Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics Letters B 595 (2004) 171–176
PHYSICS LETTERS B
www.elsevier.com/locate/physletb

Missing levels in correlated spectra

O. Bohigas [a], M.P. Pato [b]

[a] *Laboratoire de Physique Théorique et Modèles Statistiques* [1], *Bâtiment 100, Université Paris-Sud, 91405 Orsay cedex, France*
[b] *Instituto de Fisica, Universidade de Sao Paulo, C.P. 20516, 01498 Sao Paulo, S.P., Brazil*

Received 1 March 2004; received in revised form 7 April 2004; accepted 22 May 2004

Available online 22 June 2004

Editor: J.-P. Blaizot

Abstract

Complete spectroscopy (measurements of a complete sequence of consecutive levels) is often considered as a prerequisite to extract fluctuation properties of spectra. It is shown how this goal can be achieved even if only a fraction of levels are observed. The case of levels behaving as eigenvalues of random matrices, of current interest in nuclear physics, is worked out in detail.
© 2004 Elsevier B.V. Open access under CC BY license.

It is by now well established that, at least with present statistical significance limited by the relatively scarce amount of high quality data available, the statistical properties of compound nucleus resonances with fixed quantum numbers are consistent with the predictions of the Wigner–Dyson random matrix model (Gaussian orthogonal ensemble, GOE) [1]. It is presently understood, not in the form of mathematical theorems but rather from numerical experiments and from theoretical insight based on semiclassical approaches, that the origin of this (universal) behavior has to be found in the chaotic character of the underlying classical motion [2]. Analyses have been mainly performed for two degrees of freedom systems (one particle in a two-dimensional box), which can be also experimentally exemplified in studying resonances of microwave cavities with adequate boundaries or of quartz or metallic blocks (elastodynamics) (see, for instance, contributions by A. Richter, H.J. Stöckmann and C. Ellegaard [3]).

Efforts have also been devoted to find out from which excitation energy on, starting from the ground state (where, by analyzing the systematics of ground state energies, evidence for the coexistence between regular and chaotic motion has been given [4]) the 'random matrix theory' behavior holds [5]. This is a difficult task because, below neutron threshold, one can hardly be sure that all levels are detected. Work is also currently performed to investigate the statistical properties of the fine structure of analog or other giant resonant states, at excitation energies which may be below or far above the neutron threshold [6].

At different degrees, most of these studies must face the difficulty of establishing the extent to which the sequences of energies analyzed are complete (no missing levels) as well as how the missing levels may

E-mail address: mpato@if.usp.br (M.P. Pato).
[1] Unité de Recherche de l'Université Paris XI associée au CNRS.

0370-2693 © 2004 Elsevier B.V. Open access under CC BY license.
doi:10.1016/j.physletb.2004.05.065

distort the statistical properties of the otherwise ideal (complete) sequence. It is the purpose of this Letter to present some results which may be useful in this context.

The most ambitious goal may be stated as the one of detecting the location of, say, one missing level on an otherwise complete sequence. Dyson, in a recent review [7], uses information theory concepts and argues that correlations in a sequence may provide the necessary redundancy from which error correcting codes can be constructed. At one extreme where no correlations and therefore no redundancy are present (Poissonian sequence), there is no possibility of detecting one missing level. At the other extreme, a sequence of equally spaced levels (picket fence), there is a maximum redundancy and a missed level can be obviously detected as a hole in the spectrum. Eigenvalues of random matrices, which exhibit characteristic correlations, correspond to an intermediate situation between these two extremes. The attempts to locate in the last case a single missed level have remained unsuccessful so far. However, it should be mentioned that for two-dimensional chaotic systems where, besides correlations of the order of one mean spacing as described by random matrices, the presence and the role of long range correlations governed by the shortest periodic orbits and reflected in Weyl's law describing the average spectral density, is well understood. It is then possible to approximately locate, from the study of spectral fluctuations, a single missed level [8].

