Information Functionals and the Notion of (Un)Certainty: Random Matrix Theory — Inspired Case

P. Garbaczewski[*]

Institute of Physics, University of Opole

Oleska 48, 45-052 Opole, Poland

Information functionals allow one to quantify the degree of randomness of a given probability distribution, either absolutely (through min/max entropy principles) or relative to a prescribed reference one. Our primary aim is to analyze the "minimum information" assumption, which is a classic concept (R. Balian, 1968) in the random matrix theory. We put special emphasis on generic level (eigenvalue) spacing distributions and the degree of their randomness, or alternatively — information/organization deficit.

PACS numbers: 02.50.–r, 03.65.–w, 05.45.–a

1. Motivation

The statistical theory of random-matrix spectra (RMT) [1, 2] provides an ideal playground to test workings of the Shannon and Kullback–Leibler (K–L) entropies in diverse contexts. That pertains to a direct analysis of spectral data for complex quantum systems (semiclasically chaotic case included), but as well to the statistics of Gaussian matrix ensembles and random-matrix diffusion processes. Dyson's interacting Brownian motion model can be interpreted as a non--equilibrium dynamical process, whose asymptotic distribution is related to the thermodynamical equilibrium state of a Coulomb gas (RMT as equilibrium statistical mechanics). Ultimately one may pass to probability densities inferred from the ground state(s) of singular Calogero-type quantum systems: Shannon and K–L entropies prove to be proper tools in the quantum case as well.

Before embarking on these issues, let us indicate that there are ambiguities involved in the very concept of *information* and *(un)certainty*. To stay on a solid ground [3–6], we must accept a specific lore of semantic games, where baffling synonyms quite often appear and their specific meaning is under scrutiny. Examples are: information vs. entropy notions, (un)certainty and randomness vs. information deficit, entropic measures of surprise vs. information functionals,

---
[*]e-mail: pgar@uni.opole.pl

min/max entropy principle vs. effective randomness (uncertainty), uncertainty (lack of information) vs. (quantum) indeterminacy.

Since a particular definition of an entropy functional is non-unique and to high extent purpose-dependent [6] one must make suitable choices in the entropic menu ("entropic mess", with a partially random order of entries): Clausius thermodynamic entropy, Boltzmann, Gibbs, Shannon, relative, conditional, Kullback–Leibler, Renyi, Tsallis, Wehrl, information entropy, differential entropy, Kolmogorov–Sinai entropy, von Neumann entropy; the list may be continued.

We shall basically invoke the Shannon and Kullback–Leibler entropies, in conjunction with continuous probability distributions on R^+. We base our discussion on the text-book wisdom that the entropy is a measure of the degree of randomness and the tendency (in the time domain) of physical systems to become less and less organized. We extend this verbal phrase to probability densities of the functional form

$$f(x) \sim s^\beta \exp(-s^\alpha) \tag{1}$$

with $s \in R^+$, $\alpha = 1$ or 2, while $\beta = 0, 1, 2, 3, 4$.

The above formula encompasses [1] a number of "quantum chaos"-related level spacing distributions: Poisson (strictly speaking — exponential), semi-Poissonian of various types and the generic family of spacing densities, that are exact for 2×2 random matrices, and are identifiable as $n = 2, 3, 4, 5$ Bessel–Ornstein–Uhlenbeck probability laws (densities) on R^+.

The latter arise directly from Gaussian matrix ensembles and n in the exponent of s^{n-1} counts the independent 2×2 Gaussian random matrix elements, $\beta = n - 1$. With the $\langle s \rangle = 1$ normalization, we have

$$P_{\text{GOE}}(s) = s \frac{\pi}{2} \exp\left(-s^2 \frac{\pi}{4}\right), \qquad P_{\text{GUE}}(s) = s^2 \frac{32}{\pi^2} \exp\left(-s^2 \frac{4}{\pi}\right),$$

$$P_{\text{Ginibre}}(s) = s^3 \frac{3^4 \pi^2}{2^7} \exp\left(-s^2 \frac{3^2 \pi}{2^4}\right),$$

$$P_{\text{GSE}}(s) = s^4 \frac{2^{18}}{3^6 \pi^3} \exp\left(-s^2 \frac{64}{9\pi}\right). \tag{2}$$

The $\beta = 0$, $\langle s \rangle = 1$ normalized Gaussian on R^+ reads $P_0(s) = (2/\pi) \exp(-s^2/\pi)$, and has the variance $\langle (s - \langle s \rangle)^2 \rangle = (\pi - 2)/2$.

2. Random variables on R^+ and entropic measures of probability (de)localization

2.1. Entropies

Given a probability measure $\sum_{j=1}^N \mu_j = 1$. Its Shannon entropy reads $\mathcal{S}(\mu) = -\sum_{j=1}^N \mu_j \ln \mu_j$ and takes a maximum value $\ln N$ in the "most random" case of a uniform distribution: $\mu_j = 1/N$ for all $1 \leq j \leq N$. An obvious minimum at 0 appears if for any j we have $\mu_j = 1$.

We shall focus on continuous probability distributions on R^+. The corresponding Shannon entropy is introduced as follows:

$$\int \rho(s)\mathrm{d}s = 1 \rightarrow \mathcal{S}(\rho) = -\int \rho(s)\ln\rho(s)\mathrm{d}x. \tag{3}$$

At this point it is instructive to mention that in the realistic (spectral data analysis) "quantum chaos" framework, one encounters spacing histograms and definitely *not* continuous probability densities. The latter may merely be interpreted as useful continuous approximants of discrete probability measures.

The situation is more involved in the case of the corresponding Shannon entropies, where the approximation issue is delicate. Even if one follows a pedestrian reasoning, we can justify and keep under control the limiting behavior [3, 6]:

$$\sum_1^N \mu_j = 1 \rightarrow \int \rho\mathrm{d}x = 1. \tag{4}$$

An immediate question is: what can be said about the mutual relationship of $S(\mu) = -\sum_1^N \mu_j \ln\mu_j$ and $S(\rho) = -\int \rho(s)\ln\rho(s)\mathrm{d}s$?

