laplace approximation which yields the same estimate for Θ as the maximum likelihood method. For a more general discussion of missing level estimators, including unknown, energy-dependent and diffuse thresholds as well as unresolved multiplets, see Fröhner (1983).

4.2. RESONANCE-AVERAGED CROSS SECTIONS

The usual task in the unresolved resonance region is that average cross sections or cross section functionals like the average transmission are to be calculated, in an averaging interval wide enough to contain many resonances but so narrow that secular variations of level statistics and other weak energy dependences can be neglected. We may then simplify our equations by choosing boundary parameters such that locally $L_c^o = iP_c$, and by absorbing P_c in the decay widths $\gamma_{\lambda c}$. Furthermore, we shall write \mathbf{S} instead of \mathbf{U} for

the \mathbf{S} matrix, as is customary in the literature on average cross sections. The average collision matrix is then (compare Eqs. 155-160, 209-210, 284)

$$\begin{aligned}\bar{\mathbf{S}}_{ab} &= e^{-i(\varphi_a + \varphi_b)} \left[(\mathbf{1} - i\bar{\mathbf{R}})^{-1} (\mathbf{1} + i\bar{\mathbf{R}}) \right]_{ab} \\ &= e^{-i(\varphi_a + \varphi_b)} \left(\delta_{ab} + 2i \sum_{\lambda, \mu} \gamma_{\lambda a} A_{\lambda \mu} \gamma_{\lambda b} \right),\end{aligned}\quad (342)$$

with

$$(\mathbf{A}^{-1})_{\lambda \mu} = (E_\lambda - E) \delta_{\lambda \mu} - i \sum_c \gamma_{\lambda c} \gamma_{\mu c}. \quad (343)$$

4.2.1. Average Total Cross Section

In order to average the total cross section we must average the collision matrix element S_{cc} over suitable energy bands, beam profiles or resolution functions. This is easy with a Lorentzian weight function,

$$\begin{aligned}\bar{\mathbf{S}}(E) &= \int_{-\infty}^{\infty} dE' \frac{I/\pi}{(E' - E)^2 + I^2} \mathbf{S}(E') \\ &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} dE' \left(\frac{1}{E' - E - iI} - \frac{1}{E' - E + iI} \right) \mathbf{S}(E'),\end{aligned}\quad (344)$$

where $2I$ is the full width at half maximum of the Lorentzian. Due to causality the collision matrix has no poles above the real axis (see Lane and Thomas 1958), so if we close the contour by a large upper semicircle (with vanishing contribution) it encloses only the pole at $E + iI$ of the Lorentzian, and the residue is

$$\bar{\mathbf{S}}(E) = \mathbf{S}(E + iI). \quad (345)$$

As we may neglect weak energy dependences we need only replace $\mathbf{R}(E)$ by $\mathbf{R}(E + iI)$, with

$$\begin{aligned}R_{ab}(E + iI) &= \sum_{\lambda} \frac{\gamma_{\lambda a} \gamma_{\lambda b}}{E_\lambda - E - iI} \\ &\simeq \int_{-\infty}^{\infty} \frac{dE'}{D_c} \frac{\overline{\gamma_a \gamma_b}}{E' - E - iI} \simeq (R_a^\infty + i\pi s_a) \delta_{ab}.\end{aligned}\quad (346)$$

In the last approximation we exploited the fact that because of the random signs of the $\gamma_{\lambda c}$ the average matrix $\overline{\gamma_a \gamma_b}$ is practically diagonal. Furthermore, we introduced the definitions of pole strength and distant-level parameter, Eqs. 209-210, and neglected the variation of the pole strength over the peak region of the Lorentzian – exactly as in our treatment of external levels (Subsect. 3.4.1). The final result is

$$\bar{\sigma}_c = 2\pi \lambda_c^2 g_c (1 - \text{Re } \bar{\mathbf{S}}_{cc}), \quad (347)$$

$$\bar{\mathbf{S}}_{cc} = e^{-2i\varphi_c} \frac{1 + i(R_c^\infty + i\pi s_c)}{1 - i(R_c^\infty + i\pi s_c)}. \quad (348)$$

The resonance-averaged total cross section is thus expressed by the pole strength and the distant-level parameter, quantities that can be obtained either from statistical analysis

of resolved resonances or from optical-model phase shifts (after specification of a channel radius).

4.2.2. Average Partial Cross Sections: Heuristic Recipes

In contrast to the total cross section the average partial cross sections,

$$\bar{\sigma}_{ab} = \pi \lambda_a^2 g_a \overline{|\delta_{ab} - S_{ab}|^2}, \quad (349)$$

are no linear function of \mathbf{S} but require averaging over quadratic terms like $S_{ab}^* S_{cd}$. These have poles above as well as below the real axis which prevents contour integration with a Lorentzian weight function. Under the usual ergodicity and stationarity conditions of good statistics - many resonances and negligible variation of the parameter distributions within the averaging interval - one can replace the energy average by an ensemble average (i. e. expectation value) over the GOE, i. e. over the joint distribution of level energies and decay amplitudes. The ensemble average is readily obtained in the limit of widely spaced ("isolated") resonances that overlap so weakly that multi-level effects and eigenvalue correlations can be neglected. Assuming generalised Porter-Thomas (χ^2) distributions for the partial widths one obtains in many-level SLBW approximation

$$\bar{\sigma}_{ab} = \sigma_{p,a} \delta_{ab} + \pi \lambda_a^2 g_a \frac{T_a T_b}{T} \left(1 + \frac{2}{\nu_a} \delta_{ab}\right) \int_0^\infty dx \prod_c \left(1 + \frac{2T_c}{\nu_c T} x\right)^{-\delta_{ac} - \delta_{bc} - \nu_c/2} \quad (350)$$

(Dresner 1957, Lane and Lynn 1957), where $\sigma_{p,a}$ is the potential-scattering cross section, $T_c \equiv 1 - |\bar{S}_{cc}|^2$ is the transmission coefficient for channel c , $T \equiv \sum_c T_c$, and ν_c the degree of freedom for the partial widths $\Gamma_{\lambda c} = 2\gamma_{\lambda c}^2$ (remember that P_c is absorbed in $\gamma_{\lambda c}^2$). The approximation $T_c \simeq 2\pi\bar{\Gamma}_c/D_c$, valid for vanishing level overlap, was used to write the result in terms of the T_c . This is the Hauser-Feshbach formula with elastic enhancement (first pair of parentheses) and width fluctuation correction (integral) - see Moldauer (1975). We recall that $\nu_c = 1$ for single channels but that in practical applications one often uses lumped channels, with an effective $\bar{\nu}_c$ differing from unity, in order to represent e. g. all fission or capture channels or all particle channels that have the same total angular momentum and parity and thus involve the same compound levels. The number of photon channels is usually so large (except for light and magic nuclei) that one may put

$$\prod_{c \in \gamma} \left(1 + \frac{2T_c}{\nu_c T} x\right)^{-\nu_c/2} \simeq \lim_{\bar{\nu}_\gamma \rightarrow \infty} \left(1 + \frac{2T_\gamma}{\bar{\nu}_\gamma T} x\right)^{-\bar{\nu}_\gamma/2} = e^{-xT_\gamma/T}, \quad (351)$$

where $T_\gamma \equiv \sum_{c \in \gamma} T_c$. The many photon channels can thus be represented approximately by an exponential factor in the integral ("Dresner factor") of Eq. 331. Generalisation of the Hauser-Feshbach formula to arbitrary level overlap turned out to be extremely difficult. Of course one could always resort to Monte Carlo sampling of level spacings and decay amplitude from their probability distributions, with subsequent point cross section calculation and averaging. The desired cross section average is thus obtained, although with the statistical uncertainty and lack of analytical transparency typical for the Monte Carlo method. From such numerical Monte Carlo studies two practically important analytical recipes were deduced heuristically, by trial and error and educated guesswork.

The first recipe, due to Moldauer (1980), consists in using the Hauser-Feshbach formula, strictly valid only for weak level overlap, also for strong overlap, but with $\sigma_{p,a}$ interpreted as the "direct" or "shape elastic" cross section,

$$\sigma_{p,a} = \pi \lambda_a^2 g_a |1 - \bar{S}_{aa}|^2, \quad (352)$$

and with the exact expression for the particle-channel transmission coefficients,

$$T_a = \frac{4\pi s_a}{|1 - i\bar{R}_{aa}|^2}. \quad (353)$$

Furthermore, the ν_c are considered as depending on the T_c . The dependence is chosen so as to fit a large body of Monte Carlo results while giving the correct limit for small level overlap (small transmission coefficients). Moldauer's heuristic recommendation is

$$\bar{\nu}_c = [1.78 + (T_c^{1.218} - 0.78)e^{-0.228 T}] \nu_c. \quad (354)$$

The second practically important prescription is due to Hofmann, Richert, Tepel and Weidenmüller (1975) who, in the spirit of Bohr's original compound-nuclear model (no memory of compound formation), take the partial cross sections as factorisable,

$$\bar{\sigma}_{ab} = \sigma_{p,a} \delta_{ab} + \pi \lambda_a^2 g_a \frac{V_a V_b}{V} [1 + (\omega_a - 1) \delta_{ab}], \quad (355)$$

with $V \equiv \sum_c V_c$. The elastic enhancement factors ω_c are expected to approach 3 for vanishing and 2 for very strong level overlap (Satchler 1963). The authors found their Monte Carlo results adequately reproduced with

$$\omega_a = 1 + \frac{2}{1 + T_a^{0.3+1.5 T_a/T}} + 2 \left(\frac{T_a}{T} - \frac{1}{n} \right)^2, \quad (356)$$

where n is the number of open channels. With these heuristic values of the ω_a one can calculate the V_a from

$$V_a = \frac{T_a}{1 + (\omega_a - 1) V_a / V}. \quad (357)$$

by iteration, beginning with $V_c = T_c$. The last equation follows from the unitarity of \mathbf{S} .

Both prescriptions yield similar results for intermediate and strong absorption (medium and strong level overlap). Moldauer's recipe is convenient for lumped channels and, by construction, it yields the correct limit for vanishing overlap and few (nonphotonic) channels at low energies where the factorisation approximation fails. Other approximate analytic expressions were derived with picket fence models (e. g. Janeva et al. 1985) and disordered picket fence models (Müller and Harney 1987).

4.2.3. Average Partial Cross Sections: The Exact GOE Average

For decades all attempts at solving the Hauser-Feshbach problem had failed. In this situation the information-theoretic maximum entropy principle seemed to offer a possibility to bypass all "microscopic" resonance details by treating them as a kind of noise superimposed on the "macroscopic" average behaviour described by the optical model. The probability distributions of the S- and R-matrix elements were found by entropy maximisation constrained by given (optical model) averages $\bar{\mathbf{S}}$ and $\bar{\mathbf{R}}$ (Mello 1979, Mello et al. 1985, Fröhner 1986) which in principle offered the possibility to average the cross section expressions over these distributions rather than the GOE. In practice, for many channels, this looked still very difficult.

