KfK 4655
Januar 1990

Applications of the Maximum Entropy Principle in Nuclear Physics

F. H. Fröhner
Institut für Neutronenphysik und Reaktortechnik
Projekt Schneller Brüter

Kernforschungszentrum Karlsruhe

KERNFORSCHUNGSZENTRUM KARLSRUHE
Institut für Neutronenphysik und Reaktortechnik
Projekt Schneller Brüter

KfK 4655

APPLICATIONS OF THE MAXIMUM ENTROPY PRINCIPLE IN NUCLEAR PHYSICS

F.H. Fröhner

Invited paper presented at IX. Summer School on Nuclear Physics, Neutron Physics and Nuclear Energy, Varna, Bulgaria, 28 Sept. - 7. Oct. 1989

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Kernforschungszentrum Karlsruhe GmbH
Postfach 3640, 7500 Karlsruhe 1

ISSN 0303-4003

ANWENDUNGEN DES PRINZIPS DER MAXIMALEN ENTROPY IN DER KERNPHYSIK

ZUSAMMENFASSUNG

Bald nach Erscheinen der Informationstheorie wurde erkannt, daß das Prinzip der maximalen Entropie die fehlende Begründung für die bekannten Regeln der klassischen Thermodynamik liefert. Seither wurde es mit Erfolg auch in der Kernphysik angewandt. Als einfaches Beispiel leiten wir eine physikalisch sinnvolle Beschreibung des Spektrums der bei Kernspaltung emittierten Neutronen her und vergleichen das bekannte Ergebnis mit genauen Meßwerten für ^{252}Cf. Ein zweites Beispiel, Ableitung eines Ausdrucks für resonanzgemittelte Wirkungsquerschnitte von Kernreaktionen wie Streuung oder Spaltung, ist weniger einfach. Entropie-Maximierung, mit gegebenen Transmissionskoeffizienten als Nebenbedingung, liefert Wahrscheinlichkeitsverteilungen für die Elemente der R- und S-Matrix, mit deren Hilfe gemittelte Querschnitte berechnet werden können. Schränkt man nur durch eine vorgegebene Breite des Spektrums der Zwischenkernzustände ein, so erhält man das Gaußsche Orthogonal-Ensemble (GOE) von Hamilton-Matrizen, welches ebenfalls Ausdrücke für die mittleren Querschnitte liefert. Auf beiden Wegen findet man praktisch die gleichen Zahlenwerte trotz ganz verschiedener Ausdrücke für die Wirkungsquerschnitte. Diese Ergebnisse wurden bei einer neuen theoriegestützten Auswertung der Wirkungsquerschnitte von ^{238}U im nichtaufgelösten Resonanzbereich eingesetzt.

ABSTRACT

Soon after the advent of information theory the principle of maximum entropy was recognized as furnishing the missing rationale for the familiar rules of classical thermodynamics. More recently it has also been applied successfully in nuclear physics. As an elementary example we derive a physically meaningful macroscopic description of the spectrum of neutrons emitted in nuclear fission, and compare the well known result with accurate data on ^{252}Cf. A second example, derivation of an expression for resonance-averaged cross sections for nuclear reactions like scattering or fission, is less trivial. Entropy maximization, constrained by given transmission coefficients, yields probability distributions for the R- and S-matrix elements, from which average cross sections can be calculated. If constrained only by the range of the spectrum of compound-nuclear levels it produces the Gaussian Orthogonal Ensemble (GOE) of Hamiltonian matrices that again yields expressions for average cross sections. Both avenues give practically the same numbers in spite of the quite different cross section formulae. These results were employed in a new model-aided evaluation of the ^{238}U neutron cross sections in the unresolved resonance region.

1. INTRODUCTION: THE MAXIMUM ENTROPY PRINCIPLE

The main task of modern theoretical physics is prediction based on incomplete data, in other words inductive inference. For instance, one may know the laws of motion for the microscopic particles of a macroscopic thermodynamic system, but in practice it is impossible to know all the spatial and momentum coordinates needed for a deterministic prediction of its behaviour. Predictions must therefore be based on measurable and controllable macroscopic data such as temperatures, mean densities and other averages. In quantum mechanics there are not only practical reasons that prevent us from knowing all the initial phase space coordinates. The very fragility of elementary particles limits our knowledge even in principle: Heisenberg's uncertainty relations tell us that we cannot possibly know more than half the phase space coordinates exactly. We must therefore give up any deterministic description by trajectories and orbits and describe the state of a system, or rather our information about it, by probability distributions, both in thermodynamics and in quantum mechanics.

The required probabilities, although used to predict frequencies, are not frequencies themselves. They must express our information, or lack thereof, and change whenever new information becomes available, according to Bayes' theorem,

$$P(A|BC) \propto P(B|AC)P(A|C) . \tag{1}$$

Here A, B, C are different pieces of information, P(A|C) is an initial ("a-priori") probability of A being true given the condition C, P(A|BC) is the updated ("a-posteriori") probability of A given both the original condition C and new data B, while P(B|AC), the likelihood function, is the probability that the data B would have been observed if A and C were true. Note that all these probabilities are conditional, depending on either empirical or theoretical information or on assumptions. Bayes' theorem is actually a model of learning by experience: Prior knowledge of A is modified by new data B on A, under any circumstances C. (In quantum mechanics this is known as "reduction of the wave packet", about which considerable confusion exists because the distinction between the state of a system and our knowledge of it is often ignored.)

What is the most objective probability distribution to be used for predictions if a set of macroscopic data is given? Information theory gives the following answer (see e. g. Jaynes 1983). The probability distribution must be (a) consistent with the macroscopic data, and (b) it should not contain any other, spurious, misleading information. Now the unique, unambiguous measure of the missing information or indeterminacy in a probability distribution was shown by Shannon (1948) to be the information entropy, defined as

$$S_I \equiv -\sum_{i=1}^{n} p_i \ln p_i \tag{2}$$

if there are n distinct alternatives and the probabilities for their realization are $p_1, p_2, \cdots p_n$. It is nonnegative, attaining its maximum if all alternatives are equally probable (maximal indeterminacy), and vanishing if one of them is realized with certainty (no indeterminacy). For a continuous probability distribution p(x) with a priori equivalent increments dx one has

$$S_I = -\int dx\, p(x) \ln p(x) . \tag{3}$$

Let us now assume that information about the unknown distribution p(x) is given in the form of expectation values for known real functions (observables) $f_k(x)$,

$$\bar{f}_k = \int dx\, p(x)\, f_k(x)\,, \qquad k = 1, 2, \ldots K\,. \tag{4}$$

What is the probability density p(x) that satisfies these equations but does not contain any other information? Since the last requirement in fact means maximal indeterminacy of p(x) apart from the conditions (4), we must solve the variational problem $S_I = \max$ with the constraints (4) (and the additional constraints that p(x) is nonnegative and normalized to unity). The well-known solution, obtained by the method of Lagrange multipliers, is

$$p(x) = \frac{1}{Z} \exp\left(- \sum_k \lambda_k f_k(x) \right). \tag{5}$$

This is obviously positive for real Lagrange multipliers λ_k, and properly normalized with

$$Z = \int dx \exp\left(- \sum_k \lambda_k f_k(x) \right). \tag{6}$$

The K Lagrange multipliers must be found either from the K constraints (4), with p(x) given by (5) and (6), or from the equivalent equations

$$\bar{f}_k = - \frac{\partial}{\partial \lambda_k} \ln Z\,, \qquad k = 1, 2, \ldots K\,. \tag{7}$$

The maximum-entropy algorithm (5)-(7) ought to look familiar to physicists: It constitutes nothing less than Gibbs' axiomatic approach to thermodynamics. One recognizes the normalization constant Z as the partition function from which all microscopically observable ensemble averages can be obtained by suitable differentiation. Let x, for instance, denote the possible energies of the molecules in a volume of gas. Then the so-called canonical ensemble is obtained if only the average energy is given, the Lagrange multiplier being the inverse temperature. The grand-canonical ensemble results, with the chemical potential as a second Lagrange multiplier, if both the average energy and the average particle number are specified, etc. (Note that ensembles are

nothing but a frequentist visualization of probability distributions, hence ensemble averages are the same as expectation values.) The only microscopic information used is the range of the variate x: energies can vary from a suitably chosen zero to infinity. For given mean and variance (or standard deviation) the maximum-entropy distribution is Gaussian if the range of x is $-\infty$ to ∞, and lognormal if it is 0 to ∞, - which explains the ubiquity of these distributions in statistics generally and in data evaluation in particular, without any need to invoke random-noise, central-limit or ergodicity arguments. It was Jaynes (1957) who first stressed that entropy maximization is a powerful logical tool, applicable not only in thermodynamics but quite generally to all problems of inductive reasoning.

2. MACROSCOPIC DESCRIPTION OF FISSION NEUTRON SPECTRA

As a first example from nuclear physics let us consider the emission of neutrons during nuclear fission, for instance spontaneous fission of ^{252}Cf or neutron-induced fission of ^{235}U. The observed fission neutron spectrum is the statistical outcome of an enormous multitude of possible microscopic fission processes, all starting from the same excited (compound) nucleus, all leading to neutron emission but differing with regard to primary fission fragment pairs, their masses and charges, spins and excitations. In such a complex situation one can invoke, just as in thermodynamics, the maximum entropy principle, in order to find a macroscopic description in terms of only few average parameters. What is a suitable set of parameters for our problem?