We first observe that $0 \leq -\sum_1^N \mu_j \ln\mu_j \leq \ln N$ and consider an interval of length L on a line with the a priori chosen partition unit $\Delta s = L/N$. Next, we define: $\mu_j \doteq p_j\Delta s$ and notice that (formally, we bypass an issue of dimensional quantities)

$$S(\mu) = -\sum_j (\Delta s)p_j \ln p_j - \ln(\Delta s). \tag{5}$$

Let us fix L and allow N to grow, so that Δs decreases and the partition becomes finer. Then

$$\ln(\Delta s) \leq -\sum_j (\Delta s)p_j \ln p_j \leq \ln L, \tag{6}$$

where

$$S(\mu) + \ln(\Delta s) = -\sum_j (\Delta s)p_j \ln p_j \Rightarrow S(\rho) = -\int \rho(s)\ln\rho(s)\mathrm{d}s. \tag{7}$$

$S(\rho)$ is the Shannon information entropy for the probability measure on the interval L. In the infinite volume $L \rightarrow \infty$ and infinitesimal grating $\Delta s \rightarrow 0$ limits, the density functional $S(\rho)$ may be unbounded both from below and above, even non-existent, and seems to have lost any computationally useful link with its coarse-grained version $S(\mu)$.

However, the situation is not so bad, if we invoke standard methods [3, 6] to overcome a dimensional difficulty, inherent in the very definition of $S(\rho)$, if we admit dimensional units. Namely, from the start we can take a (sufficiently small) partition unit Δs to have dimension of length. We allow s to carry the length dimension as well. Then, the dimensionless expression for the Shannon entropy of a continuous probability distribution reads

$$S_\Delta(\rho) = -\int \rho(s) \ln[\Delta s \cdot \rho(s)] \mathrm{d}s \tag{8}$$

and all of a sudden, a comparison of Eqs. (5) and (8) appears to make sense. We can legitimately set estimates for $|S(\mu) - S_\Delta(\rho)|$ and directly verify the approximation validity of $S(\mu)$ in terms of $S_\Delta(\rho)$, when the partition becomes finer.

In the present paper we are interested in properties of various continuous probability distributions, and *not* their coarse-grained versions. Therefore, our further discussion will be devoid of any dimensional or partition unit connotations. Since negative values of the Shannon entropy are now admitted, instead of calling it an information measure, we prefer to tell about a "localization measure", "measure of surprise" or "measure of information deficit".

2.2. Poissonian spacing distributions

Let $X_1, X_2, \ldots$ be independent random variables on R^+, with a common for all of them exponential probability law

$$\mu(x) = \alpha \exp(-\alpha x), \tag{9}$$

where $\alpha > 0$, mean — $1/\alpha$, variance — $1/\alpha^2$. Let us denote $S_n = X_1 + X_2 + \ldots + X_n$, $n = 1, 2, \ldots$ and note that S_n has the density (Poisson probability law)

$$p_n(x) = \frac{\alpha^n x^{n-1}}{(n-1)!} \exp(-\alpha x) \tag{10}$$

coming from an $(n-1)$-fold convolution of exponential probability densities on R^+. The law is infinitely divisible

$$p_{n+m}(x) = (p_n * p_m)(x) = \int_0^x p_n(x-y) p_m(y) \mathrm{d}y \tag{11}$$

with $p_1(x) = \mu(x)$ and $n, m = 1, 2, \ldots$

In particular, $X_i + X_j$ for any $i, j \in N$ has a probability density $p_2(x) = \alpha^2 x \exp(-\alpha x)$ which upon setting $\alpha = 2$ and $x = s$ stands for an example of a semi-Poisson law

$$P(s) = 4s \exp(-2s) \tag{12}$$

known to govern the adjacent level statistics for a subclass of pseudo-integrable systems. Other (plasma-model related) semi-Poisson laws arise as well. For example, S_3 has a density $p_3(x)$, which upon setting $\alpha = 3$ and $x = s$, gives rise to

$$P(s) = \frac{27}{2} s^2 \exp(-3s). \tag{13}$$

Analogously, S_5 yields $p_5(x)$ and upon setting $\alpha = 5$ implies

$$P(s) = \frac{3125}{24} s^4 \exp(-5s). \tag{14}$$

The distribution (Eq. (10)), here identified as the Poisson probability law for the random variable S_n, in the information-theoretic literature is known as the (α, n)-Erlang distribution. Its Shannon entropy reads [3]:

$$\mathcal{S}(p_n) = \ln \Gamma(n) + (1-n)\psi(n) + n - \ln \alpha, \tag{15}$$

where the Euler gamma function $\Gamma(x) = \int_0^\infty \exp(-t)\, t^{x-1} \mathrm{d}t$ appears, together with the digamma function (logarithmic derivative of Γ) $\psi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \ln \Gamma(x)$.

We have $\Gamma(n) = (n-1)!$ and $\psi(n) = H_{n-1} - \gamma$, where $\gamma = \lim_{n\to\infty}(H_n - \ln n) \sim 0.577215$ is the Euler–Mascheroni constant, while harmonic numbers $H_n = \sum_{k=1}^n (1/k)$ take the consecutive values 1, 3/2, 11/6, 25/12, etc.

Let us notice that $\alpha = n$ should be set if one needs to address the previous $P(s)$. For the pure exponential law, we have: $\mathcal{S}(p_1) = 1 - \ln \alpha$ and the fit $\alpha = 1$ would give us $\mathcal{S}(p_1) = 0$.