Here we address a less ambitious question, namely to study how the spectral fluctuations of an ideally complete sequence are affected when a fraction f $(0 < f < 1)$ of levels are detected. It will be assumed that the sequence is infinite and stationary and that the levels are dropped at random from the complete sequence (random sampling). The last assumption is presumably justified when observing compound nucleus resonances with limited detection sensitivity. Indeed, in random matrix theory level positions and widths are uncorrelated, a feature which is confirmed by the analysis of experimental data.

We start considering the n-point correlation functions $R_n(E_1, E_2, \ldots, E_n)$ that give the joint probability density of finding levels located around each of the values $E_1, E_2, \ldots, E_n$. For the incomplete sequence these functions keep their form and they are only reduced by a factor f^n since f is the probability of observing one given level. We can therefore write the relation

$$r_n(E_1, E_2, \ldots, E_n) = f^n R_n(E_1, E_2, \ldots, E_n), \tag{1}$$

where small cases denote the actual observed quantities and capital letters those of the complete spectrum. We shall keep this notation in what follows, namely use the same letter, capital or small case, to refer to corresponding quantities for complete or incomplete sequences, respectively. For the spectral density (1) gives

$$r_1(E) = f R_1(E), \tag{2}$$

and, from (2), the unfolding procedure that maps the levels onto a new spectrum whose density is equal to one yields the relation

$$x_k = \int_{-\infty}^{E_k} r_1(E)\, dE = f X_k, \tag{3}$$

where $k = 1, 2, \ldots$ labels the levels which are observed. The normalization will always be kept the same (the mean spacing $\langle s \rangle = 1$ for the complete as for incomplete sequences). This is in contrast to some of the expressions given in [9].

One is usually interested in the n-point cluster functions $Y_n(X_1, X_2, \ldots, X_n)$ from which lower order correlations are subtracted. From their definition [10] and (1) we deduce

$$y_n(x_1, x_2, \ldots, x_n) = Y_n\left(\frac{x_1}{f}, \frac{x_2}{f}, \ldots, \frac{x_n}{f}\right). \tag{4}$$

These important equations show that the correlation functions keep their form, the only modification is a rescaling of variables. Thus, if the fraction f is known, the correlation functions can be reconstructed from the partially observed spectrum.

It is well known that the number variance $\Sigma^2(L)$ (variance of the number of levels contained in an interval of length L) and the Dyson–Mehta Δ-statistics (least square deviation, in an interval of length L, of the actual staircase counting function from its linear approximation) are related to the two-point cluster function through [10]

$$\Sigma^2(L) = L - 2\int_0^L (L - x) Y_2(x)\, dx, \tag{5}$$

$$\Delta(L) = \frac{L}{15} - \frac{1}{15L^4}\int_0^L (L-x)^3 \times \left(2L^2 - 9xL - 3x^2\right) Y_2(x)\,dx.$$

From (4) it follows immediately that

$$\sigma^2(L) = (1-f)L + f^2\Sigma^2\left(\frac{L}{f}\right) \tag{6}$$

and

$$\delta(L) = (1-f)\frac{L}{15} + f^2\Delta\left(\frac{L}{f}\right). \tag{7}$$

Notice the appearance of a linear term in both equations, if $f < 1$.

Another set of statistical measures are the $E(n,s)$ functions that give the probabilities of finding n levels, with $n = 0, 1, 2, \ldots$, inside an interval of length s. If for the complete spectrum they are given by $E(n,s)$, then their expressions if only a fraction f of levels is observed is

$$e(n,s) = \sum_{k=n}^{\infty} \frac{k!}{n!(k-n)!} f^n (1-f)^{k-n} E\left(k, \frac{s}{f}\right). \tag{8}$$

This follows from the fact that $1-f$ is the probability that one level is missed, f that it is not and $k!/n!(k-n)!$ counts the number of ways $k-n$ points can be removed from k points. From $e(n,s)$, the spacing distributions $p(n,s)$, the probability density of finding n levels ($n = 0, 1, 2, \ldots$) between two given levels separated by a distance s, can be derived using the relation

$$P(n,s) = \frac{d^2}{ds^2}\sum_{k=0}^{n}(n-k+1)E(k,s); \tag{9}$$

one obtains

$$p(n,s) = \sum_{k=n}^{\infty} \frac{k!}{n!(k-n)!} f^n (1-f)^{k-n} P\left(k, \frac{s}{f}\right). \tag{10}$$

