International Journal of Theoretical Physics, Vol. 37, No. 3, 1998

Quantum Mechanics Based on Probability Wave Functions Induced by the Minimum Mean Deviation from Statistical Equilibrium. I

Silviu Guiasu[1]

Received July 28, 1997

The maximum entropy probability distributions are generally used as probabilistic models for describing statistical equilibrium subject to given mean values of some random variables. The paper deals with the construction of probability wave functions by minimizing the mean deviation from statistical equilibrium subject to generalized correlations, induced by random fluctuations, whose values are determined looking for stationary points of the mean energy of the system. The results are applied to the study of several quantum systems (the free particle in a box, two independent particles in a box, the harmonic oscillator, and the hydrogen atom, without having to solve the corresponding Schrödinger equation.

1. INTRODUCTION

The Schrödinger equation plays a central role in nonrelativistic quantum mechanics and some textbooks include it among the postulates of the theory. Introduced in 1926, its solution is the so-called ψ-function, initially thought of by Schrödinger to represent a real disturbance, a matter wave in space. He soon abandoned these matter waves and turned to an "electromagnetic interpretation" in which $\psi^*\psi$ would be a measure of the density of electric charge, or rather a sort of "weighting function." He wrote (Schrödinger, 1926, the fourth paper): "$\psi^*\psi$ is a sort of weight function in the configuration space of the system. The wave mechanical configuration of the system is a superposition of many, strictly speaking all, the kinematically possible point-mechanical configurations. Thereby every point-mechanical configuration contributes to the true wave mechanical configuration with a certain weight,

[1] Department of Mathematics and Statistics, York University, Toronto, Ontario M3J 1P3, Canada.

0020-7748/98/0300-95 7$15.00/0

which is given precisely by $\psi^*\psi$. If one likes paradoxes, one can say that the system is found simultaneously in all conceivable kinematic locations but not in all of them in 'equal strength'." It was Born (1926) who interpreted $\psi^*\psi$ as being a probability density. As mentioned by Moore (1989), for Schrödinger, $\psi^*\psi$ pictures a continuous cloud of 'something' (presumably charge and mass), whereas Born assumes point particles (electrons) within the atom and interprets $\psi^*\psi$ as being the probability density of finding a particle at a certain location. In a letter to Einstein (June 1946), Schrödinger wrote, "God knows I am no friend of the probability theory, I have hated it from the first moment when our dear friend Max Born gave it birth. For it could be seen how easy and simple it made everything, in principle, everything ironed out and the true problems concealed . . . And actually not a year passed before it became an official credo, and it still is." It seems today that in fact Schrödinger's objections rather referred to the interpretation of the concept of probability. A frequency interpretation of probability (if the result A of an experiment has the probability p, it means that if the experiment is repeated many times then the fraction of outcomes that give the result A approximately equals p) would suggest that Born's probability density $\psi^*\psi$ rather refers to a statistical ensemble of systems. To Schrödinger, adopting such a viewpoint "we cut ourselves off from ever applying rational probability considerations to a single event." On the other hand, a purely subjective interpretation of probability, apparently more acceptable to Schrödinger, would suggest that Born's probabilistic interpretation of the wave function ψ does not relate to a system, but to our knowledge about a system. Summarizing, the interpretation of the solution ψ of the Schrödinger equation is somewhat mysterious, but $|\psi|^2 = \psi^*\psi$ has a clear meaning: it is the probability density of the particle position at given time. Details about Max Born's statistical interpretation of quantum mechanics may be found in Pais (1982). In 1954 Born was awarded the Nobel Prize "for his fundamental research, especially for his statistical interpretation of the wave function."

As the ultimate use of the Schrödinger equation is to provide us with probabilistic models for the behavior of quantum systems, a natural question comes up: Is it possible to build such probabilistic models without having to write and solve the corresponding Schrödinger equation? The question is even more justified if we take into account that the Schrödinger equation may be solved exactly only for a very limited number of quantum systems, namely the free particle in a box, the harmonic oscillator, and the hydrogen atom.

There have been attempts to deduce the Schrödinger equation from several variational principles or from prior suppositions instead of taking it as a postulate of quantum mechanics. Here are some of them:

(a) As shown in Flügge (1974, Vol. 1, pp. 2–5), the Schrödinger equation may be obtained from the Euler equations corresponding to the minimization of the energy integral.

(b) Bohm (1984) suggested, "Instead of starting from Born's probability distribution $|\psi|^2$ as an absolute and final and unexplainable property of matter, we have [to show] how his property could come out of random motions originating in a subquantum mechanical level." Several papers (Nelson, 1985; Baublitz, 1988) have attempted to derive the Schrödinger equation from classical mechanics and an assumed Markov diffusion stochastic process induced by random fluctuations of a submicroscopic medium. The difficulty in accepting such an approach resides in the fact that in quantum mechanics it is not possible to assume that the sample space on which all random variables of the stochastic process are supposed to be defined remains unique and invariable.

(c) Frieden (1989, 1990, 1991) aimed at building up a probabilistic model based on the statistical estimation theory from which the Schrödinger equation could be derived as a consequence. Dealing with the position of a particle on the real line, for instance, the Schrödinger equation is obtained in this approach by minimizing a linear combination of the Fisher (1922) information, measuring the degree of ruggedness of a probability distribution, and the mean kinetic energy of the particle.

(d) Guiasu (1992) used the minimum $\overline{\chi}^2$ deviation from the maximum entropy probability distribution subject to given mean values and the quantization rules of the old quantum theory in order to derive the same conclusions as those induced by the corresponding Schrödinger equation for a free particle in a box, the harmonic oscillator, and the hydrogen atom. Other results in the same direction of thought may be found in Slater (1994) and Preda *et al.* (1995).

The objective of this paper is to combine the approaches (a) and (d), but without using the Euler equations from (a) and the old quantization rules from (d). The new approach obtained this way is simple and straightforward. The probabilistic model obtained yields the same conclusions as those induced by solving the Schrödinger equation for a free particle in a box, two interacting particles in a box, the harmonic oscillator, or the hydrogen atom. In a personal communication made in 1985 to Moore (1989, p. 416), H. W. Peng mentioned that Schrödinger, in spite of his much discussed reservations about the probabilistic interpretation of the undulatory mechanics, once had told him that "Quantum Mechanics was born in statistics and it will end in statistics." The objective of this paper is to show that Schrödinger was right when he made such a statement.

Our program may be summarized in one sentence: we are looking for a probabilistic model of stationary energy obtained by minimizing the mean

deviation from statistical equilibrium subject to given generalized correlations induced by random internal or/and external fluctuations. Let us explain each step:

1. *Statistical Equilibrium.* We start from statistical equilibrium described by a maximum entropy probability distribution subject to constraints induced by given mean values of some random variables. This is a problem well studied in the literature. Over 125 years ago, Boltzmann (1872) introduced the famous H-function in his research on the behavior of the molecules of a gas. Seventy-six years later, Shannon (1948) showed that the discrete analogue of Boltzmann's H-function, called the probabilistic entropy, may be accepted as a measure of the amount of uncertainty contained by any probability distribution. If the probability distribution is given, then its entropy is a number showing the amount of uncertainty contained by it. If, for instance, the set of possible outcomes is given and we know nothing more, then the entropy is maximum for the uniform distribution. The inverse problem, however, is much more important: if the probability distribution is unknown and the only information available consists of one or several mean values of one or several random variables, then from the infinite set of feasible probability distributions compatible with such given constraints, we choose the one that maximizes entropy. The solution of this variational problem is the most uncertain, or unbiased (i.e. ignoring no possibility) probability distribution compatible with the given constraints represented by mean values. It is the closest probability distribution to the uniform distribution when the given mean values are known because, as we have said before, the uniform distribution maximizes entropy when there are no constraints imposed. This variational problem is known as the maximum entropy principle (MEP) and it was explicitely formulated by Jaynes (1957) and implicitly by von Neumann (1932). Eventually, it was surprising to see that well-known probability distributions could be rediscovered as solutions of the MEP in a natural way. Thus, on the positive real axis $[0, +\infty)$, the solution of the MEP subject to the mean value μ is the exponential probability distribution $\mathbf{E}(\mu)$. On the real axis $(-\infty, +\infty)$, the solution of the MEP subject to the mean μ and the variance σ^2 is the normal distribution $\mathbf{N}(\mu, \sigma^2)$. On an arbitrary interval $[a, b]$ of the real axis, if we have no constraints attached, then the probability distribution of maximum entropy is the uniform distribution $\mathbf{U}(a, b)$. As shown in Guiasu (1990), practically all the main probability distributions may be obtained as solutions of constrained variational problems similar to MEP.

2. *Systems of Orthonormal Functions.* Once the probability distribution that describes statistical equilibrium is obtained, and u is its density, we choose a sequence of orthogonal functions with the weight u as a system of generalized coordinates. Such systems of orthogonal functions are the

Laguerre polynomials, for the exponential distribution $\mathbf{E}(\mu)$, the Hermite polynomials, for the normal distribution $\mathbf{N}(\mu, \sigma^2)$, the generalized Laguerre polynomials, for the gamma distribution $\mathbf{G}(1/\alpha, \beta + 1)$, and either the trigonometric system or the Legendre (spherical) polynomials, for the uniform distribution $\mathbf{U}(a, b)$, for instance.

3. *One-Dimensional Probability Wave Functions.* As long as the statistical equilibrium is not perturbed, the mean values of the generalized coordinates remain equal to zero. If, however, random perturbations induced by internal or external interactions alter the statistical equilibrium, then some of these mean values cease to be equal to zero. We minimize the mean deviation from statistical equilibrium subject to correlations between generalized coordinates induced by the random perturbations induced by internal or/and external interactions. The mean deviation is measured using Pearson's $\overline{\chi}^2$ indicator, which is a weighted Euclidean distance. By minimizing $\overline{\chi}^2$ subject to given mean values, we force the small initial probabilities to remain basically the same, focusing on changes induced by constraints on the probabilities of the most probable outcomes. In general, a wave is just any disturbance to a field. In our context, we introduce a probability wave function to be the minimum deviation from statistical equilibrium due to random internal or/and external fluctuations. Once the probability wave function χ is obtained, the normed square of its absolute value is used as a probability density function in the set of possible configurations. Bohm (1984) argued in favor of paying attention to random fluctuations in general and at the quantum level in particular: "It is not relevant where such fluctuations come from; all that is important is to assume that they exist and to see their effects." Certainly, at the quantum level the statistical equilibrium of a system may be easily affected by interactions with other quantum systems or by the measurement process performed by an external macroscopic observer. Minimizing the $\overline{\chi}^2$ indicator, a slight generalization of the well-known least square method of Legendre (1805) and Gauss (1821/1823), we want to determine the probability wave of minimum deviation from equilibrium subject to the mean effects of the perturbations or interactions, focusing on changes induced by the given constraints on the most probable outcomes. As mentioned by Büler (1981), "The method of least squares was one of Gauss's most efficient tools in his research. . . Least squares were Gauss's indispensable theoretical tool in experimental research; increasingly, he came to see it as the most important witness to the connection between mathematics and nature." The weighted least squares generalization gives more flexibility and makes a smooth connection between the Euclidean distance and the mean deviation.