2.3. Bessel–Ornstein–Uhlenbeck processes and their invariant densities

Let $X_1, X_2, \ldots, X_n$ be independent random variables with common for all, with mean $= 0$ and variance $= 1$, Gauss (Brownian) probability law on R:

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2). \tag{16}$$

Let us consider

$$R_n \doteq (X_1^2 + \ldots + X_n^2)^{1/2}. \tag{17}$$

Let us assume the Brownian motion (Wiener process) to proceed, in n independent copies. The radial Brownian motion (Bessel process) is thereby induced on R^+. The probability density of $R \doteq R_n$, $n > 1$ at time $t \in R^+$ is denoted by $\rho(r,t)$, $r \in R^+$. We have

$$\mathrm{d}R = \left(\frac{n-1}{2R}\right) \mathrm{d}t + \mathrm{d}W \Longrightarrow \partial_t \rho = \frac{1}{2}\triangle\rho - \nabla\left[\frac{(n-1)}{2r}\rho\right]. \tag{18}$$

It is known that the point $r = 0$ is never reached with the probability 1, which models a repulsion [7]. (Here, $r = 0$ is the so-called entrance boundary.)

If we impose a restoring harmonic force (proportional to a randomly taken value of the distance R_n from the origin)

$$\mathrm{d}R = \left(\frac{n-1}{2R} - R\right) \mathrm{d}t + \mathrm{d}W \Longrightarrow \partial_t \rho = \frac{1}{2}\triangle\rho - \nabla\left[\left(\frac{n-1}{2r} - r\right)\rho\right]. \tag{19}$$

We take $\rho_0(r)$ with $r \in R^+$ as the density of distribution of the random variable R at time $t = 0$. Then the function $\rho(r,t)$, solving the Fokker–Planck (F–P) equation, is the density of $R = R(t)$ for all $t > 0$.

The $n > 1$ family of time homogeneous radial (Bessel) Ornstein–Uhlenbeck (B–OU) processes is driven by transition probability densities [8]:

$$p_t(r',r) = p(r',0,r,t) = 2r^{n-1} \exp(-r^2) \frac{1}{1 - \exp(-2t)}$$

$$\times \exp\left(-\frac{(r^2 + r'^2)\exp(-2t)}{1 - \exp(-2t)}\right) [rr'\exp(-t)]^{-\alpha} I_\alpha\left(\frac{2rr'\exp(-t)}{1 - \exp(-2t)}\right), \tag{20}$$

where $\alpha = (n-2)/2$ and $I_\alpha(z)$ is a modified Bessel function of order α:

$$I_\alpha(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k+\alpha}}{(k!)\Gamma(k+\alpha+1)}. \tag{21}$$

We recall special values of the Euler gamma function: $\Gamma(n+1) = n!$ and $\Gamma(n+1/2) = (2n)!\sqrt{\pi}/n!2^{2n}$.

Straightforwardly, one can verify that asymptotic densities of the B–OU process have the form

$$\rho_*(r) = \frac{2}{\Gamma(n/2)} r^{n-1} \exp(-r^2). \tag{22}$$

A complementary check amounts to observing that the forward drift $b(r)$ of the stationary B–OU process needs to obey $\partial_t \rho_* = (1/2)\Delta\rho_* - \nabla(b\rho_*) = 0$. The invariant (asymptotic) density reads

$$\rho_*(r) = \frac{1}{Z}\exp(-V) \tag{23}$$

with the normalization $Z = \int_{R^+} \exp(-V)\mathrm{d}r$. We have

$$V = V(r) = \frac{1}{2}[r^2 - (n-1)\ln r] \tag{24}$$

and

$$b(r) = \frac{1}{2}\nabla\ln\rho_*(r) = -\nabla V = \frac{n-1}{2r} - r. \tag{25}$$

After normalizing the mean, $\langle R \rangle = 1$, and replacing r by s we readily arrive at the previous RMT spacing formulae.

The Shannon entropy of the continuous probability distribution (B–OU family) (Eq. (22)) reads [3]:

$$\mathcal{S}(\rho_*) = \ln\Gamma\left(\frac{n}{2}\right) - \frac{n-1}{2}\psi\left(\frac{n}{2}\right) + \frac{n-1}{2}, \tag{26}$$

where for half-integer values, the digamma function ψ equals

$$\psi\left(n+\frac{1}{2}\right) = -\gamma - 2\ln 2 + \sum_{k=1}^{n}\frac{2}{2k-1}. \tag{27}$$

For the Gaussian on R^+, i.e. $\rho_*(r) = (2/\sqrt{\pi})\exp(-r^2)$, we have $\mathcal{S}(\rho_*) = (1/2)\ln\pi$. It is useful to reproduce the general Shannon entropy formula for the Gaussian on R^+:

$$\rho(r) = [2/\pi\sigma^2]^{1/2}\exp\left(-r^2/2\sigma^2\right) \Longrightarrow \mathcal{S}(\rho) = (1/2)[\ln(\sigma^2\pi/2) + 1]. \tag{28}$$

2.4. Calogero model

For general stationary diffusion processes, a formula relating forward drifts $b(x)$ of the stochastic process with potentials $\mathcal{V}(x)$ of an auxiliary conservative Hamiltonian system reads [6, 7] (we choose a diffusion coefficient to be equal to $\frac{1}{2}$, hence scaling away $\hbar$ and m):

$$\mathcal{V}(x) = \frac{1}{2}(b^2 + \nabla\cdot b). \tag{29}$$

Upon substituting

$$b(x) = \frac{\beta}{2x} - x \tag{30}$$

with $\beta = n - 1$, we arrive at

$$\mathcal{V}(x) = \frac{1}{2}\left[\frac{\beta(\beta-2)}{4x^2} + x^2\right] - \frac{1}{2}(\beta+1). \tag{31}$$

This potential function enters a standard definition of the one-particle Hamiltonian operator (no physical parameters):

$$H = -\frac{1}{2}\triangle + \mathcal{V}(x), \tag{32}$$

where $\triangle = \frac{\mathrm{d}^2}{\mathrm{d}x^2}$. The energy operator H, with the previously introduced $\mathcal{V}(x)$, is an equivalent form of a two-particle (actually two-interacting-levels) version of the Calogero–Moser Hamiltonian [1, 7, 9].

