International Journal of Theoretical Physics, Vol. 31, No. 7, 1992

Deducing the Schrödinger Equation from Minimum χ^2

Silviu Guiasu[1]

Received August 12, 1991

The most unbiased probabilistic model for the possible values of a characteristic of a quantum system subject to the constraints represented by some known mean values characterizes the system in a steady-state condition. We suppose that random fluctuations alter such a steady-state condition. The probability distribution of the possible deviations from the steady-state condition is estimated by minimizing Pearson's χ^2 subject to the mean fluctuations available. The optimum Pearson function χ^* may be interpreted as the wave function of the system and in the case of the harmonic oscillator, the free particle in a box, and the hydrogen atom, the prediction based on it is compatible with that provided by the solution of the corresponding Schrödinger equations.

1. INTRODUCTION

In classical statistical mechanics the evolution of a system of molecules was strictly deterministic but a probabilistic model was needed as an approximation of reality due to insufficient data available. In quantum mechanics the behavior of the system itself is random and the available information about the system is only partial from both practical and theoretical points of view. In both cases we are facing a similar problem: estimate the most reliable probabilistic model subject to the data available. Paraphrasing Niels Bohr (Polkinghorne, 1986), "the entire formalism is to be regarded as a tool for deriving predictions, of definite or statistical character, as regards information obtainable under experimental conditions described in classical terms."

The present model is strictly based on mean values. At the beginning, we estimate the most unbiased probabilistic model for the possible values of a characteristic X of a quantum system subject to the constraints represented by some known mean values. Such a model characterizes the behavior of the system in the steady-state condition corresponding to the given mean

[1]Department of Mathematics, York University, North York, Ontario, Canada M3J 1P3.

0020-7748/92/0700-1153$06.50/0

values. The mean values accessible to us are not sufficient for determining the probability distribution on the possible values of the characteristic in a unique way. There are in fact infinitely many probability distributions compatible with the known mean values. From all these feasible probability distributions we select the most unbiased one (i.e., the one containing the largest amount of uncertainty or, equivalently, giving no special preference to particular possible values). This problem has already been studied in the literature and Shannon's entropy, inspired by L. Boltzmann's H-function, has been used as an abstract measure of the amount of uncertainty contained by a probability distribution. By maximizing entropy subject to the given mean values, we select the most unbiased probability distribution u subject to the given constraints. We say that such a probability distribution characterizes the behavior of the characteristic of the quantum system in the steady-state condition defined by the mean values taken as constraints. The solution of such a variational problem is only a model, an estimation of the true unknown probability distribution of the values of the characteristic X of the system. But it is chosen to be the most unbiased estimation subject to the mean values accessible to us through the measurement process. When not only the mean values of some observables are given but also a reference measure induced by a certain field, then the above approach for estimating the steady state is generalized in the sense that we are looking for the probability distribution on the set of possible values of the observables that is the closest one to the given reference measure subject to the mean values accessible to us, where closeness is measured by the Kullback–Leibler divergence.

Suppose now that random fluctuations alter the initial steady-state condition, which means that the maximum entropy probability distribution u subject to the initial mean values no longer accurately describes the behavior of the characteristic of the system we are interested in. We can detect such a change in the following way. To the steady-state probability distribution u we assign a sequence of orthonormal functions with the weight u. As long as the system remains in the steady-state condition described by u, the mean value of each orthonormal function is equal to zero. When at least one of these mean values is not zero, this means that the system has deviated from the steady-state condition described by u. The mean fluctuations just detected are not sufficient to completely determine the probability distribution of the fluctuations from the steady state described by u. We estimate the probability distribution of the possible deviations from the steady-state condition described by u by minimizing the $\langle \chi^2 \rangle$ global indicator introduced by Karl Pearson in statistical inference at the beginning of this century in order to measure how different two densities are on average, subject to the mean fluctuations estimated. The optimum Pearson function χ^* behaves like the wave function of the system, its square is the probability density of the deviations from the steady-state condition described by u, and, in the case of the harmonic oscillator, the free particle in a box, and the hydrogen atom, the prediction based on it is compatible with that provided by the solution of the corresponding Schrödinger equations by simply using classical quantization rules.

Without using the formalism of this paper, 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 . . . Instead of starting from Born's probability distribution $P=|\psi|^2$ [where ψ is the wave function of the system] 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 present paper does not follow this line of thought, but aims at constructing the most unbiased probabilistic model compatible with mean fluctuations registered at the macroscopic scale, using tools from statistical inference such as Pearson's indicator, Shannon's entropy, and the Kullback–Leibler divergence.

In order to be more specific, let us distinguish three phases in building up the probabilistic model:

Phase I: Let D be the set of possible values of a characteristic X of a quantum system. Neither the value of X nor the probability density u on D is known. At the macroscopic level, by performing observations on the respective quantum system, we can get estimates of some mean values, such as the mean components of X and the corresponding variances. These mean values accessible to us do not determine, in a unique way, the probability distribution u on the set D of possible values of X. From all probability densities u on D compatible with the given mean values, we choose that one which maximizes Shannon's entropy $H(u)=-\langle \ln u|u\rangle$, a generally accepted measure of uncertainty contained by a probability distribution, inspired by Boltzmann's H-function from statistical mechanics, where $\langle \cdot|\cdot \rangle$ is the inner (scalar) product between square-integrable functions defined on D,

$$\langle f|g\rangle=\int_D f(x)g(x)\,dx$$

Such a probability density is the most random one, i.e., it takes all possible states into account in an unbiased way, without giving undue preference toward particular values, subject to the constraints provided by the mean values given by the measurement process. The principle of entropy maximization (PEM) has been amply discussed in the literature (von Neumann, 1932; Jaynes, 1957; Levine and Tribus, 1979; Guiasu and Shenitzer, 1985; Justice, 1986; Skilling, 1989). When a reference measure of density v is also given, PEM has been generalized by the principle of minimum divergence (PMD), according to which we determine the closest probability density u to the reference density v subject to the given mean values, where closeness is measured by the Kullback–Leibler divergence (Kullback and Leibler, 1951) $D(u:v)=\langle \ln(u/v)|u\rangle$. The solution u of PEM (or PMD) describes the behavior of X in the steady-state condition defined by the given mean values (and the reference measure).