For the nearest neighbor distribution (NND), i.e., $p(s) = p(0,s)$ (10) becomes

$$p(s) = \sum_{k=0}^{\infty}(1-f)^k P\left(k, \frac{s}{f}\right). \tag{11}$$

The equations derived above imply that the general relation

$$y_2(x) = 1 - \sum_{n=0}^{\infty} p(n,x)$$

is fulfilled, as it should.

Expression (11) was first proposed as an *ansatz* in Ref. [11], and in Ref. [9] it has been shown that the coefficients $f(1-f)^k$ maximize the Shannon entropy with constraints appropriately defined. Here it is obtained in a direct way.

We first apply, as a check, the previous equations to a Poissonian spectrum of uncorrelated levels in which case they must remain invariant (they should not depend on f). If, for example, we substitute in Eq. (8) the expression

$$E(n,s) = \frac{s^n}{n!}\exp(-s) \tag{12}$$

for the E-functions of a Poisson sequence, one obtains

$$e(n,s) = \sum_{k=n}^{\infty} \frac{k!}{n!(k-n)!} f^n (1-f)^{k-n} \times \frac{1}{n!}\left(\frac{s}{f}\right)^n \exp\left(-\frac{s}{f}\right) = E(n,s). \tag{13}$$

It follows that any other measure derived from the $E(n,s)$ functions will also remain invariant. Since $Y_2(X_1, X_2, \ldots, X_n) = 0$, the n-point cluster functions of the incomplete sequence also vanish and all other measures derived from them also remain invariant.

Before considering GOE spectra with a fraction of observed levels, let us first recall some of the GOE expressions ($f = 1$) to be used [10]. The two-point cluster function reads

$$Y_2(x) = \left[\frac{\sin(\pi x)}{\pi x}\right]^2 - \left[\mathrm{Si}(\pi x) - \pi\epsilon(x)\right] \times \left[\frac{\cos(\pi x)}{\pi x} - \frac{\sin(\pi x)}{(\pi x)^2}\right], \tag{14}$$

where $\mathrm{Si}(\pi x)$ is the sine-integral and $\epsilon(x) = \pm 1/2$ if $x > 0$ or $x < 0$ and $\epsilon(0) = 0$ if $x = 0$. The number variance $\Sigma^2(L)$ is given by

$$\Sigma^2(L) = \frac{2}{\pi^2}\left[\ln(2\pi L) + \gamma + 1\right] - \frac{1}{4} + O\left(L^{-1}\right), \tag{15}$$

where $\gamma = 0.577\ldots$ is Euler's constant and the Δ-statistics is given by

$$\Delta(L) = \frac{1}{\pi^2}\left[\ln(2\pi L) + \gamma + \frac{5}{4}\right] - \frac{1}{8} + O\left(L^{-1}\right). \tag{16}$$

The NND distribution is well approximated by the Wigner surmise (2×2 matrices)

$$P(s) = P(0,s) = \frac{\pi}{2} s \exp\left(-\frac{\pi}{4}s^2\right). \tag{17}$$

The next to nearest neighbor distribution $P(1,s)$ is given by the NND of the symplectic ensemble which again is well approximated by the 2×2 matrix result

$$P(1,s) = \frac{8}{3\pi^3}\left(\frac{4}{3}\right)^5 s^4 \exp\left(-\frac{16s^2}{9\pi}\right) \tag{18}$$

(care has been take to insure $\langle s \rangle = 2$). The higher ($k = 2, 3, \ldots$) spacing distributions $P(k,s)$ are well approximated by their (Gaussian) asymptotic form, centered at $k+1$ and variances $V(k)$ given by [12]

$$V^2(k) \simeq \Sigma^2(L = k) - \frac{1}{6}. \tag{19}$$

We apply now the above expressions to GOE spectra when only a fraction f of levels is observed. On Fig. 1 is displayed for different values of f the function Y_2, illustrating the scaling behavior Eq. (4) using (14).