The classic Calogero-type problem is defined by

$$H = -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}x^2 + \frac{\beta(\beta-2)}{8x^2} \tag{33}$$

with the well-known spectral solution

$$E_k(\beta) = 2k + 1 + \frac{1}{2}[1 + \beta(\beta-2)]^{1/2}, \tag{34}$$

where $k \geq 0$ and $\beta > -1$. By substituting $\beta = 1, 2, 3, 4$ we easily check that $E_0(\beta) = \frac{1}{2}(\beta+1)$.

All previously considered $n = 2, 3, 4, 5$ radial diffusion processes correspond to Calogero–Moser potentials and thence Calogero operators in the (renormalized) form $H - E_0$, where E_0 is the respective (fix n) ground state ($k = 0$) eigenvalue. These stochastic processes arise as the so-called ground state processes associated with the Calogero Hamiltonians. Let us note: we are aware of all the "fictitious time" of Dyson's model philosophy: if ψ_0 is the ground state wave function, we regard $\rho_* \doteq |\psi_0|^2$ as an invariant probability density of the stochastic B–OU process. Let us recall that the classic Ornstein–Uhlenbeck process can be regarded as the ground state process of the harmonic oscillator Hamiltonian operator.

2.5. General comments on the quantum Calogero system

The Calogero singular quantum mechanical Hamiltonian

$$H = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x^2 + \frac{\gamma}{x^2} \tag{35}$$

has the eigenvalues

$$E_n = 4n + 2 + (1 + 4\gamma)^{1/2},$$

where $n \geq 0$ and $\gamma > -\frac{1}{4}$. The eigenfunctions have the form

$$f_n(x) = x^{(2\alpha+1)/2}\exp\left(-\frac{x^2}{2}\right)L_n^\alpha(x^2)$$

with $\alpha = \frac{1}{2}(1 + 4\gamma)^{1/2}$ and

$$L_n^\alpha(x^2) = \sum_{\nu=0}^{n}\frac{(n+\alpha)!}{(n-\nu)!(\alpha+\nu)!}\frac{(-x^2)^\nu}{\nu!}. \tag{36}$$

The γ parameter range $-1/4 < \gamma < 3/4$ involves some mathematical subtleties

concerning the singularity at 0, which is not sufficiently severe to enforce the Dirichlet boundary condition.

In the range $\gamma \geq 3/4$ we deal with a double degeneracy of the ground state and of the eigenspace of the self-adjoint operator H. The singularity at $x = 0$ completely decouples $(-\infty, 0)$ from $(0, +\infty)$, so that $L^2(-\infty, 0)$ and $L^2(0, +\infty)$ are the invariant subspaces for the unitary Schrödinger evolution $\exp(-iHt)$ generated by H. The related Schrödinger probability current vanishes at $x = 0$ for all times and there is no dynamically implemented communication between those two areas, cf. [10]. The respective localization probabilities, to find a particle on a positive or negative semi-axis, are constants of motion. Because of the singularity at 0, once trapped, a particle is confined in one particular enclosure only and then cannot be detected in another one.

The (positive semi-axis) projection operator P_+ defined by $(P_+ f)(x) = \chi_{R^+}(x) f(x)$ commutes with H. It is thus tempting and (with suitable precautions) legitimate to confine the discussion to R^+ (or R^-) separately. However, we cannot tell here about two disjoint quantum problems defined respectively on R^+ and R^-. We deal with a single quantum mechanical system, though technically — with a degenerate ground state.

Let us also point out that $D(H)$ contains functions restricted to obey $f(0) = 0 = f'(0)$ and not necessarily to vanish on any of half-lines. Such functions may have support on both, positive and negative semi-axes simultaneously, excluding the origin 0. For example, a normalized linear combination (standard superposition) of the two components of the degenerate ground state of H, is a legitimate element of $D(H)$. There is no mixture here. Are they very special Schrödinger cat states? — good lurking-place for the cat-metaphysics?

2.6. Dyson's asymptotic equilibrium

Let M be a Hermitian $n \times n$ matrix with an orthogonal, unitary or symplectic invariance built-in. Then, the number of independent matrix elements equals, respectively, $N = n + \frac{1}{2}n(n-1)\beta$, $\beta = 1, 2, 4$. We introduce a Gaussian matrix ensemble: independent matrix elements are interpreted as independent Gaussian random variables with zero mean and variance [11]:

$$\mathrm{Var}(M_{ij}) = \frac{a^2}{2\beta}(1 + \delta_{ij}). \tag{37}$$

The probability density reads

$$P_*(M_1, \ldots, M_N) = c \exp(-\beta \mathrm{Tr}(MM^*)/2a^2). \tag{38}$$

The Gaussian RM joint density of (real) eigenvalues has the form

$$\Lambda_*(x_1, x_2, \ldots x_n) = C \cdot \left(\prod_{i<j} |x_i - x_j|^\beta \right) \exp\left(-\beta \left(\sum_i x_i^2 \right) / 2a^2 \right)$$

$$= C \exp\left(-\beta\left(-\sum_{i<j}\ln|x_i - x_j| + \sum_i \frac{x_i^2}{2a^2}\right)\right), \tag{39}$$

i.e. (cf. Eqs. (23) and (24))

$$\Lambda_* = \frac{1}{Z}\exp(-\beta V),$$

$$V = V(x_1, x_2, \ldots, x_n) = -\sum_{i<j}\ln|x_i - x_j| + \sum_i \frac{x_i^2}{2a^2}. \tag{40}$$

The above observations can be inferred by passing to suitable reduced densities, in the asymptotic limit of the Smoluchowski diffusion equation for the time-dependent probability density $P(M_1, M_2, \ldots, M_N, t)$:

$$\partial_t P = \sum_{M=M_{ij}}\left[\frac{1+\delta_{ij}}{\beta\nu}\frac{\partial^2}{\partial M^2}P + \frac{1}{a^2}\frac{\partial}{\partial M}(M \cdot P)\right]. \tag{41}$$

Let us fix an initial $t = 0$ condition to be M', ν is an auxiliary "friction" coefficient. We deal with independent Ornstein–Uhlenbeck processes for random matrix elements

$$P(M, t) = c(1 - q^2)^{-N/2}\exp\left(-\beta\frac{\text{Tr}(M - qM')^2}{2a^2(1 - q^2)}\right) \to P_*(M), \tag{42}$$

$$q = \exp\left(-t/a^2\nu\right),$$

which, in turn, induces the corresponding interacting Brownian motions for the eigenvalues. P_* and Λ_* stand for invariant asymptotic (unique stationary, equilibrium) densities of these processes.