Only few months after the maximum-entropy distributions of the S and R matrix had been published, Verbaarschot, Weidenmüller and Zirnbauer (1985) presented a direct

solution to the Hauser-Feshbach problem of finding an analytic expression for the average partial cross sections, i. e. to average analytically over the GOE resonance parameter distributions, with given transmission coefficients. These authors started from an expression involving a GOE Hamiltonian coupled to the channels. In our notation it reads

$$|S_{ab}|^2 = |\delta_{ab} + i \sum_{\lambda, \mu} \tilde{\gamma}_{\lambda a} A_{\lambda \mu} \tilde{\gamma}_{\mu b}|^2, \quad (358)$$

$$(\mathbf{A}^{-1})_{\lambda \mu} = H_{\lambda \mu} - E \delta_{\lambda \mu} - i \sum_c \tilde{\gamma}_{\lambda c} \tilde{\gamma}_{\mu c}, \quad (359)$$

which is a generalisation of what Eqs. 158-160 give for $|S_{ab}|^2$. The tilde indicates that the Hamiltonian has its general, nondiagonal form, so that $H_{\lambda \mu}$ and $\tilde{\gamma}_{\lambda a}$ replace $E_{\lambda} \delta_{\lambda \mu}$ and $\gamma_{\lambda a}$ of Eq. 160. By a formidable display of analytic skill the authors managed, with new tools from the many-body theory of disordered systems, to reduce the ensemble average (expectation value) of $|S_{ab}|^2$ over the GOE to a threefold integral. Making full use of the symmetries of the GOE, of a generating function involving both commuting and anticommuting (Grassmann) variables, of the Hubbard-Stratonovitch transformation to simplify the integrations, then going to the limit of infinitely many levels ($n \rightarrow \infty$ for the rank of H) by the method of steepest descent, they derived the awesome triple-integral expression

$$\begin{aligned} \overline{|S_{ab}|^2} &= |\overline{S}_{ab}|^2 + \frac{T_a T_b}{8} \int_0^\infty d\lambda_1 \int_0^\infty d\lambda_2 \int_0^1 d\lambda \frac{\lambda(1-\lambda)|\lambda_1 - \lambda_2|}{\sqrt{\lambda_1(1+\lambda_1)} \sqrt{\lambda_2(1+\lambda_2)} (\lambda + \lambda_1)^2 (\lambda + \lambda_2)^2} \\ &\times \left(\prod_c \frac{1 - T_c \lambda}{\sqrt{1 + T_c \lambda_1} \sqrt{1 + T_c \lambda_2}} \right) \left\{ \delta_{ab} (1 - T_a) \left(\frac{\lambda_1}{1 + T_a \lambda_1} + \frac{\lambda_2}{1 + T_a \lambda_2} + \frac{2\lambda}{1 - T_a \lambda} \right)^2 \right. \\ &\left. + (1 + \delta_{ab}) \left(\frac{\lambda_1(1 + \lambda_1)}{(1 + T_a \lambda_1)(1 + T_b \lambda_1)} + \frac{\lambda_2(1 + \lambda_2)}{(1 + T_a \lambda_2)(1 + T_b \lambda_2)} + \frac{2\lambda(1 - \lambda)}{(1 - T_a \lambda)(1 - T_b \lambda)} \right) \right\} \end{aligned} \quad (360)$$

for the absolute square of the collision matrix element that had caused the difficulties with its poles below and above the real axis in the complex energy plane. The channel product allows a similar treatment of the many weakly absorbing photon channels as in the Hauser-Feshbach formula:

$$\prod_c \frac{1 - T_c \lambda}{\sqrt{1 + T_c \lambda_1} \sqrt{1 + T_c \lambda_2}} \simeq e^{-(\lambda_1 + \lambda_2 + 2\lambda)T_\gamma/2} \prod_{c \notin \gamma} \frac{1 - T_c \lambda}{\sqrt{1 + T_c \lambda_1} \sqrt{1 + T_c \lambda_2}}, \quad (361)$$

with $T_\gamma \equiv \sum_{c \in \gamma} T_c$ as in Eq. 351.

Verbaarschot (1986) verified that in the limit of small level overlap the GOE triple integral (360) yields the Hauser-Feshbach formula (350) with elastic enhancement and width fluctuation correction. Thus the GOE triple integral is the long-sought rigorous solution to the Hauser-Feshbach problem, eliminating all uncertainties associated with picket fence models or heuristic analytic formulae inferred from Monte Carlo results. These uncertainties had always been bothersome because width fluctuation corrections are often quite substantial (see e. g. Lynn 1968, Gruppelaar and Reffo 1977). An important point is that above a few eV resonance-averaged cross sections are practically independent of temperature: Energy averaging involves essentially sums over peak areas, and since those are

invariant under Doppler broadening (in Kapur-Peierls, Adler-Adler, MLBW and SLBW form we have $\int dE \psi_\lambda = \pi\Gamma_\lambda/2$ and $\int dE \chi_\lambda = 0$, irrespective of temperature, see Appendix B), the same is true for average cross sections. Thus the GOE triple integral, valid for unbroadened resonances, gives also correct averages over Doppler-broadened resonances.

4.2.4. Analysis of Resonance-Averaged Data

Figs. 20-22 show average total, capture and inelastic scattering cross section data for ^{238}U and theoretical curves fitted to all these data simultaneously. The fitting was done by least-squares adjustment of average resonance parameters, viz. of s-, p-, d- and f-wave neutron strength functions (which are essentially transmission coefficients for neutron channels) and of radiation widths scaled by the mean level spacing (transmission coefficients for the lumped photon channels) with the code FITACS (Fröhner et al. 1982). The main energy dependences are introduced by the centrifugal-barrier penetration factors P_c for the neutron widths and by the employed composite level density formula of Gilbert and Cameron (1965), whereas the strength functions and radiation widths vary only little in the energy range covered. The total cross section was calculated with Eqs. 347-348, the partial cross sections with the Hauser-Feshbach formula in the form proposed by Moldauer (1980), Eqs. 350-354, and cross-checked with the GOE triple integral, Eqs. 360-361. Similar fits to many more ^{238}U data defined eventually a new evaluation for ^{238}U in the unresolved resonance region that was adopted for the evaluated data libraries JEF-2 and ENDF/B-VI (Fröhner 1989). The final adjusted average resonance parameters are fully consistent with the resolved resonance parameters determined at lower energies, and also with optical-model calculations at higher energies up to 10 MeV. The error estimates from the least-squares fits indicate that, after decades of world-wide effort, the average total and capture cross sections of ^{238}U in the resolved resonance region are finally known with about the accuracies requested for applications in nuclear technology (1-3%). For inelastic scattering this goal is not yet achieved, the uncertainties there are still of the order of 5 - 15% .

Accurate average cross sections are, however, only part of the story. The other part concerns the resonance structure, i. e. the resonance fluctuations around the average cross section curves. They are implicitly given by the level-statistical model, in particular by the GOE distributions of level spacings and partial widths together with the mean values parametrising these distributions. The presence of unresolved resonance structure manifests itself in sample-thickness and self-shielding effects. As the simplest illustration consider the relationship between the average transmission of a slab of material with thickness n (atoms/b) and average total cross section $\langle\sigma\rangle$,

$$\langle e^{-n\sigma} \rangle = e^{-n\langle\sigma\rangle} \langle e^{-n(\sigma - \langle\sigma\rangle)} \rangle = e^{-n\langle\sigma\rangle} \left(1 + \frac{n^2}{2} \text{var } \sigma - + \dots \right). \quad (362)$$

The last pair of parentheses represents a correction for resonance effects, containing the cross section variance (mean square fluctuation) and higher moments of the cross section distribution which quantify the resonance structure. Solving for the average cross section one gets

$$\langle\sigma\rangle = \frac{1}{n} \ln \langle e^{-n\sigma} \rangle + \frac{1}{n} \ln \left(1 + \frac{n^2}{2} \text{var } \sigma - + \dots \right). \quad (363)$$

Often the first term on the right-hand side is presented by measurers as the total cross section but the true average is seen to be always larger. Although the correction is small for well resolved data it becomes important for resonance-averaged data. It is dangerous to

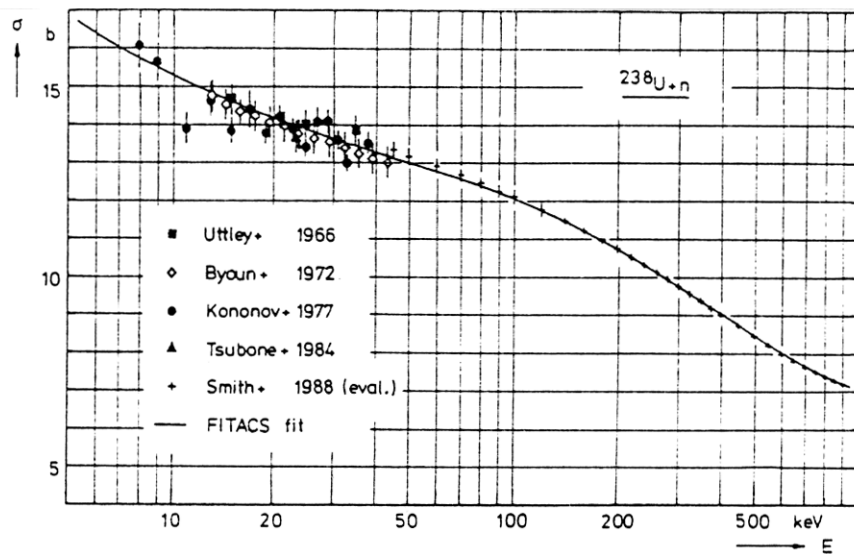


Fig. 20. Simultaneous Hauser-Feshbach fit to ^{238}U neutron data in the unresolved resonance region: total cross section (for references see Fröhner 1989)

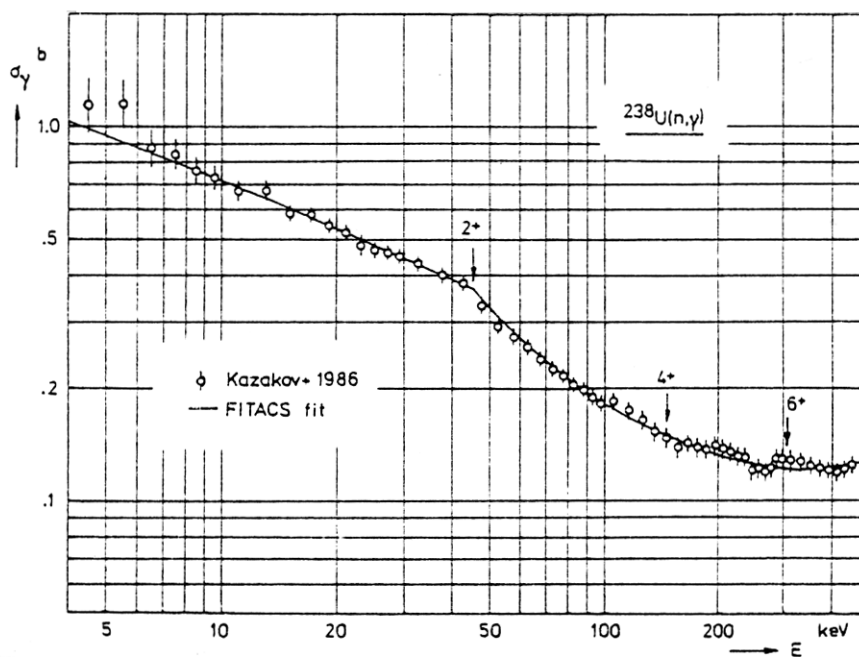


Fig. 21. Simultaneous Hauser-Feshbach fit to ^{238}U neutron data in the unresolved resonance region: capture cross section (for references see Fröhner 1989). The discontinuity (Wigner cusp) at 45 keV is due to the onset of inelastic scattering at that energy. Inelastic thresholds are indicated by spin-parity characteristics of residual levels.

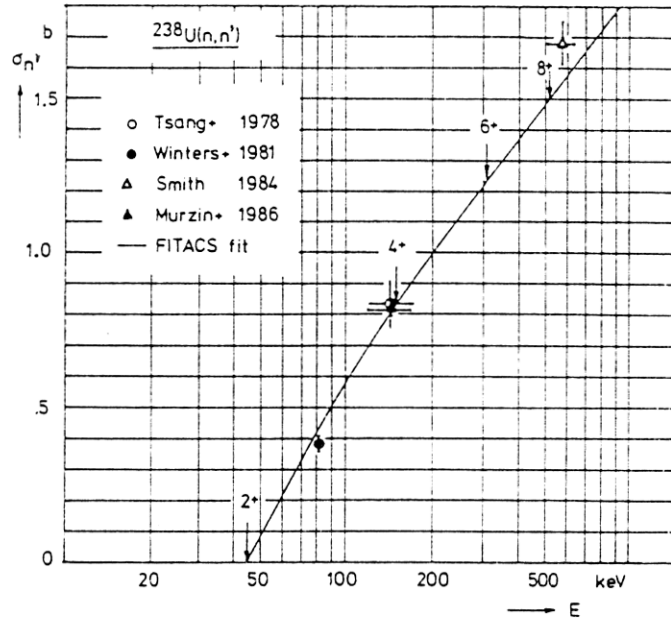


Fig. 22. Simultaneous Hauser-Feshbach fit to ^{238}U neutron data in the unresolved resonance region: inelastic-scattering cross section (for references see Fröhner 1989). Inelastic thresholds are indicated by spin-parity of residual levels.

calculate the transmission of a shield from resonance-averaged cross sections, e. g. group cross sections for infinite dilution, without correction for the cross section fluctuations. The correction is important for thick samples and strongly fluctuating cross sections such as those in the unresolved resonance region. Thick-sample transmission measurements can therefore provide information about the resonance structure. If we want to compare experiments with calculations, however, we must average the transmission over resonances. In view of the experiences with GOE averaging of partial cross sections we must expect even worse problems for the average transmission. The need to deal with Doppler-broadened cross sections makes the problem rather hopeless as far as an analytic solution is concerned.