Phase II: A this stage, we want to check whether or not the system remains in the steady-state condition described by the probability density u. Let $\mathcal{U}=\{U_n, n=0, 1, \ldots\}$ be an orthonormal, generally—but not necessarily—complete sequence of functions on D with the weight u. As long as the system remains in the steady-state condition described by u, all the mean values $\langle U_n|u\rangle$ $(n=1, 2, \ldots)$ are equal to zero. From a practical point of view, we focus on a finite number, say $U_1, \ldots, U_N$, of functions from $\mathcal{U}$, and observing M values of X, say $x_1, \ldots, x_M \in D$, we calculate the sample mean

$$\bar{U}_n^{(M)}=[U_n(x_1)+\cdots+U_n(x_M)]/M$$

for each $n=1, \ldots, N$. The number $\bar{U}_n^{(M)}$ is used as an estimate of the mean value of U_n. Due to the central limit theorem from probability theory, we are $100(1-\alpha)\%$ confident that the true mean value of U_n belongs to the confidence interval

$$[\bar{U}_n^{(M)}-Z_{\alpha/2}S_n^{(M)}/\sqrt{M},\ \bar{U}_n^{(M)}+Z_{\alpha/2}S_n^{(M)}/\sqrt{M}] \tag{1}$$

for $M>30$, where $0<\alpha<1$, $[S_n^{(M)}]^2$ is the sample variance, i.e.,

$$[S_n^{(M)}]^2=\{[U_n(x_1)-\bar{U}_n^{(M)}]^2+\cdots+[U_n(x_M)-\bar{U}_n^{(M)}]^2\}/(M-1)$$

and $Z_{\alpha/2}$ is the critical point of the standard normal distribution $N(0, 1)$ corresponding to the probability $\alpha/2$. For a 95% confidence interval, for instance, $Z_{0.025}=1.96$. Now, if the number 0 belongs to the confidence interval (1), then we do not reject the hypothesis that the probability density u accurately describes the behavior of the characteristic X. If the confidence interval does not contain the number 0, then we use the value of $\bar{U}_n^{(M)}$ as an estimate of the new mean value of U_n. This is an indication that random fluctuations have occurred along the direction U_n and the probability density u no longer describes accurately the behavior of the characteristic X. In what follows, for simplifying the writing, $\bar{U}_n^{(M)}$ will be denoted by c_n.

Phase III: At this stage, taking the mean values $c_1, \ldots, c_N$ into account, we want to determine the probability distribution of the deviations from the steady-state condition described by u due to the random fluctuations in the directions $U_1, \ldots, U_M$. We are looking for the closest density f^* to u on D subject to the mean values $c_1, \ldots, c_N$, where the closeness between a density f and u is measured by the mean relative square deviation of f from u, introduced by Pearson (1900),

$$\langle \chi^2(f: u) \rangle = \langle \chi^2(f: u) | 1 \rangle = \left\langle \frac{(f-u)^2}{u} \middle| 1 \right\rangle = \left\langle \left(\frac{f}{u} - 1 \right)^2 \middle| u \right\rangle \tag{2}$$

If we focus on the expression $\langle (f-u)^2/u | 1 \rangle$, we can see that the minimization of Pearson's indicator is not quite the same thing as the well-known (and widely applied since Legendre and Gauss) least square method, even if they are obviously connected. As the square deviations are divided by u, by minimizing $\langle \chi^2(f: u) \rangle$ we sanction more drastically the large deviations from the less probable states of the steady-state condition described by u. From a technical point of view, we take into account only densities f for which f/u is square-integrable with respect to the measure of density u. The Pearson function $\chi(f^*: u)$ corresponding to the optimum solution f^* behaves like the wave function of the system, while its square, $\chi^2(f^*: u)$ is the probability density of the sum of the square fluctuations of types $U_1, \ldots, U_N$ weighted by the mean fluctuations $c_1, \ldots, c_N$. The wave function appears to be generated by the deviations from a steady-state condition due to random fluctuations having a mean different from zero.

The formalism is applied to the harmonic oscillator, the free particle in a box, and the hydrogen atom. In all these cases, the minimization of the mean Pearson deviation $\langle \chi^2(f: u) \rangle$ from the steady-state probability density u yields a solution f^* for which the corresponding Pearson function $\chi^* = \chi(f^*: u)$ behaves like the wave function of the system satisfying a second-order differential equation that, under standard classical quantization rules, reduces to the corresponding Schrödinger equation.

A recent paper (Frieden, 1989) has relatively the same aim, namely to build up a probabilistic model based on estimation theory from which the Schrödinger equation could be derived as a consequence. The tool used, however, is different. Dealing with the position of a particle on the real line, that work obtains the Schrödinger equation by minimizing a linear combination of the Fisher information (Fisher, 1947), measuring the degree of ruggedness of a probability distribution u, and the mean kinetic energy of the particle, namely, using our notations,

$$\langle [(\ln u)']^2 | u \rangle + \lambda \langle W - V | u \rangle$$

where W is the (unknown) mean total energy, $V(x)$ is the potential energy, and λ is a fixed negative constant. Taking $u=\psi^2$ and $\lambda=-8\pi m/h^2$, where h is Planck's constant, we find that the necessary Euler–Lagrange equation of the above minimization problem gives the Schrödinger equation

$$\psi''(x)+(8\pi m/h^2)[W-V(x)]\psi(x)=0$$

In what follows, Section 2 deals with the simplest case of a random variable whose steady-state condition is perturbed by linear fluctuations. Section 3 presents the general mathematical model. Section 4 studies three special cases of steady-state conditions frequent in applications. Section 5 shows for what steady-state probability densities and types of fluctuations the corresponding wave function satisfies a differential equation of Schrödinger type. In Sections 6–8, the formalism is applied to the harmonic oscillator, the particle in a box, and the hydrogen atom, respectively. In Section 9 the formalism is applied to the nonstandard problem of determining the location of a one-dimensional particle when the only initial information available is a mean location value and a reference measure induced by a certain field. The final section contains conclusions.

2. A SIMPLE CASE

Let X be a random variable representing a characteristic of a physical system, D its range, $\langle\cdot|\cdot\rangle_u$ the scalar (inner) product with the weight u between square-integrable functions defined on D, i.e.,

$$\langle f|g\rangle_u=\int_D f(x)g(x)u(x)\,dx$$

and $\langle\cdot|\cdot\rangle$ the scalar product with the weight 1. We denote by 1 both the number 1 and the constant function identically equal to 1. If some moments of X, say $E(X^k)$, $k=i(1),\ldots,i(m)$, are given, there are infinitely many probability densities f compatible with them. The steady-state condition of X corresponding to the given moments $E(X^k)$, $k=i(1),\ldots,i(m)$, is characterized by the maximum entropy probability density u compatible with these moments. It is the solution of the variational problem

$$\max_u H(u)=-\langle\ln u|u\rangle$$

subject to

$$\langle u|1\rangle=1$$
$$\langle x^k|u\rangle=E(X^k),\qquad k=i(1),\ldots,i(m)$$

The solution of this variational problem is the most unbiased probability density subject to the constraints imposed by the given moments of X. To give an example, the maximum entropy probability density u on $D=(-\infty, +\infty)$, corresponding to the steady-state condition induced by the mean $\mu=E(X)$ and the variance $\sigma^2=E[(X-\mu)^2]$, is the normal distribution $N(\mu, \sigma^2)$. Also, when $D=(0, +\infty)$, the maximum entropy probability density corresponding to the steady-state condition induced by the mean $\lambda=E(X)$ is the density of the exponential distribution $E(\lambda)$ with parameter λ, i.e.,

$$u(x)=\frac{1}{\lambda}\,e^{-x/\lambda}$$

Finally, the maximum entropy probability density on $D=[a, b]$ describing the steady state of X when no constraint is imposed is the density of the uniform distribution $U(a, b)$, i.e., $u(x)=1/(b-a)$.