Let us notice that choosing $a^2 = \beta n$, we can write down a corresponding set of stochastic differential equations (infinitesimal increments) for the interacting Brownian motions associated with the $n \times n$ random matrix. Forward drifts read $b_j(\lambda, t) = -\nabla_j(\beta V)$ and thence

$$\mathrm{d}\lambda_j(t) = \left(-\frac{1}{2n}\lambda_j + \frac{\beta}{2}\sum_{i<j}\frac{1}{\lambda_j(t) - \lambda_i(t)}\right)\mathrm{d}t + \mathrm{d}W_j(t). \tag{43}$$

Their properties were studied in detail in the mathematical literature and despite the Coulomb (or centrifugal) singularity, the solution is known to be unique and non-explosive for all times and all n, including the $n \to \infty$ limit. The eigenvalues never cross and an initially given (time t_0) order $\lambda_1 < \lambda_2 < \ldots < \lambda_n$ of (real) eigenvalues is kept forever in the course of the diffusion process [12].

For clarity of discussion, let us illustrate a passage from P_* to spacing distributions $P(s)$ for 2×2 random matrices. Namely, given the $n = 2$ form of Eq. (24):

$$P(M_1, \ldots, M_N) = c\exp\left(-\frac{\beta}{4}(E_1^2 + E_2^2)\right) \tag{44}$$

and

$$\Lambda_*(\lambda_1, \lambda_2) = C|\lambda_1 - \lambda_2| \exp\left(-\frac{\beta}{4}(\lambda_1^2 + \lambda_2^2)\right) \tag{45}$$

yield the spacing distribution as a reduced density

$$P(s) = \mathrm{const} \int \mathrm{d}\lambda_1 \int \mathrm{d}\lambda_2 \delta(s - |\lambda_2 - \lambda_1|)\Lambda_*(\lambda_1, \lambda_2), \tag{46}$$

which upon securing the normalization of $P(s)$ on R^+ and $\langle s \rangle = 1$, gives rise to the generic RMT-spacing densities.

3. Extremum principles: what can we say about the degree of randomization?

3.1. Min/max information entropy principle in the RMT

The information-theory route, according to Balian [13, 2], begins as follows. We consider a constrained extremum for a functional of a convex function

$$I = \int \mathrm{d}\mu(M)P(M)\ln P(M) \tag{47}$$

with constraints $\int \mathrm{d}\mu(M)P(M) = 1$ and

$$\langle \mathrm{Tr}(MM^*) \rangle = \int \mathrm{d}\mu(M)\mathrm{Tr}(MM^*)P(M) = Na^2. \tag{48}$$

Next, one passes to an extended information functional with Lagrange multipliers b, d:

$$I(b, d) = \int \mathrm{d}\mu(M)[P(M)\ln P(M) + bP(M) + d\mathrm{Tr}(MM^*)P(M)], \tag{49}$$

and looks for an extremum of I (actually for a minimum information measure, cf. convexity property), under the imposed constraints. The outcome is

$$P_*(M) = \exp(-(1 + b + d\mathrm{Tr}(MM^*))) \tag{50}$$

with $d = 1/2a^2$ and $\exp(-(1 + a)) = c = (1/2\pi a^2)^{N/2}$. We have thus arrived at the invariant (Gaussian) probability measure $P_*(M)$ for the Gaussian matrix ensemble, cf. also [1, 14–16].

3.2. Maximum randomness issue: constrained extremum

To have a better insight into the extremum principles at work, let us recall the standard maximum (information/Shannon) entropy principle: consider $[a, b] \in R$, assume that everything you know about the a priori unknown probability measure are (possibly) its moments

$$\int_a^b x^k \rho(x)\mathrm{d}x = m_k \tag{51}$$

with $k = 0, 1, \ldots, M$ and $m_0 = 1$ — the normalization condition.

We look for densities that maximize the Shannon entropy of a continuous probability distribution (now we encounter a functional of a concave function):

$$\mathcal{S}[\rho] = -\int_a^b \rho \ln \rho \mathrm{d}x \tag{52}$$

under the constraint of M fixed moments [5, 17].

The extremum of a functional

$$\tilde{\mathcal{S}} = -\int_a^b \rho \ln \rho \, \mathrm{d}x + \sum_0^M \lambda_k \left(\int_a^b x^k \rho \, \mathrm{d}x - m_k \right) \tag{53}$$

(a concavity property of $\mathcal{S}$ needs to be remembered) sets the functional form of ρ which maximizes the entropy

$$\rho_*(x) = C \exp\left(-\sum_a^b \lambda_k x^k \right), \tag{54}$$

where $C = \exp(-\lambda_0 - 1)$ is the normalization constant and λ_k's are fixed by identities

$$\int_a^b x^k \exp\left(-\sum_a^b \lambda^k x^k \right) \mathrm{d}x = m_k. \tag{55}$$

If there is a unique solution in terms of $\lambda_1, \ldots, \lambda_M$, we say that an entropy maximizing (under the m_k "circumstances") density does exist.

For reference, let us reproduce some pieces of a standard wisdom:

(i) If a and b are finite, there exists a unique maximum entropy density.

(ii) In R^+, e.g. $[0, +\infty)$, a maximizing density exists if $m_1^2 \leq m_2 \leq 2m_1^2$.