Only the cross section variance can be calculated theoretically in the case that Doppler broadening is not too important (as for light nuclei and structural materials like iron or nickel). Averaging with a Lorentzian weight function and using causality, as in the derivation of the total cross section expression (347-348), one finds $\langle S_{cc}^2 \rangle = \langle S_{cc} \rangle$ and

$$\text{var } \sigma = 2\pi\lambda^2 \sum_c g_c (\sigma_c^{\text{CN}} - \langle \sigma_c^{\text{non}} \rangle), \quad (364)$$

where $\sigma_c^{\text{CN}} = \pi\lambda^2 g_c (1 - |\langle S_{cc}^2 \rangle|)$ is the usual compound nucleus formation cross section of the optical model and $\langle \sigma_c^{\text{non}} \rangle = \pi\lambda^2 g_c (1 - \langle |S_{cc}|^2 \rangle)$ the resonance-averaged nonelastic cross section of Hauser-Feshbach theory that must be calculated from optical-model transmission coefficients either in many-level SLBW approximation, Eq. 350, or more rigorously

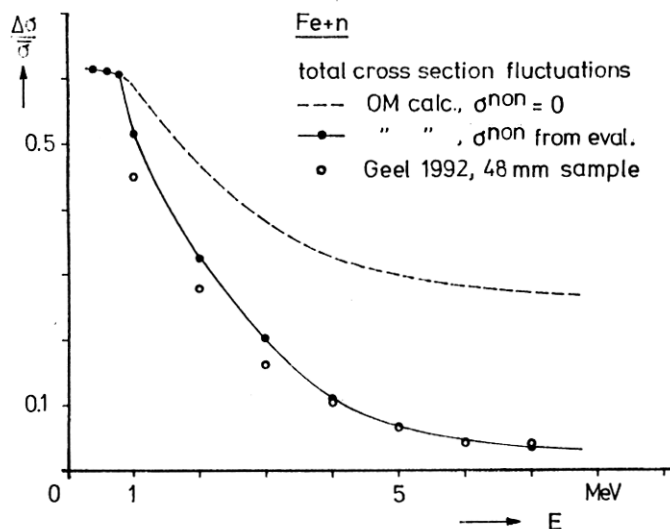


Fig. 23. Relative root-mean-square fluctuations of the total cross section of iron ($\Delta\sigma = \sqrt{\text{var } \sigma}$). Solid line: calculation with Eq. 364; broken line: calculation with nonelastic processes neglected; open circles: values obtained directly from data taken with very high resolution (Berthold et al. 1994). From Fröhner (1994).

with the GOE triple integral, Eq. 360. Fig. 23 shows relative root-mean-square fluctuations calculated in this way for ^{56}Fe together with values obtained directly from recent high-resolution transmission data. Agreement is reasonable in view of the fact that Eq. 364 is exact only for pure compound reactions and for negligible Doppler and resolution broadening, all of which tend to reduce the resonance fluctuations.

Although general analytical expressions for resonance-averaged cross section functionals are not available, it is perfectly straightforward to calculate them by Monte Carlo methods. An example is shown in Fig. 24, where thick-sample transmission data measured at room temperature are plotted together with Monte Carlo calculated curves. "Ladders" of resonances were defined by sampling resonance spacings from the Wigner distribution and of partial widths from Porter-Thomas distributions, with average resonance parameters taken from the JEF-2 evaluation. The corresponding total cross section was calculated, Doppler broadened, exponentiated, and averaged. In this way 100 000 transmission values were sampled and averaged for each data point, so that the statistical error of the Monte Carlo results was negligible compared to the uncertainties of the data. The good agreement between experiment and calculation in Fig. 24 indicates that the JEF-2 evaluation describes not only the total cross section well (see Fig. 20) but also its resonance structure, e. g. the ratio of resonance (compound) to potential scattering (direct) cross sections, and the "windows".

More information about the cross section structure comes from self-indication measurements. Those differ from transmission measurements only insofar as the detector measuring the transmitted fraction of the neutron beam consists of a thin sample ("radiator"), made of the same material as the thick transmitting sample ("filter") and viewed by gamma ray detectors. From "filter in" and "filter out" runs one obtains the self-indication ratio

$$\frac{\langle e^{-n\sigma} \sigma_\gamma \rangle}{\langle \sigma_\gamma \rangle} = e^{-n\bar{\sigma}} \left(1 - n \frac{\text{cov}(\sigma, \sigma_\gamma)}{\langle \sigma_\gamma \rangle} + \dots \right) \quad (365)$$

which involves the covariance between the total and capture cross section structure,

$$\text{cov}(\sigma, \sigma_\gamma) \equiv \langle (\sigma - \langle \sigma \rangle)(\sigma_\gamma - \langle \sigma_\gamma \rangle) \rangle = \langle \sigma \sigma_\gamma \rangle - \langle \sigma \rangle \langle \sigma_\gamma \rangle . \quad (366)$$

(For positive covariance the two arguments tend to vary in the same sense - if one increases, the other one is likely to increase too - for negative covariance they tend to vary in opposite directions.) In practice the radiators are not ideally thin, so that the capture cross section σ_γ ought to be replaced by the capture yield y_γ that includes self-shielding and multiple-collision capture. Both effects require Monte Carlo techniques, in addition to ladder sampling one must now also simulate multiple-collision events in the radiator (for details see Fröhner 1989a). Fig. 25 shows that the measured data and the Monte Carlo results are in good agreement again, indicating that also the capture cross section structure is adequately represented by the average resonance parameters of the JEF-2 evaluation.

4.3. GROUP CONSTANTS

We saw that for a given average total cross section the average transmission (in some finite energy interval containing many resonances) of a thick sample is larger if the cross section fluctuates than if it is smooth (see Eq. 362). This means that the sample becomes less transparent as the temperature rises, due to the smoothing effect of Doppler broadening. (Thermal expansion of the sample counteracts this effect to some degree.) In a reactor region filled with a mixture of materials a temperature increase means that (n,x) processes, e. g. (n, γ) reactions in ^{238}U , become more probable with increasing temperature because the flux depletion across the resonances (cf. Fig. 9) becomes weaker as the resonance structure is smoothed out. In order to calculate these complicated effects one simplifies by using group constants, i. e. suitably defined cross section averages. The (n,x) reaction rate for a given nuclide, averaged over the region and over a finite (group) interval ΔE , can be written as

$$\langle \varphi \sigma_x \rangle = f_x \langle \sigma_x \rangle \langle \varphi \rangle \quad \text{with} \quad \langle \dots \rangle \equiv \int_{\Delta E} \frac{dE}{\Delta E} \dots \quad (367)$$

The group boundaries are usually taken as equidistant on a logarithmic energy scale, i. e. on a linear lethargy scale, so that there is always the same number of groups per energy decade. The cross section σ_x is to be understood as Doppler broadened. Since $\langle \sigma_x \rangle$ does not depend on temperature (apart from edge effects at group boundaries which become negligible if the group interval contains many resonances) the main temperature dependence for given average flux is contained in the so-called self-shielding or Bondarenko factor f_x .

4.3.1. Bondarenko Factors

The self-shielding factor depends not only on temperature but also on the cross sections of all other nuclides in the mixture, the so-called dilution. The data filed in group constant sets for technological applications are (cf. e. g. Bondarenko et al. 1964)

- cross sections for infinite dilution $\langle \sigma_x \rangle$,
- self-shielding factors $f_x = \frac{\langle \varphi \sigma_x \rangle}{\langle \varphi \rangle \langle \sigma_x \rangle}$,

stored for each nuclide on a grid of temperatures and dilution cross sections σ_d , e. g.

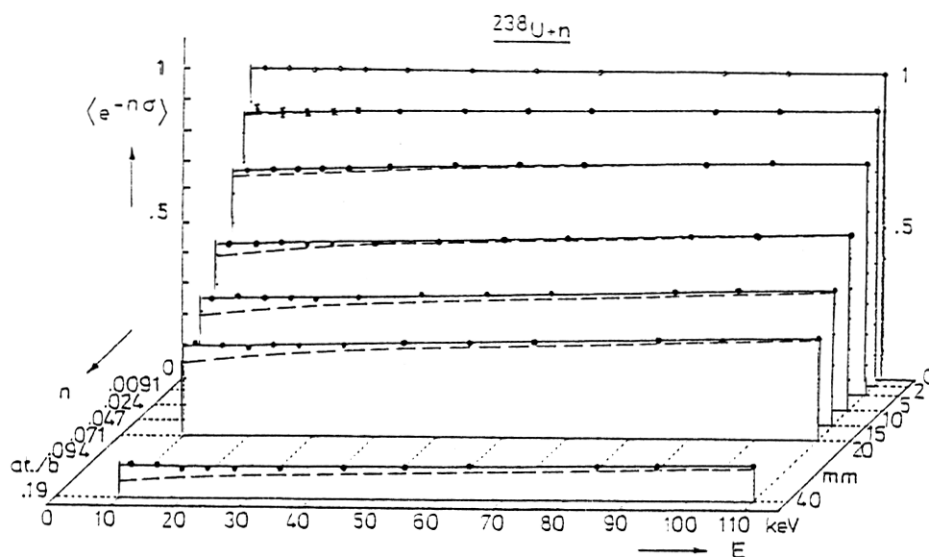


Fig. 24. Thick-sample transmission data of Bokhovko et al. (1988) (point symbols) and curves generated with Monte Carlo techniques from JEF-2 average resonance parameters (solid lines). Also shown are the transmission curves obtained without correction for resonance structure (broken lines). From Fröhner (1991).