4. *Multidimensional Probability Waves.* For incompatible entities (like position and momentum, for instance) it is senseless to talk about a joint probability distribution. For independent compatible entities (like the three components of the position of a particle in a three-dimensional Euclidean space, for instance), the joint probability density is simply the product of the probability densities (i.e., the marginals) of the corresponding entities. For dependent compatible entities (like the positions of several interacting particles, for instance) the joint probability distribution is not uniquely determined by the marginals and the partial information available about their dependence (like covariances, or correlations, for instance). We construct the multidimensional probability wave obtained by minimizing the mean $\overline{\chi}^2$ deviation from the direct independent product of the marginals subject to the given mixed moments (or generalized covariances). The solution of such a variational problem assumes nothing about the dependence among the components beyond the given mixed moments.

5. *Stationary Values of the Mean Energy.* Once the one- or multidimensional probability wave is obtained by solving the above-mentioned variational problem, we determine the stationary points of the corresponding mean energy. In this way we determine the stationary probability waves whose normed square may be used as a probability density for predicting the mean values of some observables of interest.

6. *Applications.* The formalism applied to the study of a free particle in a box, of two independent particles in a box, of the harmonic oscillator, and of the hydrogen atom. The second part of this paper will show that the formalism presented here may be extended to the study of quantum systems for which the corresponding Schrödinger equation cannot be solved exactly, such as the helium and lithium atoms, for instance.

2. THE MATHEMATICAL MODEL

Let D be a domain in the k-dimensional Euclidean space $\mathbf{R}^k$. We denote

$$\langle f \rangle = \int_D f(x)dx, \qquad \langle f^* g \rangle = \int_D f^*(x)g(x)dx, \qquad \|f\| = \langle f^* f \rangle^{1/2}$$

provided that the above integrals exist, where f^* means the complex conjugate of f.

2.1. Statistical Equilibrium

Long ago, Boltzmann (1872) defined the expression

$$\mathbf{E} = \int_0^{+\infty} f(x, t) \left\{ \ln \left[\frac{f(x, t)}{\sqrt{x}} \right] - 1 \right\} dx$$

and showed that it possessed the properties of a negative entropy: thus $-\mathbf{E}$ always increased monotonically with time t if the distribution function $f(x,t)$ deviated from the Maxwellian distribution $\sqrt{x} \exp\{-x/(kT)\}$, with x denoting the kinetic energy, T the absolute temperature, and k a positive constant; it stayed constant if the latter distribution was assumed. This result, i.e.,

$$-\frac{d\mathbf{E}}{dt} \geq 0$$

was later called the "H-theorem."

Let u be an arbitrary probability density function defined on $D \subseteq \mathbf{R}^1$. The amount of uncertainty contained by u is measured by the (absolute) entropy $\mathbf{H}(u) = -\langle u \ln u \rangle$, introduced by Shannon (1948) by analogy with Boltzmann's function from statistical mechanics. When u is given, then $\mathbf{H}(u)$ is a number. In many applications, however, u is not known and the only information available is provided by mean values of some random variables. In an everchanging world, the mean values seem to be the only kind of data accessible to us on which we can rely in building up a relatively stable model of reality. Generally, however, there are several (very often infinitely many) probability density functions u compatible with the given mean values. According to the principle of maximum entropy, used implicitly by von Neumann (1932) and formulated explicitly and generally by Jaynes (1957), we choose the probability density function u which maximizes $\mathbf{H}(u)$ subject to the constraints induced by the given mean values. Such a solution u is the most unbiased probability density function (i.e. the most uncertain, ignoring no possibility) compatible with the given constraints. The principle of maximum entropy is an objective criterion for constructing a subjective probabilistic model when some mean values of some random variables are known. A probability density function u, solution of the principle of maximum entropy, describes the statistical equilibrium corresponding to the given mean values used as constraints of this nonlinear variational problem. The following results are known in the literature and may be obtained without difficulty by applying the classic Lagrange multipliers method from the calculus of variation.

Proposition 1. If $D = [a, b]$, then the solution of the nonlinear program $\max_u \mathbf{H}(u)$ subject to $\langle u \rangle = 1$ is the uniform distribution $\mathbf{U}(a, b)$ with parameters a and b, whose density is

$$u(x) = \frac{1}{b-a}, \qquad x \in [a, b] \tag{1}$$

Proposition 2. If $D = [0, +\infty)$, then the solution of the nonlinear program $\max_u \mathbf{H}(u)$ subject to

$$\langle u \rangle = 1, \qquad \langle xu(x) \rangle = \mu$$

is the exponential distribution $\mathbf{E}(\mu)$ whose density is

$$u(x) = \frac{1}{\mu} e^{-x/\mu}, \qquad x \in [0, +\infty) \tag{2}$$

Proposition 3. If $D = (-\infty, +\infty)$, then the solution of the nonlinear program $\max_u \mathbf{H}(u)$ subject to

$$\langle u \rangle = 1, \qquad \langle xu(x) \rangle = \mu, \qquad \langle (x-\mu)^2 u(x) \rangle = \sigma^2$$

is the normal distribution $\mathbf{N}(\mu, \sigma^2)$ with parameters μ and σ^2, whose density is

$$u(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, \qquad x \in (-\infty, +\infty) \tag{3}$$

Let u be a probability density function and v a nonnegative function, both defined on $D \subseteq \mathbf{R}^1$, such that u is absolutely continuous with respect to v, which means that $u(x) = 0$ if $v(x) = 0$. The relative entropy of u with respect to the reference measure of density v [also called the Kullback–Leibler (1951) indicator, or the divergence of u with respect to v] is defined by

$$\mathbf{H}(u|v) = \left\langle u \ln \frac{u}{v} \right\rangle = -\langle u \ln v \rangle - \mathbf{H}(u)$$

provided that the integrals exist. We have $\mathbf{H}(u|v) \geq 0$ with equality if and only if $u = v$, u−almost everywhere. Obviously, if the reference function v is constant on D, i.e., $v(x) = c$ $(c \neq 1)$, then the relative entropy is just the negative absolute entropy up to an additive constant, namely

$$\mathbf{H}(u|v) = -\ln c - \mathbf{H}(u)$$

In particular, if $D = [a, b]$ and v is the probability density of the uniform distribution on $[a, b]$, then u is absolutely continuous with respect to v and

$$\mathbf{H}(u|v) = \ln(b-a) - \mathbf{H}(u)$$

which shows that $\mathbf{H}(u)$ measures how much the probability density u differs from the uniform distribution with respect to the logarithmic mean. According to the principle of minimum relative entropy, we determine the probability density u which is the closest one to the reference measure of density v

subject to given mean values of some random variables. Thus, the principle of maximum (absolute) entropy may be viewed as a special case of the principle of minimum relative entropy when the reference measure is just the uniform distribution. The next proposition (Guiasu (1990)) may be proved using the standard Lagrange multipliers method:

Proposition 4. If $D = [0, +\infty)$, and $v(x) = x^{\beta}$ $(\beta > -1)$, then the solution of the nonlinear program $\min_u \mathbf{H}(u|v)$ subject to

$$\langle u \rangle = 1, \qquad \langle xu(x) \rangle = \mu$$

is the gamma distribution $\mathbf{G}(1/\alpha, \beta + 1)$, with parameters α and $\beta + 1$, whose density is

$$u(x) = \frac{1}{\alpha^{\beta+1}\Gamma(\beta + 1)} x^{\beta} e^{-x/\alpha}, \qquad x \in [0, +\infty) \tag{4}$$

where $\alpha = \mu/(\beta + 1)$.

2.2. Systems of Orthonormal Functions

Let u be a probability density function on the set $D \subseteq \mathbf{R}^1$, and $\{U_n, n = 0, 1, \ldots\}$, $U_0 \equiv 1$, a sequence of orthonormal functions with the weight u, i.e.,

$$\langle U_n U_l u \rangle = \begin{cases} 1 & \text{if } l = n \\ 0 & \text{if } l \neq n \end{cases}$$

which implies that $\|U_n\| = 1$, for all n $(n = 0, 1, \ldots)$, with $\langle U_0 u \rangle = 1$, and $\langle U_n u \rangle = 0$ for all n $(n = 1,2,\ldots)$. We call $\{U_n, n = 0, 1, \ldots\}$ a system of generalized coordinates associated to the probability density function u.

The four probability density functions describing statistical equilibrium listed above have well-known corresponding systems of orthonormal functions:

(a) A system of orthonormal polynomials with the weight (1) is

$$U_n(x) = (2n + 1)^{1/2} P_n\left(\frac{2}{b - a} x - \frac{a + b}{b - a}\right) \qquad (n = 0, 1, \ldots) \tag{5}$$

where $P_n(x)$ is the Legendre (spherical) polynomial of degree n. The first ones are

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x), \qquad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

A system of orthonormal functions with the weight $u(x) = 1/a$ on $D = [0, a]$ is the trigonometric system

$$U_0 \equiv 1, \qquad U_n(x) = \sqrt{2}\sin\frac{n\pi}{a}x \qquad (n = 1, 2, \ldots) \tag{6}$$

A system of orthonormal functions with the weight $u(x) = 1/a$ on $D = [-a/2, a/2]$ is the trigonometric system

$$U_0(x) = 1, \qquad U_n(x) = \begin{cases} \sqrt{2}\sin\dfrac{n\pi}{a}x & \text{if } n \text{ is even} \\ \sqrt{2}\cos\dfrac{n\pi}{a}x & \text{if } n \text{ is odd} \end{cases}$$

for $n = 1, 2, \ldots$.