By comparing Eq. (6) with (15) and (7) with (16) one immediately sees one major effect of the incompleteness of a sequence. Instead of a logarithmic increase of the number variance or of the Dyson–Mehta statistic ($f = 1$), one has a linear increase for $f < 1$ of slope $(1-f)$ or $(1-f)/15$, respectively.

Consider finally $p(s)$, the NND very often discussed in the literature. It is given by (11), (17), (18), (19). The range of the argument determines the number of terms to be included in (11). The slope at the origin given by the first term in (11) is increased by a factor $1/f$ with respect to the GOE value. Let us also mention that, in fact, Eq. (11) was already used to analyze high quality data of ^{238}U neutron resonances [13] and of ^{48}Ti proton resonances [9]. It was concluded that a fraction of about 10% of the resonances is missing in both cases and that the correction is important in investigating the parity dependence of nuclear level densities [14].

It is convenient, in particular to exhibit the asymptotic behavior of $p(s)$, to rewrite Eq. (11) in a different

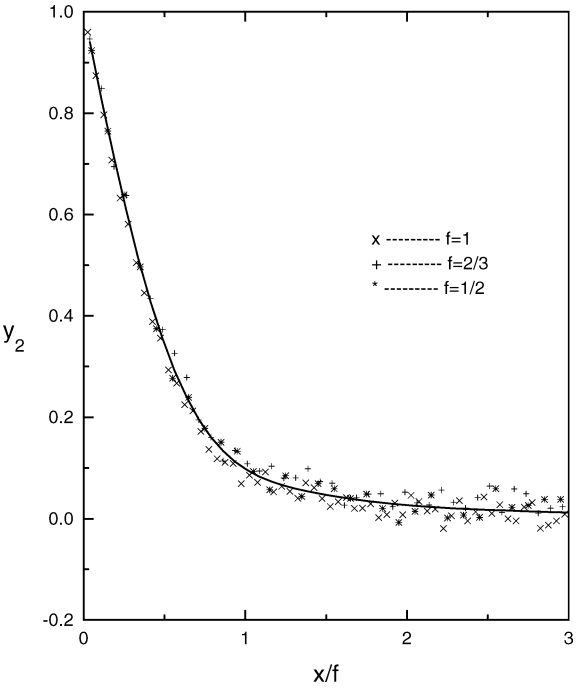

Fig. 1. The two-level cluster function Y_2. Full line: Eqs. (4) and (14). Points: obtained with numerical simulation for different values of the fraction f of levels observed.

form. Let us first separate the first K terms in the sum and approximate the rest (the infinite sum from $k = K$ to $k = \infty$) by an integral in which k is treated as a continuous variable

$$p(s) = \sum_{k=0}^{k=K-1} (1-f)^k P\left(k, \frac{s}{f}\right) + \int_K^\infty \frac{dk}{\sqrt{2\pi V^2(k)}} \exp\left[k\ln(1-f) - \frac{1}{2V^2(k)}\left(\frac{s}{f} - k - 1\right)^2\right]. \tag{20}$$

A steepest descent approximation to the integral can be worked out by mapping the exponent in the integrand onto a parabola as

$$F(k) = -k\ln(1-f) + \frac{1}{2V^2(k)}\left(\frac{s}{f} - k - 1\right)^2 = F(k_s) + t^2, \tag{21}$$

where k_s, the stationary point, is the root of the equation $F'(k) = 0$. This transcendent equation can

be solved approximately if one considers the variance $V^2(k)$ as a slowly varying function of its argument, in which case one can write

$$k_s = \frac{s}{f} - 1 + V^2\left(\frac{s}{f} - 1\right)\ln(1-f). \tag{22}$$

Replacing then the coefficient $1/\sqrt{2\pi V^2(k)}(dk/dt)$ by its value at the stationary point, we obtain the closed expression

$$p(s) = \sum_{k=0}^{k=K-1} (1-f)^k P\left(k, \frac{s}{f}\right) + \frac{1}{2}\exp\left[-F(k_s)\right]\left[1 - \operatorname{erf}(t_s)\right], \tag{23}$$

where

$$t_s = \begin{cases} \sqrt{F(K) - F(k_s)}, & \text{for } k_s < K, \\ -\sqrt{F(K) - F(k_s)}, & \text{for } k_s > K. \end{cases} \tag{24}$$