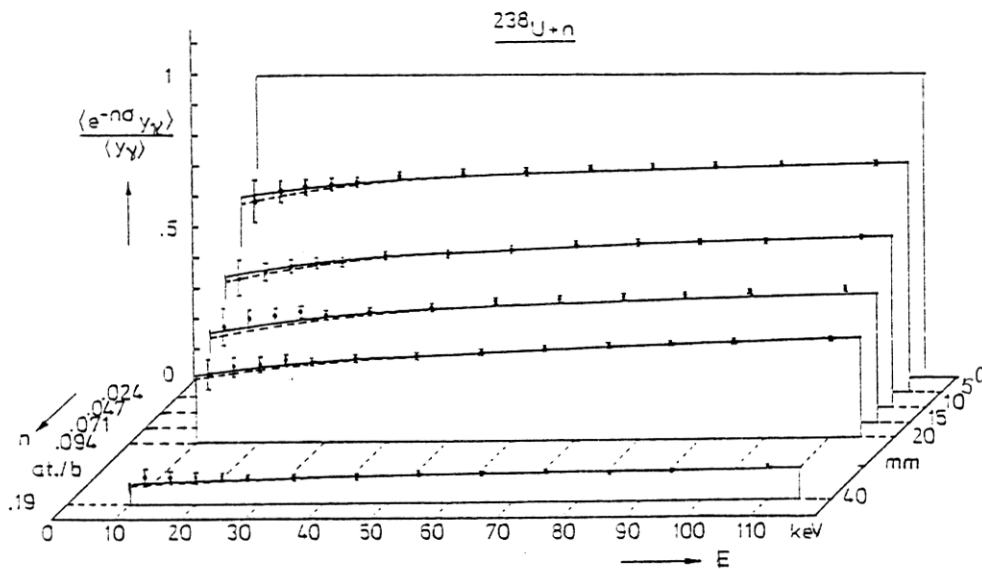


Fig. 25. Self-indication ratios measured by Bokhovko et al. (1988) (point symbols) and curves generated with Monte Carlo techniques from JEF-2 average resonance parameters (solid lines). Also shown are ratios calculated without corrections for resonance self-shielding and multiple scattering (broken lines). From Fröhner (1991).

$$\begin{aligned} T &= 300, 900, 1500, 3100 \text{ K} , \\ \sigma_d &= 0, 1, 10, 100, 1000, 10\,000, 100\,000, 1\,000\,000 \text{ b} . \end{aligned}$$

The self-shielded group cross section

$$\bar{\sigma}_x \equiv f_x \langle \sigma_x \rangle \quad (368)$$

is defined so that multiplication with the group-averaged flux φ gives the correct reaction rate. With the definition of the covariance one can write

$$f_x = 1 + \text{cov} \left(\frac{\varphi}{\langle \varphi \rangle}, \frac{\sigma_x}{\langle \sigma_x \rangle} \right) . \quad (369)$$

Now the flux is low where the cross section is high, so the two are anticorrelated, the covariance is negative, hence $f_x < 1$. On the other hand f_x must be positive since otherwise the average reaction rate would become negative. It follows (at least in the case of many levels within the group interval) that one has $0 < f_x < 1$. We can be more explicit by invoking the narrow-resonance approximation, valid in the important case that the resonances are narrow as compared to the mean energy loss of scattered neutrons. In this approximation the flux is proportional to the reciprocal macroscopic total cross section, $\varphi \propto 1/(\sigma + \sigma_d)$, where $\sigma = \sum_x \sigma_x$ is the total cross section of the nuclide considered. One has then in narrow-resonance approximation

$$f_x = \frac{\langle \sigma_x / (\sigma + \sigma_d) \rangle}{\langle \sigma_x \rangle \langle 1 / (\sigma + \sigma_d) \rangle} = \frac{\int_0^\infty dn e^{-n\sigma_d} \langle e^{-n\sigma} \sigma_x \rangle / \langle \sigma_x \rangle}{\int_0^\infty dn e^{-n\sigma_d} \langle e^{-n\sigma} \rangle} , \quad (370)$$

Since σ_d is a constant in the Bondarenko scheme one recognises that $f_x \rightarrow 1$ if either $T \rightarrow \infty$ (smooth total cross section) or $\sigma_d \rightarrow \infty$ (infinite dilution). Therefore $\langle \sigma_x \rangle$ is called the group cross section for infinite dilution (or the unshielded group cross section). In groups containing many resonances it is just the average cross section in the usual sense.

The last expression shows how self-shielding factors are related to self-indication ratios, Eq. 365, and average transmissions, Eq. 362. If those latter quantities can be predicted accurately for thick samples the self-shielding factor, too, can be predicted well. With the results shown in Figs. 15 and 16, and because of the positive correlation between numerator and denominator in the last equation, it is concluded that the self-shielding factors for the unresolved resonance region of ^{238}U can be calculated to 1-2% accuracy from the JEF-2 average resonance parameters.

4.3.2. Analytic and Monte Carlo Methods for Group Constant Generation

The practically most important technique for group constant generation is the analytic method (Frölich 1965, Hwang 1965). The averages in the last equation are calculated on the basis of level statistics in narrow-resonance approximation. The simplest version includes the following additional approximations:

- Cross sections are written as sums over SLBW terms ("many-level" Breit-Wigner approximation).
- Doppler broadening is described by the symmetric and asymmetric Voigt profiles ψ and χ .
- Interference between resonance and potential scattering (terms with χ) are neglected
- Level-statistical averages are calculated for each level sequence with the other sequences approximately represented by a smooth cross section included in σ_d .

The result can be written in the form

$$\bar{\sigma}_x = f_x \langle \sigma_x \rangle = (\sigma_p + \sigma_d) \left(1 - \sum_s \frac{\langle \Gamma J \rangle_s}{D_s} \right)^{-1} \sum_s \frac{\langle \Gamma_x J \rangle_s}{D_s \cos 2\varphi_s}, \quad (371)$$

where σ_p is the potential scattering cross section of the nuclide considered, $\langle \dots \rangle_s$ denotes an average over all partial widths for the s -th level sequence, the summations are over all sequences, and J is the integral

$$J(\kappa, \beta) \equiv \int_0^\infty dx \frac{\psi(x, \beta)}{\psi(x, \beta) + \kappa} \quad (372)$$

introduced by Dresner (1960). It involves the symmetric Voigt profiles (compare Eq. 244 and Appendix B)

$$\psi(x, \beta) = \frac{1}{\beta\sqrt{\pi}} \int_{-\infty}^{\infty} \exp \left[- \left(\frac{x-y}{\beta} \right)^2 \right] \frac{dy}{1+y^2}, \quad (373)$$

where $\beta \equiv 2\Delta/\Gamma$ is the Doppler width in units of the natural half width at half maximum, $\Gamma/2$, and $x \equiv 2(E - E_0)/\Gamma$ is the distance to the resonance peak at E_0 in the same units. Furthermore,

$$\kappa = \frac{\sigma_d + r}{\hat{\sigma}}, \quad (374) \quad \hat{\sigma} = 4\pi\lambda^2 g \frac{\Gamma_n}{\Gamma} \cos 2\varphi, \quad (375)$$

with r describing eigenvalue repulsion in approximate form. This is the fastest method available for group constant generation. It is employed in many widely used codes, e. g. ETOX (Schenter et al. 1969), MIGROS (Broeders and Krieg 1977), NJOY (MacFarlane et al. 1982), and GRUCON (Sinitza 1983).

The slowing-down method uses Monte Carlo sampled resonance ladders so that the calculation of average reaction rates can be reduced to the case of resolved resonances. The TIMS code (Takano et al. 1980) is an example. Monte Carlo sampled resonance ladders are also used in the subgroup/multiband methods pioneered by Nikolaev et al. (1970) and Cullen (1974) (see also Ribon and Maillard 1986). One stores, for each of few (e. g. four) subgroups/bands, the weights α_i and the band averages σ_i , σ_{xi} representing in a crude way the cross section distribution within an energy group. They must be found by matching averages obtained from ladder cross sections as follows,

$$\langle \sigma_x \rangle = \sum_i \alpha_i \sigma_{xi}, \quad (376) \quad \langle \sigma \rangle = \sum_i \alpha_i \sigma_i, \quad (377)$$

$$\left\langle \frac{\sigma_x}{\sigma + \sigma_d} \right\rangle = \sum_i \frac{\alpha_i \sigma_{xi}}{\sigma_i + \sigma_d}, \quad (378) \quad \left\langle \frac{1}{\sigma + \sigma_d} \right\rangle = \sum_i \frac{\alpha_i}{\sigma_i + \sigma_d}. \quad (379)$$

The subgroup/multiband method is essentially a coarse but efficient variant of the probability table method (Levitt 1972) where one generates from sampled resonance ladders the whole multivariate probability density

$$p(\sigma, \sigma_n, \sigma_\gamma, \dots) = p(\sigma)p(\sigma_n|\sigma)p(\sigma_\gamma|\sigma, \sigma_n) \dots \quad (380)$$

The distribution of the total cross section, $p(\sigma)$, is stored together with the conditional probabilities $p(\sigma_n|\sigma)$, $p(\sigma_\gamma|\sigma, \sigma_n)$ etc. in suitably discretised form, so that macroscopic

(isotope-weighted, Doppler broadened) cross sections rather than resonance parameters may be sampled directly.

5. Concluding remarks

The first part of the present overview is devoted to the probabilistic tools of data evaluators, with the objective to make readers aware of how modern probability theory has advanced and simplified the treatment of data uncertainties, in particular with respect to utilisation of prior knowledge, assignment of probabilities representing vague or global information, and handling of systematic errors and the correlations induced by them – areas where conventional statistics had not much to offer. The Bayesian scheme of learning from observations is illustrated with typical applications to counting statistics, data evaluation, and model fitting. Bayesian parameter estimation leaves no room for guesswork about “estimators” and their more or less desirable properties (absence of bias, sufficiency, efficiency, admissibility etc.) once prior, statistical model and loss function are specified. Under quadratic loss the optimal estimate consists of (posterior) mean value and mean square error (or, for multivariate problems, mean vector and covariance matrix) rather than modes or medians and frequentist confidence intervals. The least-squares method emerges as a natural consequence of the information-theoretic maximum-entropy principle in the practically most important cases where the given input consists of data with standard (root-mean-square) errors and, perhaps, correlations. Generalisation to nonlinear models is straightforward. There has also been some progress with respect to discrepant data but there is need for more work. The ENDF format used world-wide for evaluated nuclear reaction data is not yet satisfactory for the storage of uncertainty information. Storage of the intuitively clear standard errors and correlation coefficients instead of the inconvenient and error-prone variances and covariances should be admitted and encouraged.

The remainder of the overview is specifically devoted to the evaluation of resolved and unresolved resonance data. The close relationship between R-matrix theory, the statistical model of resonance reactions, Hauser-Feshbach theory with width-fluctuation corrections and the optical model is explained. Doppler broadening is treated in some detail for the practically important resonance formalisms and for tabulated cross sections. Experimental effects like resolution broadening, self-shielding and multiple-scattering, backgrounds and impurities are also discussed. The entire exposition is brief by necessity but it is hoped that enough material and references are presented to newcomers and non-specialists to have an adequate starting base for further study and professional work. Especially in the unresolved resonance region there is considerable need for methods development and creative programming, in both the fission and the fusion reactor field. Missing level estimation is still based essentially on the maximum likelihood approximation as reviewed by Fröhner (1983). A more rigorous Bayesian approach seems feasible and worthwhile but has to be worked out. Those who are particularly interested in analysis and evaluation of resonance data or in a rigorous R-matrix treatment of the spherical optical model will find additional material in the 1988 ICTP proceedings (Fröhner 1989a). Nuclear data evaluation methods and procedures were also discussed in considerable depth, with emphasis on practical experience, by Poenitz (1981) and Bhat (1981).

ACKNOWLEDGMENTS. My sincere thanks are due to P. Finck (now at ANL) who invited me to Cadarache, to R. Jacqmin (CEA Cadarache) and C. Nordborg (NEADB) who organised a four months visit, to P. Bioux (EdF) who urged me to write this paper, and to O. Bouland and E. Fort and the other colleagues who made my visit to Cadarache scientifically challenging and rewarding and generally enjoyable.

Appendices

APPENDIX A: PRACTICALLY IMPORTANT PROBABILITY DISTRIBUTIONS

In this appendix we summarise briefly the probability distributions which are most important in nuclear data evaluation and analysis, together with their parameter estimates under quadratic loss (means and variances). The notation is as follows:

| | |
|-------------|---|
| $P(A I)$ | probability of A given information I |
| $p(x I) dx$ | infinitesimal probability of x in dx given I , with probability density $p(x I)$ |

univariate distributions

| | |
|--|--|
| $\langle x \rangle$ | mean (expectation value) |
| $\text{var } x \equiv \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$ | variance (mean square error) |
| $\Delta x \equiv \sqrt{\text{var } x}$ | standard deviation (root-mean-square or standard error) |

multivariate distributions

| | |
|---|--|
| $\langle \mathbf{x} \rangle$ | mean vector |
| $d(\mathbf{x}) = \prod_{\nu} dx_{\nu}$ | volume element in \mathbf{x} -space |
| $\mathbf{C} = \langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^{\dagger} \rangle$ | covariance matrix |
| $C_{\mu\nu} = \Delta x_{\mu} \rho_{\mu\nu} \Delta x_{\nu}$ | covariance matrix element |
| $\rho_{\mu\nu} = \rho_{\nu\mu}$ | correlation coefficient |
| | $(-1 \leq \rho_{\mu\nu} \leq +1, \quad \rho_{\nu\nu} = 1)$ |

The dagger indicates transposition (Hermitean conjugation of real vectors and matrices). Expectation values are denoted by angular brackets, $\langle x \rangle$, sample averages by overbars, \bar{x} .