(b) A system of orthonormal polynomials with the weight (2) is

$$U_n(x) = L_n\left(\frac{x}{\mu}\right) \qquad (n = 0, 1, \ldots) \tag{7}$$

where $L_n(x)$ is the Laguerre polynomial of degree n. The first ones are

$$L_0(x) = 1, \qquad L_1(x) = -x + 1, \qquad L_2(x) = \frac{1}{2}(x^2 - 4x + 2)$$

$$L_3(x) = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6)$$

$$L_4(x) = \frac{1}{24}(x^4 - 16x^3 + 72x^2 - 96x + 24)$$

$$L_5(x) = \frac{1}{120}(-x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120)$$

(c) A system of orthonormal polynomials with the weight (3) is

$$U_n(x) = (n!)^{-1/2}He_n\left(\frac{x-\mu}{\sigma}\right), \qquad He_n(x) = 2^{-n/2}H_n\left(\frac{x}{\sqrt{2}}\right) \tag{8}$$

where $n = 0, 1, \ldots$ and $H_n(x)$ is the Hermite polynomial of degree n. The first ones are

$$H_0(x) = 1, \qquad H_1(x) = 2x, \qquad H_2(x) = 4x^2 - 2, \qquad H_3(x) = 8x^3 - 12x$$
$$H_4(x) = 16x^4 - 48x^2 + 12, \qquad H_5(x) = 32x^5 - 160x^3 + 120x$$

(d) A system of orthonormal polynomials with the weight (4) is

$$U_n(x) = \left(\frac{n!\,\Gamma(\beta + 1)}{\Gamma(\beta + n + 1)}\right)^{1/2} L_n^{(\beta)}\left(\frac{x}{\alpha}\right) \qquad (n = 0, 1, \ldots) \tag{9}$$

where $L_n^{(\beta)}(x)$ is the generalized Laguerre polynomial of degree n and order β. We have

$$L_n^{(k)}(x) = \sum_{l=0}^{n} (-1)^l \binom{n + k}{n - l} \frac{1}{l!} x^l, \qquad L_n(x) = L_n^{(0)}(x)$$

Some such polynomials are

$$L_n^{(k)}(0) = \binom{n + k}{n}; \qquad L_0^{(k)}(x) = 1; \qquad L_1^{(k)}(x) = 1 + k - x$$

$$L_2^{(1)}(x) = \frac{1}{2}(6 - 6x + x^2); \qquad L_3^{(1)}(x) = \frac{1}{6}(24 - 36x + 12x^2 - x^3)$$

$$L_4^{(1)}(x) = \frac{1}{24}(120 - 240x + 120x^2 - 20x^3 + x^4)$$

$$L_2^{(2)}(x) = \frac{1}{2}(12 - 8x + x^2); \qquad L_3^{(2)}(x) = \frac{1}{6}(60 - 60x + 15x^2 - x^3)$$

$$L_4^{(2)}(x) = \frac{1}{24}(360 - 480x + 180x^2 - 24x^3 + x^4)$$

The generalized Laguerre polynomials satisfy the following equalities:

$$L_n^{(\beta-1)}(x) = L_n^{(\beta)}(x) - L_{n-1}^{(\beta)}(x) \tag{10}$$

$$x \frac{d^2}{dx^2} L_n^{(\beta)}(x) + (\beta + 1 - x) \frac{d}{dx} L_n^{(\beta)}(x) + nL_n^{(\beta)}(x) = 0 \tag{11}$$

Details about the orthonormal polynomials mentioned above may be found in Abramowitz and Stegun (1972) or in Gradshteyn and Ryzhik (1980); they may be easily generated using the computer package MATHEMATICA (Wolfram, 1991).

2.3. One-Dimensional Probability Waves

Let u be a probability density on $D \subseteq \mathbf{R}^1$ and let $\{U_n, n = 0, 1, \ldots\}$, $U_0 \equiv 1$, be a sequence of orthonormal polynomials with the weight u. As

mentioned before, as long as nothing alters the statistical equilibrium described by the probability density function u, we have $\langle U_n u\rangle = 0$ ($n = 1, 2, \ldots$) and $\langle U_0 u\rangle = \langle u\rangle = 1$. Very often, however, random fluctuations alter such a statistical equilibrium and the probability density function u has to be replaced by another probability density function. If the sequence $\{U_n, n = 0, 1, \ldots\}$ is complete with respect to the weight u, then another probability density function f on D may be written as

$$f = u\left(1 + \sum_{n=1}^{+\infty} c_n U_n\right) \tag{12}$$

where $c_n = \langle U_n f\rangle$ is the U_n-moment of f or the mean fluctuation in the direction U_n, and the sum is taken with respect to the values of n ($n = 1, 2, \ldots$).

The weighted deviation of f from u is the *probability wave function*

$$\chi = \frac{f - u}{\sqrt{u}} = \sqrt{u} \sum_{n=1}^{+\infty} c_n U_n \tag{13}$$

The probability density function generated by the probability wave function of the minimum weighted deviation from u induced by the mean fluctuations $\{c_n, n = 1, 2, \ldots\}$ is

$$\left(\sum_{n=1}^{+\infty} c_n^2\right)^{-1} \chi^2$$

The elementary probability wave functions are

$$\psi_0 = \sqrt{u}, \qquad \chi_n = \sqrt{u} U_n \qquad (n = 1, 2, \ldots)$$

where ψ_0 is the ground probability wave function induced by the probability density function u, and χ_n is the elementary wave function of level n. Therefore the probability wave χ is a linear combination of the elementary probability wave functions, namely,

$$\chi = \sum_{n=1}^{+\infty} c_n \chi_n$$

The jump from level l to level n is defined by $\psi_{l,n} = \chi_n - \chi_l$. The elementary jumps are

$$\psi_n = \psi_{n-1,n} = \chi_n - \chi_{n-1}, \qquad \text{where} \qquad \chi_0 = \psi_0$$

Using the Lagrange multipliers method, we obtain:

Proposition 5. The solution of the quadratic program

$$\min_f \overline{\chi}^2 = \langle \chi^2 \rangle = \left\langle \left(\frac{f}{u} - 1 \right)^2 u \right\rangle$$

subject to

$$\langle U_n f \rangle = c_n \qquad (n = 1, 2, \ldots, N) \tag{14}$$

is

$$f = u\left(1 + \sum_{n=1}^{N} c_n U_n \right) \tag{15}$$

The least $\overline{\chi}^2$ technique is a weighted variant of the least squares method of Legendre and Gauss. Minimizing $\overline{\chi}^2$ subject to given mean values, we force the small initial probabilities to remain basically the same, focusing on changes induced by the constraints on the most probable outcomes.

Practically, c_n is estimated by calculating a confidence interval for a mean value. If $\{x_1, \ldots, x_M\}$ is a random sample of size N from the population to which both u and f refer, we calculate the sample mean

$$\overline{U}_n^{(M)} = \frac{1}{\mathrm{M}} [U_n(x_1) + \cdots + U_n(x_M)]$$

An $100(1 - \alpha)\%$ confidence interval for c_n is

$$\left(\overline{U}_n^{(M)} - t_{\alpha/2, M-1} \frac{s_n^{(M)}}{\sqrt{M}}, \overline{U}_n^{(M)} + t_{\alpha/2, M-1} \frac{s_n^{(M)}}{\sqrt{M}} \right)$$

where $t_{\alpha/2, M-1}$ is the critical point of the t-distribution with $M - 1$ degrees of freedom, corresponding to the significance level $\alpha/2$, and $s_n^{(M)}$ is the sample standard deviation

$$s_n^{(M)} = \frac{1}{M - 1} \sum_{k=1}^{M} \left[U_n(x_k) - \overline{U}_n^{(M)} \right]^2$$

Introducing (15) into (13), we obtain the one-dimensional probability wave

$$\chi = \sqrt{u} \sum_{n=1}^{N} c_n U_n \tag{16}$$

generated by the deviation from u due to the generalized moments (14).

Introducing $U_n(x)$ given by (5), (7), or (8) into (12) [or into (15), as an approximation of (12)] we obtain the closest probability density function [or an approximation of it when using (15)] to u given by (1), (2), or (3), respectively, when the generalized moments c_n are given, where closeness

is measured using Pearson's $\overline{\chi}^2$ indicator, expressed by the squared amplitude of the probability wave (13). As long as the statistical equilibrium described by the probability density function u remains unchanged, no probability wave is generated. When the occurrence of random fluctuations is detected by estimating the generalized moments (14), then the probability wave (16) shows the deviation from the statistical equilibrium described by u. An external observer can generally observe only a change in an equilibrium state.

2.4. Multidimensional Probability Wave

Let u and v be two probability density functions on $D_1 \subseteq \mathbf{R}^1$ and $D_2 \subseteq \mathbf{R}^1$, respectively, and let $\{U_n, n = 0, 1, \ldots\}$ and $\{V_l, l = 0, 1, \ldots\}$ ($U_0 \equiv 1$, $V_0 \equiv 1$) be two complete systems of orthonormal functions on D_1 and D_2 with the weights u and v, respectively. If there is independence between marginals, then the joint probability density on $D_1 \times D_2$ is simply the direct product uv. But what happens if there is interdependence between the two components?

Consider the system of functions $\{U_n V_l; n, l = 0, 1, \ldots\}$ on $D_1 \times D_2$ with the weight uv. A joint probability density function f on $D_1 \times D_2$ has the form

$$f = uv\left(1 + \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} c_{nl} U_n V_l\right) \tag{17}$$

Such a joint probability density function is the closest one, in the $\overline{\chi}^2$ sense, to the direct independent product uv subject to the generalized mixed moments (or correlations)

$$c_{nl} = \langle U_n V_l f \rangle$$

as shown by the next proposition, whose proof is a standard application of the calculus of variation.

Proposition 6. Let

$$\chi = \frac{f - uv}{\sqrt{uv}} = \left(\frac{f}{uv} - 1\right)\sqrt{uv} \tag{18}$$

be the weighted deviation of f from the independent direct product uv. The solution of the quadratic program

$$\min_f \overline{\chi}^2 = \langle \chi^2 \rangle$$

subject to

$$\langle U_n V_l f \rangle = c_{nl}; \qquad (n = 0, 1, \ldots, N; \quad l = 0, 1, \ldots, L; \quad (n, l) \neq (0, 0)) \tag{19}$$

is the density

$$f = uv \left(1 + \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl} U_n V_l \right) \tag{20}$$

Introducing (20) into (18), we obtain the two-dimensional probability wave

$$\chi = \sqrt{uv} \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl} U_n V_l \tag{21}$$

generated by the deviation from independence of u and v due to the generalized correlations (19). The elementary probability wave functions are

$$\psi_{00} = \sqrt{uv}, \qquad \chi_{nl} = \sqrt{uv}\, U_n V_l,$$
$$(n = 0, 1, \ldots; \quad l = 0, 1, \ldots; \quad (n, l) \neq (0, 0))$$

where ψ_{00} is the ground probability wave function and χ_{nl} is the elementary wave function of level (n, l). The jump from level (i, j) to level (k, l) is defined by $\psi_{ij,kl} = \chi_{kl} - \chi_{ij}$. Thus, the probability wave function is a linear combination of the elementary probability wave functions, namely

$$\chi = \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{+\infty} \sum_{l=0}^{+\infty} c_{nl} \chi_{nl}$$

The generalization to more than two components is straightforward.