In the asymptotic region, s/f large, (23) takes the simple expression

$$p(s) \sim \frac{1}{1-f}\exp\left[\frac{s}{f}\ln(1-f)\right] \tag{25}$$

illustrating that the spacing distribution approaches a Poisson distribution as f tends to zero. We have therefore a family of NND distributions, parametrized by the fraction f of the observed levels, that interpolates between the GOE and the Poisson statistics.

In Fig. 2 comparison is made with numerical simulations. It can be seen that for $f = 0.90$ the asymptotic regime Eq. (25) is not yet reached in the range considered whereas for $f = 0.50$ it is already reached for $s \gtrsim 2$.

In conclusion, we have shown how the problem of missing, at random, a fraction of levels of a correlated spectrum can be solved. From the observation of an incomplete spectrum the results presented here have a twofold application: (i) if the missing fraction is known, one can recover the statistical properties of the complete spectrum, (ii) if the statistical nature of the complete spectrum is known, an estimate of the fraction of missed levels can be obtained. This may at first sight appear as too constraining. Remember, however, that one of the first concerns in a given experimental setup is to have the efficiency of the detectors under control, without which little statistically significant can be extracted. In this respect

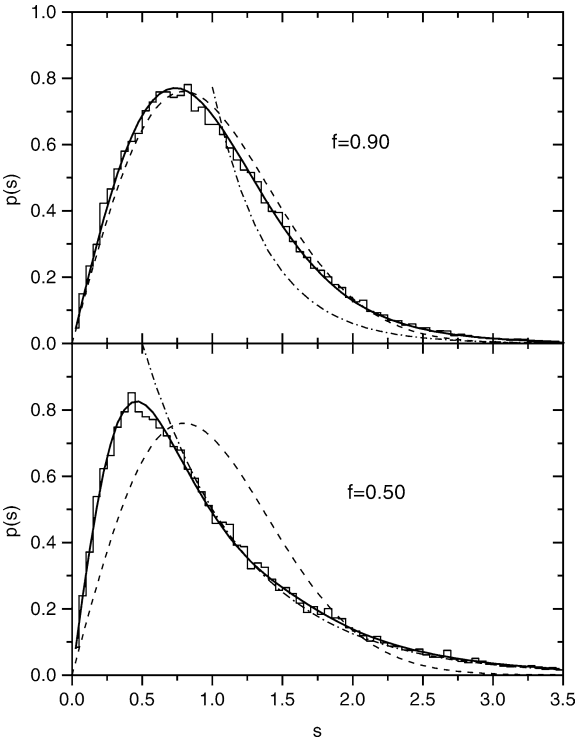

Fig. 2. Nearest neighbor spacing distribution $p(s)$ for two values of f. Full line: theory Eq. (11) or (23). Dashed line: Wigner surmise Eq. (17). Dotted line: asymptotic behavior as given by Eq. (25). Histograms: numerical simulation.

let us emphasize that the linear increase of the number variance and of the Δ-statistics accompanied by an exponential decrease of the NND can be a signature of the incompleteness of the spectrum. It is also possible in some cases to obtain independent complementary information. In the case discussed here the analysis of resonance intensities [9] may provide an independent information, to be crosschecked with the results of spectral fluctuations.

Some final remarks are in order: (i) a straightforward extension of the present Letter can be performed in order to include effects not of missed but of spurious levels; (ii) in order to extract parameters from the data, effects due to the finiteness of experimental samples must be added; (iii) with the advent of the neutron nTOF facility at CERN [15] a new generation of measurements of compound nucleus resonances is likely to take place, for which the methods discussed here may be particularly adapted; (iv) besides its prac-

tical interest, the one parameter (f) family interpolating spectral fluctuations between Wigner–Dyson and Poissonian correlations is interesting in its own.