Empirically it is well established that practically all fission neutrons are emitted from fragments fully accelerated by Coulomb repulsion, and that neutron emission is practically isotropic in each fragment's rest system (Budtz-Jørgensen and Knitter, 1988). We can therefore take each momentum coordinate, for instance p_x', as symmetrically distributed around zero, so that the lowest nonvanishing moment of the p_x'-distribution is $\langle p_x'^2 \rangle$ (The prime denotes centre-of-mass quantities.) Let us see how far we get if we neglect all higher moments. From the maximum-entropy algorithm (5)-(7) we find immediately that the most objective probability distribution for given $\langle p_x'^2 \rangle$ is a Gaussian with mean zero and variance $\langle p_x'^2 \rangle$. With the same reasoning for

the y and z directions, and with the notation $\langle p_x'^2 \rangle = \langle p'^2 \rangle/3 = p_T^2/2$ we find Boltzmann's momentum distribution, the three-dimensional Gaussian

$$\chi(\vec{p}')d^3p' = \frac{1}{\pi^{3/2}} e^{-p'^2/p_T^2} \frac{d^3p'}{p_T^3}, \qquad -\infty < p_x', p_y', p_z' < \infty . \tag{8}$$

Upon solid angle integration (with $d^3p' = p'^2dp'd\Omega$), and after rewriting everything in terms of energies (with $E' \equiv p'^2/(2m)$ and $T \equiv p_T^2/(2m)$, where m denotes the neutron mass), one gets the Maxwell energy spectrum

$$N_M(E'|T)dE' = \frac{2}{\sqrt{\pi}} e^{-E'/T} \sqrt{\frac{E'}{T}} \frac{dE'}{T}, \quad 0 < E' < \infty . \tag{9}$$

The single parameter characterizing this spectrum, the temperature T, is related to the mean energy by $\langle E' \rangle = 3T/2$. This emission spectrum, the simplest choice imaginable, works already quite well as comparison with the more physical evaporation formula of Weisskopf (1952),

$$N(E'|T)dE' \propto e^{-E'/T} E' \sigma_C(E')dE', \qquad 0 < E' < \infty . \tag{10}$$

shows. The Maxwell spectrum corresponds to a $1/\sqrt{E'}$ shape of the compound formation cross section σ_C, and this is in fact the shape of the dominant s-wave (inverse) cross sections of the residual fragments up to several MeV.

Next we go from the fragment rest system to the lab system by the Galileo transformation $\vec{p}' = \vec{p} - m\vec{u} \equiv \vec{p} - \vec{q}$, where $\vec{u}$ is the fragment velocity and $\vec{q}$ the fragment momentum per nucleon. Inserting this in (1) and integrating over all (equiprobable) directions of $\vec{u}$, for fixed $\vec{p}$, one obtains

$$\chi(p)dp = \frac{1}{\sqrt{\pi}} \left[e^{-(p-q)^2/p_T^2} - e^{-(p+q)^2/p_T^2} \right] \frac{p}{q} \frac{dp}{p_T},$$

$$0 < p < \infty . \tag{11}$$

In terms of energies this is the distribution proposed by Watt (1952),

$$N_W(E|T)dE$$

$$= \frac{1}{2\sqrt{\pi}}\left[e^{-(\sqrt{E}-\sqrt{E_W})^2/T} - e^{-(\sqrt{E}+\sqrt{E_W})^2/T}\right]\frac{dE}{\sqrt{E_W T}}$$

$$= e^{-E_W/T}\,\frac{\sinh\sqrt{4EE_W/T^2}}{\sqrt{4EE_W/T^2}}\,N_M(E|T)dE\,, \qquad 0<E<\infty\,. \tag{12}$$

In addition to the temperature T there is now a second parameter, the fragment kinetic energy per nucleon, $E_W \equiv mu^2/2$, and the mean energy is $\langle E\rangle = 3T/2 + E_W$. So the Watt spectrum is just a Maxwellian emission spectrum transformed to the lab system, approaching it for $E_W \to 0$ and being proportional to it but lower for small E_W. We shall write T_M and T_W to distinguish the temperatures obtained by fitting Maxwell or Watt spectra to the same data.

In order to find the spectrum for the entire ensemble of fission modes one ought to average the Watt spectrum over all possible values of T_W and E_W for the various fission fragments. As a first step one can consider a representative pair of fragments, using the empirical facts that the nuclear temperatures and the numbers of emitted neutrons are similar (on average) for both fragments. The corresponding superposition of two Watt spectra is found to differ only slightly from a single Watt spectrum. It may therefore be expected that already a single Watt distribution, with T_W and E_W interpreted as effective parameters, describes the true spectrum reasonably well.

Figs. 1 and 2 show that this expectation is fulfilled. A Watt distribution was fitted to ten recent data sets for ^{252}Cf, covering the range from 25 keV to 20 MeV. (Neutron emission by the spontaneously fissioning nuclide ^{252}Cf is particularly well studied because of its importance as a standard in neutron metrology and in fission and fusion reactor dosimetry.) All ten data sets are of good quality, and their consistency is impressive. The (small) corrections for instrumental resolution were easy since the Watt distribution is formally identical with the free-gas kernel familiar from Doppler broadening of compound- nuclear resonances. Convolution of two such kernels for temperatures T_1 and T_2 yields a kernel for $T_1 + T_2$ (Fröhner 1967). Thus a Watt distribution

can be resolution-broadened by simply raising the temperature slightly, by an amount corresponding to the instrumental resolution. The fit gave

$$T_W = 1.175 \pm 0.005 \text{ MeV}, \quad E_W = 0.359 \pm 0.009 \text{ MeV}, \quad \rho(T_W,E_W) = -0.984,$$

which corresponds to a mean energy of $<E> = 2.122 \pm 0.017$ MeV, the correlation coefficient ρ indicating almost complete anticorrelation between the two parameters. In view of the extreme simplicity of our model the fit is surprisingly good, over five decades of intensity. Only one of the ten data sets deviates significantly above 17 MeV, but its uncertainties appear rather low in view of the experimental difficulties caused by the extremely weak intensities at these energies. Actually, the overall chi-square achieved, 489 for 470 degrees of freedom, does not indicate any need for a better model, nor is the fit improved by refinements such as superposition of two Watt distributions (for a representative fragment pair) or relativistic corrections (Fröhner 1987).

How does the fitted Watt spectrum compare with available microscopic spectrum calculations that are based on detailed mass and charge distributions of fission fragments and on evaporation theory involving the statistical model of compound-nuclear reactions and the theory of level densities (e. g. Madland and Nix 1982, Märten and Seeliger 1986)? Fig. 3 shows that in the technologically and metrologically most important region up to 15 MeV it fits at least equally well, and below the peak at about 0.7 MeV even better than the microscopic calculations. At the highest energies it seems to underpredict, but not as much as the Madland-Nix model, and the general situation there is obscured anyway by the large uncertainties and the scatter of the data. Thus the overall performance of the Watt distribution is not bad at all. Furthermore, the fact that it can be easily calculated from just two parameters, T_W and E_W, makes it much more attractive for routine applications than its more demanding microscopic competitors.

Should it turn out that the misfit data in the high-energy tail are correct one could easily encorporate this information in the maximum entropy treatment. Instead of the Maxwellian emission spectrum, equivalent to a $1/\sqrt{E'}$ behaviour of the compound nuclear formation cross section, one could use a more general power law to simulate the deviations at high energies expected for p-, d- ... wave contributions. This would

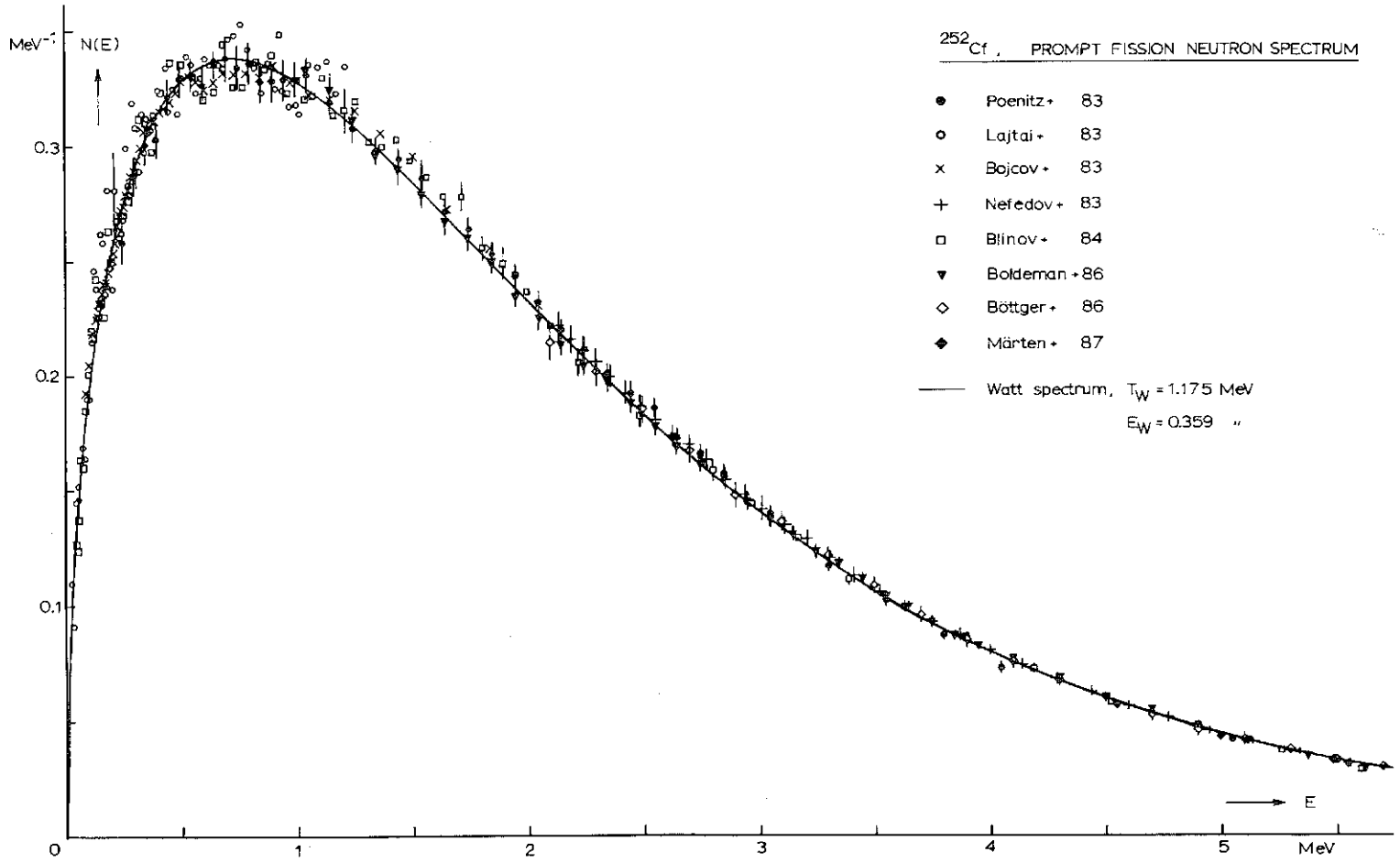

Fig. 1 - Utilized experimental data below 5.7 MeV (most error bars are omitted for clarity) and adjusted Watt distribution, showing the quality of the fit in the practically most important part of the fission neutron spectrum. For references see Fröhner (1987).