2.5 Stationary Values of the Mean Energy

If χ is a probability wave function and $\hat{A}$ an operator, then the mean value of $\hat{A}$ is defined by

$$\langle \hat{A} \rangle = \frac{\langle \chi^* \hat{A} \chi \rangle}{\langle \chi^* \chi \rangle}$$

The formalism discussed so far may be applied for building up a probabilistic model for describing the behavior of an arbitrary system. In order to apply it to a quantum system, we take into account that the Hamiltonian operator $\hat{H}$ is the starting point of all applications of quantum mechanics. When the one-dimensional (or multidimensional) wave function χ is given, like (16)

[or (21)], for instance, we look for the stationary points of the mean energy of the quantum system defined by

$$\langle \hat{E} \rangle = \frac{\langle \chi^* \hat{H} \chi \rangle}{\langle \chi^* \chi \rangle} \tag{22}$$

For the probability wave functions (16) and (21), the corresponding mean energy becomes $\langle E \rangle (c_1, c_2, \ldots)$ and $\langle E \rangle$ $(c_{01}, c_{10}, c_{11}, \ldots)$, respectively. We determine the stationary points of the mean energy by analyzing the system of equations

$$\frac{\partial \langle E \rangle}{\partial c_n} = 0 \qquad (n = 1, \ldots, N),$$

or the system of equations

$$\frac{\partial \langle E \rangle}{\partial c_{nl}} = 0; \qquad (n = 0, 1, \ldots, N; \quad l = 0, 1, \ldots, L; \quad (n, l) \neq (0, 0))$$

respectively. Obviously, if the functions $\sqrt{u}\, U_n$ are eigenfunctions of the operator $\hat{H}$, i.e.,

$$\hat{H}(\sqrt{u} U_n) = E_n(\sqrt{u} U_n)$$

then the mean value of energy is

$$\langle \hat{E} \rangle = \sum_{n=1}^{N} E_n \frac{c_n^2}{\sum_{l=1}^{N} c_l^2}$$

Similarly, in the two-dimensional case, if the functions $\sqrt{uv}\, U_n V_l$ are eigenfunctions of the operator $\hat{H}$, i.e.,

$$\hat{H}(\sqrt{uv} U_n V_l) = E_{nl}(\sqrt{uv} U_n V_l)$$

then the mean value of energy is

$$\langle \hat{E} \rangle = \sum_{\substack{n=0 \\ (n,l) \neq (0,0)}}^{N} \sum_{l=0}^{L} E_{nl} P_{nl}$$

where the probability of the energy value E_{nl} is

$$P_{nl} = c_{nl}^2 \Bigg/ \left(\sum_{\substack{n'=0 \\ (n',l') \neq (0,0)}}^{N} \sum_{l'=0}^{L} c_{n'l'}^2 \right)$$

Therefore, another approach is to see whether the functions $\sqrt{u} U_n$ in the one-

dimensional case, or $\sqrt{uv}U_nV_l$ in the multidimensional case, are eigenfunctions of the Hamiltonian $\hat{H}$ of the system and find the corresponding eigenvalues as possible values of the mean energy of the respective quantum system.

3. APPLICATION 1: A FREE PARTICLE IN A ONE-DIMENSIONAL BOX

The quantum system consists of a free particle in a box $D = [0, a]$. Having no constraints except the box itself, the statistical equilibrium is given by the uniform probability density function $u(x) = 1/a$ in D; let us take

$$U_0 \equiv 1, \qquad U_n(x) = \sqrt{2}\sin\frac{n\pi}{a}x \qquad (n = 1,2,\ldots)$$

as a system of orthonormal functions with the weight u. The probability wave function is approximated by

$$\chi(x) = \sqrt{u(x)}\sum_{n=1}^{N} c_n U_n(x) = \left(\frac{2}{a}\right)^{1/2}\sum_{n=1}^{N} c_n \sin\frac{n\pi}{a}x$$

with at least one coefficient c_n different from zero. As the potential energy is equal to zero, the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}$$

where m is the mass of the particle and $\hbar = h/2\pi$, h being the Planck constant. We have

$$\hat{H}\chi(x) = \frac{\hbar^2}{2m}\left(\frac{2}{a}\right)^{1/2}\left(\frac{\pi}{a}\right)^2\sum_{n=1}^{N} c_n n^2 \sin\frac{n\pi}{a}x$$

Standard integration of trigonometric functions gives

$$\langle\chi\hat{H}\chi\rangle = \int_0^a \chi(x)\hat{H}\chi(x)\,dx = \frac{\hbar^2}{2m}\left(\frac{\pi}{a}\right)^2\sum_{n=1}^{N} c_n^2 n^2$$

$$\langle\chi\chi\rangle = \int_0^a \chi^2(x)\,dx = \sum_{n=1}^{N} c_n^2$$

Introducing these two results in

$$\langle\hat{E}\rangle = \frac{\langle\chi\hat{H}\chi\rangle}{\langle\chi\chi\rangle}$$

we get

$$\langle \hat{E} \rangle \sum_{n=1}^{N} c_n^2 = \frac{\hbar^2}{2m} \left(\frac{\pi}{a} \right)^2 \sum_{n=1}^{N} c_n^2 n^2 \tag{23}$$

Taking the partial derivatives with respect to c_n ($n = 1, \ldots, N$), and looking for a stationary point of the mean energy, i.e.,

$$\frac{\partial \langle \hat{E} \rangle}{\partial c_n} = 0 \qquad (n = 1, \ldots, N)$$

we get

$$2c_n \left[\frac{\hbar^2}{2m} \left(\frac{\pi}{a} \right)^2 n^2 - \langle \hat{E} \rangle \right] = 0$$

which implies that either $c_n = 0$ or

$$\langle \hat{E} \rangle = \frac{\hbar^2}{2m} \left(\frac{\pi}{a} \right)^2 n^2 = \frac{h^2 n^2}{8ma^2}$$

Therefore, the possible values of the energy are

$$E_n = \frac{h^2 n^2}{8ma^2} \qquad (n = 1, 2, \ldots, N) \tag{24}$$

Introducing these values into (23), we get an equivalent expression for the mean energy, namely,

$$\langle \hat{E} \rangle = \sum_{n=1}^{N} E_n \frac{c_n^2}{\sum_{l=1}^{N} c_l^2} \tag{25}$$

which shows that

$$\frac{c_n^2}{\sum_{l=1}^{N} c_l^2}$$

may be interpreted as being the probability that the value of the energy of the system is E_n. Obviously, if only one coefficient c_n is different from zero, then $\langle \hat{E} \rangle = E_n$, with certainty.

Remark 1. The above conclusions may be obtained without looking for the stationary points of the energy by simply noticing that $\sqrt{u(x)}U_n(x)$ is an eigenfunction of $\hat{H}$ and the corresponding eigenvalue is just E_n. Indeed,

$$\begin{aligned}
&\hat{H}(\sqrt{\overline{u(x)}}U_n(x)) \\
&\quad = \left(\frac{2}{a}\right)^{1/2} \hat{H}\left(\sin\frac{n\pi}{a}x\right) \\
&\quad = -\left(\frac{2}{a}\right)^{1/2} \frac{\hbar^2}{2m}\frac{d^2}{dx^2}\left(\sin\frac{n\pi}{a}x\right) \\
&\quad = \frac{\hbar^2\pi^2n^2}{2ma^2}\left(\frac{2}{a}\right)^{1/2}\sin\frac{n\pi}{a}x = E_n\sqrt{\overline{u(x)}}U_n(x)
\end{aligned}$$

Then, from (22) we get (25) because

$$\begin{aligned}
\langle\hat{E}\rangle &= \frac{\left\langle\left(\sqrt{\overline{u}}\sum_{l=1}^{N} c_l U_l\right)\left(\sqrt{\overline{u}}\sum_{n=1}^{N} c_n E_n U_n\right)\right\rangle}{\left\langle\left(\sqrt{\overline{u}}\sum_{l=1}^{N} c_l U_l\right)\left(\sqrt{\overline{u}}\sum_{n=1}^{N} c_n U_n\right)\right\rangle} \\
&= \sum_{n=1}^{N} E_n \frac{c_n^2}{\sum_{l=1}^{N} c_l^2}
\end{aligned}$$

Remark 2. The above results have been obtained without solving the corresponding Schrödinger equation. It is somehow unexpected to see that the eigenfunctions of the Hamiltonian of this quantum system are just generalized coordinates of a maximum-entropy probability distribution. But even more unexpected is to subsequently see that this happens with other quantum systems as well, which eventually will allow us to apply the same approach even to quantum systems for which the corresponding Schrödinger equation cannot be solved exactly.

4. APPLICATION 2: A FREE PARTICLE IN A THREE-DIMENSIONAL BOX

The quantum system consists of a free particle in a three-dimensional box $D = [0, a] \times [0, b] \times [0, c]$. Having no constraints except the box itself, the statistical equilibrium is given by the independent product of the uniform marginals, and therefore the joint probability density function is

$$u(x)v(y)w(z) = \frac{1}{a}\frac{1}{b}\frac{1}{c} \qquad (0 \le x \le a,\ 0 \le y \le b,\ 0 \le z \le c) \tag{26}$$

We take again the trigonometric system (6) as the generalized coordinates whose weights are the uniform marginals on $[0, a]$, $[0, b]$, and $[0, c]$, respectively. The corresponding probability three-dimensional wave function induced by the minimum χ^2 deviation from statistical equilibrium described by (26) is

$$\begin{aligned}\chi(x, y, z) &= \sqrt{u(x)v(y)w(z)} \\ &\quad \times \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} c_{nlk} U_n(x) V_l(y) W_k(z) \\ &= \left(\frac{8}{abc}\right)^{1/2} \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} c_{nlk} \sin\frac{n\pi}{a} x \sin\frac{l\pi}{b} y \sin\frac{k\pi}{c} z \qquad (27)\end{aligned}$$

with at least one coefficient c_{nlk} different from zero. As the potential energy is equal to zero in D, the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2$$

where ∇^2 is the Laplacian operator. We have

$$\begin{aligned}&\hat{H}\chi(x, y, z) \\ &\quad = C \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} c_{nlk} \left[\left(\frac{n}{a}\right)^2 + \left(\frac{l}{b}\right)^2 + \left(\frac{k}{c}\right)^2\right] \\ &\qquad \times \sin\frac{n\pi}{a} x \sin\frac{l\pi}{b} y \sin\frac{k\pi}{c} z\end{aligned}$$

where

$$C = \frac{\hbar^2\pi^2}{2m} \left(\frac{8}{abc}\right)^{1/2}$$

Standard integration of trigonometric functions gives

$$\langle \chi \hat{H} \chi \rangle = \frac{\hbar^2\pi^2}{2m} \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} c_{nlk}^2 \left[\left(\frac{n}{a}\right)^2 + \left(\frac{l}{b}\right)^2 + \left(\frac{k}{c}\right)^2\right]$$

$$\langle \chi\chi \rangle = \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} c_{nlk}^2$$