Let u be the maximum entropy probability density on D corresponding to the steady-state condition induced by some given moments of X. We call u the maximum entropy steady-state probability density on D. Let $\mu=E(X)=\langle x|u\rangle$, $\sigma^2=E[(X-\mu)^2]=\langle (x-\mu)^2|u\rangle$, and $\mathscr{U}=\{U_0, U_1\}$, where $U_0=1$, $U_1=(x-\mu)/\sigma$. Obviously, $\mathscr{U}$ is an orthonormal set of square-integrable functions on D with the weight u. The function U_1 represents a linear fluctuation. In the steady-state condition characterized by u, the mean linear fluctuation of type U_1 is zero, $\langle U_1|u\rangle=0$. Suppose that, due to random fluctuations of linear type U_1, the system gets out of the steady-state condition characterized by u and the mean linear fluctuation becomes $\langle U_1|f\rangle=c$, where c is a given real constant and f the unknown probability density describing the behavior of X for the new condition the system is in. The mean Pearson deviation $\langle\chi^2(f\colon u)\rangle$ of f from u is defined by (2).

As f is unknown, we approximate it by the solution f^* of the variational problem $\min_f\langle\chi^2(f\colon u)\rangle$ subject to

$$\left\langle \frac{x-\mu}{\sigma}\middle| f\right\rangle=c \tag{3}$$

As a more general variational problem of this type will be solved in the next section, we simply give here the solution, which is

$$f^*(x)=u(x)\left[1+c\,\frac{x-\mu}{\sigma}\right]$$

It is a signed density satisfying (3), together with $\langle f^*|u\rangle = 1$, and $\langle \chi^{*2}\rangle = \langle \chi^2(f^*: u)\rangle = c^2$. The corresponding wave function is

$$\chi^*(x) = \chi(f^*: u)(x) = \left(\frac{f^*(x)}{u(x)} - 1\right)[u(x)]^{1/2}$$

$$= cU_1(x)[u(x)]^{1/2} = c\frac{x-\mu}{\sigma}[u(x)]^{1/2}$$

The normed wave function is $\psi^*(x) = c^{-1}\chi^*(x)$. Its square,

$$\psi^{*2}(x) = c^{-2}\left(\frac{f^*(x)}{u(x)} - 1\right)^2 u(x) = \left(\frac{x-\mu}{\sigma}\right)^2 u(x)$$

may be interpreted, in a natural way, as the probability density of the standardized square deviations from the mean, showing the intensity of the deviation from the steady-state condition.

3. THE MATHEMATICAL MODEL

Let X be a random variable representing a characteristic of a physical system, D its range, and u the probability density on D describing the behavior of X in a steady-state condition. Suppose $u > 0$ on D. Let $\mathcal{U} = \{U_n, n = 0, 1, \ldots\}$ be an orthonormal sequence of functions on D with the weight u, with $U_0 = 1$. Then, $\langle U_m|U_n\rangle_u = \delta_{m,n}$, where $\delta_{m,m} = 1$, and $\delta_{m,n} = 0$ if $m \neq n$. Due to random fluctuations, the system gets out of the initial steady-state condition and the behavior of X is no longer described by u. We consider only fluctuations given by square-integrable functions. Any such fluctuation V may be approximated by finite sums of the form $\sum \langle V|U_n\rangle U_n$ in the topology induced by the scalar product $\langle\cdot|\cdot\rangle_u$. Thus, it is sufficient to take into account fluctuations of type $U_1, \ldots, U_N$.

Suppose that random fluctuations of type $U_1, \ldots, U_N$ have occurred and let f be the new probability density that describes the behavior of X. The only information available is represented by the mean fluctuations (1). These constraints do not determine uniquely the probability density f. We approximate f by the density f^* that is the closest one to the steady-state probability density u subject to the mean fluctuations (1), where closeness is measured by the mean Pearson deviation from u given by (2). Thus, we are looking for the solution f^* of the variational problem

$$\min_f \langle \chi^2(f: u)\rangle \tag{4}$$

subject to the constraints (1). Introducing the Lagrange multipliers α_n and the function

$$F=(f-u)^2/u-\sum_{n=1}^{N}\alpha_n f U_n$$

we have

$$\langle\chi^2(f:u)\rangle-\sum_{n=1}^{N}\alpha_n\langle f|U_n\rangle=\langle F|1\rangle$$

The inner product is stationary if the Euler–Lagrange equation $\partial F/\partial f=0$ is satisfied. The solution of this equation is

$$f^*=u\left[1+\sum_{n=1}^{N}\tfrac{1}{2}\alpha_n U_n\right]\tag{5}$$

Introducing this expression in (1), we get $\frac{1}{2}\alpha_n=c_n$ and (5) becomes

$$f^*=u\left[1+\sum_{n=1}^{N}c_n U_n\right]\tag{6}$$

The density (6) gives a minimum of $\langle\chi^2(f:u)\rangle$. Indeed, for any square-integrable function f for which f/u is also square-integrable, if $\mathscr{U}$ is a complete system, we have

$$\frac{f}{u}=\sum_{n=0}^{\infty}\left\langle\frac{f}{u}\middle|U_n\right\rangle_u U_n$$

which gives

$$f=u\sum_{n=0}^{\infty}\langle f|U_n\rangle U_n=u\left[\sum_{n=1}^{N}c_n U_n+W\right]=f^*+u(W-1)$$

where

$$W=\sum_n{}^{*}\langle f|U_n\rangle U_n$$

the sum being taken for all nonnegative integers $n\notin\{1,\ldots,N\}$. Thus, as $\langle f^*-u|W\rangle=0$, we have

$$\langle\chi^2(f:u)\rangle=\langle\chi^2(f^*:u)\rangle+\langle(W-1)^2|u\rangle\geq\langle\chi^2(f^*:u)\rangle$$

The mean Pearson deviation of the solution f^* from the steady-state probability density u is

$$\langle \chi^{*2} \rangle = \langle \chi^2(f^*:u) \rangle = \sum_{n=1}^{N} c_n^2$$

The wave function is just the Pearson function

$$\chi^* = \chi(f^*:u) = \left(\frac{f^*}{u} - 1\right)\sqrt{u} = \sqrt{u} \sum_{n=1}^{N} c_n U_n \tag{7}$$

It may be interpreted as being the minimum deviation from the steady-state condition described by the probability density u due to the random fluctuations of types $U_1, \ldots, U_N$, with the means $c_1, \ldots, c_N$, respectively. The normed wave function is

$$\psi^* = \langle \chi^{*2} \rangle^{-1/2} \chi^* = \left(\sum_{n=1}^{N} c_n^2 \right)^{-1/2} \chi^* \tag{8}$$

Its square,

$$\psi^{*2} = \left(\sum_{n=1}^{N} c_n^2 \right)^{-1} \left(\sum_{n=1}^{N} c_n U_n \right)^2 u$$

may be interpreted as being the probability density on D induced by the minimum mean relative square deviations from the steady-state condition described by u due to the fluctuations of type $U_1, \ldots, U_N$ having the mean values $c_1, \ldots, c_N$, respectively.