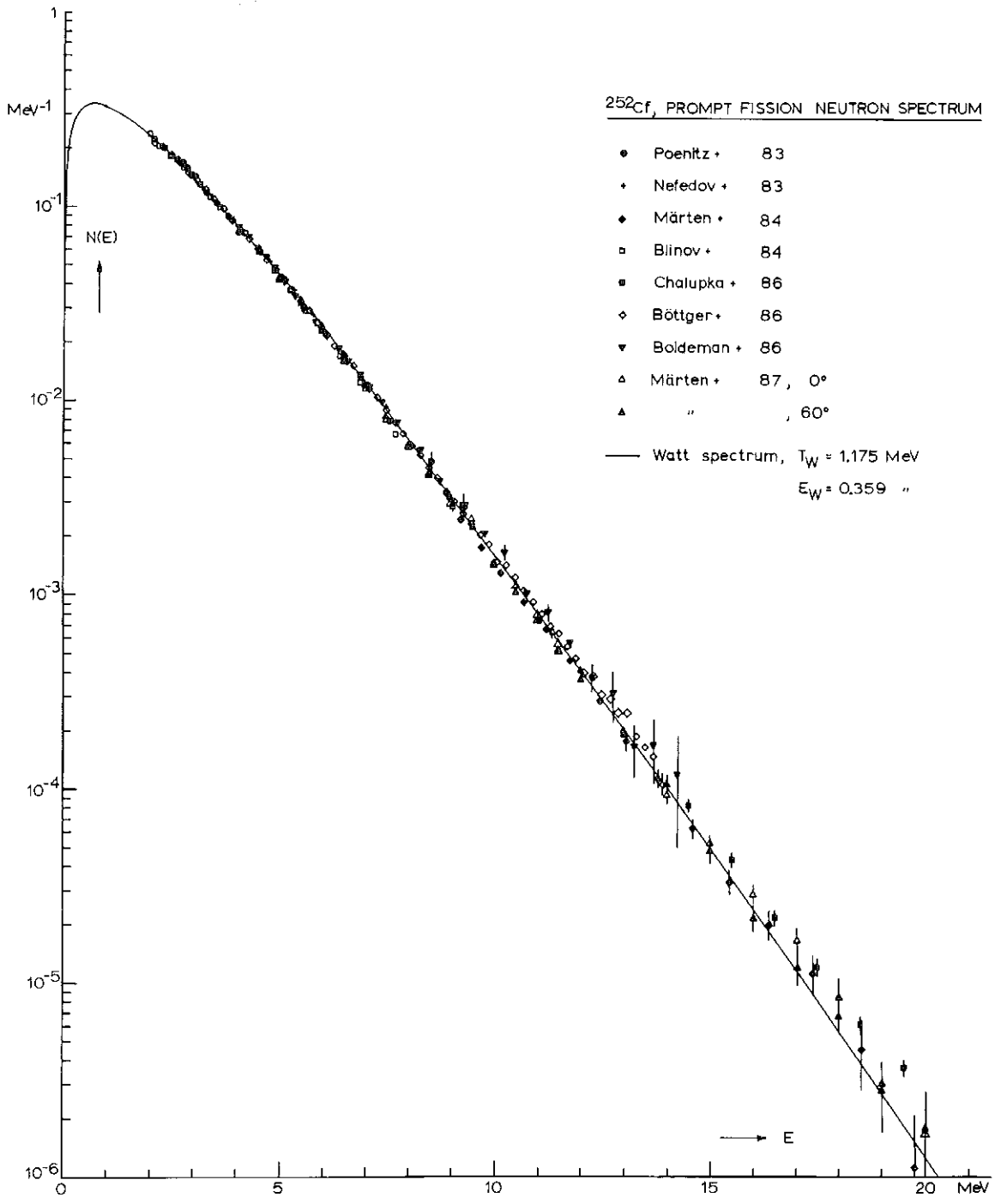

Fig. 2 - Utilized experimental data above 3 MeV (most error bars are omitted for clarity) and adjusted Watt distribution, showing the quality of the fit in the tail of the fission neutron spectrum. For references see Fröhner (1987).

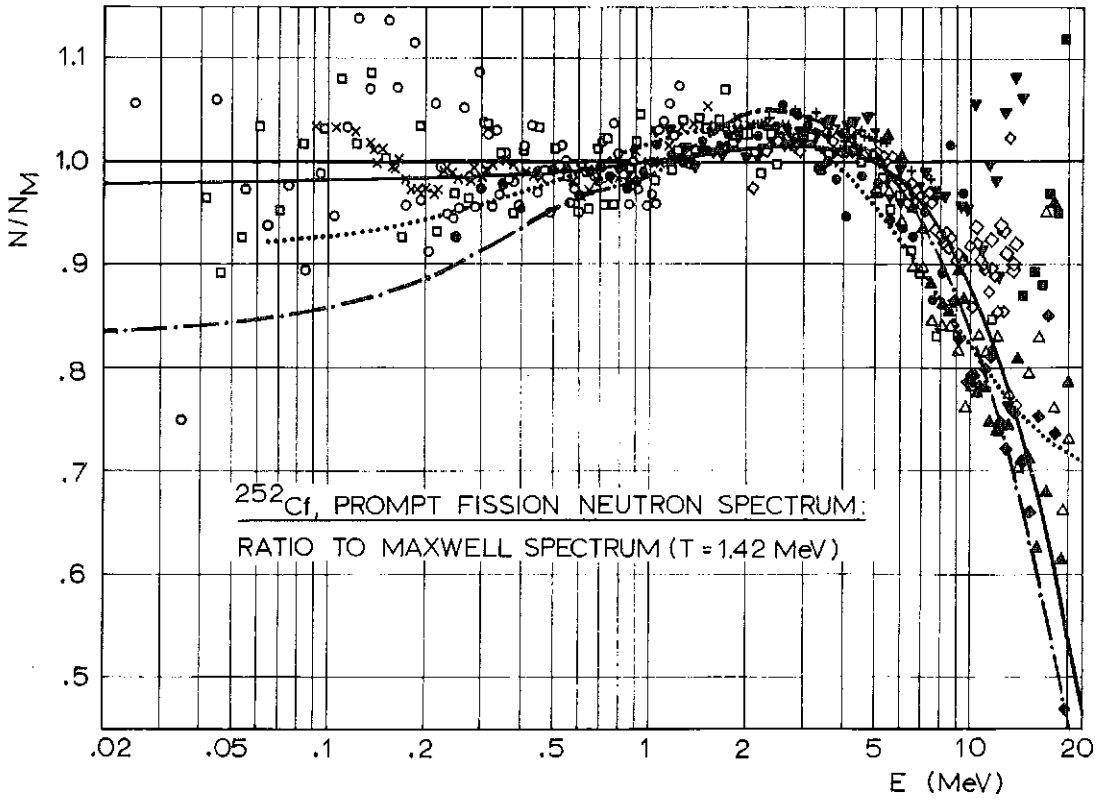

Fig. 3 - All utilized experimental data (error bars omitted for clarity) together with adjusted Watt distribution (solid line) and microscopic-model curves (dash-dotted: Madland-Nix model, calculation by Walsh (1989), dotted: generalized Madland-Nix model, calculation by Märten and Seeliger (1986)), plotted relative to a Maxwellian distribution with T_M = 1.42 MeV. (This has become a standard way of representing the ^{252}Cf fission neutron spectrum.) Point symbols have the same meaning as in Figs. 1 and 2.

result in a more general gamma distribution instead of the Maxwellian, with an exponent λ instead of 1/2 in Eq. 9 as a third fitting parameter. With the present chi-square, however, there seems to be no need for such a refinement.

3. MAXIMUM-ENTROPY SOLUTIONS TO THE HAUSER-FESHBACH PROBLEM

If nuclei are bombarded by neutrons or protons the observed cross sections have pronounced resonance structure at low energies that becomes more and more washed out as energy increases, mainly because of decreasing instrumental resolution, but also because of increasing resonance widths and decreasing resonance spacings. Thus there is usually a so-called resolved resonance range at low energies, an unresolved range at intermediate energies, and a high-energy range where cross sections are rather smooth because of complete level overlap. The resonances in the unresolved range, although invisible, give rise to quite noticeable effects like resonance self-shielding and temperature-dependent absorption. These must be taken into account in data analysis and especially in reactor technology: The only inherent safety feature in fast reactors that counteracts a sudden reactivity rise is enhanced neutron absorption that occurs if self-shielding of unresolved resonances (mainly those of ^{238}U) decreases as a consequence of increasing Doppler broadening. Calculations of resonance-averaged cross sections, self-shielding factors and other cross section functionals of interest in the unresolved resonance range must be based on the statistical theory of compound-nuclear reactions (Hauser-Feshbach theory) that has evolved on two levels.

3.1 S- and R-Matrix Expressions for Partial Cross Sections

The "microscopic" level, involving compound resonances and their statistics, was explored first. One considers energy averages over intervals containing so many resonances that a level-statistical treatment is possible. R-matrix theory provides the simplest formal description. As a reminder and in order to establish the notation we begin with some basic cross section formulae of nuclear reaction theory (see Lane and Thomas 1958; we essentially adopt their notation that has become standard in applied nu-

clear physics). The partial cross section σ_{ab} for a reaction leading from an entrance channel a to an exit channel b is determined by the collision matrix element (transition amplitude) S_{ab},

$$\sigma_{ab} = \pi \lambda\!\!\!^{-2}_{a} g_a |\delta_{ab} - S_{ab}|^2 , \quad (13)$$

where $2\pi\lambda\!\!\!^{-}_{a}$ is the de-Broglie wave length of relative motion in the entrance channel and g_a the spin factor. Each channel label stands for two collision partners and the orbital, spin and total angular momenta. The collision matrix S connects channels with the same parity and total angular momentum. Its general features are

- unitarity (from conservation of overall probability),
- symmetry (from invariance under time reversal),
- absence of poles above the real axis in the complex energy plane (from causality).