Acknowledgements

Fruitful discussions with A. Heine, G.E. Mitchell and A. Richter are acknowledged. This work is supported by a project CAPES-COFECUB. M.P.P. is partially supported by the Brazilian agency Conselho Nacional de Pesquisas (CNPq).

References

[1] R.U. Haq, A. Pandey, O. Bohigas, Phys. Rev. Lett. 48 (1982) 1086;
O. Bohigas, R.U. Haq, A. Pandey, in: K.H. Böchhoff (Ed.), Nuclear Data for Science and Technology, Reidel, Dordrecht, 1983, p. 809;
O. Bohigas, R.U. Haq, A. Pandey, Phys. Rev. Lett. 54 (1985) 1645;
M. Lombardi, O. Bohigas, T. Seligman, Phys. Lett. B 324 (1994) 263.

[2] O. Bohigas, M.J. Giannoni, C. Schmit, Phys. Rev. Lett. 52 (1984) 1;
M.V. Berry, Proc. R. Soc. London A 400 (1985) 229;
S. Müller, S. Heusler, P. Braun, F. Haake, A. Altland, nlin.CD/0401021;
For reviews see, O. Bohigas, Random matrix theories and chaotic dynamics, in: M.-J. Giannoni, et al. (Eds.), Chaos and Quantum Physics, Proceedings Les Houches Summer School 1989, Elsevier, Amsterdam, 1991;
T. Guhr, A. Müller-Groeling, H.A. Weidenmüller, Phys. Rep. 299 (1998) 189.

[3] K.-F. Berggren, S. Aberg (Eds.), Nobel Symposium Quantum Chaos Y2K, Physica Scripta T90 (2001).

[4] O. Bohigas, P. Leboeuf, Phys. Rev. Lett. 88 (2002) 09502;
O. Bohigas, P. Leboeuf, Phys. Rev. Lett. 88 (2002) 129903, Erratum.

[5] T. von Egidy, A.N. Behkami, H.H. Schmidt, Nucl. Phys. A 454 (1986) 109;
T. von Egidy, H.H. Schmidt, A.N. Behkami, Nucl. Phys. A 481 (1988) 189;
J.F. Schriner Jr., G.E. Mitchell, T. von Egidy, Z. Phys. A 338 (1991) 309;
J.D. Garrett, J.Q. Robinson, A.J. Foglia, H.Q. Jin, Phys. Lett. B 392 (1997) 24.

[6] S. Strauch, P. von Neumann-Cosel, C. Rangacharyulu, A. Richter, G. Schrieder, K. Schweda, J. Wambach, Phys. Rev. Lett. 85 (2000) 2913.

[7] F.J. Dyson, Lecture at Workshop on Random Matrix Theory, Berkeley, CA, 2002 unpublished.

[8] C. Schmit, M.J. Giannoni, O. Bohigas, in: M. di Toro, W. Nörenberg, M. Rosina, S. Stringari (Eds.), Phase Space Approach to Nuclear Dynamics, Proceedings of Trieste, 1985, World Scientific, Singapore, 1986.

[9] U. Agvaanluvsan, G.E. Mitchell, J.F. Shriner Jr., M.P. Pato, Nucl. Instrum. Methods A 498 (2003) 459.

[10] M.L. Mehta, Random Matrices, second ed., Academic Press, New York, 1991.

[11] W.A. Watson, E.G. Bilpuch, G.E. Mitchell, Nucl. Instrum. Methods 188 (1981) 571.

[12] J.B. French, P.A. Mello, A. Pandey, Ann. Phys. (N.Y.) B 18 (1978) 277;
O. Bohigas, P. Leboeuf, M.-J. Sanchez, Physica D 131 (1999) 186.

[13] H.I. Liou, H.S. Camarda, F. Rahn, Phys. Rev. C 5 (1972) 1002.

[14] U. Agvaanluvsan, G.E. Mitchell, J.F. Shriner Jr., M. Pato, Phys. Rev. C 67 (2003) 064608.

[15] CERN/INTC/2000-004, 2000-008.