Introducing these two results into

$$\langle \hat{E} \rangle = \frac{\langle \chi \hat{H} \chi \rangle}{\langle \chi \chi \rangle} \tag{28}$$

we get

$$\sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} \left\{ \frac{h^2}{8m} \left[\left(\frac{n}{a} \right)^2 + \left(\frac{l}{b} \right)^2 + \left(\frac{k}{c} \right)^2 \right] - \langle \hat{E} \rangle \right\} c_{nlk}^2 = 0$$

Taking the partial derivatives with respect to c_{nlk}, and looking for a stationary point of the mean energy, i.e.,

$$\frac{\partial \langle \hat{E} \rangle}{\partial c_{nlk}} = 0$$

for ($n = 0, 1, \ldots, N$; $l = 0, 1, \ldots, L$; $k = 0, 1, \ldots, K$; $(n, l, k) \neq (0, 0, 0)$), we get either $c_{nlk} = 0$ or

$$\langle \hat{E} \rangle = E_{nlk} = \frac{h^2}{8m} \left[\left(\frac{n}{a} \right)^2 + \left(\frac{l}{b} \right)^2 + \left(\frac{k}{c} \right)^2 \right]$$

which are the possible values of the energy of the system. Introducing these values into (28), we get an equivalent expression for the mean energy, namely,

$$\langle \hat{E} \rangle = \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{N} \sum_{l=0}^{L} \sum_{k=0}^{K} E_{nlk} p_{nlk} \tag{29}$$

which shows that

$$p_{nlk} = c_{nlk}^2 \Bigg/ \left(\sum_{\substack{n'=0 \\ (n',l',k')\neq(0,0,0)}}^{N} \sum_{l'=0}^{L} \sum_{k'=0}^{K} c_{n'l'k'}^2 \right)$$

may be interpreted as being the probability that the value of the energy of the system is E_{nlk}. Obviously, if only one coefficient c_{nlk} is different from zero, then $\langle \hat{E} \rangle = E_{nlk}$, with certainty.

Remark 3. The above conclusions may be obtained without looking for the stationary points of the energy by simply noticing that $\sqrt{uvw}\, U_n V_l W_k$ is an eigenfunction of $\hat{H}$ and the corresponding eigenvalue is just E_{nlk}. Then, from (28) we get (29).

5. APPLICATION 3: TWO NONINTERACTING FREE PARTICLES IN A ONE-DIMENSIONAL BOX

The quantum system consists of two noninteracting free particles in a one-dimensional box $[0, a]$. The domain is $D = [0, a] \times [0, a]$. Since the

two particles are independent and having no constraints except the box itself, the statistical equilibrium is given by the independent product of the uniform marginals, and therefore the joint probability density function is

$$u(x)v(y) = \frac{1}{a}\frac{1}{a} \qquad (0 \le x \le a, \quad 0 \le y \le a) \tag{30}$$

We take again the trigonometric system (6) as the generalized coordinates whose weights are the uniform marginals on $[0, a]$. The corresponding probability two-dimensional wave function induced by the minimum χ^2 deviation from statistical equilibrium described by (30) is

$$\begin{aligned}\chi(x, y) &= \sqrt{u(x)v(y)} \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl} U_n(x) V_l(y) \\ &= \frac{2}{a} \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl} \sin\frac{n\pi}{a} x \sin\frac{l\pi}{a} y \end{aligned} \tag{31}$$

with at least one coefficient c_{nl} different from zero. As the potential energy is equal to zero in D, the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial y^2}$$

where m_1 and m_2 are the masses of the two particles. We have

$$\hat{H}\chi(x, y) = \frac{\hbar^2\pi^2}{a^3} \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl}\left(\frac{n^2}{m_1} + \frac{l^2}{m_2}\right) \sin\frac{n\pi}{a} x \sin\frac{l\pi}{a} y$$

Standard integration of trigonometric functions gives

$$\langle \chi \hat{H} \chi \rangle = \left(\frac{\hbar\pi}{a}\right)^2 \frac{1}{2} \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl}^2 \left(\frac{n^2}{m_1} + \frac{l^2}{m_2}\right)$$

$$\langle \chi\chi \rangle = \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} c_{nl}^2$$

Introducing these two results into

$$\langle \hat{E} \rangle = \frac{\langle \chi \hat{H} \chi \rangle}{\langle \chi\chi \rangle} \tag{32}$$

we get

$$\sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} \left[\frac{1}{2} \left(\frac{\hbar\pi}{a} \right)^2 \left(\frac{n^2}{m_1} + \frac{l^2}{m_2} \right) - \langle \hat{E} \rangle \right] c_{nl}^2 = 0$$

Taking the partial derivatives with respect to c_{nl}, and looking for a stationary point of the mean energy, i.e.,

$$\frac{\partial \langle \hat{E} \rangle}{\partial c_{nl}} = 0; \qquad n = 0, 1, \ldots, N; \; l = 0, 1, \ldots, L; \qquad (n, l) \neq (0, 0)$$

we get either $c_{nl} = 0$ or

$$\langle \hat{E} \rangle = E_{nl} = \frac{1}{2} \left(\frac{\hbar\pi}{a} \right)^2 \left(\frac{n^2}{m_1} + \frac{l^2}{m_2} \right)$$

which are the possible values of the energy of the system. Introducing these values into (32), we get an equivalent expression for the mean energy, namely,

$$\langle \hat{E} \rangle = \sum_{\substack{n=0 \\ (n,l)\neq(0,0)}}^{N} \sum_{l=0}^{L} E_{nl} p_{nl} \tag{33}$$

which shows that

$$p_{nl} = c_{nl}^2 \Bigg/ \left(\sum_{\substack{n'=0 \\ (n',l')\neq(0,0)}}^{N} \sum_{l'=0}^{L} c_{n'l'}^2 \right)$$

may be interpreted as being the probability that the value of the energy of the system is E_{nl}. Obviously, if only one coefficient c_{nl} is different from zero, then $\langle \hat{E} \rangle = E_{nl}$, with certainty.

Remark 4. The above conclusions may be obtained without looking for the stationary points of the energy, by simply noticing that $\sqrt{uv}\, U_n V_l$ is an eigenfunction of $\hat{H}$ and the corresponding eigenvalue is just E_{nl}. Then, from (32) we get (33).

6. APPLICATION 4: THE HARMONIC OSCILLATOR

This application deals with a quantum system which randomly oscillates around the origin on the real axis with the variance σ^2. The domain is $D = (-\infty, +\infty)$. As shown by Proposition 3, the statistical equilibrium on the real axis corresponding to the mean value $\mu = 0$ and variance σ^2 is described by the normal probability distribution $N(0, \sigma^2)$ whose density is

$$u(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \qquad (-\infty < x < +\infty)$$

We take (8) with $\mu = 0$ as system of generalized coordinates associated to u, namely

$$U_n(x) = (2^n n!)^{-1/2} H_n\left(\frac{x}{\sigma\sqrt{2}}\right) \qquad (n = 0, 1, \ldots)$$

The probability wave function is approximated by

$$\begin{aligned}\chi(x) &= \sqrt{u(x)} \sum_{n=1}^{N} c_n U_n(x) \\ &= (2\pi\sigma^2)^{-1/4} \sum_{n=1}^{N} (2^n n!)^{-1/2} c_n e^{-x^2/4\sigma^2} H_n\left(\frac{x}{\sigma\sqrt{2}}\right)\end{aligned} \tag{34}$$

with at least one coefficient c_n different from zero. The Hamiltonian operator of the harmonic oscillator is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 \tag{35}$$

where m is the reduced mass and k is the force constant. Hermite polynomials satisfy the differential equation (Abramowitz and Stegun, 1972, p. 781)

$$\frac{d^2}{dx^2}[e^{-x^2/2} H_n(x)] + (2n + 1 - x^2) e^{-x^2/2} H_n(x) = 0 \tag{36}$$

Applying the operator (35) to the probability wave function (34) and taking (36) into account, we get

$$\begin{aligned}\hat{H}\chi(x) &= -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\chi(x) + \frac{1}{2}kx^2\chi(x) \\ &= \sum_{n=1}^{N} (2^n n!)^{-1/2} c_n \\ &\quad \times \left[\frac{\hbar^2}{2\sigma^2 m}\left(n + \frac{1}{2}\right) - \left(\frac{\hbar^2}{8\sigma^4 m} - \frac{k}{2}\right)x^2\right] \\ &\quad \times (2\pi\sigma^2)^{-1/4} e^{-x^2/(4\sigma^2)} H_n\left(\frac{x}{\sigma\sqrt{2}}\right)\end{aligned} \tag{37}$$

The Hermite polynomials satisfy the following recurrence formula (Abramowitz and Stegun, 1972, p. 782):

$$\xi H_n(\xi) = nH_{n-1}(\xi) + \frac{1}{2} H_{n+1}(\xi)$$

from which we get

$$\begin{aligned}\xi^2 H_n(\xi) &= n\xi H_{n-1}(\xi) + \frac{1}{2}\xi H_{n+1}(\xi) \\ &= n(n-1)\, H_{n-2}(\xi) + \left(n + \frac{1}{2}\right) H_n(\xi) + \frac{1}{4} H_{n+2}(\xi) \qquad (38)\end{aligned}$$

Denote

$$\chi_n(x) = (2^n n!)^{-1/2}(2\pi\sigma^2)^{-1/4} e^{-x^2/4\sigma^2} H_n\left(\frac{x}{\sigma\sqrt{2}}\right)$$

Obviously, $\langle \chi_l \chi_n \rangle$ is equal to 1 if $l = n$ and to 0 if $l \neq n$. Then, (34) and (37) may be written as

$$\chi(x) = \sum_{n=1}^{N} c_n \chi_n(x), \qquad \langle \chi\chi \rangle = \sum_{n=1}^{N} c_n^2 \qquad (39)$$

$$\hat{H}\chi(x) = \sum_{n=1}^{N} c_n \left[\frac{\hbar^2}{2\sigma^2 m}\left(n + \frac{1}{2}\right) - \left(\frac{\hbar^2}{8\sigma^4 m} - \frac{k}{2}\right)x^2 \right] \chi_n(x) \qquad (40)$$

Applying the recurrence relation (38), we get

$$\begin{aligned}x^2\chi_n(x) &= (2^n n!)^{-1/2}(2\pi\sigma^2)^{-1/4} e^{-x^2/4\sigma^2} 2\sigma^2 \left(\frac{x}{\sigma\sqrt{2}}\right)^2 H_n\left(\frac{x}{\sigma\sqrt{2}}\right) \\ &= \sigma^2\sqrt{n(n+1)}\,\chi_{n-2}(x) + 2\sigma^2\left(n + \frac{1}{2}\right) \\ &\quad \times \chi_n(x) + \sigma^2\sqrt{(n+1)(n+2)}\,\chi_{n+2}(x)\end{aligned}$$

which implies

$$\langle \chi_l x^2 \chi_n \rangle = \begin{cases} \sigma^2\sqrt{n(n-1)} & \text{if} \quad l = n-2 \\ 2\sigma^2\left(n + \frac{1}{2}\right) & \text{if} \quad l = n \\ \sigma^2\sqrt{(n+1)(n+2)}, & \text{if} \quad l = n+2 \\ 0 & \text{otherwise} \end{cases}$$