Summing over all exit channels and using the unitarity one gets the total cross section as a linear function of S,

$$\sigma_a \equiv \sum_b \sigma_{ab} = 2\pi \lambda\!\!\!^{-2}_{a} g_a (1 - \mathrm{Re}\, S_{aa}) . \quad (14)$$

The unitary matrix S can be expressed in terms of the real symmetric matrix R as follows,

$$S_{ab} = e^{-i(\phi_a + \phi_b)} [(1 - iR)^{-1}(1 + iR)]_{ab} \quad (15)$$

The elements of the R matrix,

$$R_{ab} = \sum_\lambda \frac{\gamma_{\lambda a}\gamma_{\lambda b}}{E_\lambda - E} , \quad (16)$$

are sums over all resonances with given spin and parity. All the resonance parameters are neatly wrapped up in the R-matrix: For each resonance there is a real level energy E_λ and a set of real decay amplitudes $\gamma_{\lambda c}$ for all energetically accessible exit channels. The exponential factor, containing hard-sphere phase shifts ϕ_c , is known once the

(arbitrary) channel radii are fixed. We simplified by choosing the R-matrix boundary parameters so that $L_c^0 = iP_c$ (in standard notation), and by absorbing the centrifugal barrier penetrabilities P_c in the $\gamma_{\lambda c}$. This is appropriate for the unresolved resonance region where one typically considers an averaging interval centred at E which is so narrow that the energy dependence of L_c^0 and P_c as well as that of ϕ_c and $\lambda\!\!\!^{-2}_c$ can be neglected. The collision matrix can also be written in terms of the level matrix A as

$$S_{ab} = e^{-i(\phi_a + \phi_b)} [\delta_{ab} + 2i \sum_{\lambda, \mu} \gamma_{\lambda a} A_{\lambda\mu} \gamma_{\mu b}] , \tag{17}$$

where

$$(A^{-1})_{\lambda\mu} = (E_\lambda - E)\delta_{\lambda\mu} - i \sum_c \gamma_{\lambda c} \gamma_{\mu c} . \tag{18}$$

The E_λ and $\gamma_{\lambda c}$ may be identified with the eigenvalues and eigenvector coordinates of the nuclear Hamiltonian which, to the extent that nuclear interactions are invariant under time reversal, can be represented as a real and symmetric $n \times n$ matrix, $H = H^+ = H^*$. As suggested by Wigner (1957) a simple statistical model of resonance reactions is obtained if H is considered as a member of the Gaussian Orthogonal Ensemble (GOE). It is derivable from the requirements that the $H_{\lambda\mu} = H_{\mu\lambda}$ be uncorrelated random variables, and that the form of their joint distribution must be invariant under rotations in the space of the eigenvectors, all orthogonal bases being equivalent. The eigenvector components $\gamma_{\lambda c}$ have (in the limit of very many resonances, $n \to \infty$) the normal distribution around zero (Porter and Rosenzweig 1960) hypothesized by Porter and Thomas (1956). At low energies, in the absence of direct reactions, there are no correlations between decay amplitudes for different channels or different levels, or between decay amplitudes and level energies, so that the average matrix $\overline{S}$ is diagonal. The eigenvalues E_λ , on the other hand, are highly correlated, forming a remarkably regular ("stiff") sequence exhibiting eigenvalue repulsion (Dyson and Mehta 1963). The case with direct reactions can be formally reduced to the pure compound case by means of a transformation found by Engelbrecht and Weidenmüller (1973).

Resonance-averaged cross sections can in principle be obtained by averaging the cross section expressions (13) and (14), with S and R given by (15) and (16), either over

energies, for given resonance parameters, or over resonance parameters, for given energy. The energy average over the total cross section is easily carried out with a Lorentzian weight function centred at E, with a half width of I. Contour integration yields immediately $\overline{S} = S(E+iI)$ since there are no poles above the real energy axis except for the pole at

$$\mathcal{E} \equiv E + iI \tag{19}$$

that is contributed by the Lorentzian. Thus the average total cross section is

$$\overline{\sigma}_a = 2\pi\lambda\!\!\!^{-2}_a g_a [1 - \mathrm{Re}\, S(\mathcal{E})_{aa}] , \tag{20}$$

(Thomas 1955). To get $S(\mathcal{E})$ we need simply replace R(E) by $R(\mathcal{E})$, neglecting weak energy dependences again. For partial cross sections, however, absolute squares of S-matrix elements must be averaged. These absolute squares have poles both below and above the real axis which precludes contour integration. The other possibility, averaging over the GOE, is easy only for the limiting case of widely spaced, narrow resonances ("isolated" levels) for which level overlap und thus eigenvalue correlations can be neglected. Assuming χ^2 distributions with ν_c degrees of freedom for the partial widths $\Gamma_{\lambda c} \equiv 2\gamma^2_{\lambda c}$ Dresner (1957) found

$$\overline{\sigma}_{ab} \simeq \sigma_{pa}\delta_{ab}$$

$$+ \pi\lambda\!\!\!^{-2}_a g_a \frac{T_a T_b}{T}\left(1 + \frac{2}{\nu_a}\delta_{ab}\right)\int_0^\infty dx \prod_c \left(1 + \frac{2T_c}{\nu_c T}x\right)^{-\delta_{ac} - \delta_{bc} - \nu_c/2} , \tag{21}$$

where $\sigma_{pa} = 4\pi\lambda\!\!\!^{-2}_a g_a \sin^2\phi_a$ is the potential-scattering cross section, $T_c \equiv 1 - |\overline{S}_{cc}|^2$ the transmission coefficient for channel c and $T \equiv \Sigma T_c$. The approximation $T_c \simeq 2\pi\overline{\Gamma}_c/D$ (where D is the mean level spacing), valid for vanishing level overlap, was used to write the result in terms of the transmission coefficients. This is the Hauser-Feshbach formula with elastic enhancement (first pair of parentheses) and width fluctuation correction (integral). The transmission coefficients on which it depends are macroscopic quantities, obtainable without invoking resonance behaviour. Those for particle channels can be obtained from the optical model, those for photon channels from the giant dipole resonance model, and those for fission channels from the potential barriers of the channel theory of fission. For single channels one has $\nu_c = 1$, but

in practical applications one uses often lumped channels with an effective $\bar{\nu}_c$ differing from unity, for example to represent all fission or photon or particle channels with the given total angular momentum and parity. The number of photon channels is usually large (except for light and magic compound nuclei) so that one may put

$$\prod_{c\in\gamma}\left(1+\frac{2T_c}{\nu_c T}x\right)^{-\nu_c/2} \simeq \lim_{\bar{\nu}_\gamma\to\infty}\left(1+\frac{2T_\gamma}{\bar{\nu}_\gamma T}x\right)^{-\bar{\nu}_\gamma/2} = e^{-xT_\gamma/T}, \tag{22}$$

with $T_\gamma \equiv \sum_{c\in\gamma} T_c$, whence for $a,b \notin \gamma$

$$\prod_{c}\left(1+\frac{2T_c}{\nu_c T}x\right)^{-\delta_{ac}-\delta_{bc}-\nu_c/2} \simeq e^{-xT_\gamma/T}\prod_{c\notin\gamma}\left(1+\frac{2T_c}{\nu_c T}x\right)^{-\delta_{ac}-\delta_{bc}-\nu_c/2} \tag{23}$$

Generalization to arbitrary level overlap (which invariably occurs at higher energies) turned out to be extremely difficult. In spite of much effort this so-called Hauser-Feshbach problem remained unsolved for almost three decades. Accurate average cross sections could only be obtained as Monte Carlo results, from which heuristic analytical expressions were extracted e. g. by Hofmann, Richert, Tepel and Weidenmüller (1975) and by Moldauer (1980). Other approximate analytic expressions were derived with picket fence models (e. g. Janeva et al. 1985) or disordered picket fence models (Müller and Harney 1987).

In this situation information theory seemed to offer the possibility to bypass the "microscopic" resonance details completely by treating them as a kind of noise superimposed on the "macroscopic" average cross sections obtained e. g. from optical-model calculations. This approach, first tried by Bloch (1968), was pursued vigorously by Mello (1979) who pointed out that, given an average S-matrix, for instance from an optical model calculation, one can obtain the corresponding distribution of S matrices by entropy maximization. Mello, Pereyra and Seligman (1985) found that the Poisson kernel defined in the domain of unitary symmetric matrices (Hua 1963) appears to be the required distribution, having all the properties that ergodicity and the analytic (causal) structure of S demand, while its form implies maximal information entropy for given transmission coefficients. Fröhner (1985) arrived at the same result more directly,

determining first the distribution of R-matrix elements by straightforward entropy maximization, then rewriting the result in terms of the S matrix.

While the S- and R-matrix distributions were established on the macroscopic level, Verbaarschot, Weidenmüller and Zirnbauer (1985) succeeded on the microscopic level. With new tools from the theory of disordered systems they finally managed to average the R-matrix expression for the partial cross sections over the Gaussian Orthogonal Ensemble (GOE) of Hamilton matrices (which is tantamount to averaging over the Dyson-Mehta distribution of level positions and over the Porter-Thomas distributions of all partial widths, see Porter (1965)).

The following sections contain more explicit details of these new techniques, some of them unpublished so far.

3.2 Invariant Volume Elements for Matrices

Application of the maximum entropy principle to matrix ensembles was discussed in particularly lucid fashion by Balian (1968). The main complication is the need to generalize the equiprobable increments dx of a univariate distribution p(x)dx suitably if the random variate is a matrix rather than a scalar. Since this problem is not treated in appropriate detail in the level-statistical literature, and since there has been already some confusion (about "Dyson's measure"), we give here a short account of the work of Hua (1963) as far as it is relevant to our topic.