A.1. Binomial and Beta Distributions

Applications: Bernoulli trials with two possible outcomes (success or failure, positive or negative parity, fermion level occupied or not, ...).

Sampling distribution (probability of success, in 1 trial):

$$P(1|1, \theta) = \theta . \tag{A1}$$

Likelihood function for s successes in n trials (binomial distribution):

$$P(s|n, \theta) = \binom{n}{s} \theta^s (1 - \theta)^{n-s} \quad s = 0, 1, 2, \dots, n \geq 1 . \tag{A2}$$

Case 1 – Total ignorance about parameter θ , admitting even the possibility that there is one alternative only, $\theta = 0$ or 1.

Least informative (group-theoretic) prior (Haldane's rule, see Jaynes 1968):

$$p(\theta)d\theta \propto \frac{d\theta}{\theta(1-\theta)} , \quad 0 \leq \theta \leq 1 . \tag{A3}$$

Posterior (beta distribution):

$$p(\theta|s, n)d\theta = B(s, n-s)^{-1} \theta^{s-1} (1-\theta)^{n-s-1} d\theta , \quad 0 \leq \theta \leq 1 , \tag{A4}$$

with $B(x, y) \equiv \Gamma(x)\Gamma(y)/\Gamma(x+y)$ (beta function) and, for $x = n$ integer, $\Gamma(n) = (n-1)!$ (gamma function). Parameter estimates under quadratic loss:

$$\langle \theta \rangle = \frac{s}{n}, \quad (A5) \quad \text{var } \theta = \frac{1}{n+1} \frac{s}{n} \left(1 - \frac{s}{n}\right). \quad (A6)$$

As long as only successes or only failures occur the probability remains overwhelmingly concentrated at $\theta = 1$ or $\theta = 0$, with zero variance. As soon as there is at least 1 success ($s \geq 1$) and one failure ($n-s \geq 1$) other expectation values are obtained and the variance becomes finite.

Case 2 – If there is no doubt a priori that one has genuine Bernoulli trials with two alternatives, the appropriate prior is equal to what Eq. (A4) would give after observation of one success and one failure (Bayes-Laplace rule):

$$p(\theta)d\theta = d\theta, \quad 0 < \theta < 1. \quad (A7)$$

Posterior (beta distribution):

$$p(\theta|s, n)d\theta = \frac{(n+1)!}{(n-s)!s!} \theta^s (1-\theta)^{n-s} d\theta, \quad 0 < \theta < 1. \quad (A8)$$

Estimate under quadratic loss:

$$\langle \theta \rangle = \frac{s+1}{n+2}, \quad (A9) \quad \text{var } \theta = \frac{1}{n+3} \frac{s+1}{n+2} \left(1 - \frac{s+1}{n+2}\right). \quad (A10)$$

(A9) is Laplace's rule of succession.

A.2. Poisson and Gamma Distributions

Applications: Radioactive decay, counting statistics, rare events with constant average rate of occurrence.

Average time interval

$$\langle t \rangle = \tau \quad (A11)$$

Maximum entropy distribution for this constraint, Lagrange parameter λ :

$$p(t|\tau) dt = e^{-\lambda t} \lambda dt, \quad 0 < t < \infty \quad (A12)$$

(exponential interval distribution) with

$$\langle t \rangle = \tau = 1/\lambda = \Delta t. \quad (A13)$$

Probability for n events between 0 and t , in time-ordered but else arbitrary intervals dt_1, dt_2, \dots, dt_n (Poisson distribution):

$$\begin{aligned} P(n|\lambda, t) &= e^{-\lambda t} \lambda^n \int_0^t dt_n \int_0^{t_n} dt_{n-1} \dots \int_0^{t_2} dt_1 \\ &= e^{-\lambda t} \lambda^n \frac{1}{n!} \int_0^t dt_n \int_0^{t_n} dt_{n-1} \dots \int_0^{t_2} dt_1 \\ &= \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots, \end{aligned} \quad (A14)$$

with expectation values

$$\langle n \rangle = \lambda t, \quad (A15) \quad \text{var } n = \lambda t. \quad (A16)$$

Prior for rate λ (Jeffreys' prior for scale parameters)

$$p(\lambda) d\lambda \propto d \ln \lambda = \frac{d\lambda}{\lambda}, \quad 0 < \lambda < \infty. \quad (A17)$$

The likelihood function for n events during observation time t is the Poisson distribution (A14).

Posterior (gamma distribution):

$$p(\lambda|t, n) d\lambda = \frac{e^{-\lambda t} (\lambda t)^n}{\Gamma(n)} \frac{d\lambda}{\lambda}, \quad 0 < \lambda < \infty. \quad (A18)$$

Estimate under quadratic loss:

$$\langle \lambda \rangle = \text{var } \lambda = \frac{n}{t}, \quad (A19) \quad \frac{\Delta \lambda}{\lambda} = \frac{1}{\sqrt{n}}. \quad (A20)$$

Note: Posterior and estimate are the same if number of events n and observation time t were accumulated during several distinct measurements (runs).

A.3. Univariate Gaussian

Applications: Unknown errors, uncontrollable fluctuations, error propagation, combination of data from various sources, etc., valid if many independent components act together (central limit theorem) or if only means and standard deviations are known (maximum entropy principle) with possible errors or deviations between $-\infty$ and $+\infty$.

Sampling distribution (probability of possible deviation or error $x - \mu$, given true value μ and standard deviation σ):

$$p(x|\mu, \sigma) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] dx, \quad -\infty < x < \infty. \quad (A21)$$

Likelihood for x -values in dx_1, \dots, dx_n at x_1, \dots, x_n :

$$\begin{aligned} & p(x_1, \dots, x_n | \mu, \sigma) dx_1 \dots dx_n \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left[-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2 \right] dx_1 \dots dx_n \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{ns'^2}{2\sigma^2} \left[1 + \left(\frac{\bar{x} - \mu}{s'} \right)^2 \right] \right\} dx_1 \dots dx_n. \end{aligned} \quad (A22)$$

This depends on the sample only through the sample mean and the sample variance,

$$\bar{x} \equiv \frac{1}{n} \sum_{j=1}^n x_j, \quad (A23) \quad s'^2 \equiv \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 = \overline{x^2} - \bar{x}^2. \quad (A24)$$

Case 1 – Location parameter μ and scale parameter σ both unknown.

Least informative (group-theoretic) prior:

$$p(\mu, \sigma) d\mu d\sigma \propto \frac{d\mu d\sigma}{\sigma}, \quad -\infty < \mu < \infty, \quad 0 < \sigma < \infty. \quad (\text{A25})$$

Posterior:

$$\begin{aligned} p(\mu, \sigma | \bar{x}, s', n) \\ &= p(u|v, n) du p(v|n) dv = \frac{e^{-vu^2}}{\sqrt{\pi}} \sqrt{v} du \frac{e^{-v} v^{(n-1)/2}}{\Gamma(\frac{n-1}{2})} \frac{dv}{v} \\ &= p(v|u, n) dv p(u|n) du = \frac{e^{-(1+u^2)v} [(1+u^2)v]^{n/2}}{\Gamma(\frac{n}{2})} \frac{dv}{v} \frac{du}{\text{B}(\frac{1}{2}, \frac{n-1}{2})(1+u^2)^{n/2}}, \\ &-\infty < u \equiv \frac{\mu - \bar{x}}{s'} < \infty, \quad 0 < v \equiv \frac{ns'^2}{2\sigma^2} < \infty, \end{aligned} \quad (\text{A26})$$

The two factorisations correspond to the two forms of the fundamental product rule for joint probabilities.

Marginal distribution for u (Student's t distribution with $t = u\sqrt{n-1}$, $n-1$ degrees of freedom):

$$p(u|n) du = \frac{du}{\text{B}(\frac{1}{2}, \frac{n-1}{2})(1+u^2)^{n/2}}, \quad -\infty < u < \infty, \quad (\text{A27})$$

and for v (gamma or chi-square distribution with $\chi^2 = 2v$, $n-1$ degrees of freedom):

$$p(v|n) dv = \frac{e^{-v} v^{(n-1)/2}}{\Gamma(\frac{n-1}{2})} \frac{dv}{v}, \quad 0 < v < \infty. \quad (\text{A28})$$

Parameter estimate under quadratic loss: With the marginal distributions one finds readily $\langle u \rangle$, $\langle v \rangle$, $\text{var } u$, $\text{var } v$ and the estimates

$$\langle \mu \rangle = \bar{x}, \quad (\text{A29}) \quad \text{var } \mu = \frac{s'^2}{n-3}, \quad (\text{A30})$$

$$\langle \sigma^{-2} \rangle = \frac{n-1}{n} s'^{-2}, \quad (\text{A31}) \quad \text{var } \sigma^{-2} = 2 \frac{n-1}{n^2} s'^{-4}. \quad (\text{A32})$$

The estimated parameters are uncorrelated, $\text{cov}(u, v) = \langle uv \rangle - \langle u \rangle \langle v \rangle = 0$ implies

$$\text{cov}(\mu, \sigma^{-2}) = 0. \quad (\text{A33})$$

Case 2 – If σ is known the posterior is simply

$$p(\mu|\sigma, n) d\mu = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\mu - \bar{x})^2}{2\sigma^2/n}\right] d\mu, \quad -\infty < \mu < \infty. \quad (\text{A34})$$

Estimate under quadratic loss:

$$\langle \mu \rangle = \bar{x}, \quad (\text{A35}) \quad \text{var } \mu = \frac{\sigma^2}{n}. \quad (\text{A36})$$

Case 3 – Repeated uncorrelated measurements of μ , results reported as $x_1 \pm \sigma_1, \dots, x_n \pm \sigma_n$.

Sampling distribution with maximal entropy:

$$p(x_j | \sigma_j) dx_j = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp \left[-\frac{1}{2} \left(\frac{x_j - \mu}{\sigma_j} \right)^2 \right] dx_j, \quad -\infty < x_j < \infty. \quad (A37)$$

Likelihood function:

$$p(x_1, \dots, x_n | \mu, \sigma_1, \dots, \sigma_n) dx_1 \dots dx_n \propto \exp \left[-\frac{1}{2} \sum_{j=1}^n \left(\frac{x_j - \mu}{\sigma_j} \right)^2 \right] dx_1 \dots dx_n \quad (A38)$$

Posterior:

$$p(\mu | \{x_j, \sigma_j\}) d\mu = \frac{1}{\sqrt{2\pi\bar{\sigma}^2/n}} \exp \left[-\frac{(\mu - \bar{x})^2}{2\bar{\sigma}^2/n} \right] d\mu, \quad -\infty < \mu < \infty, \quad (A39)$$

where the overbars denote weighted averages over the sample (over measurements),

$$\bar{x} \equiv \frac{\sum_j \sigma_j^{-2} x_j}{\sum_j \sigma_j^{-2}}, \quad (A40) \quad \bar{\sigma}^2 \equiv \frac{\sum_j \sigma_j^{-2} \sigma_j^2}{\sum_j \sigma_j^{-2}} = \frac{n}{\sum_j \sigma_j^{-2}}. \quad (A41)$$

Estimate under quadratic loss:

$$\langle \mu \rangle = \bar{x}, \quad (A42) \quad \text{var } \mu = \frac{\bar{\sigma}^2}{n}, \quad (A43)$$

A.4. Multivariate Gaussian

Applications: Propagation of correlated errors, least-squares fitting in multidimensional sample and parameter spaces. Derivations follow closely those for the univariate Gaussian, resulting in formally similar vector and matrix expressions. Roman subscripts (j, k) denote trials or sample members, Greek subscripts (ν, κ) denote parameters (Cartesian coordinates in parameter space).