Notes: if there is no constraint, there is no maximizing density; if only the mean $m_1 = 1/\alpha$ is given, we get the exponential one: $\rho_*(x) = \alpha \exp(-\alpha x)$; for the Gaussian on R^+, like e.g. $\rho(r) = (2/\sqrt{\pi}) \exp(-r^2)$, we have $\mathcal{S}(\rho) = (\ln \pi + 1)/2$, which is a maximum of the Shannon entropy under the moment constraints $m_1 = \langle r \rangle = 1/\sqrt{\pi}$ and $m_2 = \langle r^2 \rangle = 1/2$; for another Gaussian on R^+, $P_0(s) = (2/\pi) \exp(-s^2/\pi)$, we have $m_1 = \langle s \rangle = 1$, $m_2 = \pi/2$ and $\mathcal{S}(\rho) = [\ln(\pi^2/4) + 1]/2$.

(iii) In R, with no moment prescribed, or given the mean only, there is no maximum entropy density.

Notes: if m_1 and m_2 are given, the maximum entropy distribution is the normal (Gaussian) one, with the variance $\sigma^2 = m_2 - m_1^2$, i.e. $\rho(x) = \exp(-(x - m_1)^2/2\sigma^2)/(\sqrt{2\pi}\sigma)$ and the Shannon entropy value is $\mathcal{S}(\rho) = \ln(2\pi e \sigma^2)/2$. That is to be compared with the previous outcome, Eq. (28), for the Gaussian on R^+.

3.3. Extremum principles: statement of purpose

Our RMT/Bessel–OU (spacing) distributions fall into the above category (ii), where an absolute (least constrained) entropy maximum on R^+ is set by the exponential density. With quite a variety of probability densities on R^+ in hands, can we quantify their "randomness level" [18] absolutely, or relative to any of reference densities?

Possible indications towards this end read as follows:

(*) compare absolute values of the respective Shannon entropies for densities on R^+ (Cover–Thomas differential entropy tables);

(**) invoke Kullback–Leibler relative entropies and look for a minimum of the relative entropy; relative with respect to the chosen reference density;

(***) look for min/max principles (like e.g. Helmholtz free energy properties) that govern the time evolution of standard diffusion-type processes (it is immaterial whether the time label is real or "fictitious");

(****) investigate the "organization level" of ground state densities of relevant quantum systems in comparison to all possible eigenstate-related probability densities. What about their "randomization" behavior with the growth of energy eigenvalues?

In fact, we have established before direct comparison tools, e.g. Shannon entropy values for the Poissonian and generalized Gaussian (B–OU, with a polynomial repulsion factor) densities, cf. Eqs. (15), (26), and (28). Hence the above point (*) has received due attention.

3.4. Kullback relative entropy route

Concerning the point (**), let us define the relative entropy for densities on R^+:

$$I(\rho : \rho_{\mathrm{ref}}) = \int \rho \ln \frac{\rho}{\rho_{\mathrm{ref}}} \mathrm{d}x \qquad (56)$$

with the a priori prescribed ρ_{ref}.

Let us investigate [4] a minimum of the Kullback entropy $I(\rho : \rho_{\mathrm{ref}})$ under the constraint imposed with the aid of an auxiliary function $T(x)$:

$$\int T(x)\rho(x)\mathrm{d}x = \theta = \mathrm{const.} \qquad (57)$$

A conditional extremum (minimum, in view of the convexity property) of a functional

$$\tilde{I} = \int \left[\rho \ln \frac{\rho}{\rho_{\mathrm{ref}}} + \lambda T(x)\rho + \lambda_0 \rho \right) \mathrm{d}x \qquad (58)$$

is reached at $\rho \to \rho_*$ with the constraint $\int T(x)\rho_*(x)\mathrm{d}x = \theta = \mathrm{const}$:

$$\rho_*(x) = C\rho_{\mathrm{ref}}(x) \exp\left(-\lambda T(x)\right). \qquad (59)$$

We need an overall normalization of ρ_* on R^+. It is thus useful to demand that we have fixed an integral

$$\frac{1}{C(\lambda)} = \int \rho_{\mathrm{ref}}(x) \exp\left(-\lambda T(x)\right) \mathrm{d}x. \qquad (60)$$

Our guess: choose $\rho_{\mathrm{ref}}(x)$ and $T(x)$, next adjust the values of θ and $C(\lambda)$ to fit either of $\lambda = 1, 2, 3 \ldots$

Explicit examples are:

(i) Exponential family (semi-Poisson case) on R^+:

$$\rho_{\mathrm{ref}} = \alpha \exp(-\alpha x), \qquad T(x) = -\ln x. \qquad (61)$$

(ii) Gaussian family (Wigner surmise) on R^+:

$$\rho_{\mathrm{ref}} = (1/\sqrt{\pi}) \exp(-x^2), \qquad T(x) = -\ln x. \qquad (62)$$

The only delicate computational issue is the constraint (Eq. (57)), where logarithmic integrations are to be carried out. The pertinent (reference) integrals read

$$\int_0^\infty \exp(-\alpha x)\ln x \mathrm{d}x = -\frac{1}{\alpha}(\gamma + \ln\alpha), \tag{63}$$

and

$$\int_0^\infty \exp(-\alpha x^2)\ln x \mathrm{d}x = -\left(\frac{\pi}{16\alpha}\right)^{1/2}[\gamma + \ln(4\alpha)], \tag{64}$$

where γ is the Euler–Mascheroni constant, $\alpha > 0$.

3.5. Thermodynamical extremum principles in the Smoluchowski processes

Now we can pass to the point (***) and discuss the role of the Helmholtz extremum principle which often takes the role of more familiar min/max entropy principle in random motion, cf. [6, 19, 20]. Given a probability density $\rho(x,t)$ solving the F–P equation

$$\partial_t \rho = D\triangle\rho - \nabla\cdot(b\rho). \tag{65}$$

We introduce $u = D\ln\rho$ and $v = b - u$, which obeys $\partial_t\rho = -\nabla(\rho v)$.