We consider the independent real and imaginary parts of the elements $Z_{ab} = X_{ab} + iY_{ab}$ of an arbitrary complex matrix Z as Cartesian coordinates, $X_{11} = x_1$, $Y_{11} = x_2$, $X_{12} = x_3$ etc., in a space with as many dimensions as there are independent real parameters specifying the matrix Z. A metric in this space is introduced if we define the line element ds by

$$ds^2 = \sum_{a,\,b}\left[(dX_{ab})^2 + (dY_{ab})^2\right] = \mathrm{tr}(dZ^{+}dZ)$$

$$= \sum_{\mu,\nu} dx_\mu g_{\mu\nu} dx_\nu \,. \tag{24}$$

where $g_{\mu\nu}$ is the metric tensor. According to the rules of Riemannian geometry the corresponding volume element is

$$d[Z] = \sqrt{\det\, g} \prod_\mu dx_\mu \,. \tag{25}$$

It is obviously invariant under translations and rotations in the space of the x_μ just as dx above was invariant under translations along the real axis. For an arbitrary complex matrix Z the metric tensor is simply the unit tensor. A real symmetric $n \times n$ matrix R, however, has n^2 real elements, but because of $R_{ab} = R_{ba}$ only $n(n+1)/2$ of these are independent. Therefore,

$$ds^2 = \sum_a (dR_{aa})^2 + 2 \sum_{a<b} (dR_{ab})^2 = \sum_{\mu,\nu} dx_\mu g_{\mu\nu} dx_\nu \,, \tag{26}$$

and the $n(n+1)/2$-dimensional volume element in the parameter space is

$$d[R] = 2^{n(n-1)/4} \prod_{a \le b} dR_{ab} \,. \tag{27}$$

For unitary symmetric matrices S there are $2n^2$ real and imaginary parts of matrix elements, with n^2 unitarity and $n(n-1)/2$ symmetry relations existing between them. Hence the number of independent parameters is the same as for real symmetric matrices, $n(n+1)/2$. This is consistent with the possibility to express each S-matrix unambiguously in terms of an R-matrix once the hard-sphere phases are fixed. In fact, R is essentially what Hua (1963) calls the parameter of S. Differentiation of Eq. 15 yields

$$dS_{ab} = 2i\, e^{-i(\phi_a + \phi_b)} [(1 - iR)^{-1} dR (1 - iR)^{-1}]_{ab} \,, \tag{28}$$

whence

$$ds^2 = 4\, \mathrm{tr}[(1 + R^2)^{-1} dR (1 + R^2)^{-1} dR] \,, \tag{29}$$

the hard-sphere phases cancelling due to the cyclic invariance of the trace, tr(AB) = tr(BA). The determinant of the metric tensor is then

$$\det\, g = 2^{n(3n+1)/2}\, \det(1 + R^2)^{-(n+1)/2}\,, \tag{30}$$

as is readily seen in a coordinate system in which $1 + R^2$ is diagonal, and the invariant volume element is

$$d[S] = 2^{n(n+1)/2}\, \det(1 + R^2)^{-(n+1)/2}\, d[R] \tag{31}$$

with d[R] given by (27). This relationship between d[S] and d[R] is invariant under transformations of the form $S \to USU^+$ with a unitary matrix U.

Instead of Cartesian coordinates one can use "polar coordinates" (Hua 1963). These are obtained by explicit introduction of the eigenvalues of the matrix. Denoting the diagonal form of a matrix by the subscript D we have $R = OR_DO^+$ with $O^+O = 1$ and $\det O = +1$, i. e. the real symmetric matrix R is diagonalized by a real orthogonal matrix O representing a pure rotation. (Note that the Hermitean conjugate O^+ of the real matrix O is just the transpose of O.) The eigenvalues R_D can, without loss of generality, be considered as ordered, $R_1 < R_2 < \ldots < R_n$, because the matrices O include those which produce reordering. Differentiating OR_DO^+ one gets, with $dO^+O = -O^+dO$,

$$dR = O(dR_D + \delta O\, R_D - R_D \delta O)O^+\,, \tag{32}$$

where

$$\delta O \equiv O^+dO = -\delta O^+\,, \tag{33}$$

whence

$$ds^2 = \sum_c (dR_c)^2 + 2 \sum_{a<b} (R_a - R_b)^2 \delta O_{ab}^2\,. \tag{34}$$

Via the metric tensor for the polar-coordinate increments dR_c and δO_{ab} one finds

$$d[R] = 2^{n(n-1)/4} \prod_c dR_c \prod_{a<b} |R_a - R_b| \delta O_{ab} ,$$

$$-\infty < R_1 < R_2 < \ldots < R_n < \infty . \qquad (35)$$

The eigenvalues R_c of R imply eigenvalues $(1 + iR_c)/(1 - iR_c) \equiv \exp(i\theta_c)$ for the random part of S, $(1 - iR)^{-1}(1 + iR)$. Substituting with $R_c = \tan(\theta_c/2)$ in Eq. 31 one finds the polar coordinate form of d[S],

$$d[S] = 2^{n(n-1)/4} \prod_c d\theta_c \prod_{a<b} |e^{i\theta_a} - e^{i\theta_b}| \delta O_{ab} ,$$

$$-\pi < \theta_1 < \theta_2 < \ldots < \theta_n < \pi . \qquad (36)$$

The absolute values of eigenvalue differences in Eqs. 35 and 36 represent eigenvalue repulsion: the smaller the difference, the smaller is its a-priori probability. We emphasize that eigenvalue repulsion is a universal feature of matrix ensembles. It is encountered whenever "polar" coordinates are introduced. Replacing R_{ab} by $H_{\lambda\mu}$ and R_c by E_λ one recognizes that Eq. 35 in fact describes the familiar Wigner repulsion between compound-nuclear levels (see Porter 1965), i. e. between the eigenvalues of the nuclear Hamiltonian H.

3.3 Derivation of the GOE by Entropy Maximization

We are now ready to derive the Gaussian Orthogonal Ensemble by entropy maximization as suggested by Porter (1965). This is much simpler than the derivation outlined above. With the volume element $d[H] \propto \prod_{\mu \leq \nu} dH_{\mu\nu}$ we write the information entropy for the joint probability density p(H) as

$$S_I = -\int d[H]\, p(H) \ln p(H) . \qquad (37)$$

If we only assume that the eigenvalues $E_1, E_2, \ldots E_n$ of H are distributed around some central energy E_0 with some finite variance σ^2 we have to maximize the entropy subject to the constraint

$$\overline{\sum_{\mu=1}^{n} (E_\mu - E_0)^2} = \overline{\mathrm{tr}(H - E_0 1)^2} = N\sigma^2 , \tag{38}$$

where 1 denotes the unit matrix. Since

$$\mathrm{tr}(H - E_0 1)^2 = \sum_{\mu} (H_{\mu\mu} - E_0)^2 + 2 \sum_{\mu<\nu} H_{\mu\nu}^2 \tag{39}$$

the integrand of Z factorizes into independent Gaussians which are readily integrated:

$$Z = \left(\int_{-\infty}^{\infty} dx\, e^{-\lambda x^2} \right)^n \left(\int_{-\infty}^{\infty} dx\, e^{-2\lambda x^2} \right)^{n(n-1)/2} = \left(\frac{\pi}{\lambda} \right)^{n/2} \left(\frac{\pi}{2\lambda} \right)^{n(n-1)/4} , \tag{40}$$

whereupon differentiation of ln Z with the constraint (38) yields $\lambda = (n+1)/(4\sigma^2)$. With $\sigma^2 \equiv (n+1)a^2$ one gets the customary form of the GOE distribution,

$$p(H)d[H] = \frac{2^{n(n-1)/4}}{(4\pi a^2)^{n(n+1)/4}} \exp\left(- \frac{\mathrm{tr}(H - E_0 1)^2}{4a^2} \right) \prod_{\mu \le \nu} dH_{\mu\nu} . \tag{41}$$

One small step further yields the Wigner distribution of level spacings. We specialize to n = 2 and introduce polar coordinates by means of (33) and (35) to get

$$p(H)d[H] = \frac{1}{(4\pi a^2)^{3/2}} \exp\left(- \frac{(E_1 - E_0)^2 + (E_2 - E_0)^2}{4a^2} \right)$$

$$\cdot \sqrt{2}\, dE_1 dE_2 |E_2 - E_1|\, d\phi . \tag{42}$$

Next we substitute E_1 and E_2 by the level spacing $D = (E_2 - E_1)$ and the mean level energy $\bar{E} = (E_1 + E_2)/2$ and integrate over ϕ and $\bar{E}$. The result is the Wigner distribution

$$p(D)dD = 2x e^{-x^2} dx , \qquad 0 < x \equiv \frac{\sqrt{\pi}\, D}{2\bar{D}} < \infty , \tag{43}$$

where the average width is $\overline{D} = \sqrt{2\pi a^2}$ (and a factor of 1/2 is due to the condition $D > 0$, i. e. $E_2 > E_1$). This distribution, although strictly valid only for real symmetric matrices of rank 2, is an excellent approximation also for such matrices of higher rank (Gaudin 1961).

As we specified only mean and variance of the eigenvalue spectrum the Gaussian Orthogonal Ensemble of real symmetric random matrices is seen to be a rather direct generalization of the univariate Gaussian, the maximum entropy distribution for given mean and variance in the case of real random numbers. Of course, we cannot expect the GOE, constrained only by the second moment, to reproduce model-dependent secular features of nuclear level sequences such as Fermi-gas level densities or shell effects. In fact, the semicircular GOE level density law (see Porter 1965) looks rather unphysical. Moreover, the assumed mutual independence of the Hamiltonian matrix elements is at variance with the predominantly two-body nature of internucleon forces (see Brody et al. 1981). Nevertheless, the GOE gives an excellent description of local level statistics (level spacing and width distributions) obtained experimentally in the resolved resonance region or theoretically with more general matrix ensembles (Dyson and Mehta 1963) or in shell model calculations (Brody et al. 1981). Seeming deviations from GOE distributions usually vanish if the secular variation of average parameters is properly taken into account (Verbaarschot and Brussaard 1979).