Sampling distribution (probabilities of possible vector deviations or errors $\mathbf{x} - \boldsymbol{\mu}$, given true vector $\boldsymbol{\mu}$ and covariance matrix \mathbf{C}):

$$p(\mathbf{x} | \boldsymbol{\mu}, \mathbf{C}) d(\mathbf{x}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\dagger \mathbf{C}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] d(\mathbf{x}), \quad -\infty < x_\nu < \infty. \quad (A44)$$

Likelihood for \mathbf{x} -vectors in $d(\mathbf{x}_1), \dots, d(\mathbf{x}_n)$ at $\mathbf{x}_1, \dots, \mathbf{x}_n$:

$$\begin{aligned} p(\mathbf{x}_1, \dots, \mathbf{x}_n | \boldsymbol{\mu}, \mathbf{C}) d(\mathbf{x}_1) \dots d(\mathbf{x}_n) \\ &= \frac{1}{\det(2\pi\mathbf{C})^{n/2}} \exp \left[-\frac{1}{2} \sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu})^\dagger \mathbf{C}^{-1} (\mathbf{x}_j - \boldsymbol{\mu}) \right] d(\mathbf{x}_1) \dots d(\mathbf{x}_n) \\ &= \frac{1}{\det(2\pi\mathbf{C})^{n/2}} \exp \left[-\frac{n}{2} \left[\text{tr}(\bar{\mathbf{C}}\mathbf{C}^{-1}) + (\bar{\mathbf{x}} - \boldsymbol{\mu})^\dagger \mathbf{C}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \right] \right] d(\mathbf{x}_1) \dots d(\mathbf{x}_n). \end{aligned} \quad (A45)$$

This depends on the sample only through the sample mean vector and the sample covariance matrix,

$$\bar{\mathbf{x}} \equiv \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j, \quad (A46) \quad \bar{\mathbf{C}} \equiv \frac{1}{n} \sum_{j=1}^n (\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})^\dagger = \overline{\mathbf{x}\mathbf{x}^\dagger} - \bar{\mathbf{x}}\bar{\mathbf{x}}^\dagger. \quad (A47)$$

Case 1 – Location parameters μ_ν and covariance matrix elements $C_{\nu\nu'} = \sigma_\nu \rho_{\nu\nu'} \sigma_{\nu'}$ unknown.

Least informative (group-theoretic) prior in m -dimensional parameter space ($\nu = 1, \dots, m$):

$$p(\boldsymbol{\mu}, \mathbf{C}) d(\boldsymbol{\mu}) d(\mathbf{C}) \propto \frac{d(\boldsymbol{\mu}) d(\mathbf{C})}{\det \mathbf{C}^{(m+1)/2}} = \frac{d(\boldsymbol{\mu}) d(\mathbf{C}^{-1})}{\det \mathbf{C}^{-(m+1)/2}},$$

$$-\infty < \mu_\nu < \infty, \quad 0 < \sigma_\nu^2 \equiv C_{\nu\nu} < \infty, \quad -1 < \rho_{\nu\nu'} \equiv \frac{C_{\nu\nu'}}{\sqrt{C_{\nu\nu} C_{\nu'\nu'}}} < +1. \quad (\text{A48})$$

Posterior: With the identities $\exp(-\text{tr } \mathbf{V}) = \det \{\exp(-\mathbf{V})\}$ and $\det(\mathbf{1} + \mathbf{u}\mathbf{u}^\dagger) = 1 + \mathbf{u}^2$ one finds

$$\begin{aligned} p(\boldsymbol{\mu}, \mathbf{C} | \bar{\mathbf{x}}, \bar{\mathbf{C}}, n) \\ &= p(\mathbf{u} | \mathbf{V}, n) d(\mathbf{u}) p(\mathbf{V} | n) d(\mathbf{V}) \\ &= \frac{\exp(-\mathbf{u}^\dagger \mathbf{V} \mathbf{u}) d(\mathbf{u})}{\sqrt{\det(\pi \mathbf{V}^{-1})}} \frac{\det \{\exp(-\mathbf{V}) \mathbf{V}^{(n-1)/2}\}}{\pi^{m(m-1)/4} \prod_{\nu=1}^m \Gamma(\frac{n-\nu}{2})} \frac{d(\mathbf{V})}{\det \mathbf{V}^{(m+1)/2}} \\ &= p(\mathbf{u} | n) d(\mathbf{u}) p(\mathbf{V} | \mathbf{u}, n) d(\mathbf{V}) \\ &= \frac{(1 + \mathbf{u}^2)^{-n/2} d(\mathbf{u})}{\prod_{\nu=1}^m \text{B}(\frac{1}{2}, \frac{n-1}{2})} \frac{\det \{\exp[-(\mathbf{1} + \mathbf{u}\mathbf{u}^\dagger) \mathbf{V}] [(\mathbf{1} + \mathbf{u}\mathbf{u}^\dagger) \mathbf{V}]^{n/2}\}}{\pi^{m(m-1)/4} \prod_{\nu=1}^m \Gamma(\frac{n+1-\nu}{2})} \frac{d(\mathbf{V})}{\det \mathbf{V}^{(m+1)/2}}, \end{aligned} \quad (\text{A49})$$

where, in close analogy to the univariate case, we defined the vector

$$\mathbf{u} \equiv \mathbf{S}'^{-1}(\boldsymbol{\mu} - \bar{\mathbf{x}}), \quad -\infty < u_\nu < \infty, \quad (\text{A50})$$

and the positive definite, real symmetric matrix (with real positive eigenvalues)

$$\mathbf{V} \equiv \frac{n}{2} \mathbf{S}' \mathbf{C}^{-1} \mathbf{S}', \quad 0 < \det \mathbf{V} < \infty, \quad (\text{A51})$$

with the positive definite, real symmetric matrix \mathbf{S}' defined by

$$\bar{\mathbf{C}} = \mathbf{S}'^2 \quad (\text{A52})$$

(in the principal-axes system of $\bar{\mathbf{C}}$). The two factorisations correspond to the two forms of the fundamental product rule for joint probabilities.

Marginal distribution for \mathbf{u} (multivariate t distribution)

$$p(\mathbf{u} | n) d(\mathbf{u}) = \frac{(1 + \mathbf{u}^2)^{-n/2} d(\mathbf{u})}{\prod_{\nu=1}^m \text{B}(\frac{1}{2}, \frac{n-1}{2})} \quad (\text{A53})$$

and for \mathbf{V} (multivariate gamma distribution or Wishart distribution)

$$p(\mathbf{V} | n) d(\mathbf{V}) = \frac{\det \{\exp(-\mathbf{V}) \mathbf{V}^{(n-1)/2}\}}{\pi^{m(m-1)/4} \prod_{\nu=1}^m \Gamma(\frac{n-\nu}{2})} \frac{d(\mathbf{V})}{\det \mathbf{V}^{(m+1)/2}}. \quad (\text{A54})$$

Parameter estimate under quadratic loss: With the marginal distributions one finds $\langle \mathbf{u} \rangle$, $\langle \mathbf{u}\mathbf{u}^\dagger \rangle$, $\langle \mathbf{V} \rangle$, and finally

$$\langle \boldsymbol{\mu} \rangle = \bar{\mathbf{x}}, \quad (A55) \quad \langle (\boldsymbol{\mu} - \bar{\mathbf{x}})(\boldsymbol{\mu} - \bar{\mathbf{x}})^\dagger \rangle = \frac{\bar{\mathbf{C}}}{n - m - 2}, \quad (A56)$$

$$\langle \mathbf{C}^{-1} \rangle = \frac{n-1}{n} \bar{\mathbf{C}}^{-1}. \quad (A57)$$

Case 2 – If \mathbf{C} is known the posterior is simply

$$p(\boldsymbol{\mu} | \mathbf{C}, n) d(\boldsymbol{\mu}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C}/n)}} \exp \left[-\frac{n}{2} (\boldsymbol{\mu} - \bar{\mathbf{x}})^\dagger \mathbf{C}^{-1} (\boldsymbol{\mu} - \bar{\mathbf{x}}) \right] d(\boldsymbol{\mu}), \quad (A58)$$

$-\infty < \mu_\nu < \infty.$

Estimate under quadratic loss:

$$\langle \boldsymbol{\mu} \rangle = \bar{\mathbf{x}}, \quad (A59) \quad \langle (\boldsymbol{\mu} - \langle \boldsymbol{\mu} \rangle)(\boldsymbol{\mu} - \langle \boldsymbol{\mu} \rangle)^\dagger \rangle = \frac{\mathbf{C}}{n}. \quad (A60)$$

Case 3 – Repeated correlated measurements of $\boldsymbol{\mu}$, results reported as vectors \mathbf{x}_j and covariance matrices \mathbf{C}_{jk} ($j, k = 1, \dots, n$).

Definitions:

$$\mathbf{x} \equiv \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \quad (A61) \quad \mathbf{C} \equiv \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \dots & \mathbf{C}_{1n} \\ \mathbf{C}_{12} & \mathbf{C}_{22} & \dots & \mathbf{C}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{1n} & \mathbf{C}_{2n} & \dots & \mathbf{C}_{nn} \end{pmatrix}. \quad (A62)$$

Likelihood:

$$p(\mathbf{x} | \boldsymbol{\mu}, \mathbf{C}) d(\mathbf{x}) \propto e^{-Q/2} d(\mathbf{x}), \quad (A63)$$

where

$$\begin{aligned} Q &= \sum_{j,k} (\mathbf{x}_j - \boldsymbol{\mu})^\dagger (\mathbf{C}^{-1})_{jl} (\mathbf{x}_k - \boldsymbol{\mu}) \\ &= \sum_{j,k} \text{tr} [(\mathbf{C}^{-1})_{jk} (\boldsymbol{\mu} - \mathbf{x}_k)(\boldsymbol{\mu} - \mathbf{x}_j)^\dagger] \\ &= n \text{tr} [(\overline{\mathbf{x}\mathbf{x}^\dagger} - \overline{\mathbf{x}\mathbf{x}}^\dagger) \bar{\mathbf{C}}^{-1}] + n(\boldsymbol{\mu} - \bar{\mathbf{x}})^\dagger \bar{\mathbf{C}}^{-1} (\boldsymbol{\mu} - \bar{\mathbf{x}}). \end{aligned} \quad (A64)$$

The overbars denote weighted averages over the sample (over measurements),

$$\bar{\mathbf{x}} \equiv \left(\sum_{j,k} (\mathbf{C}^{-1})_{jk} \right)^{-1} \sum_{j,k} (\mathbf{C}^{-1})_{jk} \mathbf{x}_k, \quad (A65)$$

$$\overline{\mathbf{x}\mathbf{x}^\dagger} \equiv \left(\sum_{j,k} (\mathbf{C}^{-1})_{jk} \right)^{-1} \sum_{j,k} (\mathbf{C}^{-1})_{jk} \mathbf{x}_k \mathbf{x}_j^\dagger, \quad (A66)$$

$$\bar{\mathbf{C}} \equiv \left(\sum_{j,k} (\mathbf{C}^{-1})_{jk} \right)^{-1} \sum_{j,k} (\mathbf{C}^{-1})_{jk} (\mathbf{C}^{-1})_{jk} \mathbf{C}_{kj} = n \left(\sum_{j,k} (\mathbf{C}^{-1})_{jk} \right)^{-1} \quad (A67)$$

Posterior:

$$p(\boldsymbol{\mu}|\mathbf{x}, \mathbf{C}) d(\boldsymbol{\mu}) = \frac{1}{\sqrt{\det(2\pi\overline{\mathbf{C}}/n)}} \exp\left[-\frac{n}{2}(\boldsymbol{\mu} - \overline{\mathbf{x}})^\dagger \overline{\mathbf{C}}^{-1}(\boldsymbol{\mu} - \overline{\mathbf{x}})\right] d(\boldsymbol{\mu})$$

$$-\infty < \mu_\nu < \infty, \quad (A68)$$

Estimate under quadratic loss:

$$\langle \boldsymbol{\mu} \rangle = \overline{\mathbf{x}}, \quad (A69) \quad \langle (\boldsymbol{\mu} - \langle \boldsymbol{\mu} \rangle)(\boldsymbol{\mu} - \langle \boldsymbol{\mu} \rangle)^\dagger \rangle = \frac{\overline{\mathbf{C}}}{n} \quad (A70)$$

APPENDIX B: MATHEMATICAL PROPERTIES OF THE VOIGT PROFILES ψ AND χ

The shapes of Doppler broadened isolated resonances can be described by the symmetric and asymmetric Voigt profiles $\psi(x, \beta)$ and $\chi(x, \beta)$. The arguments

$$x \equiv \frac{E - E_0}{\Gamma/2}, \quad (B1) \quad \beta \equiv \frac{\Delta}{\Gamma/2} \quad (B2)$$

depend on the resonance energy E_0 , the total width Γ , the Doppler width Δ (see Eq. 110), and the bombarding energy E (all in the laboratory system).