Let now X_1 and X_2 be two independent random variables describing two characteristics of a physical system, D_1 and D_2 their ranges, u_1 and u_2 their steady-state probability densities on D_1 and D_2 respectively, and $\mathscr{U}^{(i)} = \{U_n^{(i)}, n=0, 1, \ldots\}$, with $U_0^{(i)} = 1, i=1, 2$, two orthonormal sets of functions with the weights u_1 and u_2, respectively. Then,

$$\mathscr{U} = \{U_m^{(1)} U_n^{(2)}, m=0, 1, \ldots; n=0, 1, \ldots\}$$

is an orthonormal set of functions with the weight $u_1 u_2$. If, due to random fluctuations, X_1 and X_2 become dependent, then their behavior is no longer described by the independent product density $u_1 u_2$, but by a joint probability density f. Suppose that we know the generalized correlations between X_1 and X_2 induced by fluctuations of type $U_m^{(1)} U_n^{(2)}$ $(m=1, \ldots, M; n=1, \ldots, N)$, i.e.,

$$\langle U_m^{(1)} | U_n^{(2)} \rangle_f = c_{m,n}, \qquad m=1, \ldots, M; n=1, \ldots, N \tag{9}$$

These generalized correlations between X_1 and X_2 do not determine uniquely the joint probability density f. We approximate f by the density f^* that is the closest one to the independent product probability density u_1u_2, subject to the constraints (9). Thus, we solve the variational problem

$$\min_f \chi^2(f: u_1u_2)$$

subject to the constraints (9). The solution f^* of this problem is the most independent product density subject to the given generalized correlations (9). This is a kind of principle of maximum independence with constraints. By applying it, we want to approximate the joint probability density by supposing nothing more about the dependence between X_1 and X_2 than that it is contained in the given generalized correlations (9). The solution is similar to that discussed in the one-dimensional case above, namely,

$$f^* = u_1u_2\left[1 + \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n} U_m^{(1)} U_n^{(2)}\right]$$

We have $\langle f^*|1\rangle = 1$, $\langle f^*|1\rangle^{(1)} = u_2$, and $\langle f^*|1\rangle^{(2)} = u_1$, where $\langle\cdot|\cdot\rangle^{(i)}$ is the inner product in D_i, $i = 1, 2$. The deviation from independence is

$$\phi^* = f^* - u_1u_2 = u_1u_2 \sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n} U_m^{(1)} U_n^{(2)}$$

The wave function is the minimum deviation from independence, i.e.,

$$\chi^* = \chi(f^*: u_1u_2) = (u_1u_2)^{1/2}\left(\frac{f^*}{u_1u_2} - 1\right)$$

The normed wave function is

$$\psi^* = \left(\sum_{m=1}^{M} \sum_{n=1}^{N} c_{m,n} U_m^{(1)} U_n^{(2)}\right)^{-1/2} \chi^*$$

and ψ^{*2} may be interpreted as the probability density of the minimum deviations from independence.

4. SPECIAL CASES OF STEADY-STATE CONDITIONS

Four types of steady-state conditions prove to be important for the topic of this paper:

(a) If the range is $D = (-\infty, +\infty)$, then the maximum entropy condition defined by a known mean value μ and variance σ^2 is described by the normal distribution $N(\mu, \sigma^2)$, having the density

$$u(x) = [\sigma(2\pi)^{1/2}]^{-1}\, e^{-(x-\mu)^2/(2\sigma^2)}, \qquad -\infty < x < +\infty$$

which maximizes the entropy

$$H(u) = -\int_{-\infty}^{+\infty} u(x) \ln u(x)\, dx$$

subject to the constraints μ and σ^2. In such a case,

$$U_n(x) = (2^n n!)^{-1/2} H_n[(x-\mu)/(\sigma\sqrt{2})], \qquad n = 0, 1, \ldots$$

where $H_n(x)$ is the Hermite polynomial (Abramowitz and Stegun, 1972) of degree n.

(b) If the range is $D = (0, +\infty)$, then the maximum entropy condition corresponding to the given mean value μ is characterized by the exponential distribution $E(\mu)$ having the density $u(x) = (1/\mu)\, e^{-x/\mu}$, $x > 0$. In such a case, $U_n(x) = L_n(x/\mu)$, $n = 0, 1, \ldots$, where $L_n(x)$ is the Laguerre polynomial (Abramowitz and Stegun, 1972) of degree n.

(c) If the range is $D = [a, b]$ and no constraint is imposed, then the maximization of the entropy $H(u)$ is achieved by the uniform distribution $u = U([a, b])$, with the density $u(x) = 1/(b-a)$, $a \leq x \leq b$. In such a case,

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

where $P_n(x)$ is the Legendre (spherical) polynomial (Abramowitz and Stegun, 1972) of degree n.

(d) In general, the logarithmic deviation (Kullback and Leibler, 1951; Guiasu, 1977, 1987) of the probability density u from the reference measure with the density $q > 0$ is

$$D(u\colon q) = \left\langle \ln \frac{u}{q} \middle| u \right\rangle$$

If the reference measure is uniform (i.e., $q = 1$), then

$$D(u\colon q) = D(u\colon 1) = \langle \ln u | u \rangle = -H(u)$$

where $H(u)$ is the entropy of u. In such a case, $\min_u D(u\colon q)$ is equivalent to $\max_u H(u)$. In cases (a)–(c) mentioned above, the reference measure was supposed to be uniform and the maximum entropy probability density u was the closest probability density to the uniform (i.e., Lebesgue) measure subject to the respective constraints given by mean values. If the reference measure is not uniform with respect to the Lebesgue measure and has the density q, then the steady-state probability density u is obtained by minimizing $D(u\colon q)$ subject to the given mean values of X.

If the range is $D=(0, +\infty)$, the constraint is the mean μ, and the reference measure has the density $q(x)=x^\alpha$ $(\alpha>-1)$, then

$$u(x)=\beta^{-(\alpha+1)}[\Gamma(\alpha+1)]^{-1}x^\alpha e^{-x/\beta}, \qquad x>0$$

where $\beta=\mu/(\alpha+1)$, minimizes $D(u: q)$. This is just the density of the gamma distribution with parameters $\alpha+1$ and β, and we write $u=\Gamma(\alpha+1, \beta)$. In this case

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

where $L_n^\alpha(x)$ is the generalized Laguerre polynomial (Abramowitz and Stegun, 1972) of degree $n-\alpha$ and order α.

(e) Of course, there are several orthonormal sets of functions $\mathscr{U}$ with the same weight u. Thus, if $D=[0, a]$ and no constraint is given, the maximum entropy probability distribution is the uniform one $U(0, a)$, with the density $u(x)=1/a$, $0\le x\le a$. Then, both $U_n(x)=\sqrt{2}\cos(n\pi x/a)$ $(n=0, 1, \ldots)$ and $U_n(x)=\sqrt{2}\sin(n\pi x/a)$ $(n=1, 2, \ldots)$, and

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

are complete orthonormal sets of functions with the weight u.