The Shannon entropy of ρ:

$$\mathcal{S}(t) = -\langle\ln\rho\rangle \tag{66}$$

typically is not a conserved quantity. We impose boundary restrictions that ρ, $v\rho$, $b\rho$ vanish at spatial infinities or other integration interval borders. We consider

$$D\dot{\mathcal{S}} = \langle v^2\rangle - \langle b\cdot v\rangle. \tag{67}$$

We may pass to time-independent drift fields and set $b = f/m\beta$, $j \doteq v\rho$, $f = -\nabla V$ plus $D = k_\mathrm{B}T/m\beta$. Then

$$\dot{\mathcal{S}} = \dot{\mathcal{S}}_\mathrm{int} + \dot{\mathcal{S}}_\mathrm{ext}, \tag{68}$$

where

$$k_\mathrm{B}T\dot{\mathcal{S}}_\mathrm{int} \doteq m\beta\langle v^2\rangle \geq 0 \tag{69}$$

stands for the entropy production rate, while

$$k_\mathrm{B}T\dot{\mathcal{S}}_\mathrm{ext} = \dot{\mathcal{Q}} = -\int f\cdot j\mathrm{d}x - m\beta\langle b\cdot v\rangle \tag{70}$$

(as long as negative which is not a must) may be interpreted as the heat dissipation rate: that in view of $\dot{\mathcal{Q}} = -\int f\cdot j\mathrm{d}x$.

Let us notice that because of $T\dot{S} \doteq k_\mathrm{B}T\dot{\mathcal{S}}$ we do have

$$T\dot{S}_\mathrm{int} = T\dot{S} - \dot{Q} \geq 0 \Rightarrow T\dot{S} \geq \dot{Q}. \tag{71}$$

In view of $j = \rho v = \frac{\rho}{m\beta}(f - k_\mathrm{B}T\nabla\ln\rho) \doteq -\frac{\rho}{m\beta}\nabla\Psi$, i.e. $v = -(1/m\beta)\nabla\Psi$ and $f = -\nabla V$, we can introduce

$$\Psi = V + k_\mathrm{B}T\ln\rho, \tag{72}$$

whose mean value stands for the Helmholtz free energy of the random motion

$$F \doteq \langle\Psi\rangle = U - TS. \tag{73}$$

Here $S \doteq k_\mathrm{B}\mathcal{S}$ and an internal energy is $U = \langle V\rangle$. Since we assume ρ and $\rho V v$ to vanish at the integration volume boundaries, we get

$$\dot{F} = \dot{Q} - T\dot{S} = -(m\beta)\langle v^2 \rangle = -k_{\mathrm{B}}T\dot{S}_{\mathrm{int}} \le 0. \tag{74}$$

Clearly, F decreases as a function of time towards its minimum, or remains constant.

Let us consider the stationary regime $\dot{S} = 0$ associated with an invariant density ρ_*. Then

$$b = u = D\nabla \ln \rho_*$$

and

$$-(1/k_{\mathrm{B}}T)\nabla V = \nabla \ln \rho_* \implies \rho_* = \frac{1}{Z}\exp(-V/k_{\mathrm{B}}T). \tag{75}$$

Hence

$$\Psi_* = V + k_{\mathrm{B}}T \ln \rho_* \implies \langle \Psi_* \rangle - k_{\mathrm{B}}T \ln Z \doteq F_* \tag{76}$$

with $Z = \int \exp(-V/k_{\mathrm{B}}T)\mathrm{d}x$. F_* stands for a minimum of the Helmholtz free energy F. Because of

$$Z = \exp(-F_*/k_{\mathrm{B}}T), \tag{77}$$

we have

$$\rho_* = \exp((F_* - V)/k_{\mathrm{B}}T). \tag{78}$$

Therefore, the conditional Kullback–Leibler entropy, of the density ρ relative to an equilibrium density ρ_* acquires the form

$$k_{\mathrm{B}}T\mathcal{H}_{\mathrm{c}} \doteq -k_{\mathrm{B}}T \int \rho \ln\left(\frac{\rho}{\rho_*}\right)\mathrm{d}x = F_* - F. \tag{79}$$

In view of the concavity property of the function $f(w) = -w \ln w$, $\mathcal{H}_{\mathrm{c}}$ takes only negative values, with a maximum at 0. We have $F_* \le F$ and $k_{\mathrm{B}}T\dot{\mathcal{H}}_{\mathrm{c}} = -\dot{F} \ge 0$. $\mathcal{H}_{\mathrm{c}}$ is bound to grow monotonically towards 0, while F drops down to its minimum F_* which is reached for ρ_* of Eqs. (75) and (78).

3.6. Shannon entropy in quantum systems

Presently, we pass to the point (****) raised before in Sect. 3.3. For probability distributions $p(x)$ on R, with any finite mean value, whose variance is fixed at the prescribed value σ^2, we have $S(p) \le \frac{1}{2}\ln(2\pi e\sigma^2)$. $S(p)$ becomes maximized if and only if p is a Gaussian with that variance.