Using this equality and (40), we obtain

$$\langle \chi \hat{H} \chi \rangle = \left(\frac{\hbar^2}{4\sigma^2 m} + k\sigma^2 \right) \sum_{n=1}^{N} c_n^2 \left(n + \frac{1}{2} \right) - \frac{1}{2} \left(\frac{\hbar^2}{4\sigma^2 m} - k\sigma^2 \right) \times \left(\sum_{n=3}^{N} c_{n-2} c_n \sqrt{n(n-1)} + \sum_{n=1}^{N-2} c_n c_{n+2} \sqrt{(n+1)(n+2)} \right) \quad (41)$$

Introducing (41) and (39) into

$$\langle \hat{E} \rangle = \frac{\langle \chi \hat{H} \chi \rangle}{\langle \chi \chi \rangle} \quad (42)$$

we get

$$\sum_{n=1}^{N} \left[\left(\frac{\hbar^2}{4\sigma^2 m} + k\sigma^2 \right) \left(n + \frac{1}{2} \right) - \langle \hat{E} \rangle \right] c_n^2 - \frac{1}{2} \left(\frac{\hbar^2}{4\sigma^2 m} - k\sigma^2 \right) \times \left(\sum_{n=3}^{N} c_{n-2} c_n \sqrt{n(n-1)} + \sum_{n=1}^{N-2} c_n c_{n+2} \sqrt{(n+1)(n+2)} \right) = 0 \quad (43)$$

The expected value $\langle \hat{E} \rangle$ of energy depends on the probability wave function χ, i.e., it depends both on the generalized coordinates $\{H_n,\ n = 0, 1, \ldots\}$ and on the coefficients $\{c_n,\ n = 1, 2, \ldots\}$ associated to χ. The stationary values of the mean energy are obtained from the system of equations

$$\frac{\partial \langle \hat{E} \rangle}{\partial c_n} = 0 \qquad (n = 1, 2, \ldots, N)$$

Using these equations and taking in (43) partial derivatives with respect to each c_n, we get

$$2A_n c_n - \frac{1}{2} B c_{n+2} \sqrt{(n+1)(n+2)} = 0 \quad (n = 1, 2)$$

$$2A_n c_n - \frac{1}{2} B (c_{n-2} \sqrt{n(n-1)} + c_{n+2} \sqrt{(n+1)(n+2)}) = 0 \quad (n = 3, 4, \ldots, N-2)$$

$$2A_n c_n - \frac{1}{2} B c_{n-2} \sqrt{n(n-1)} = 0 \quad (n = N-1, N)$$

where

$$A_n = \left(\frac{\hbar^2}{4\sigma^2 m} + k\sigma^2 \right) \left(n + \frac{1}{2} \right) - \langle E \rangle, \qquad B = \frac{\hbar^2}{4\sigma^2 m} - k\sigma^2$$

This homogeneous system consists of N linear equations with N unknowns $c_1, \ldots, c_N$. Either $c_1 = \ldots = c_N = 0$, which corresponds to the statistical equilibrium described by the probability distribution $N(0, \sigma^2)$, or $B = 0$ and $A_n = 0$ for the values of n for which $c_n \neq 0$. But $B = 0$ implies

$$\sigma^2 = \frac{\hbar}{2\sqrt{mk}} \tag{44}$$

which is a mathematical expression of the correspondence principle: If we neglect $\hbar$ (i.e., $\hbar \to 0$), or if the relative mass m is large (i.e., $m \to +\infty$), or if the force k is strong (i.e., $k \to +\infty$), then the variance σ^2 becomes negligible (i.e., $\sigma^2 \to 0$) and we obtain the classical harmonic oscillator whose behavior has nothing random in it and for which equilibrium means having the deviation from origin equal to zero. As for the equalities $A_n = 0$, if $c_n \neq 0$, they show the possible values of the energy. Thus, if $\chi = \chi_n$, then $A_n = 0$ becomes

$$\langle \hat{E} \rangle = E_n = \left(\frac{\hbar^2}{4\sigma^2 m} + k\sigma^2 \right) \left(n + \frac{1}{2} \right)$$

which, taking into account (44), becomes

$$E_n = \hbar \sqrt{\frac{k}{m}} \left(n + \frac{1}{2} \right) \tag{45}$$

Introducing (39), (41), and (44) into (42) and taking (45) into account, we get

$$\langle \hat{E} \rangle = \sum_{n=1}^{N} E_n \frac{c_n^2}{\sum_{l=1}^{N} c_l^2}$$

which shows that

$$\frac{c_n^2}{\sum_{l=1}^{N} c_l^2}$$

may be interpreted as being the probability that the value of energy is E_n.

7. APPLICATION 5: THE HYDROGEN ATOM

The hydrogen atom consists of a proton fixed at the origin and an electron of reduced mass m interacting with the proton through a Coulomb potential:

$$U(r) = -\frac{e^2}{4\pi\varepsilon_0 r}$$

where e is the charge on the proton, ε_0 is the permitivity of free space, and r is the distance between the electron and the proton, namely, $r = (x^2+y^2+z^2)^{1/2}$.

The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + U(r)$$

where ∇^2 is the Laplacian.

7.1. The Ground State

As the Coulomb potential depends only on

$$r(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

the statistical equilibrium should also depend only on $r(x, y, z)$. As the range of $r(x, y, z)$ is $[0, +\infty)$, according to Proposition 2, the most unbiased probability distribution on $[0, +\infty)$ subject to the mean value μ is the exponential distribution $\mathbf{E}(\mu)$, with the density

$$g(x, y, z) = M\frac{1}{\mu}e^{-r(x,y,z)/\mu}$$

where M is a positive constant. The model itself suggests that we use a spherical coordinate system with the proton at the origin, namely,

$$x = r\sin\theta\cos\omega, \qquad y = r\sin\theta, \sin\omega, \qquad z = r\cos\theta$$

Taking $s = \cos\theta$, the element of volume in the three-dimensional Euclidean space becomes

$$dx\,dy\,dz = r^2\sin\theta\,dr\,d\theta\,d\omega = r^2 dr\,ds\,d\omega$$

The three variables r, s, ω are independent and their ranges are $0 \le r < +\infty$, $-1 \le s \le 1$, $0 \le \omega \le 2\pi$. Therefore, the statistical equilibrium of the hydrogen atom is described by the probability density function

$$g(x, y, z)\, dx\, dy\, dz = M \frac{1}{\mu} \exp(-\sqrt{x^2 + y^2 + z^2}/\mu)\, dx\, dy\, dz$$

$$= M \frac{1}{\mu} e^{-r/u} r^2\, dr\, ds\, d\omega$$

where the positive constant M is determined from

$$1 = \int_{\mathbf{R}^3} g(x, y, z)\, dx\, dy\, dz$$

$$= M \int_0^{+\infty} \frac{1}{\mu} r^2 e^{-r/\mu}\, dr \int_{-1}^{1} ds \int_0^{2\pi} d\omega = M(2\mu^3)(2)(2\pi) = 8\pi\mu^3 M$$

which implies $M = 1/(8\pi\mu^3)$, and

$$g(x, y, z) = \frac{1}{8\pi\mu^3} \exp(-\sqrt{x^2 + y^2 + z^2}/\mu) = \frac{1}{8\pi\mu^3} e^{-r(x,y,z)/\mu}$$

Denoting

$$u(r) = \frac{1}{2\mu^3} r^2 e^{-r/\mu} \qquad (0 \le r < +\infty)$$

$$v(s) = \frac{1}{2} \quad (-1 \le s \le 1); \qquad w(\omega) = \frac{1}{2\pi} \quad (0 \le \omega \le 2\pi)$$

we can write

$$g(x, y, z)\, dx\, dy\, dz = \frac{1}{8\pi\mu^3} \exp(-\sqrt{x^2 + y^2 + z^2}/\mu) dx\, dy\, dz$$

$$= \frac{1}{8\pi\mu^3} e^{-r/\mu} r^2\, dr\, ds\, d\omega$$

$$= u(r)v(s)w(\omega) dr\, ds\, d\omega \tag{46}$$

Statistical equilibrium is therefore described in the space $[0, +\infty) \times [-1, +1] \times [0, 2\pi]$ by the uniform probability distribution $\mathbf{U}(0, 2\pi)$ on $[0, 2\pi]$ for ω, the uniform probability distribution $\mathbf{U}(-1, +1)$ on $[-1, 1]$ for s, and the gamma distribution $\mathbf{G}(1/\mu, 3)$ with parameters $1/\mu$ and 3, on $[0, +\infty)$, for the radial variable r. The ground probability wave function induced by the probability density function $g(x, y, z)$ is

$$\psi_0(x, y, z) = \sqrt{g(x, y, z)} = \frac{1}{\sqrt{\pi}} \left(\frac{1}{2\mu}\right)^{3/2} e^{-r(x,y,z)/(2\mu)} \tag{47}$$

As ψ_0 depends only on the radial distance $r(x, y, z)$, we have

$$\nabla^2\psi_0 = \frac{d^2\psi_0}{dr^2} + \frac{2}{r}\frac{d\psi_0}{dr} = \left(\frac{1}{4\mu^2} - \frac{1}{\mu r}\right)\psi_0 \tag{48}$$

In general, if the function g depends only on $r = (\Sigma_{i=1}^{n} x_i^2)^{1/2}$ in $\mathbf{R}^n$, then g is integrable in $\mathbf{R}^n$ if and only if $g(r)\, r^{n-1}$ is integrable on $[0, +\infty)$, and we have

$$\int_{\mathbf{R}^n} g(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = S_n \int_0^{+\infty} g(r) r^{n-1}\, dr$$

where S_n is the area of the sphere in $\mathbf{R}^n$ with radius 1, namely, $S_1 = 2$, $S_2 = 2\pi$, $S_3 = 4\pi$, etc. Consequently, using (47) and (48), we get

$$\int_{\mathbf{R}^3} \psi_0\, \nabla^2\psi_0\, dx\, dy\, dz = -\frac{1}{4\mu^2}$$

and, similarly,

$$\langle \hat{E} \rangle = \int_{\mathbf{R}^3} \psi_0 \hat{H} \psi_0\, dx\, dy\, dz = \frac{\hbar^2}{8m\mu^2} - \frac{e^2}{8\pi\varepsilon_0\mu} \tag{49}$$

Using $\alpha = 1/\mu$, the mean energy becomes

$$\langle \hat{E} \rangle = \frac{\hbar^2}{8m}\alpha^2 - \frac{e^2}{8\pi\varepsilon_0}\alpha$$