Definition:

$$\psi(x, \beta) = \frac{1}{\beta\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x-x')^2/\beta^2} \frac{dx'}{1+x'^2} = \psi(-x, \beta) \quad (B3)$$

$$\chi(x, \beta) = \frac{1}{\beta\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x-x')^2/\beta^2} \frac{x'dx'}{1+x'^2} = -\chi(-x, \beta) \quad (B4)$$

Special arguments:

at resonance energy, $E = E_0$

$$\psi(0, \beta) = \frac{\sqrt{\pi}}{\beta} e^{1/\beta^2} \operatorname{erfc} \frac{1}{\beta} \quad (B5) \quad \chi(0, \beta) = 0 \quad (B6)$$

for zero temperature, $T = 0$

$$\psi(x, 0) = \frac{1}{1+x^2} \quad (B7) \quad \chi(x, 0) = \frac{x}{1+x^2} \quad (B8)$$

Convergent series:

$$\psi(x, \beta) = \frac{1}{\beta} e^{(1-x^2)/\beta^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{x}{\beta^2}\right)^{2n} \Gamma\left(-n + \frac{1}{2}, \frac{1}{\beta^2}\right) \quad (B9)$$

$$\chi(x, \beta) = \frac{1}{\beta} e^{(1-x^2)/\beta^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{x}{\beta^2}\right)^{2n+1} \Gamma\left(-n - \frac{1}{2}, \frac{1}{\beta^2}\right) \quad (B10)$$

where $\Gamma(a, t)$ is the incomplete gamma function, with

$$\Gamma(a + 1, t) = a\Gamma(a, t) + e^{-t}t^a = \int_t^\infty dt' e^{-t'} t'^a \quad (B11)$$

$$\Gamma\left(\frac{1}{2}, t\right) = \sqrt{\pi} \operatorname{erfc} \sqrt{t} \quad (B12)$$

Asymptotic series for low temperatures (small β):

$$\psi(x, \beta) = \sum_{n=0}^{\infty} \frac{(2n+1)!!}{2n+1} \left(-\frac{\beta^2}{2}\right)^n \left(\frac{1}{1+x^2}\right)^{n+1/2} \cos[(2n+1) \arctan x] \quad (B13)$$

$$\chi(x, \beta) = \sum_{n=0}^{\infty} \frac{(2n+1)!!}{2n+1} \left(-\frac{\beta^2}{2}\right)^n \left(\frac{1}{1+x^2}\right)^{n+1/2} \sin[(2n+1) \arctan x] \quad (B14)$$

whence

$$\psi(x, \beta) + i\chi(x, \beta) = \sum_{n=0}^{\infty} \frac{(2n+1)!!}{2n+1} \left(-\frac{\beta^2}{2}\right)^n \left(\frac{1}{1+x^2} + \frac{ix}{1+x^2}\right)^{2n+1} \quad (B15)$$

Relationship with complex probability integral:

$$\psi(x, \beta) + i\chi(x, \beta) = \frac{\sqrt{\pi}}{\beta} W\left(\frac{x+i}{\beta}\right), \quad (B16)$$

where

$$W(z) = \frac{1}{\pi i} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{t-z} dt = e^{z^2} \left(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{-t^2} dt\right) \quad (B17)$$

Derivatives:

$$\frac{\partial \psi}{\partial x} = \frac{2}{\beta^2} (\chi - x\psi) \quad (B18)$$

$$\frac{\partial \chi}{\partial x} = \frac{2}{\beta^2} (1 - \psi - x\chi) \quad (B19)$$

Integrals:

$$\int_{-\infty}^{\infty} \psi(x, \beta) dx = \pi \quad (B20)$$

$$\int_{-\infty}^{\infty} \chi(x, \beta) dx = 0 \quad (B21)$$

REFERENCES

- V.M. Adamov, *Conf. on Nucl. Cross Sections and Technol.*, NBS SP 594, p. 995, Knoxville (1979)
- F.T. Adler and D.B. Adler, *Nucl. Data for Reactors*, IAEA Vienna (1970) p. 777
- A. Anzaldo-Meneses, *Z. Physik A* **353** (1995) 295
- R. Arlt et al., *6-th All Union Conf. on Neutron Physics, Kiev*, vol. 2, p. 129 (1983)
- T. Bayes, *Phil. Trans. Roy. Soc.* **53** (1763) 370; reprinted in E.S. Pearson and M.G. Kendall, *Studies in the History of Statistics and Probability*, Hafner, Darien, Conn. (1970)
- J.M. Bernardo and A.F.M. Smith, *Bayesian Theory*, Wiley, Chichester (1994)
- J.O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer, New York (1985)
- J. Bernoulli, *Ars Conjectandi*, Thurnisiorum, Basel (1713); reprinted in *Die Werke von Jakob Bernoulli*, Birkhäuser, Basel (1975)
- K. Berthold, C. Nazareth, G. Rohr and H. Weigmann, *Proc. Int. Conf. on Nucl. Data for Sci. and Technol.*, J.K. Dickens (ed.), ANS, La Grange Park (1994) p. 218
- H.A. Bethe, *Rev. Mod. Phys.* **9** (1937) 69
- M.R. Bhat, *Proc. Conf. Nucl. Data Eval. Meth. and Proc.*, B.A. Magurno and S. Pearlstein (eds.), Brookhaven report BNL-NCS-51363, (1981) vol. I, p. 291
- M.R. Bhat and G.E. Lee-Whiting, *Nucl. Instr. Meth.* **47** (1967) 277
- L.C. Biedenharn, Oak Ridge report ORNL-1501 (1953)
- J.M. Blatt and L.C. Biedenharn, *Rev. Mod. Phys.* **24** (1952) 258
- C. Bloch, *Nucl. Phys.* **A112** (1968) 257, 273
- Å. Bohr, *Conf. on Peaceful Uses of Atomic Energy*, Geneva (1955) vol. 2, p. 151
- M.V. Bokhovko, V.N. Kononov, G.N. Manturov, E.D. Poletaev, V.V. Sinitsa, A.A. Vovodskij, *Yad. Konst.* **3** (1988) 11; Engl. transl.: IAEA report INDC(CCP)-322 (1990) p. 5
- L.M. Bollinger and G.E. Thomas, *Phys. Rev.* **171** (1968) 1293
- I.I. Bondarenko et al., *Group Constants for Nuclear Reactor Calculations*, Consultant Bureau Enterprises Inc., New York (1964)
- O. Bouland, private communication, C. E. Cadarache (1999)
- M. Born, *Optik*, Springer, Berlin (1933)
- T.A. Brody, J. Flores, J.B. French, P.A. Mello, A. Pandey and S.S. Wong, *Rev. Mod. Phys.* **53** (1981) 385
- A. Brusegan et al., IRMM Geel, private communication (1992)
- I. Broeders and B. Krieg, Karlsruhe report KfK 2388 (1977)
- A.G.W. Cameron and R.M. Elgin, *Can. J. Phys.* **43** (1965) 1288
- M. Cancé and G. Grenier, *Nucl. Sci. Eng.* **68** (1978) 197
- E.R. Cohen, K.M. Crowe, J.W.M. DuMond, *Fundamental Constants of Physics*, Interscience, New York (1957); see also E.R. Cohen and B.N. Taylor, *Phys. Today* **BG9** (Aug. 1992)
- R.T. Cox, *Am. J. Physics* **14** (1946) 1
- C.E. Cullen and C.R. Weisbin, *Nucl. Sci. Eng.* **60** (1976) 199
- C.E. Cullen, *Nucl. Sci. Eng.* **55** (1974) 387
- M.H. DeGroot, *Optimal Statistical Decisions*, McGraw-Hill, New York (1970)
- H. Derrien, G. de Saussure, N.M. Larson, L.C. Leal and R.B. Perez, *Nucl. Data for Sci. and Technol.*, S. Igarasi (ed.), JAERI (1988) p. 83
- G. de Saussure and R.B. Perez, Oak Ridge report ORNL-TM-2599 (1969)
- A. de-Shalit and I. Talmi, *Nuclear Shell Theory*, Acad. Press, New York – London (1963) ch. 15 and Appendix

- L. Dresner, *Proc. Int. Conf. on Neutron Reactions with the Nucleus*, Columbia U. 1957, Report CU-157 (1957) p. 71
- L. Dresner, *Resonance Absorption in Nuclear Reactors*, Pergamon, Oxford (1960)
- F.G. Dunnington, *Rev. Mod. Phys.* **11** (1939) 65
- B. Efron and C. Morris, *J. Roy. Statist. Soc.* **B 35** (1973) 379
- C.A. Engelbrecht and H.A. Weidenmüller, *Phys. Rev.* **C8** (1973) 859; N. Nishioka and H.A. Weidenmüller, *Phys. Letters* **157B** (1985) 101
- T. Ericson, *Adv. Phys.* **9** (1960) 425
- U. Fano and G. Racah, *Irreducible Tensor Sets*, Acad. Press, New York (1959)
- R.A. Fisher, *Annals of Eugenics* **6** (1935) 391; reprinted in R.A. Fisher, *Contributions to Mathematical Statistics*, J. Tukey (ed.), J. Wiley & Sons, New York (1950)
- R. Frölich, Karlsruhe report KfK 367 (1965)
- F.H. Fröhner, General Atomic report GA-6909 (1966)
- F.H. Fröhner, unpublished (1970)
- F.H. Fröhner, Karlsruhe report KfK 2145 (1977)
- F.H. Fröhner, *Conf. on Neutron Phys. and Nucl. Data*, Harwell (1978) p. 306
- F.H. Fröhner, Karlsruhe Report KfK-2669 (1978); reprinted in *Nuclear Theory for Applications*, Report IAEA-SMR-43, ICTP Trieste (1980) p. 59
- F.H. Fröhner, B. Goel, U. Fischer, *Proc. Meet. on Fast-Neutron Capture Cross Sections*, Argonne, ANL-83-4 (1982) p. 116
- F.H. Fröhner, *Nucl. Data for Sci. and Technol.*, K.H. Böckhoff (ed.), Reidel, Dordrecht (1983) p. 623; cf. also Karlsruhe report KfK 3553 (1983)
- F.H. Fröhner, *Proc Int. Conf. Nucl. Data for Basic and Appl. Sci., Santa Fe 1985*, New York etc. (1986) p. 1541; reprinted in *Rad. Effects* **96** (1986) 199
- F.H. Fröhner, *Nucl. Sci. Eng.* **103** (1989) 119
- F.H. Fröhner, *Applied Nuclear Theory and Nuclear Model Calculations for Nuclear Technology Applications*, M.K. Mehta and J.J. Schmidt (eds.), World Scientific, Singapore etc. (1989a) p. 170
- F.H. Fröhner, *Nuclear Physics, Neutron Physics and Nuclear Energy*, W. Andrejtscheff and D. Elenkov (eds.), World Scientific, Singapore (1990) p. 333; separately available as Karlsruhe report KfK 4655 (1990)
- F.H. Fröhner, Karlsruhe report KfK 4911 (1991); *Nucl. Sci. Eng.* **111** (1992) 404
- F.H. Fröhner, in *Maximum Entropy and Bayesian Methods*, W.T. Grandy Jr. and L.H. Schick (eds.), Kluwer, Dordrecht (1991a) p. 93
- F.H. Fröhner, Karlsruhe report KfK 5073 (1992)
- F.H. Fröhner, *Proc. Int. Conf. on Nucl. Data for Sci. and Technol.*, J.K. Dickens (ed.), ANS, La Grange Park (1994) p. 597
- F.H. Fröhner, *Nucl. Sci. Eng.* **126** (1997) 1
- F.H. Fröhner, *Z. Naturforsch.* **53 a** (1998) 637
- F.H. Fröhner, *Nucl. Reaction Data and Nuclear Reactors*, A. Gandini and G. Reffo (eds.), World Scient., Singapore etc. (1998a), vol. 1, p. 54
- M. Gaudin, *Nucl. Phys.* **25** (1961) 447; reprinted in Porter (1965)
- A. Gilbert and A.G.W. Cameron, *Can. J. Phys.* **43** (1965) 1446
- I.J. Good, *The Estimation of Probabilities*, Cambridge, Mass. (1965)
- S.M. Grimes, *Conf. on Moment Methods in Many Fermion Systems*, B.J. Dalton et al. (eds.), Plenum Press, New York (1980), p. 17
- H. Gruppelaar and G. Reffo, *Nucl. Sci. Eng.* **62** (1977) 756
- J.A. Harvey, private communication, Oak Ridge National Laboratory (1995)
- H.M. Hofmann, J. Richert, J.W. Tepel, H.A. Weidenmüller, *Ann. Phys.* **90** (1975) 403