Given an $L^2(R)$-normalized function $\psi(x)$, $(\mathcal{F}\psi)(p)$ is its Fourier transform. The corresponding probability densities: $\rho(x) = |\psi(x)|^2$ and $\tilde{\rho}(p) = |(\mathcal{F}\psi)(p)|^2$ give rise to position and momentum information (differential, e.g. Shannon) entropies

$$\mathcal{S}(\rho) \doteq S_q = -\langle \ln \rho \rangle = -\int \rho(x)\ln \rho(x)\mathrm{d}x,$$

$$\mathcal{S}(\tilde{\rho}) \doteq S_p = -\langle \ln \tilde{\rho} \rangle = -\int \tilde{\rho}(p)\ln \tilde{\rho}(p)\mathrm{d}p. \tag{80}$$

For the sake of clarity, we use dimensionless quantities, although there exists a consistent procedure for handling dimensional quantities in the Shannon entropy definition, cf. Eq. (8) [6, 20]. We assume both entropies to take finite values to

yield an entropic uncertainty relation [21, 22]:

$$S_q + S_p \geq (1 + \ln \pi). \tag{81}$$

If we define the squared standard deviation value for an observable A in a pure state ψ as $(\Delta A)^2 = (\psi, [A - \langle A \rangle]^2 \psi)$ with $\langle A \rangle = (\psi, A\psi)$, then for the position X and momentum P operators we have ($\hbar \equiv 1$):

$$\Delta X \cdot \Delta P \geq \frac{1}{2\pi e} \exp\left(\mathcal{S}(\rho) + \mathcal{S}(\tilde{\rho})\right) \geq \frac{1}{2}. \tag{82}$$

After the Fourier transformation, taking into account the entropic uncertainty relation, we have (σ^2 stands for the variance)

$$4\tilde{\sigma}^2 \geq 2(e\pi)^{-1} \exp\left(-2\langle \ln \tilde{\rho} \rangle\right) \geq (2e\pi) \exp\left(2\langle \ln \rho \rangle\right) \geq \sigma^{-2}. \tag{83}$$

For the momentum operator P that is conjugate to the position operator X in the adopted dimensional convention $\hbar \equiv 1$. Setting $P = -\mathrm{id}/\mathrm{d}x$ and presuming that all averages are finite, we get: $[\langle P^2 \rangle - \langle P \rangle^2] = (\Delta P)^2 = \tilde{\sigma}^2$. The standard indeterminacy relationship $\sigma \cdot \tilde{\sigma} \geq 1/2$ follows. In the stationary state $\langle P \rangle = 0$.

There is a subtlety to be mentioned, if we wish to extend the above reasoning to Calogero-type models [10]. We have *effectively* considered them in $L^2(R^+)$ instead of $L^2(R)$, to establish links with RMT-spacing distributions. Such restriction of the degenerate quantum model to its non-degenerate projection on R^+ is not an innocent step in the fully-fledged quantum formalism. For example, there is no standard momentum observable (self-adjoint operator of the form $-\mathrm{i}\hbar\partial_x$) on $L^2(R^+)$ alone.

A discussion of this issue can be found in Ref. [10]. We point out that the degeneracy-induced $R = R^- \oplus R^+$ decomposition of $L^2(R)$ into $L^2(R^-) \oplus L^2(R^+)$, for the Calogero quantum system, makes legitimate the usage of the standard momentum observable. To this end, the sufficient and necessary condition is that the Hamiltonian H decomposes as well $H = H_- \oplus H_+$ where H_- and H_+ are self-adjoint on their domains in $L^2(R^-)$ and $L^2(R^+)$, respectively. We have here fulfilled the conditions for a permanent dynamical confinement on $L^2(R^+)$, as secured by H_+, set in coexistence with the standard $L^2(R)$ momentum observable.

Another possible suggestion is to accept that a momentum-type operator needs not to be self-adjoint but merely symmetric. The uncertainty relations are known to hold true in this case (by the way there are many mathematical references to closely related issue of the time-frequency indeterminacy).

The Shannon entropy in the position space, and the so-called Leipnik entropy, which coincides with the sum $S_q + S_p$, were investigated numerically for excited states of various quantum systems [23–25]. The typical observation is that the entropy values, evaluated for the ground state probability densities, are minimal. An explicit check has never been made for the Calogero system, but we expect this generic pattern of behavior to be respected.

References

[1] F. Haake, *Quantum Signatures of Chaos*, Springer-Verlag, Berlin 2000.

[2] M.L. Mehta, *Random Matrices*, Academic Press, Boston 1991.

[3] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley, New York 1991.

[4] S. Kullback, *Information Theory and Statistics*, Wiley, New York 1959.

[5] K. Sobczyk, *Mech. Syst. Sign. Proc.* **15**, 475 (2001).

[6] P. Garbaczewski, *J. Stat. Phys.* **123**, 315 (2006).

[7] P. Garbaczewski, *Phys. Lett. A* **299**, 447 (2002).

[8] S. Karlin, M.H. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York 1981.

[9] F. Calogero, *J. Math. Phys.* **10**, 2191 (1969).

[10] P. Garbaczewski, W. Karwowski, *Am. J. Phys.* **72**, 924 (2004).

[11] F.J. Dyson, *J. Math. Phys.* **3**, 1191 (1962).

[12] L.C.G. Rogers, Z. Shi, *Probab. Theory Relat. Fields* **95**, 555 (1993).

[13] R. Balian, *Nuovo Cimento B* **57**, 183 (1968).

[14] T. Yukawa, *Phys. Rev. Lett.* **54**, 1883 (1985).

[15] J.Z. Ma, H. Hasegawa, *Z. Phys. B* **93**, 529 (1994).

[16] K.A. Muttalib, *Phys. Rev. Lett.* **65**, 745 (1990).

[17] L.R. Mead, N. Papanicolau, *J. Math. Phys.* **25**, 2404 (1984).

[18] P. Garbaczewski, *Acta Phys. Pol. B* **33**, 1001 (2002).

[19] M.C. Mackey, M. Tyran-Kamińska, *Physica A* **365**, 360 (2006).

[20] P. Garbaczewski, *Appl. Math. Information Sci.* **1**, 1 (2007).

[21] A.J. Stam, *Inf. Control,* **2**, 101 (1959).

[22] I. Białynicki-Birula, J. Mycielski, *Commun. Math. Phys.* **44**, 129 (1975).

[23] R.J. Yañez, W. Van Assche, J.S. Dehesa, *Phys. Rev. A* **50**, 3065 (1994).

[24] R.J. Yañez, W. Van Assche, R. González-Férez, Jesús S. Dehesa, *J. Math. Phys.* **40**, 5675 (1999).

[25] S.E. Massen, *Phys. Rev. C* **67**, 014314 (2003).