Looking for a value of α for which the mean energy $\langle \hat{E} \rangle$ is stationary, we have

$$\frac{\partial\langle \hat{E} \rangle}{\partial\alpha} = 0 \quad \text{if and only if} \quad \alpha = \frac{e^2 m}{2\hbar^2\pi\varepsilon_0} \quad \text{or} \quad \mu = \frac{2\hbar^2\pi\varepsilon_0}{me^2} = \frac{a_0}{2}$$

where $a_0 = 4\pi\varepsilon_0\hbar^2/(me^2)$ is the Bohr radius. Introducing this value of μ into (49), we get

$$\langle \hat{E} \rangle = -\frac{1}{2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{m}{\hbar^2}$$

The mean location of the electron, i.e., the mean distance from the nucleus, is

$$\int_{\mathbf{R}^3} r(x, y, z)g(x, y, z)\, dx\, dy\, dz = \int_{[0,+\infty)\times[-1,1]\times[0,2\pi]} r\, \frac{1}{8\pi\mu^3}\, e^{-r/\mu} r^2\, dr\, ds\, d\omega$$

$$= \int_0^{+\infty} \frac{1}{8\pi\mu^3}\, r^3 e^{-r/\mu}\, dr \int_{-1}^{1} ds \int_0^{2\pi} d\omega$$

$$= 3\mu = \frac{3}{2}\, a_0$$

Denoting by $u(r)$ the radial probability density, we have

$$u(r) = \int_{[-1,1]\times[0,2\pi]} \frac{1}{8\pi\mu^3}\, e^{-r/\mu} r^2\, ds\, d\omega$$

$$= \frac{1}{8\pi\mu^3}\, r^2 e^{-r/\mu}\, dr \int_{-1}^{1} ds \int_0^{2\pi} d\omega = \frac{1}{2\mu^3}\, r^2 e^{-r/\mu}$$

which is the gamma probability distribution with parameters $1/\mu$ and 3. Its mean value and variance are 3μ and $3\mu^2 = (3/4)a_0^2$, respectively. The most probable location of the electron, i.e., the most probable distance from the nucleus, is given by the root of the equation

$$u'(r) = \frac{1}{2\mu^3}\, re^{-r/\mu}\left(2 - \frac{r}{\mu}\right) = 0$$

which gives $r = 2\mu = a_0$, i.e. the Bohr radius.

7.2. Jumps from the Ground State

As we have just seen, the probability density corresponding to statistical equilibrium is

$$g(x, y, z)\, dx\, dy\, dz = \frac{1}{8\pi\mu^3}\, e^{-r(x,y,z)/\mu}\, dx\, dy\, dz$$

$$= \frac{1}{8\pi\mu^3} \exp(-\sqrt{x^2 + y^2 + z^2}/\mu)\, dx\, dy\, dz$$

$$= g(x(r, \theta, \omega), y(r, \theta, \omega), z(r, \theta, \omega))r^2 \sin\theta\, dr\, d\theta\, d\omega$$

$$= g(r \sin\theta \cos\omega, r \sin\theta \sin\omega, r \cos\theta)r^2 \sin\theta\, dr\, d\theta\, d\omega$$

$$= \frac{1}{8\pi\mu^3}\, r^2 e^{-r/\mu} \sin\theta\, dr\, d\theta\, d\omega$$

$$= \frac{1}{2\mu^3}\, r^2 e^{-r/\mu}\, \frac{1}{2}\frac{1}{2\pi}\, dr\, ds\, d\omega = u(r)v(s)w(\omega)\, dr\, ds\, d\omega$$

According to (5), (6), (9), and (17), any probability density function q in the space $[0, +\infty) \times [-1, 1] \times [0, 2\pi]$ may be written as

$$\begin{aligned} & q(r, s, \omega)\, dr\, ds\, d\omega \\ & \quad = u(r)v(s)w(\omega) \\ & \qquad \times \left[1 + \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{+\infty} \sum_{l=0}^{+\infty} \sum_{k=0}^{+\infty} c_{nlk} U_n(r) V_l(s) W_k(\omega) \right] dr\, ds\, d\omega \end{aligned} \tag{50}$$

where $U_n(r)$ is the generalized Laguerre polynomial $L_n^{(2)}(r/\mu)$, $V_l(s)$ is the Legendre polynomial $P_l(s)$, and $W_k(\omega)$ is the trigonometric function $\sqrt{2}\ \sin(k\omega/2)$ with $L_0^{(2)} \equiv 1$, $P_0 \equiv 1$, and $W_0 \equiv 1$. The marginal probability density corresponding to the radial variable r is

$$\begin{aligned} & \int_{-1}^{1} \int_{0}^{2\pi} q(r, s, \omega)\, ds\, d\omega = \frac{1}{2\mu^3} r^2 e^{-r/\mu} \left[1 + \sum_{n=1}^{+\infty} c_{n00} L_n^{(2)}\left(\frac{r}{\mu}\right) \right] dr \\ & \quad = \frac{1}{8\pi\mu^3} e^{-r(x,y,z)/\mu} \left[1 + \sum_{n=1}^{+\infty} c_{n00} L_n^{(2)}\left(\frac{r(x, y, z)}{\mu}\right) \right] dx\, dy\, dz \\ & \quad = f(x, y, z)\, dx\, dy\, dz \end{aligned} \tag{51}$$

Focusing on the radial variable r, the probability wave function induced by the deviation of the probability density $f(x, y, z)$ from the ground probability density $g(x, y, z)$ which describes the statistical equilibrium is

$$\chi(r) = \frac{f - g}{\sqrt{g}} = \frac{1}{\sqrt{8\mu^3\pi}} e^{-r/(2\mu)} \sum_{n=1}^{+\infty} c_{n00} L_n^{(2)}(r/\mu) = \sum_{n=1}^{+\infty} c_{n00} \chi_n$$

The ground probability wave function is

$$\psi_0(r) = \frac{1}{\sqrt{8\pi\mu^3}} e^{-r/(2\mu)}$$

For $\beta = 2$, equality (10) becomes

$$L_n^{(2)}(x) - L_{n-1}^{(2)}(x) = L_n^{(1)}(x) \tag{52}$$

The simple jump from the ground state is

$$\psi_1(r) = \chi_1 - \chi_0 = \frac{1}{\sqrt{8\pi\mu^3}} e^{-r/(2\mu)} [L_1^{(2)}(r/\mu) - L_0^{(2)}(r/\mu)]$$

which, taking into account (52), gives

$$\psi_1(r) = \frac{1}{\sqrt{8\pi\mu^3}}\, e^{-r/(2\mu)} L_1^{(1)}(r/\mu) = \psi_0(r) L_1^{(1)}(r/\mu)$$

Similarly, the simple jump from the elementary probability wave function χ_{n-1} to the next elementary probability wave function χ_n is

$$\psi_n(r) = \chi_n - \chi_{n-1} = \frac{1}{\sqrt{8\pi\mu^3}}\, e^{-r/(2\mu)} [L_n^{(2)}(r/\mu) - L_{n-1}^{(2)}(r/\mu)]$$

which, taking into account (52), may be written as

$$\psi_n(r) = \frac{1}{\sqrt{8\pi\mu^3}}\, e^{-r/(2\mu)} L_n^{(1)}(r/\mu) = \psi_0(r) L_n^{(1)}(r/\mu) \tag{53}$$

For $\beta = 1$, equality (11) becomes

$$x \frac{d^2}{dx^2} L_n^{(1)}(x) + (2 - x) \frac{d}{dx} L_n^{(1)}(x) + n L_n^{(1)}(x) = 0 \tag{54}$$

From (53), we get

$$\frac{d}{dr} \psi_n(r) = \psi_0(r) \left[-\frac{1}{2\mu} L_n^{(1)}(r/\mu) + \frac{d}{dr} L_n^{(1)}(r/\mu) \right] \tag{55}$$

and also

$$\frac{d^2}{dr^2} \psi_n(r) = \psi_0(r) \left[\frac{1}{4\mu^2} L_n^{(1)}(r/\mu) - \frac{1}{\mu} \frac{d}{dr} L_n^{(1)}(r/\mu) + \frac{d^2}{dr^2} L_n^{(1)}(r/\mu) \right] \tag{56}$$

Replacing x by r/μ in (54), we get

$$\frac{r}{\mu} \frac{d^2}{d(r/\mu)^2} L_n^{(1)}(r/\mu) + \left(2 - \frac{r}{\mu}\right) \frac{d}{d(r/\mu)} L_n^{(1)}(r/\mu) + n L_n^{(1)}(r/\mu) = 0$$

which, dividing by $r/\mu > 0$ and taking into account that

$$d(r/\mu) = \frac{1}{\mu} dr \qquad \text{and} \qquad d(r/\mu)^2 = \frac{1}{\mu^2} dr^2$$

becomes

$$\frac{d^2}{dr^2} L_n^{(1)}(r/\mu) + \left(\frac{2}{r} - \frac{1}{\mu}\right) \frac{d}{dr} L_n^{(1)}(r/\mu) = -\frac{n}{r\mu} L_n^{(1)}(r/\mu) \tag{57}$$

Introducing (56) and (55) into

$$\nabla^2\psi_n = \frac{d^2\psi_n}{dr^2} + \frac{2}{r}\frac{d\psi_n}{dr}$$

using (57), and taking (53) into account, we get

$$\nabla^2\psi_n = \psi_0(r)\left(\frac{1}{4\mu^2} - \frac{1}{\mu r} - \frac{n}{\mu r}\right)L_n^{(1)}(r/\mu) = \left(\frac{1}{4\mu^2} - \frac{n+1}{\mu r}\right)\psi_n(r)$$

Taking this result into account, the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r}$$

applied to the probability wave function ψ_n gives

$$\hat{H}\psi_n(r) = \left[-\frac{\hbar^2}{8m\mu^2} + \left(\frac{\hbar^2(n+1)}{2m\mu} - \frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{r}\right]\psi_n(r) \tag{58}$$

The probability wave function ψ_n is an eigenfunction of the Hamiltonian $\hat{H}$ if the coefficient of ψ_n in (58) does not depend on $1/r$, which happens if

$$\mu = \frac{4\pi\varepsilon_0\hbar^2(n+1)}{2me^2} \tag{59}$$

Introducing (59) into (58), we get

$$\hat{H}\psi_n(r) = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2(n+1)^2}\psi_n(r) \tag{60}$$

Therefore, the value of energy corresponding to ψ_n is

$$E_n = -\frac{1}{2}\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\frac{m}{\hbar^2(n+1)^2} \qquad (n = 0, 1, \ldots) \tag{61}$$