5. THE WAVE EQUATION

A simple calculation based on (7) and (8) shows that the wave function ψ^* satisfies an equation of Schrödinger type

$$(\psi^*)''-(2\pi/h)^2K\psi^*=0 \tag{10}$$

if

$$u^{5/2}U_n''+uu'U_n'+[\tfrac{1}{2}uu''-\tfrac{1}{4}(u')^2-(2\pi/h)^2Ku^{5/2}]U_n=0$$

If $\mathscr{U}=\{U_n, n=0, 1, \ldots\}$ is an orthonormal set of polynomials with the weight u, then U_n satisfies a second-order differential equation (Abramowitz and Stegun, 1972),

$$g_2(x)U_n''(x)+g_1(x)U_n'(x)+g_0(x)U_n(x)=0$$

Thus, ψ^* satisfies the Schrödinger equation if

$$g_2=u^{5/2}, \qquad g_1=uu', \qquad g_0=\tfrac{1}{2}uu''-\tfrac{1}{4}(u')^2-(2\pi/h)^2Ku^{5/2}$$

This is just what happens in the examples from the next two sections.

Let also notice that substituting $\psi^* = \exp(2\pi S/h)$ in (10), where S has the dimensions of an action, we get

$$(h/2\pi)S'' + (S')^2 - K = 0$$

which at the macroscopic level (i.e., taking $h \to 0$) becomes the Hamilton–Jacobi equation

$$(S')^2 - K = 0$$

6. THE HARMONIC OSCILLATOR

Let X be the displacement of a one-dimensional harmonic oscillator. It is a random variable with the range $D = (-\infty, +\infty)$. Let μ and σ^2 be the mean and variance of X. In the steady-state condition corresponding to the mean μ and variance σ^2, according to the special case (a) from Section 4, the maximum entropy probability distribution is $u = N(\mu, \sigma^2)$ and the corresponding orthonormal system with the weight u is

$$U_n(x) = (2^n n!)^{-1/2} H_n[(x-\mu)/(\sigma\sqrt{2})], \qquad n = 0, 1, \ldots$$

where H_n is the Hermite polynomial of degree n. As long as the harmonic oscillator is in a steady-state condition described by u, the mean fluctuations are $\langle U_n | u \rangle = 0$, $n = 1, 2, \ldots$. If random fluctuations of types $U_1, \ldots, U_N$ occur and the mean fluctuations are $c_1, \ldots, c_N$, then the steady-state probability density u is replaced by the probability density f such that $\langle U_n | f \rangle = c_n$, $n = 1, \ldots, N$. According to (6), its approximation given by the minimum Pearson deviation from the steady-state condition u is

$$f^*(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/(2\sigma^2)} \left\{ 1 + \sum_{n=1}^{N} c_n (2^n n!)^{-1/2} H_n[(x-\mu)/(\sigma\sqrt{2})] \right\}$$

According to (7) and (8), the corresponding normed wave function is

$$\psi^*(x) = \left(\sum_{n=1}^{N} c_n^2 \right)^{-1/2} (2\pi\sigma^2)^{-1/4} e^{-(x-\mu)^2/(4\sigma^2)} \times \sum_{n=1}^{N} c_n (2^n n!)^{-1/2} H_n[(x-\mu)/(\sigma\sqrt{2})]$$

In particular, if U_n is the only type of fluctuation, then the normed Pearson function is

$$\psi^*(x) = (2\pi\sigma^2)^{-1/4} e^{-(x-\mu)^2/(4\sigma^2)} (2^n n!)^{-1/2} H_n[(x-\mu)/(\sigma\sqrt{2})]$$

It is well known (Abramowitz and Stegun, 1972) that the function $e^{-x^2/2}H_n(x)$ satisfies the differential equation

$$[e^{-x^2/2}H_n(x)]'' + (2n+1-x^2)\, e^{-x^2/2}H_n(x) = 0$$

from which we get that the wave function ψ_n^* given above is the solution of the equation

$$[\psi_n^*(x)]'' + \sigma^{-2}[(n+\tfrac{1}{2}) - (x-\mu)^2/(4\sigma^2)]\psi_n^*(x) = 0$$

In particular, if the mean $\mu = 0$, we get the equation

$$[\psi_n^*(x)]'' + \sigma^{-2}[(n+\tfrac{1}{2}) - x^2/(4\sigma^2)]\psi_n^*(x) = 0 \tag{11}$$

Notice that up to this point neither the Schrödinger equation nor any quantization rules have been used. The Schrödinger equation for the one-dimensional harmonic oscillator in state n is (Pauling and Wilson, 1935; McQuarrie, 1983)

$$[\psi_n^*(x)]'' + (8\pi^2\theta/h^2)[E_n - \tfrac{1}{2}kx^2]\psi_n^*(x) = 0 \tag{12}$$

where k is the force constant between two masses m_1 and m_2, $\theta = m_1m_2/(m_1+m_2)$ is the reduced mass, E_n is the energy, and h is Planck's constant. Equation (11) becomes the Schrödinger equation (12), assuming the quantization rules

$$\sigma^2 = \frac{h}{4\pi(\theta k)^{1/2}}, \qquad E_n = \frac{h}{2\pi}\left(\frac{k}{\theta}\right)^{1/2}\left(n+\frac{1}{2}\right)$$

The second quantization rule is the standard way of defining the energy of the harmonic oscillator as a function of h, k, θ, and n. The first quantization rule throws a new light on the correspondence principle: if the force k is very strong or/and the masses m_1, m_2 are very big, then the variance σ^2 is negligible and nothing is random in the displacement of the harmonic oscillator; on the other hand, if Planck's constant h is neglected (as at the macroscopic level), then the variance σ^2 is again negligible and the displacement is given by the mean value μ.

7. A FREE PARTICLE IN A ONE-DIMENSIONAL BOX [0, a]

Let X be the location of a particle in the box $[0, a]$. X is a random variable with the range $D = [0, a]$. As the particle is supposed to be free, there are no constraints imposed on X except the range itself. Therefore, according to the special case (c) from Section 4, the steady-state condition is characterized by the uniform distribution $U(0, a)$ having the probability density $u(x)=1/a$, $0\leq x\leq a$, and an orthonormal system with the weight u is

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

In the steady-state condition described by u, the mean fluctuations are $\langle U_n|u\rangle=0$, $n=1,2,\ldots$. If random fluctuations of types $U_1,\ldots,U_N$ occur and the mean fluctuations are $c_1,\ldots,c_N$, then the steady-state probability density u is replaced by the probability density f such that

$$\left\langle \sqrt{2}\sin\frac{n\pi x}{a}\middle| f\right\rangle = c_n, \qquad n=1,\ldots,N$$

Then, with the notations from Section 3, f is approximated by

$$f^*(x)=\frac{1}{a}\left(1+\sqrt{2}\sum_{n=1}^{N} c_n\sin\frac{n\pi x}{a}\right), \qquad 0\leq x\leq a$$

and the corresponding normed wave function is

$$\psi^*(x)=\left(\frac{a}{2}\sum_{m=1}^{N} c_m^2\right)^{-1/2}\sum_{n=1}^{N} c_n\sin\frac{n\pi x}{a}$$

If only $c_n\neq 0$, which means that the random fluctuation have occurred only in the direction U_n, then

$$f^*(x)=\frac{1}{a}\left(1+\sqrt{2}\,c_n\sin\frac{n\pi x}{a}\right)$$

and the normed Pearson function is

$$\psi_n^*=(2/a)^{1/2}\sin\frac{n\pi x}{a}$$

which satisfies the wave equation

$$[\psi_n^*(x)]''+(n^2\pi^2/a^2)\psi_n^*(x)=0, \qquad 0\leq x\leq a \tag{13}$$

Notice that up to this point neither the Schrödinger equation nor any quantization rules have been used. The Schrödinger equation for a free particle of mass m in the box $[0, a]$ is (Pauling and Wilson, 1935; McQuarrie, 1983)

$$[\psi_n^*(x)]''+(8\pi^2 mE_n/h^2)\psi_n^*(x)=0, \qquad 0\leq x\leq a \tag{14}$$

where E_n is the energy of the particle in state n and h is Planck's constant. Equation (13) becomes the Schrödinger equation (14) with the standard quantization rule $E_n = h^2 n^2/(8ma^2)$.