3.4 Maximum-Entropy Distribution for the S- and R-Matrix

We begin with the single-channel case. For pure elastic scattering one has simply $S = \exp(i\theta)$ with real θ, and $d[S] = 2d\theta$, $-\pi < \theta < \pi$ (see Eq. 36) so that all phases are equiprobable a priori. Now we assume that we know $\overline{S}$, the average S function, from an optical model calculation. Since S has no poles above the real energy axis one knows with $\overline{S}$ also $\overline{S^2} = \overline{S}^2$, $\overline{S^3} = \overline{S}^3$ etc., where the overbars denote energy averages with Lorentzian weight functions, obtainable by contour integration. Since the averages of all powers of S are known, one knows also the averages of all analytic functions of S (matrix functions defined by their expansions in powers of S) and of their complex conjugates. Particularly convenient matrix functions for the calculation of Z, Eq. 6, are logarithms, for instance

$$f(S) = \ln(|S - \bar{S}|^2) = \ln(1 - \bar{S}^*S) + \text{c.c.} \tag{44}$$

The expansion of the last logarithm involves all powers of S, hence the average

$$\overline{f(S)} = 2 \ln(1 - \bar{S}^*\bar{S}) = 2 \ln T \tag{45}$$

utilizes all the input data, $T \equiv 1 - |\bar{S}|^2$ being the usual definition of the transmission coefficient. Maximizing the information entropy for the distribution $p(\theta)$ with the last equation as constraint one gets

$$p(\theta)d\theta = \frac{1}{Z} |1 - \bar{S}^* e^{i\theta}|^{-2\lambda} d\theta , \qquad -\pi < \theta \leq \pi , \tag{46}$$

$$Z = \int_{-\pi}^{\pi} |1 - \bar{S}^* e^{i\theta}|^{-2\lambda} d\theta . \tag{47}$$

The last integral is essentially a Legendre function (see e. g. Whittaker and Watson 1951). Using the properties of these functions one finds $\lambda = 1$, so that the maximum entropy distribution for given $\bar{S}$ is

$$p(\theta|\bar{S})d\theta = \frac{T}{|1 - \bar{S}^* e^{i\theta}|^2} \frac{d\theta}{2\pi} , \qquad -\pi < \theta \leq \pi . \tag{48}$$

(Fröhner 1985). This distribution, known as the Poisson kernel in potential theory, was first obtained by Lopez, Mello and Seligman (1981) without entropy maximisation. They recognized that knowledge of all powers of $\bar{S}$ and $\bar{S}^*$ means knowledge of all Fourier coefficients of the probability density $p(\theta)$, whence $p(\theta)$ itself is readily obtained. We note that microscopic information about the exact distribution of level energies does not appear in (48). Actually de los Reyes, Mello and Seligman (1980) verified via Monte Carlo sampling of resonance ladders that both the Poisson distribution (no level repulsion) and the Wigner distribution (the model case of level repulsion), lead to this θ-distribution, at least within the statistical accuracy of their Monte Carlo calculation.

Rewriting the Poisson kernel in terms of the R-function one gets a Lorentzian (Cauchy distribution),

$$p(R\,|\,R(\mathcal{E}))dR = \frac{1}{\pi}\,\frac{dx}{1+x^2}\,, \qquad -\infty < x \equiv \frac{R-R^{\infty}}{\pi s} \leq \infty\,, \tag{49}$$

(Fröhner 1985). The distant-level parameter R^{∞} and the pole strength s, defined by $R(\mathcal{E}) \equiv R^{\infty} + i\pi s$, are macroscopic quantities, uniquely determined once $\overline{S}$ and the channel radius are given.

If more than one channel is open the Fourier approach does not work. The reason is that the Fourier coefficients include averages over products involving both S and S* such as $S_{ab}{}^{*}S_{cd}$. Since their poles occur both above and below the real axis, energy averaging by contour integration is not possible. Entropy maximization, on the other hand, always works. In order to find the joint distribution of all S we maximize the information entropy subject to the constraint

$$\overline{\ln\,|\,\det(S-\overline{S})\,|^2} = \overline{\ln\,\det(1-\overline{S}^{+}S)} + \text{c.c.} = 2\,\ln\det\,T\,, \tag{50}$$

where $T = 1 - \overline{S}^{+}\overline{S}$ is the transmission matrix, the multi-channel generalization of the transmission coefficient, introduced by Satchler (1963). This means replacement of the function $1 - \overline{S}^{*}S$ in (44) by the scalar $\det(1 - \overline{S}^{+}S)$, where 1 is again the unit matrix. The determinant ensures utilization of all possible averaged products of the S_{ab} , i. e. of all given macroscopic information. With the constraint (50) one gets

$$\begin{aligned} p(S\,|\,\overline{S})d[S] &\propto |\,\det(S-\overline{S})\,|^{2\lambda}\,d[S] \\ &\propto |\,\det(R-R(\mathcal{E}))\,|^{2\lambda}\,\det(1+R^2)^{-\lambda-(n+1)/2}\,d[R]\,, \end{aligned} \tag{51}$$

where the relationship (15) between S and R and the relationship (31) between d[S] and d[R] have been used. Because the decay widths for different channels are independently distributed around zero (for pure compound-nuclear reactions) the matrix $R(\mathcal{E})$ is practically diagonal,

$$R(\mathcal{E})_{ab} = \sum_{\lambda} \frac{\gamma_{\lambda a}\gamma_{\lambda b}}{E_{\lambda}-\mathcal{E}} = (R_a^{\infty} + i\pi s_a)\delta_{ab}\,. \tag{52}$$

Expressing the distribution in terms of the real symmetric matrix R instead of the unitary symmetric matrix S simplifies things considerably (as is usual in R-matrix theory).

Eq. 51 suggests the choice $\lambda = -(n+1)/2$, which yields the joint distribution of R-matrix elements

$$p(R \mid R(\mathcal{E}))d[R] = C_n \frac{\prod_{a \le b} dX_{ab}}{\det(1+X^2)^{(n+1)/2}}, \quad -\infty < X_{ab} \equiv \frac{R_{ab} - R_a^{\infty}\delta_{ab}}{\pi\sqrt{s_a s_b}} < \infty,$$

$$C_n = \pi^{-n(n+1)/4} \frac{\Gamma((n+1)/2)}{\Gamma(1/2)} \prod_{c=1}^{n} \frac{\Gamma(n-c-1)}{\Gamma((n-c)/2+1)}. \tag{53}$$

This is a matrix generalization of the scalar t-distribution of ordinary statistics. The normalization is obtained by repeated application of the recursion relation (see Hua 1963)

$$I_n(\lambda) \equiv \int_{-\infty}^{\infty} \frac{d[X]}{\det(1+X^2)^{\lambda}}$$

$$= 2^{(n-1)/2}\pi^{n/2} \frac{\Gamma(\lambda - 1/2)\Gamma(2\lambda - (n+1)/2)}{\Gamma(1/2)\Gamma(2\lambda - 1)} I_{n-1}(\lambda - 1/2), \tag{54}$$

which permits reduction of the Integral I_n involving a real symmetric n×n matrix X to the analogous integral I_{n-1} involving the (n-1)×(n-1) matrix obtained from X by deletion of the n-th row and n-th column. The proof that the chosen λ is correct can be based on the observation that the distribution of R in the single-channel case is also the marginal distribution of each diagonal element R_{cc} in the multi-level case: A given set of level energies E_λ and decay amplitudes γ_λ defines the R function completely. Since the distribution functions for level energies and decay amplitudes are not changed by the presence of other open channels, the very same set could equally well occur in the multi-channel case for some channel, c say, if $\gamma_\lambda = \gamma_{\lambda c}$, $s = s_c$ and $R = R_{cc}$. At each energy E the diagonal element R_{cc} is then equal to R, hence the marginal distribution for R_{cc} must again be the Cauchy distribution. Our choice of λ is confirmed by the fact that one actually gets the Cauchy distribution in the course of the normalization (in the second but last step). Thus (53) is the correct maximum-entropy distribution for the constraint (50). Rewriting it in terms of S one finds the Poisson kernel defined in the domain of unitary symmetric matrices (Hua 1963)

$$p(S|\bar{S})d[S] = C_{n'}\left(\frac{\det T}{\det|1-\bar{S}^{+}S|^{2}}\right)^{(n+1)/2} d[S],$$

$$C_{n'} = 2^{-n(3n+1)/4} C_n . \tag{55}$$

"Polar" coordinates can be introduced with Eq. 36.

Knowing the distributions of S- or R-matrix elements one can, in principle, calculate average cross sections rigorously. The dimensionality of the integrals to be computed for n open channels is $n(n+1)/2$. With polar coordinates and integration over angles (i. e. over the δO_{ab} in Eq. 36) one can reduce the dimensionality to n. In the simplest case of two equivalent channels ($T_1 = T_2 \equiv 1 - r^2$) one gets, for instance,

$$\overline{|S_{12}|^2} = \frac{\pi}{4}\int_{-\pi}^{\pi}\frac{d\theta_1}{2\pi}\int_{-\pi}^{\pi}\frac{d\theta_2}{2\pi}\left(\sin^2\frac{\theta_1-\theta_2}{2}\prod_{c=1}^{2}\frac{1-r^2}{1+r^2+2r\cos\theta_c}\right)^{3/2} \tag{56}$$

In general, however, first experience with the Poisson kernel or the R-distribution shows that their compact determinantal structure makes them quite intractable. The next order of business should be to find practical ways to handle them, i. e. to find suitable expansions, to reduce the dimensionality of the integrals, and to deal with the very many weakly absorbing photon channels in a way similar to the recipe in conventional Hauser-Feshbach theory, where one lumps them to get an essentially non-fluctuating radiation width. At this point, however, the GOE approach has stolen the show again.