Remark 5. In this section we have used the spherical coordinates (r, θ, ω) in the space $[0, +\infty) \times [0, \pi] \times [0, 2\pi]$, or, equivalently, the coordinates (r, s, ω) in the space $[0, +\infty) \times [-1, 1] \times [0, 2\pi]$, where $s = \cos\theta$. As shown in (46), in obtaining a probabilistic model for the behavior of the hydrogen atom we started from a statistical equilibrium described by the gamma distribution $\mathbf{G}(1/\mu, 3)$ for the radial variable r on $[0, +\infty)$, the uniform distribution for s on $[-1, 1]$, and the uniform distribution for ω on $[0, 2\pi]$. In its general form, the probability wave function has the form

$$\chi(r, s, \omega) = \psi_0(r) \sum_{\substack{n=0 \\ (n,l,k)\neq(0,0,0)}}^{+\infty} \sum_{l=0}^{+\infty} \sum_{k=0}^{+\infty} c_{nlk} U_n(r) V_l(s) W_k(\omega) \tag{62}$$

where $\psi_0(r)$ is the probability wave function (47) of the ground state, $U_n(r)$ is the generalized Laguerre polynomial $L_n^{(2)}(r/\mu)$, $V_l(s)$ is the Legendre polynomial $P_l(s)$, and $W_k(\omega)$ is the trigonometric function $\sqrt{2}\sin(k\omega/2)$, as discussed in Section 2. In all the considerations made in this section we focused on the radial variable r and, instead of (62), we dealt with the simpler probability wave function

$$\chi(r) = \psi_0(r) \sum_{n=1}^{+\infty} c_{n00} U_n(r) V_0(s) W_0(\omega)$$

where $V_0 \equiv 1$, $W_0 \equiv 1$. In the general context, the jumps (53) should be replaced by

$$\psi_{nlk,n'l'k'}(r, s, \omega) = \psi_0(r)[U_n(r)V_l(s)W_k(\omega) - U_{n'}(r)V_{l'}(s)W_{k'}(\omega)]$$

Remark 6. The probabilistic model for the hydrogen atom, just discussed, started from the statistical equilibrium described by the exponential distribution

$$g(x, y, z) = M \frac{1}{\mu} e^{-r(x,y,z)/\mu}$$

which is the solution of the principle of maximum entropy $\max_g \mathbf{H}(g)$ subject to the mean value μ, or, equivalently, the solution of the principle of minimum relative entropy $\min_g \mathbf{H}(g|1)$ on $[0, +\infty)$, where $1 = r^0$ (x, y, z), in which case the radial probability density on $0 \leq r < +\infty$ is

$$u(r) = \frac{1}{2\mu^3} r^2 e^{-r/\mu}$$

i.e., the gamma distribution $\mathbf{G}(1/\mu, 3)$, and the system of orthogonal polynomials with the weight u is the sequence of generalized Laguerre polynomials $\{L_n^{(2)}(r/\mu),\ n = 0, 1, \ldots\}$. This proved to be enough for getting the entire energy spectrum (61) of the hydrogen atom. The approach can be generalized if we start from the statistical equilibrium described by the solution of the principle of minimum relative entropy $\min_g \mathbf{H}(g|r^{2l}(x, y, z))$, which is the gamma probability distribution $\mathbf{G}(1/\mu, 2l + 1)$, whose density is

$$g(x, y, z) = \frac{1}{\Gamma(2l + 1)} \left(\frac{1}{\mu}\right)^{2l+1} r^{2l}(x, y, z) e^{-r(x,y,z)/\mu}$$

in which case the radial probability density on $0 \leq r < +\infty$ is

$$u(r) = M\left(\frac{1}{\mu}\right)^{2l+3} r^{2l+2}e^{-r/\mu} \qquad (M \text{ constant})$$

i.e., the gamma distribution $\mathbf{G}(1/\mu, 2l + 3)$, and the system of orthogonal polynomials with the weight u is the sequence of generalized Laguerre polynomials $\{L_n^{(2l+2)}(r/\mu), n = 0, 1, \ldots\}$. The simple jumps are in this case described by the probability wave functions

$$\psi_n(r) = \sqrt{u(r)}[L_n^{(2l+2)}(r/\mu) - L_{n-1}^{(2l+2)}(r/\mu)] = \sqrt{u(r)}L_n^{(2l+1)}(r/\mu)$$

8. CONCLUSION

In standard nonrelativistic quantum mechanics, the Schrödinger equation is taken as a postulate and the squared absolute value of its solution is interpreted as being a probability density function used for making predictions about the behavior of quantum systems. This paper deals with a nonstandard approach. Given a quantum system, we determine the probability wave function whose corresponding probability distribution (i.e., the square of its absolute value) is the closest one to statistical equilibrium subject to generalized correlation coefficients whose values are obtained by looking for the stationary points of the mean energy of the system. Statistical equilibrium, determined by using the principle of maximum entropy, gives the most unbiased probability distribution on the possible states of the system subject to given mean values. The closest probability distribution to statistical equilibrium is obtained by minimizing Pearson's mean deviation subject to given generalized correlation coefficients whose values are obtained, as said before, by looking for stationary points of the mean energy of the system.

Whether nature is acting according to the principle of minimum mean deviation from statistical equilibrium is a matter of philosophy. This paper uses this principle only as a tool for constructing a mathematical model for the behavior of quantum systems. Surprisingly enough, this variational method recovers, in a unitary way, the exact solutions for the harmonic oscillator, the free particle in a box, two independent particles in a box, and the hydrogen atom, without using the Schrödinger equation, and is in agreement with Ernst Mach's economy of thought principle. The method used in this paper may be briefly called the *miniminimax* model, which is an abbreviation for 'minimizing the mean (ground or transition) energy corresponding to the probability wave function obtained by minimizing the mean deviation from maximum entropy condition subject to generalized correlations.' According to McQuarrie (1983, p. 297), "the inclusion of electron correlations in atomic and molecular wave functions is a problem of current and active interest." The

second part of this paper will show that the formalism presented above gives excellent simple approximations when it is applied to the ground state of the helium and lithium atoms.

REFERENCES

Abramowitz, M., and Stegun, I.A. (eds.) (1972). *Handbook of Mathematical Functions,* Dover, New York, pp. 773–783.

Baublitz, Jr., M. (1988). Derivation of the Schrödinger equation from a stochastic theory, *Progress of Theoretical Physics,* **80**, 232–244.

Bohm, D. (1984). *Causality and Chance in Modern Physics,* 3th ed., University of Pennsylvania Press, Philadelphia, Pennsylvania.

Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, *Sitzungsberichte der Akademie der Wissenschaften(Wien),* **66**, 275–370.

Born, M. (1926). Zur Quantenmechanik der Stossvorgänge, *Zeitschrift für Physik,* **37**, 863–867.

Born, M. (1963). *Ausgewählte Aghandlungen,* Vandenhoeck & Rupprecht, Göttingen.

Born, M. (1969). *Atomic Physics,* 8th ed. Blackie, London.

Bühler, W. K. (1981). *Gauss,* Springer-Verlag, Berlin.

Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics, *Philosophical Transactions of Royal Society, London, Series A,* **222**, 309–368.

Flügge, S. (1974). *Practical Quantum Mechanics,* Springer-Verlag, Berlin.

Frieden, B. R. (1989). Fisher information as the basis for the Schrödinger wave equation, *American Journal of Physics,* **57**(11), 1004–1008.

Frieden, B. R. (1990). Fisher information, disorder, and the equilibrium distributions of physics, *Physical Review A,* **15**, 4265–4276.

Frieden, B. R. (1991). Fisher information and the complex nature of the Schrödinger wave equation, *Foundations of Physics,* **21**, 757–771.

Gauss, C.F. (1821/1823). Theoria combinationis observationum erroribus minimis obnoxiae. (Part I and Part II). In Gauss, C. F. (1973). *Werke,* Vol. 4, *Wahrscheinlichkeitsrechn ung und Geometrie,* Georg Olms Verlag, Hildesheim.

Gradshteyn, I. S., and Ryzhik, I. M. (1980). *Table of Integrals, Series, and Products,* Academic Press, Orlando, Florida.

Guiasu, S. (1977). *Information Theory with Applications,* McGraw-Hill, New York.

Guiasu, S. (1987). Probability space of wave functions, *Physical Review A,* **36**, 1971–1977.

Guiasu, S. (1990). A classification of the main probability distributions by minimizing the weighted logarithmic measure of deviation, *Annals of the Institute of Statistical Mathematics,* **42**, 269–279.

Guiasu, S. (1992). Deducing the Schrödinger equation from minimum χ^2, *International Journal Theoretical Physics,* **31**, 1153–1176.

Guiasu, S., and Shenitzer, A. (1985). The principle of maximum entropy, *Mathematical Intelligencer,* **7**, 42–48.

Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Review,* **106**, 620–630; **108**, 171–190.

Jaynes, E. T. (1979). Where do we stand on maximum entropy? In *The Maximum Entropy Formalism,* R. Levin and M. Tribus, eds., MIT Press, Cambridge, Massachusetts, pp. 15–118.

Kullback, S., and Leibler, R. A. (1951). On information and sufficiency, *Annals of Mathematical Statistics,* **22**, 79–86.

Legendre, A. (1805). *Nouvelles Méthodes pour la Détermination des Orbites des Comètes,* Courcier, Paris.

McQuarrie, D. A. (1983). *Quantum Chemistry,* University Science Books, Mill Valley, California.

Mehra, J., and Rechenberg, H. (1987). *The Historical Development of Quantum Theory,* Vol. 5, Springer-Verlag, New York.

Moore, W. (1989). *Schrödinger. Life and Thought,* Cambridge University Press, Cambridge, England.

Nelson, E. (1985). *Quantum Fluctuations,* Princeton University Press, Princeton, New Jersey.

Pais, A. (1982). Max Born's statistical interpretation of quantum mechanics, *Science,* **218**, 1193–1198.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine* (5th Series), **50**, 157–175.

Preda, V., Bulacu, F., and Bulacu, M. (1995). Pearson global indicator and the quantum mechanics of p-dimensional systems, manuscript.

Schrödinger, E. (1926). Quantisierung als Eigenwertproblem, *Annalen der Physik* (4), **79**, 361–376, 489–527; **80**, 437–490; **81**, 109–139.

Schrödinger, E. (1928). *Collected Papers on Wave Mechanics,* Blackie, London.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal,* **27**, 379–423, 623–656.

Slater, P. B. (1994). Schrödinger-like equations for spin measurement states, and wave functions, based on the statistical approach of Guiasi, *Canadian Journal of Physics,* **72**, 130–133.

Stigler, S. M. (1981). Gauss and the invention of least squares, *Annals Statistics,* **9**, 465–474.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik,* Springer-Verlag, Berlin.

Wolfram, S. (1991). *Mathematica. A System for Doing Mathematics by Computer,* 2nd ed., Addison-Wesley, Redwood City, California.