As in any undulatory phenomenon, the presence of the particle in the box is detected by the probability density ψ_n^{*2} of the deviations from the steady-state condition due to random fluctuations of type $\sqrt{2}\sin(n\pi x/a)$.

8. THE HYDROGEN ATOM

The hydrogen atom describes the interaction between two point particles (the proton and the electron) due to the Coulomb attraction of their electrical charges. Using the spherical polar coordinates (R, Θ, Φ), where R is the distance between the electron and the proton fixed at the origin, we obtain that for the Cartesian coordinates

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

the Jacobian of the transformation is $r^2\sin\theta$, where the ranges of R, Θ, and Φ are $0 < r < \infty$, $0 \leq \theta \leq \pi$, and $0 \leq \phi \leq 2\pi$, respectively. The whole interaction is a radial one. Consequently, there are no restrictions imposed on Θ and Φ. As the range of Φ is $[0, 2\pi]$ and there are no constraints on Φ, its steady-state probability density is uniform, $U(0, 2\pi)$, i.e.,

$$u^{(3)}(\phi) = 1/(2\pi), \qquad 0 \leq \phi \leq 2\pi \tag{15}$$

The standard set $\mathcal{U}$ of orthonormal functions with the weight $u^{(3)}$ is the trigonometric system:

$$U_0^{(3)}(\phi) = 1, \qquad U_m^{(3)}(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) \\ \sqrt{2}\sin(m\phi) \end{cases} \qquad (m = 1, 2, \ldots) \tag{16}$$

The range of Θ is $[0, \pi]$. Let us introduce the random variable $V = \cos\Theta$. Its range is $[-1, 1]$. Having no constraints on Θ, the steady-state probability distribution of V is the uniform distribution on $[-1, 1]$ [see the special case (c) in Section 4]. The Jacobian of this transformation is $\sin\theta$. Therefore, the steady-state probability density is

$$\tilde{u}^{(2)}(v) = \tfrac{1}{2}, \qquad -1 \leq v \leq 1$$

and as system $\mathcal{U}$ of normed functions with the weight $\tilde{u}^{(2)}$ we can take (Abramowitz and Stegun, 1972) either the sequence of normed Legendre polynomials $\{(2l+1)^{1/2}P_l(v), l = 0, 1, \ldots\}$ or, more generally, the normed associated Legendre polynomials

$$\left(\frac{(2l+1)(l-t)!}{(l+t)!}\right)^{1/2} P_l^t(v), \qquad l=0,1,\ldots, \qquad 0\leq t\leq l$$

Returning to the range $[0, \pi]$ of Θ, we obtain for the steady-state probability density

$$u^{(2)}(\theta)=\tilde{u}^{(2)}(\cos\theta)\sin\theta=\tfrac{1}{2}\sin\theta, \qquad 0\leq\theta\leq\pi \tag{17}$$

and the corresponding orthonormal system $\mathscr{U}$ with the weight $u^{(2)}(\theta)$ is

$$U_{l,t}^{(2)}(\theta)=\left(\frac{(2l+1)(l-t)!}{(l+t)!}\right)^{1/2} P_l^t(\cos\theta), \qquad l=0,1,\ldots, \qquad 0\leq t\leq l \tag{18}$$

The orthogonality of the sequence (18) is with respect to the subscript l.

The range of R is $[0, \infty)$ and let μ be the mean value of R estimated at the macroscopic scale. In describing the random variable R we are looking for the most random probability distribution on $[0, \infty)$ subject to the mean value μ. The reference measure on $[0, \infty)$ has as density the r-component of the Jacobian mentioned above, i.e., r^2. Therefore, we are looking for the closest probability density $u^{(1)}(r)$ to r^2 on $[0, \infty)$ subject to the mean value μ, where closeness is measured by the Kullback–Leibler divergence from statistical inference. Taking $\alpha=2$ in the result mentioned in paragraph (d) of Section 4, we find that the solution of this variational problem is the gamma distribution $\Gamma(3, \mu/3)$, with parameters 3 and $\mu/3$, having the density

$$u^{(1)}(r)=\frac{\beta^3}{2} r^2 e^{-\beta r}, \qquad r\geq 0 \tag{19}$$

with $\beta=3/\mu$. The variance is $\sigma^2=\mu^2/3$. As a set $\mathscr{U}$ of normed functions with the weight $u^{(1)}$ we can take either

$$U_0^{(1)}(r)=1, \qquad U_k^{(1)}(r)=\left(\frac{(k-1)!}{k(k!)^3}\right)^{1/2} L_k^1(\beta r), \qquad k=1,2,\ldots$$

or the richer set

$$U_{0,0}^{(1)}(r)=1, \qquad U_{k,s}^{(1)}(r)=\left(\frac{2\beta^{s-1}(k-s)!}{(k!)^3(2k-s+1)}\right)^{1/2} r^{(s-1)/2} L_k^s(\beta r) \tag{20}$$

with $s=1, \ldots, k$; $k=1, 2, \ldots$; where the associate Laguerre polynomial of degree $k-s$ and order s is

$$L_k^s(x)=\frac{d^s}{dx^s}\left(e^x \frac{d^k}{dx^k}(x^k e^{-x})\right)$$

The orthogonality of the sequence (20) is with respect to the subscript k.