3.5 The GOE Triple Integral

Only a few months after the maximum-entropy distributions of the S and R matrix had been published by Mello, Pereyra and Seligman (1985) and by Fröhner (1985), Verbaarschot, Weidenmüller and Zirnbauer (1985) presented an analytic solution to the problem of averaging partial cross sections over the Gaussian Orthogonal Ensemble. They started from the expressions

$$|S_{ab}|^2 = |\delta ab + i \sum_{\lambda,\mu} \tilde{\gamma}_{\lambda a} A_{\lambda,\mu} \tilde{\gamma}_{\mu b}|^2, \tag{57}$$

$$(A^{-1})_{\lambda\mu} = H_{\lambda\mu} - E\delta_{\lambda\mu} - i \sum_{c} \tilde{\gamma}_{\lambda c} \tilde{\gamma}_{\mu c}, \tag{58}$$

a generalization of what Eqs. 17 and 18 give for $|S_{ab}|^2$. The tilde indicates that the Hamiltonian is left in its general nondiagonal form, so that $H_{\lambda\mu}$ and $\tilde{\gamma}_{\lambda c}$ replace the more familiar $E_\lambda \delta_{\lambda\mu}$ and $\gamma_{\lambda c}$ of Eqs. 17 and 18. Assuming H to belong to the GOE they managed, by a formidable display of analytic skill and with new tools from the many-body theory of disordered systems, to reduce the ensemble average of $|S_{ab}|^2$ over the GOE to a threefold integral. Fully exploiting the symmetries of the GOE, using a generating function (analogous to the partition function of the maximum-entropy approach) of both commuting and anticommuting (Grassmann) variables, simplifying the resulting integrations with the Hubbard-Stratonovitch transformation, and going to the limit of infinitely many levels by the method of steepest descent, they derived the expression

$$\overline{|S_{ab}|^2} = |\bar{S}_{ab}|^2$$

$$+ \frac{T_a T_b}{8} \int_0^\infty d\lambda_1 \int_0^\infty d\lambda_2 \int_0^1 d\lambda \frac{\lambda(1-\lambda)|\lambda_1 - \lambda_2|}{\sqrt{\lambda_1(1+\lambda_1)\lambda_2(1+\lambda_2)}\,(\lambda+\lambda_1)^2(\lambda+\lambda_2)^2}$$

$$\cdot \prod_c \frac{1 - T_c\lambda}{\sqrt{1+T_c\lambda_1}\sqrt{1+T_c\lambda_2}} \left\{ \delta_{ab}(1-T_a)\left(\frac{\lambda_1}{1+T_a\lambda_1} + \frac{\lambda_2}{1+T_a\lambda_2} + \frac{2\lambda}{1-T_a\lambda} \right)^2 \right.$$

$$+ (1+\delta_{ab})\left(\frac{\lambda_1(1+\lambda_1)}{(1+T_a\lambda_1)(1+T_b\lambda_1)} + \frac{\lambda_2(1+\lambda_2)}{(1+T_a\lambda_2)(1+T_b\lambda_2)} \right.$$

$$\left. \left. + \frac{2\lambda(1-\lambda)}{(1-T_a\lambda)(1-T_b\lambda)} \right) \right\} \tag{59}$$

Note that the analogous maximum-entropy expressions available so far involve at least n-dimensional integrals for n open reaction channels, whereas here we have a three-

dimensional integral no matter how many open channels there are. A particularly useful feature is the product over channels that permits a similar treatment of lumped photon channels as we used before (Eqs. 22, 23): If the number of photon channels is large we may approximate with

$$\prod_c \frac{1 - T_c\lambda}{\sqrt{1 + T_c\lambda_1}\sqrt{1 + T_c\lambda_2}}$$

$$\simeq e^{-(2\lambda + \lambda_1 + \lambda_2)T_\gamma/2} \prod_{c \notin \gamma} \frac{1 - T_c\lambda}{\sqrt{1 + T_c\lambda_1}\sqrt{1 + T_c\lambda_2}} \quad . \tag{60}$$

Verbaarschot (1986) verified that in the limit of vanishing level overlap the cross section expression involving the triple integral reduces to the old Hauser-Feshbach formula (21) with elastic enhancement and width fluctuation correction. He also chekked averages computed with the triple integral against averages over the Poisson kernel. In spite of the utterly different appearance of the multiple integrals in the two approaches he found the same numbers up to 3 or 4 digits, i.e. agreement within the numerical accuracy of the two computations. This constitutes another, quite stringent verification of the irrelevance of resonance details for the average behaviour of compound-nuclear cross sections. The GOE triple integral, which fully accounts for width fluctuation corrections and elastic enhancement, eliminates all uncertainties associated with picket fence approximations for average cross sections or heuristic analytic formulae derived from Monte Carlo calculations. These methodical uncertainties had always been bothersome as width fluctuation corrections are usually quite substantial (see e. g. Lynn 1968, Gruppelaar and Reffo 1977).

3.6 Application: Evaluation of ^{238}U in the Unresolved Resonance Region

^{238}U cross sections for neutron-induced reactions have, since the very beginning of nuclear technology, always been among the data requested with the highest accuracy. The unresolved resonance range, or more precisely, the range between 10 and 300 keV, has recently been reevaluated for JEF-2, the 2nd version of the Joint Evaluated File, with the new theoretical tools described above, in particular the GOE triple inte-

gral for partial cross sections (Fröhner 1989). The main reactions in this range are radiative capture, elastic scattering, and inelastic scattering with thresholds at 45, 149 and 310 keV corresponding to excitation of ^{238}U levels with spin-parity characteristics 2+, 4+ and 6+. The minute, technologically unimportant amount of subthreshold fission can be neglected in the present context. The capture and scattering cross sections are mostly needed for fast-reactor applications such as neutron transport calculations, breeding ratios, reactivity response to temperature changes or to voids in the coolant, while accurate total cross sections are needed for shielding calculations and as a valuable constraint on the partial cross sections in data evaluation.

The role of nuclear theory in the evaluation of resonance-averaged cross section data is twofold. First it provides smooth average curves where experimental data show real fluctuations due to partially resolved resonance structure (at lower energies) or spurious fluctuations due to experimental effects (throughout the unresolved range). In the latter case theory reduces experimental errors. Furthermore, theory permits simultaneous utilization of information from all open reaction channels, and also from other energy ranges, e. g. from the resolved resonance range or from optical-model studies at higher energies. Since it relates the average cross sections for all open channels, theory-aided coherent evaluation of all available information provides powerful constraints and reduces uncertainties drastically. This means one should fit the total and partial cross section formulae (Eqs. 20 and 59) to all data simultaneously which yields optimal estimates for the Hauser-Feshbach parameters (transmission coefficients or equivalent strength functions and distant-level parameters). The uncertainties and correlations obtained in the fits can then be used to establish confidence bands about the fitted curves.

A recent example for such a coherent fit to resonance-averaged data is shown in Figs. 4 - 6. Three types of neutron data for ^{238}U (total, capture and recent high-precision inelastic scattering cross sections) were fitted simultaneously in the energy range between 4 and 500 keV with the Hauser-Feshbach code FITACS, by Bayesian least-squares fitting which utilizes not only the a-priori values of the parameters (as first guesses) but also their uncertainties. (We note in passing that the maximum entropy principle provides also a very simple and completely general derivation of the least-squares formalism, see Fröhner 1986). A-priori values for the s wave were taken from

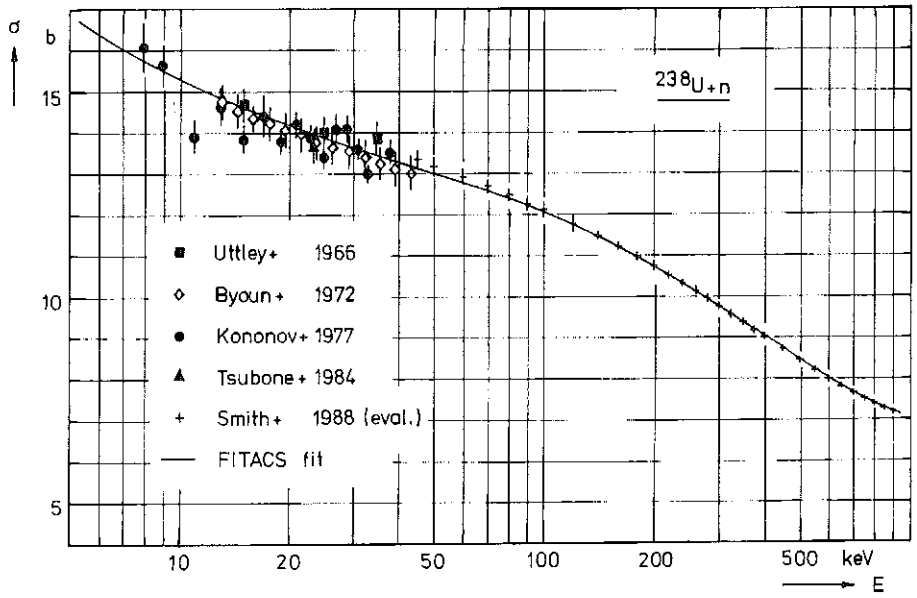

Fig. 4 - Coherent fit to ^{238}U data in the unresolved-resonance region: total cross section (for references see Fröhner 1989)

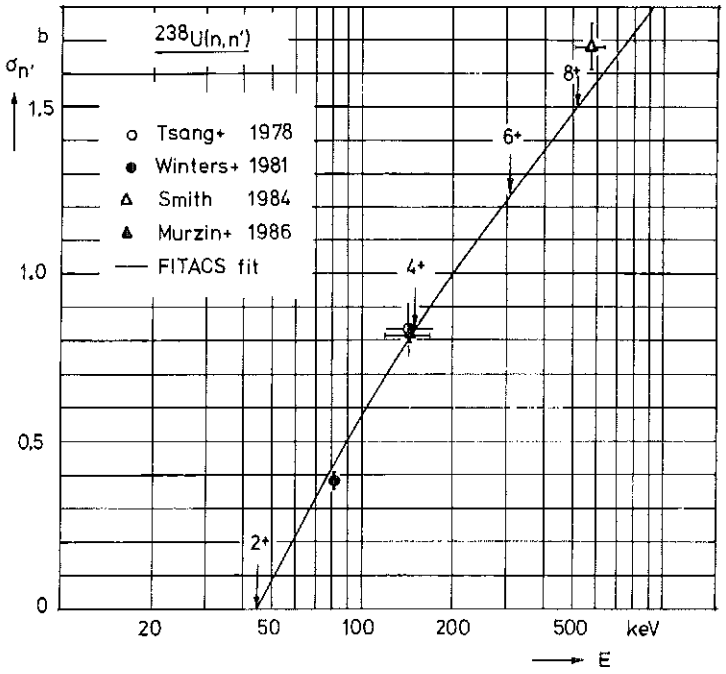

Fig. 5 - Coherent fit to ^{238}U data in the unresolved-resonance region: inelastic-scattering cross section (for references see Fröhner 1989). Inelastic thresholds are denoted by spin-parity characteristics of residual levels.