- L.K. Hua, *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*, Am. Math. Soc., Providence, R.I. (1963)
- H.H. Hummel and D. Okrent, *Reactivity Coefficients in Large Fast Power Reactors*, Monogr. Ser. Nucl. Sci. Technol., Am. Nucl. Soc., Hindsdale, Ill. (1970)
- R.N. Hwang, *Nucl. Sci. Eng.* **21** (1965) 523; **52** (1973) 157
- N. Janeva, N. Koyumdjieva, A. Lukyanov, S. Toshkov, *Proc. Int. Conf. Nucl. Data for Basic and Appl. Sci., Santa Fe 1985*, New York etc. (1986) p. 1615; N. Koyumdjieva, N. Savova, N. Janeva, A.A. Lukyanov, *Bulg. J. Phys.* **16** (1989) 1
- E.T. Jaynes, *IEEE Trans. Syst. Cybern.* **SSC-4** (1968) 227; reprinted in Jaynes (1983) p. 114
- E.T. Jaynes, *Found. Phys.* **3** (1973) 477; reprinted in Jaynes (1983) p. 131
- E.T. Jaynes, *Statistical Inference and Statistical Theories of Science*, W.L. Harper and C.A. Hooker (eds.), Reidel, Dordrecht (1976); reprinted in Jaynes (1983) p. 151
- E.T. Jaynes, *The Maximum Entropy Formalism*, R.D. Levine and M. Tribus (eds.), M.I.T. Press, Cambridge, Mass. (1978); reprinted in Jaynes (1983) p. 210
- E.T. Jaynes, *Bayesian Analysis in Econometrics and Statistics*, A. Zellner (ed.), North-Holland, Amsterdam (1980); reprinted in Jaynes (1983) p. 337
- E.T. Jaynes, *Papers on Probability, Statistics and Statistical Physics*, R.D. Rosenkrantz (ed.), Reidel, Dordrecht (1983)
- H. Jeffreys, *Theory of Probability*, Clarendon Press, Oxford (1939)
- J.H.D. Jensen and J.M. Luttinger. *Phys. Rev.* **86** (1952) 907
- K. Kari, Karlsruhe report KfK 2673 (1978)
- P.L. Kapur and R.E. Peierls, *Proc. Roy. Soc. (London)* **A166** (1938) 277
- R.A. Knief, *Nuclear Energy Technology*, McGraw-Hill, New York etc. (1981)
- G.A. Korn and T.M. Korn, *Math. Handbook for Scientists and Engineers*, McGraw - Hill, New York etc. (1968)
- W.E. Lamb, *Phys. Rev.* **55** (1939) 750
- A.M. Lane and J.E. Lynn, *Proc. Phys. Soc.* **A70** (1957) 557
- A.M. Lane and R.G. Thomas, *Rev. Mod. Phys.* **30** (1958) 257
- K. Lange, *Numerical Analysis for Statisticians*, New York etc. (1999)
- P.S. Laplace, *Théorie Analytique des Probabilités*, Courcier, Paris (1812); reprinted in *Œuvres complètes*, Gauthiers-Villars, Paris (1878-1912)
- N.M. Larson, *Conf. on Nucl. Data for Basic and Appl. Sci.*, Gordon & Breach, New York (1986) p. 1593
- N.M. Larson and F.G. Perey, Oak Ridge report ORNL/TM-7485 (1980); N.M. Larson, ORNL/TM-9179 (1984) with revisions ORNL/TM-9179/R1 (1985), /R2 (1989), /R3 (1996), /R4 (1998)
- C.M. Lederer, V.S. Shirley (eds.), *Table of Isotopes*, 7th ed., Wiley, New York (1978)
- H.D. Lemmel, *Conf. on Nucl. Cross Sections and Technol.*, NBS SP 425, Washington (1975) vol. 1, p. 286
- L.B. Levitt, *Nucl. Sci. Eng.* **49** (1972) 450
- Li Jingwen et al., *Conf. on Nucl. Data for Sci. and Technol.*, K.H. Böckhoff (ed.), Reidel, Dordrecht (1982) p. 55
- J.E. Lynn, *The Theory of Neutron Resonance Reactions*, Clarendon Press, Oxford (1968)
- J.E. Lynn, Harwell report AERE-R7468 (1974)
- R.E. MacFarlane, D.W. Muir, R.M. Boicourt, Los Alamos report LA-9303-MS, vols. I and II (1982); R.E. MacFarlane and D.W. Muir, vol. III (1987); D.W. Muir and R.E. MacFarlane, vol. IV (1985)

- M. Mahdawi and G.F. Knoll, *Conf. on Nucl. Data for Sci. and Technol.*, K.H. Böckhoff (ed.), Reidel, Dordrecht (1982) p. 58
- M.L. Mehta, *Nucl. Phys.* **18** (1960) 395; reprinted in Porter (1965)
- M.L. Mehta, *Random Matrices*, 2nd ed., Acad. Press, Boston (1991) P.A. Mello, *Phys. Lett.* **B82** (1979) 103; cf. also P.A. Mello and T.H. Seligman, *Nucl. Phys.* **A344** (1980) 489
- P.A. Mello, P. Pereyra, T.H. Seligman, *Ann. Phys.* **161** (1985) 254
- P.A. Moldauer, *Phys. Rev.* **C11** (1975) 426; **C12** (1975) 744
- P.A. Moldauer, *Nucl. Phys.* **A344** (1980) 185
- S.F. Mughabghab, M. Divadeenam, N.E. Holden, *Neutron Cross Sections, Neutron Resonance Parameters and Thermal Cross Sections*, Vol. A: Z = 1 - 60, Acad. Press, New York etc. (1981)
- S.F. Mughabghab, *Neutron Cross Sections, Neutron Resonance Parameters and Thermal Cross Sections*, Vol. B: Z = 61 - 100, Acad. Press, New York etc. (1984)
- A. Müller and H.L. Harney, *Phys. Rev.* **C35** (1987) 1231
- M.N. Nikolaev et al., *At. Energiya* **29** (1970) 11
- William of Ockham (Occam, 1285-1349) propounded the principle of parsimony “entia non multiplicanda praeter necessitatem” (do not introduce more entities than necessary)
- F.G. Perey, *Int. Conf. on Neutron Physics and Nucl. Data*, Harwell (1978) p. 104
- C.M. Perey, F.G. Perey, J.A. Harvey, N.W. Hill, N.M. Larson, ⁵⁶Fe Resonance Parameters for Neutron Energies up to 850 keV, Oak Ridge National Laboratory, ORNL/TM-11742 (1990)
- W.P. Poenitz, *Proc. Conf. Nucl. Data Eval. Mth. and Proc.*, B.A. Magurno and S. Pearlstein (eds.), Brookhaven report BNL-NCS-51363, (1981) vol. I, p. 249
- W.P. Poenitz, J.F. Whalen, A.B. Smith, *Nucl. Sci. Eng.* **78** (1981) 333
- C.E. Porter and R.G. Thomas, *Phys. Rev.* **104** (1956) 483; reprinted in Porter (1965)
- C.E. Porter (ed.) *Statistical Theory of Spectra: Fluctuations*, Acad. Press, New York - London (1965)
- C.E. Porter and N. Rosenzweig, *Suomal. Tiedeakad. Toimit. (Ann. Acad. Sci. Fenn.)* **AVI** 44 (1960); reprinted in Porter (1965)
- C.W. Reich and M.S. Moore, *Phys. Rev.* **111** (1958) 929
- A. Rényi, *Valoszinűsegszámítás* (= Probability Calculus), Budapest (1954); slightly generalised proof by J. Aczel, *Lectures on Functional Equations and Their Applications*, Academic Press, New York (1978) p. 104
- P. Ribon and J.M. Maillard, Proc. ANS Meeting on Adv. in Reactor Physics and Safety, Saratoga Springs 1986, NUREG/CP-0080 (1986) vol. 2, p. 280
- P.F. Rose and C.L. Dunford (eds.) , Report ENDF-102, NNDC Brookhaven (1990)
- G.R. Satchler, *Phys. Letters* **7** (1963) 55
- R.E. Schenter, J.L. Baker, R.B. Kidman, Batelle North West report BNWL-1002 (1969)
- E. Schrödinger, *Proc. R. Irish Acad.* **51 A** (1947) 51, **51 A** (1947) 141; reprinted in *Gesammelte Abhandlungen*, Wien (1984) vol. 1, p. 463 and p. 479
- C.E. Shannon, *Bell Syst. Tech. J.* **27** (1948) 379 and 623; cf. also C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, U. of Illinois Press, Urbana (1949)
- V.V. Sinitsa, *Yad. Konst.* **5(54)** (1983); Engl. transl. IAEA report INDC(CCP)-225/G (1985)
- A.B. Smith, D.L. Smith, R.J. Howerton, Argonne report ARNL/NDM-88 (1985)
- A.W. Solbrig, *Nucl. Sci. Eng.* **10** (1961) 167

- C. Stein, *Third Berkeley Symp. on Math. Statistics and Probab.*, U. of Calif. Press, vol. I, p. 197 (1956); see also W. James and E. Stein, *Fourth Berkeley Symp. on Math. Statistics and Probab.*, U. of Calif. Press, vol. I, p. 361 (1961)
- H. Takano, Y. Ishiguro, Y. Matsui, report JAERI-1267 (1980)
- R.G. Thomas, *Phys. Rev.* **97** (1955) 224
- A.M. Turing, *Proc. London Math. Soc.*, Ser. 2, **48** (1943) 180
- R. Vandenbosch and J.R. Huizenga, *Nuclear Fission*, Acad. Press, New York (1973)
- J.M. Verbaarschot, H.A. Weidenmüller, M.R. Zirnbauer, *Phys. Reports* **129** (1985) 367
- J.M. Verbaarschot, *Ann. Phys.* **168** (1986) 368
- E. Voigt, *Sitz.-Ber. Bayer. Akad. Wiss.* (1912) p. 603
- A. Wald, *Statistical Decision Functions*, Wiley, New York (1950)
- A.E. Waltar and A.B. Reynolds, *Fast Breeder Reactors*, Pergamon, New York etc. (1981)
- C.H. Westcott, K. Ekberg, G.C. Hanna, N.S. Pattenden and S. Sanatani, *At. En. Rev.* **3** (1965) Issue 2, p. 3
- E.P. Wigner and L. Eisenbud, *Phys. Rev.* **72** (1947) 29; E.P. Wigner, *J. Am. Phys. Soc.* **17** (1947) 99
- E.P. Wigner, *Can. Math. Congr. Proc.*, Toronto (1957) p. 174; *Ann. Math.* **67** (1958) 325; both reprinted in Porter (1965)
- E.P. Wigner, *Conf. on Neutron Phys. by Time-of-Flight*, Gatlinburg 1956, Oak Ridge report ORNL-2309 (1957) p. 59; reprinted in Porter (1965)