As the random variables R, Θ, and Φ are independent, the joint steady-state probability density is, according to what was discussed in the last part of Section 3, $u(r, \theta, \phi)=u^{(1)}(r)u^{(2)}(\theta)u^{(3)}(\phi)$. In order to allow a subsequent dependence between the systems of orthogonal functions (16), (18), and (20) we can assume any relationship $s=i(l, m)$, $t=j(k, m)$, where i and j are arbitrary nonnegative integer-valued functions. As long as the system is in a steady-state condition characterized by the density u, the mean fluctuation of type $U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)$ is equal to zero. If due to random fluctuations of this type, the system gets out of the steady state characterized by u, then the new probability density f for which some covariances $\langle U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)|f\rangle=c_{klm}$ are known to be different from zero may be approximated by (see Section 3)

$$f^*(r, \theta, \phi)=u(r, \theta, \phi)\left[1+\sum_{k,l,m} c_{klm}U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)\right]$$

If c_{klm} is the only correlation known to be different from zero, then

$$f^*(r, \theta, \phi)=u(r, \theta, \phi)[1+c_{klm}U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)]$$

and the corresponding normed wave function is

$$\psi_{klm}^*(r, \theta, \phi)=[u(r, \theta, \phi)]^{1/2}U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)$$

As $s=i(l, m)$ and $t=j(k, m)$, if $m\neq m'$, then $\langle\psi_{klm}^*|\psi_{k'l'm'}^*\rangle=0$ due to the orthogonality of the functions $U_m^{(3)}$ with respect to m. If $m=m'$ and $l\neq l'$, then $\langle\psi_{klm}^*|\psi_{k'l'm}^*\rangle=0$ due to the orthogonality of the functions $U_{l,t}^{(2)}$ with respect to l. Finally, if $m=m'$, $l=l'$, but $k\neq k'$, then $\langle\psi_{klm}^*|\psi_{k'lm}^*\rangle=0$ due to the orthogonality of the functions $U_{k,s}^{(1)}$ with respect to k.

Now, $\psi_{klm}^{*2}(r, \theta, \phi)$ may be interpreted as being the probability density induced by the minimum mean relative square deviation from the steady-state condition due to random fluctuations of type $U_{k,s}^{(1)}(r)U_{l,t}^{(2)}(\theta)U_m^{(3)}(\phi)$ with known mean. The larger the deviation from the steady-state condition in a subdomain, the larger is the probability of having the electron located in that subdomain.

Up to this point neither the Schrödinger equation nor any quantization rules have been used. The classical results are obtained if we assume that the mean value μ is equal to $\mu=\frac{3}{2}na_0$, or, equivalently, if the most probable radial value [i.e., the value of r for which the steady-state probability density (19) is maximum] is $r_{mp}=na_0$, where n is the total (or principal) quantum number and a_0 is the Bohr radius or the distance of the electron from the nucleus in the first (circular) orbit. In such a case $\beta=2/r_{mp}=3/\mu=2/na_0$ and, as a common practice, k and s in (20) are replaced by $n+l$ and $2l+1$, respectively, while t from (18) and m from (16) are replaced by $|m|$. In such a case, l is the azimuthal (or angular momentum) quantum number, m is the magnetic quantum number, and the possible values of the quantum numbers are $n=1, 2, \ldots$; $l=0, 1, \ldots, n-1$; $m=0, \pm 1, \ldots, \pm l$. Therefore, the whole quantum mechanics formalism for the hydrogen atom is obtained from minimizing the mean χ^2 deviation from the steady-state condition of maximum uncertainty subject to the only quantization rule according to which the most probable radial value of the electron orbit corresponding to the principal quantum number n is na_0, where a_0 is the Bohr radius or the minimum distance of the electron from the nucleus in the old quantum mechanics.

9. A NONSTANDARD EXAMPLE

The cases analyzed above are standard. Let us give the nonstandard example of determining the location X of a particle on $[1, \infty)$ knowing only the mean location $\mu=1.7137$, estimated at the macroscopic scale, when the reference measure on $[1, \infty)$ has the density x^{-2}, which means that due to an existing field on $[1, \infty)$ we expect the probability distribution of X to be similar to x^{-2}. Following the same steps explained in Section 3, the steady-state probability density u of X is obtained by minimizing the Kullback–Leibler divergence from the reference measure with density x^{-2} subject to the mean $\mu=1.7137$. Using the Lagrange multiplier technique, we get

$$u(x)=[E_2(\beta)]^{-1}x^{-2}\,e^{-\beta x}, \qquad x\geq 1$$

where

$$E_n(z)=\int_1^{\infty} t^{-n}\,e^{-zt}\,dt, \qquad n=0, 1, \ldots; \quad z>0$$

is the exponential integral and β the solution of the exponential equation $\mu=E_1(\beta)/E_2(\beta)$, which for $\mu=1.7137$ gives $\beta=0.5$. From tables for the exponential integral (Abramowitz and Stegun, 1972), $E_1(0.5)=0.5597736$ and $E_2(0.5)=0.3266439$. Thus,

$$u(x)=3.061437853x^{-2}\,e^{-x/2}, \qquad x\geq 1 \tag{21}$$

The Gram–Schmidt orthogonalization technique may be applied for getting the sequence of orthonormal polynomials with the weight (21). The first ones are (see Figure 1):

$$U_0(x)=1, \qquad U_1(x)=1.1346x-1.9443$$
$$U_2(x)=0.4421x^2-2.7184x+3.0166$$
$$U_3(x)=0.0994x^3-1.4097x^2+4.9331x-4.3259$$

As long as the system remains in the steady state described by u, the mean values $\langle U_n|u\rangle$ $(n\geq 1)$ are equal to zero. If random fluctuations occur along the "direction" U_2, for instance, the mean value $c_2=\langle U_2|u\rangle$, estimated at the macroscopic scale by the sample mean

$$N^{-1}\sum_{i=1}^{N} U_2(x_i)$$

where N is the sample size, is no longer equal to zero. The normed wave function of the system corresponding to the random fluctuation of type U_2 is $\phi_2(x)=[u(x)]^{1/2}U_2(x)$ and the probability density of the deviations from the steady-state condition due to the random fluctuations of type U_2 is $\Phi_2^2(x)$. The larger the deviation from the steady-state condition in an arbitrary subinterval of $[1, \infty)$, the larger is the probability of having the particle located in that subinterval. The graphs of u (marked with stars), $\phi_1=\sqrt{u}\,U_1$ (marked with squares), and ϕ_1^2 (marked with triangles) are

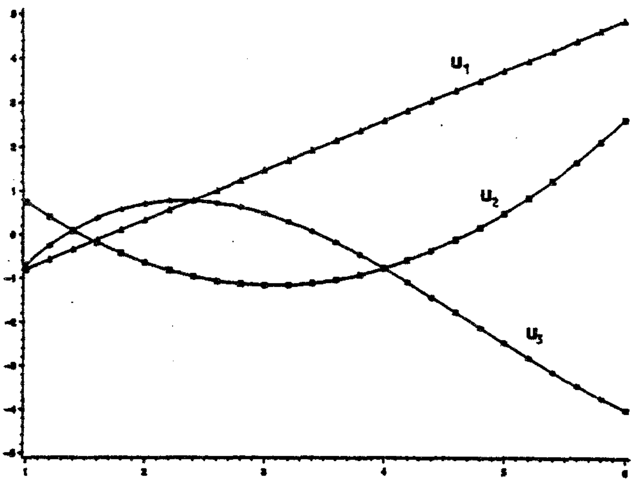

Fig. 1. The first three orthogonal polynomials.