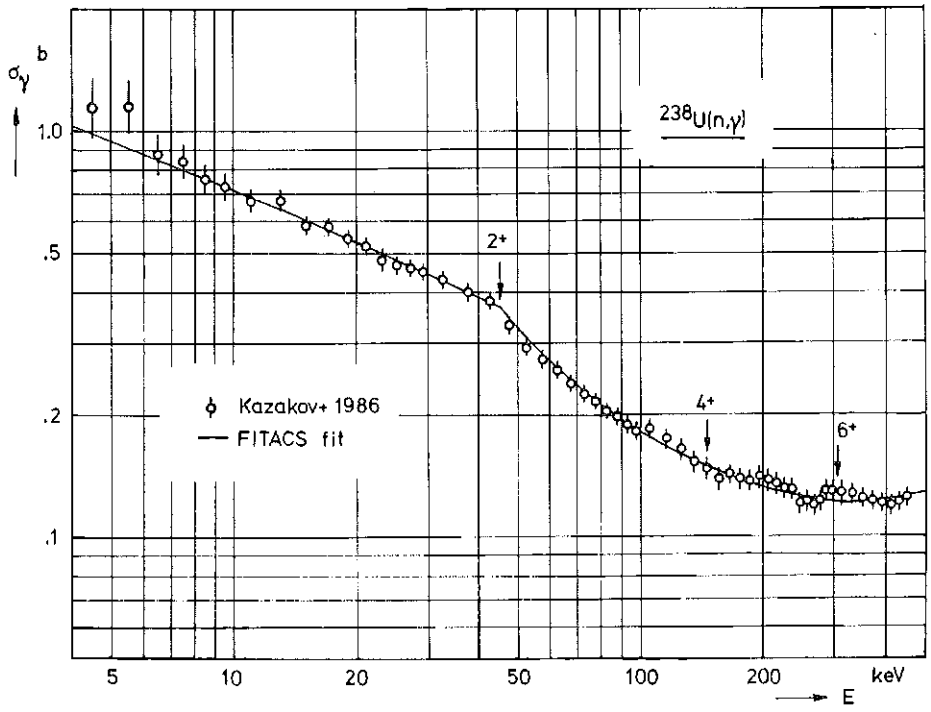

Fig. 6 - Coherent fit to ^{238}U data in the unresolved-resonance region: capture cross section (for references see Fröhner 1989). The discontinuity (Wigner cusp) at 45 keV is due to competition by inelastic scattering above that energy. Inelastic thresholds are denoted by spin-parity characteristics of residual levels.

4. SUMMARY

Entropy maximization, as a general logical tool for inductive reasoning, has important applications in statistical physics. It permits to find objective probability distributions for predictions from a few global characteristics, in particular mean values. Examples from nuclear physics were given: The simplest case of a scalar variate was illustrated by derivation of an analytical expression for the spectrum of neutrons emitted in nuclear fission, whereas matrix variates were involved in the derivation of the Gaussian Orthogonal Ensemble of nuclear Hamilton matrices and of the distributions of the R and S matrix in Hauser-Feshbach theory. In all cases the maximum-entropy expressions permit excellent reproduction of recent experimental data.

ACKNOWLEDGEMENTS

It is a pleasure to thank the organizers of the IX. International Summer School on Nuclear Physics, Neutron Physics and Nuclear Energy, Varna, October 1989, and in particular Prof. W. Andrejtscheff, for the opportunity to present this work. Support by the German Fast Breeder Project and by G. Kessler and H. Küsters (KfK) is gratefully acknowledged.

REFERENCES

- R. Balian, Nuovo Cim. **57B** (1968) 183
- C. Bloch, Nucl. Phys. **A112** (1968) 257, 273
- T.A. Brody, J.Flores, J.B. French, P.A. Mello, A. Pandey, S.S.M. Wong, Rev. Mod. Phys. **53** (1981) 385
- C. Budtz-Jørgensen and H.-H. Knitter, Nucl. Phys. **A490** (1988) 307
- J. de los Reyes, P.A. Mello and T.H. Seligman, Z. Physik **A247** (1980) 247
- L. Dresner, Proc. Int. Conf. on Neutron Interactions with the Nucleus, Columbia U., 1957, Report CU-175 (1957) 71, A.M. Lane and J.E. Lynn, Report AERE T/R 2210, Harwell (1957)
- F.J. Dyson and M.L. Mehta, J. Math. Phys. **4** (1963) 701,

reprinted in Porter (1965) p. 489
- C.A. Engelbrecht and H.A. Weidenmüller, Phys. Rev. **C8** (1973) 859; H. Nishioka and H.A. Weidenmüller, Phys. Lett. **157B** (1985) 101
- F.H. Fröhner, Nucl. Instr. Meth. **49** (1967) 89
- F.H. Fröhner, Proc. Internat. Conf. on Nucl. Data for Basic and Applied Sci., Santa Fe, N. M., 1985 (Gordon and Breach, New York, 1986) p. 1541; reprinted in Rad. Effects **96** (1986) 199
- F.H. Fröhner, Report KfK 4099 (1986)
- F.H. Fröhner, Proc. Meet. Nucl. Theory for Fast Neutron Nuclear Data Evaluation, Beijing, 12-16 Oct. 1987, Report IAEA-TECDOC-483 (1988) p. 160; Nucl. Sci. Eng. (in print)
- F.H. Fröhner, Nucl. Sci. Eng.**103** (1989) 119
- M. Gaudin, Nucl. Phys. **25** (1961) 447
- H. Gruppelaar and G. Reffo, Nucl. Sci. Eng. **62** (1977) 756
- H.M. Hofmann, J. Richert, J.W. Tepel and H.A. Weidenmüller, Ann. Phys. **90** (1975) 403
- L.K. Hua, Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains, (Amer. Math. Soc., Providence, R. I., 1963)
- N. Janeva, N. Koyumdjieva, A. Lukyanov, S. Toshkov, Proc. Internat. Conf. on Nucl. Data for Basic and Applied Sci., Santa Fe, N. M., 1985 (Gordon and Breach, New York, 1986) p. 1615; N. Koyumdjieva, N. Savova, N. Janeva, A.A. Lukyanov, Bulg. J. Phys. **16** (1989) **1**
- E.T. Jaynes, Phys. Rev. **106** (1957) 620, **108** (1957) 171
- E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics, edited by R.D. Rosenkrantz (D. Reidel Publishing Co., Dordrecht, 1983) ch. 7 and 10
- A.M. Lane and R.G. Thomas, Rev. Mod. Phys. **30** (1958) 257;
- G. Lopez, P.A.Mello, T.H.Seligman, Z. Physik **A302** (1981) 351
- J.E. Lynn, "The Theory of Neutron Resonance Reactions", Clarendon Press, Oxford (1968) p. 229
- D.G. Madland, Proc. Int. Conf. Nucl. Data for Sci. and Technol., Mito, May 30-June 3, 1988, JAERI (1988) p. 759

- D.G. Madland and J.R. Nix, Nucl. Sci. Eng. **81** (1982) 213,
- H. Märten and D. Seeliger, Nucl. Sci. Eng. **93** (1986) 370
- P.A. Mello, Phys. Lett. **B81** (1979) 103;
  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, Nucl. Phys. **A344** (1980) 185; see also
  P.A. Moldauer, Phys. Rev. **C11** (1975) 426, **C12** (1975) 744
- A. Müller and H.L. Harney, Phys. Rev. **C35** (1987) 1231
- C.E. Porter ed., Statistical Theory of Spectra: Fluctuations,
  (Academic Press, New York and London, 1965)
- C.E. Porter and N. Rosenzweig, Suomal. Tiedeakat.Toimit.
  (Ann. Acad. Sci. Fennicae) **AVI**, No. 44 (1960), reprinted in
  Porter (1965) p. 235, see also p. 2
- C.E. Porter and R.G. Thomas, Phys. Rev. **104** (1956) 483,
  reprinted in Porter (1965) p. 167
- G.R. Satchler, Phys. Letters 7 (1963) 55
- C.E. Shannon, Bell Syst. Tech. J. **27** (1948) 379 and 623,
  reprinted in C.E. Shannon and W. Weaver, The Mathematical
  Theory of Communication (Univ. of Illinois Press, Urbana, 1949)
- R.G. Thomas, Phys. Rev. **97** (1955) 224
- R.L. Walsh, Nucl. Sci. Eng. **102** (1989) 119
- J.M. Verbaarschot, Ann. Phys. **168** (1986) 368
- J.M. Verbaarschot and P.J. Brussaard, Phys. Lett. **87B** (1979) 155
- J.M. Verbaarschot, H.A. Weidenmüller and M.R. Zirnbauer,
  Phys. Repts. **129** (1985) 367;
- B.E. Watt, Phys. Rev. **87** (1952) 1952
- V. Weisskopf, Phys. Rev. **52** (1952), 295
- E.T. Whittaker and G.N. Watson, Modern Analysis,
  (University Press, Cambridge, 1951)
- E.P. Wigner, Can. Math. Congr. Proc. (U. of Toronto Press,
  Toronto, 1957) p. 174; Ann. Math. **67** (1958) 325;
  both reprinted in Porter (1965) pp. 188, 226