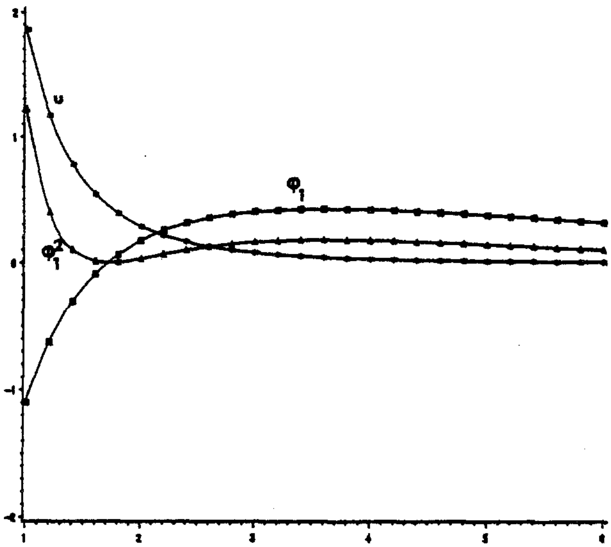

Fig. 2. The graphs of u, ϕ_1, and ϕ_1^2.

shown in Figure 2. The similar graphs of u, $\phi_2 = \sqrt{u}\, U_2$, and ϕ_2^2 are given in Figure 3.

10. CONCLUSION

It is almost unanimously agreed that the solution of the Schrödinger equation, considered as a fundamental postulate of quantum mechanics, is a probability wave. This wave function describes a physical system in the sense that it gives information concerning the probabilities of the results of various observations which can be made on the system. In this paper the probability wave function is not deduced from the Schrödinger equation, but from a variational problem involving the minimization of the mean relative square deviation used by Pearson in statistical inference at the beginning of this century. Instead of starting from the Schrödinger equation and eventually interpreting the square of the module of its solution as being a probability density, an interpretation Schrödinger himself never fully agreed with (Mehra and Rechenberg, 1987), we estimate the most unbiased probability density subject to the mean values and mean fluctuations accessible to us through the measurement process and eventually get the Schrödinger equation as a consequence. A system in a steady-state condition is described by a probability density u depending only on a set of mean values determined by the measurement process. Such a probability density is usually determined

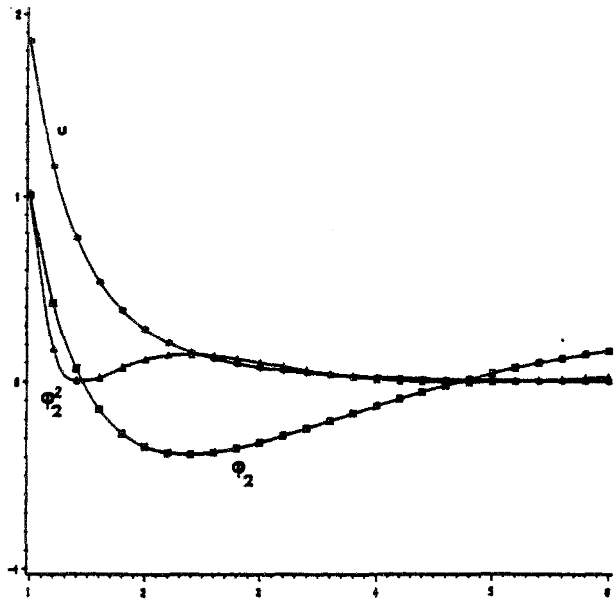

Fig. 3. The graphs of u, ϕ_2, and ϕ_2^2.

by applying the principle of maximum entropy or, more generally, the principle of minimum divergence from a reference measure, subject to the known mean values that characterize the respective steady-state condition. We attach to u a finite or countable set $\mathscr{U}=\{U_n, n=0, 1, \ldots\}$ of orthonormal functions with the weight u, where $U_0=1$. As long as the system remains in the steady-state condition described by u, the mean fluctuations of type U_n ($n=1, 2, \ldots$) are all equal to zero, i.e., $\langle U_n|u\rangle=0$ ($n=1, 2, \ldots$). Suppose that due to random fluctuations the system gets out of the steady-state condition and the previous probability density u has to be replaced by a new probability density f compatible with the new mean fluctuations $\langle U_n|f\rangle = c_n \neq 0$ ($n=1, \ldots, N$) obtained at the macroscopic scale. We approximate f by the solution f^* of the variational problem $\min\langle\chi^2(f: u)\rangle$ subject to the mean fluctuations $\langle U_n|f\rangle=c_n$ ($n=1, \ldots, N$), where $\langle\chi^2(f: u)\rangle$ is the Pearson mean relative square deviation of f from u defined by (2). The wave function ψ^* of the system is just the normed optimum Pearson function $\chi(f^*: u)$, and its square ψ^{*2} is naturally interpreted as the probability density of minimum mean relative square deviation from the steady-state condition described by u, due to random fluctuations of type U_n, $n=1, 2, \ldots$. The formalism is applied to the harmonic oscillator, the free particle in a one-dimensional box, and the hydrogen atom. The standard results are obtained from the above formalism by using only classical quantization rules. The quantization rules seem to be the bridge between the unbiased probabilistic model built up and some physical characteristics (energy, for instance) of the quantum system involved.

ACKNOWLEDGMENT

Support from the Natural Sciences and Engineering Research Council (NSERC Canada) Grant No. A-5712 is acknowledged.

REFERENCES

Abramowitz, M., and Stegun, I. A., eds. (1972). *Handbook of Mathematical Functions*, Dover, New York, pp. 773–783.
Baublitz, Jr., M. (1988). *Progress of Theoretical Physics*, **80**, 232.
Bohm, D. (1984). *Causality and Chance in Modern Physics*, 3rd ed., University of Pennsylvania Press, Philadelphia, Pennsylvania.
Fisher, R. A. (1947). *The Design of Experiments*, Oliver and Boyd, Edinburgh.
Frieden, B. R. (1989). *American Journal of Physics*, **57**, 1004.
Guiasu, S. (1977). *Information Theory with Applications*, McGraw-Hill, New York.
Guiasu, S. (1987). *Physical Review A*, **36**, 1971.
Guiasu, S., and Shenitzer, A. (1985). *Mathematical Intelligencer*, **7**, 42.
Jaynes, E. T. (1957). *Physical Review*, **106**, 620.
Justice, J. H., ed. (1986). *Maximum Entropy and Bayesian Methods in Applied Statistics*, Cambridge University Press, Cambridge.
Kullback, S., and Leibler, R. A. (1951). *Annals of Mathematical Statistics*, **32**, 79.
Levine, R. D., and Tribus, M., eds. (1979). *The Maximum Entropy Formalism*, MIT Press, Cambridge, Massachusetts.
McQuarrie, D. A. (1983). *Quantum Chemistry*, University Science Books, Mill Valley, California.
Mehra, J., and Rechenberg, H. (1987). *The Historical Development of Quantum Theory*, Springer-Verlag, New York, Vol. 5, Part 2, pp. 827–828.
Nelson, E. (1985). *Quantum Fluctuations*, Princeton University Press, Princeton, New Jersey.
Pauling, L., and Wilson, Jr., E. B. (1935). *Introduction to Quantum Mechanics*, McGraw-Hill, New York.
Pearson, K. (1900). *Philosophical Magazine* (5th Series), **50**, 157.
Polkinghorne, J. C. (1986). *The Quantum World*, Penguin Books, Middlesex.
Skilling, J., ed. (1989). *Maximum Entropy and Bayesian Methods*, Kluwer, Dordrecht, Holland.
Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin.