Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information

Marcel Reginatto

U.S. Department of Energy, Environmental Measurements Laboratory, New York, New York 10014-4811

(Received 2 March 1998)

The many-particle time-dependent Schrödinger equation is derived using the principle of minimum Fisher information. This application of information theory leads to a physically well motivated derivation of the Schrödinger equation, which distinguishes between subjective and objective elements of the theory. [S1050-2947(98)07809-3]

PACS number(s): 03.65.Bz, 03.65.Ca, 89.70.+c

INTRODUCTION

The main result of this paper is a derivation of the many-particle time-dependent Schrödinger equation using the principle of minimum Fisher information. There are two basic assumptions that enter into this derivation: that one can associate a wave front with the motion of particles, and that the probability distribution that describes the position of particles should satisfy the principle of minimum Fisher information. This information-theoretic approach is of interest in that it provides a physically well motivated derivation of the Schrödinger equation that isolates the subjective and objective aspects of the formalism. From this point of view, the epistemological content of the theory lies in the prescription to minimize the Fisher information associated with the probability distribution that describes the position of particles. The physical content of the theory is contained in the assumption that one can associate a wave front with the motion of particles. Although the principle of minimum Fisher information has been used before to arrive at the equations of nonrelativistic quantum mechanics (see, for example, the discussion in Frieden and Soffer [1], and references therein), these previous derivations of the Schrödinger equation can be criticized in that they only permit solutions that are real and time independent. The derivation presented here leads directly to the time-dependent Schrödinger equation for a complex wave function.

I start with a brief discussion of the Fisher information and the Fisher information matrix. Consider the problem of estimating a parameter θ in the presence of unknown added noise x. A measurement y of the parameter will be related to x and θ by

$$y=\theta+x. \tag{1}$$

For example, y might be a measurement of the position of a particle, while θ is the actual position of the particle. The probability distribution $P(x)$ for the noise x will be related to $P(y|\theta)$ by

$$P(y|\theta)=P(y-\theta)=P(x), \tag{2}$$

where the first equality is a consequence of the invariance of $P(y|\theta)$ under translations. In this case, the Fisher information I is given by [2–4]

$$I\equiv\int \frac{1}{P(y|\theta)}\left(\frac{\partial P(y|\theta)}{\partial\theta}\right)^2 dy=\int \frac{1}{P(x)}\left(\frac{dP(x)}{dx}\right)^2 dx. \tag{3}$$

It can be shown that the mean-square error in any unbiased estimate of θ must exceed $1/I$, which is known as the Cramer-Rao bound [3,4].

The principle of minimum Fisher information asserts that one should choose the probability distribution which minimizes the Fisher information subject to the constraints known about the system. Note that the probability distribution that minimizes the Fisher information will also maximize the mean square error. In this sense, it will be as noninformative as possible while still satisfying the constraints.

The Fisher information matrix I_{kl} is a generalization of the Fisher information to the case where the probability distribution is a function of an n-dimensional parameter θ^i and an n-dimensional vector random variable x^i. When $y^i=\theta^i+x^i$, the Fisher information matrix can be written as

$$I_{kl}\equiv\int P(y^i|\theta^i)\left(\frac{\partial}{\partial\theta^k}[\ln P(y^i|\theta^i)]\,\frac{\partial}{\partial\theta^l}[\ln P(y^i|\theta^i)]\right)d\mu(y^i)=\int P(x^i)\left(\frac{\partial}{\partial x^k}[\ln P(x^i)]\,\frac{\partial}{\partial x^l}[\ln P(x^i)]\right)d\mu(x^i). \tag{4}$$

The relationship that exists between the mean-square error and the Fisher information can be extended to an inequality relating the covariance matrix and the inverse of the Fisher information matrix [5].

The Fisher information is closely related to the Kullback discrimination information (also known as the cross entropy or Kullback-Liebler distance). Suppose that θ^i and $\theta^i+\Delta\theta^i$ are neighboring points in the parameter space. Then it can be shown [5] that under certain regularity conditions and to within the second-order terms,

$$D(\theta^i:\theta^i+\Delta\theta^i)\equiv\int P(y^i|\theta^i)\ln\left(\frac{P(y^i|\theta^i)}{P(y^i|\theta^i+\Delta\theta^i)}\right)d\mu(y^i)$$

$$=\tfrac{1}{2}\sum_{j,k} I_{jk}\Delta\theta^j\Delta\theta^k. \qquad (5)$$

Because of this relationship, the Fisher information and the Kullback discrimination information share similar properties.

DERIVATION OF THE SCHRÖDINGER EQUATION FROM THE PRINCIPLE OF MINIMUM FISHER INFORMATION

I first derive the Schrödinger equation for the case of a single free particle of mass m, and later generalize to the case of many particles moving in a potential.

Consider the case of an ensemble of particles described by a normalized probability density $P(x^i,t)$. Assume that the set of particle trajectories forms a coherent system, or equivalently, that one can associate a wave front with the motion of the particles. Then the velocity vector $v^j(x^i,t)$ of a particle at point x^i can be related to a real function $S(x^i,t)$ by an expression of the form

$$v^j=\sum_{k=1}^{3} g^{jk}\frac{\partial S}{\partial x^k}, \qquad (6)$$

where the inverse metric $g^{jk}=\mathrm{diag}(1/m,1/m,1/m)$ is the one commonly used to define the kinematical line element in configuration space [6]. It follows that the probability distribution must satisfy a conservation law of the form

$$\frac{\partial P}{\partial t}+\sum_{i,k=1}^{3} g^{ik}\frac{\partial}{\partial x^i}\left(P\frac{\partial S}{\partial x^k}\right)=0. \qquad (7)$$

Equation (7) can be derived from a variational principle, by minimization of the expression

$$\Phi_A=\int P\left(\frac{\partial S}{\partial t}+\tfrac{1}{2}\sum_{i,k=1}^{3} g^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k}\right)d^nx\,dt \qquad (8)$$

with respect to S.

Note that variation of Φ_A with respect to P leads trivially to the Hamilton-Jacobi equation for the free particle. Therefore, minimization of Φ_A and with respect to both S and P will lead to the equations of motion of an ensemble of particles in classical mechanics. There is still considerable freedom in the choice of probability density that can be used to describe the ensemble, since it is only subject to Eq. (7). To arrive at the equations of quantum mechanics, one needs to restrict the choice of probability densities using the principle of minimum Fisher information.

One can define the amount of information in P using the Fisher information matrix. Without introducing additional structure, there is only one natural definition of the amount of information in P, obtained by contracting the metric g^{ik} with the elements of the Fisher information matrix,

$$\Phi_B=\sum_{i,k=1}^{3} g^{ik}\int\frac{1}{P}\frac{\partial P}{\partial x^i}\frac{\partial P}{\partial x^k}d^nx\,dt=\sum_{i,k=1}^{3} g^{ik}I_{ik}. \qquad (9)$$

Note that the asymmetry between space and time in Φ_B is a consequence of the different role that they play in nonrelativistic mechanics. Such a distinction is not present in the derivation of the Klein-Gordon equation (see the Appendix).

I now consider the consequences of applying the principle of minimum Fisher information with the constraint that Eq. (7) must be satisfied. While not the only way, the simplest way to ensure consistency is to minimize (with respect to both P and S) a function Φ that is a linear combination of Φ_A and Φ_B,

$$\Phi=\Phi_A+\lambda\Phi_B=\int P\left(\frac{\partial S}{\partial t}+\frac{1}{2}\sum_{i,k=1}^{3} g^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k}\right)d^nx\,dt$$

$$+\lambda\sum_{i,k=1}^{3} g^{ik}\int\frac{1}{P}\frac{\partial P}{\partial x^i}\frac{\partial P}{\partial x^k}d^nx\,dt. \qquad (10)$$

The parameter λ, which is assumed fixed, has units of action squared. It determines the relative weight of the two terms that enter into the minimization.

For variations that vanish at the boundary, the variation of Φ with respect to S and P leads to the two equations

$$\frac{\partial P}{\partial t}+\sum_{i,k=1}^{3} g^{ik}\frac{\partial}{\partial x^i}\left(P\frac{\partial S}{\partial x^k}\right)=0, \qquad (11)$$

$$\frac{\partial S}{\partial t}+\sum_{i,k=1}^{3}\tfrac{1}{2}g^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k}-\lambda\sum_{i,k=1}^{3} g^{ik}\left(\frac{2}{P}\frac{\partial^2 P}{\partial x^i\partial x^k}\right.$$

$$\left.-\frac{1}{P^2}\frac{\partial P}{\partial x^i}\frac{\partial P}{\partial x^k}\right)=0. \qquad (12)$$

Equations (11) and (12) are identical to the free-particle Schrödinger equation,

$$i\hbar\frac{\partial\psi}{\partial t}=\frac{-\hbar^2}{2}\sum_{i,k=1}^{3} g^{ik}\frac{\partial^2\psi}{\partial x^i\partial x^k} \qquad (13)$$

provided that the parameter λ is set equal to

$$\lambda=\frac{\hbar^2}{8} \qquad (14)$$

and the wave function ψ is expressed in terms of P and S as

$$\psi=P^{1/2}\exp(iS/\hbar). \qquad (15)$$

In the limit where $\hbar\to 0$, Eq. (12) becomes the classical Hamilton-Jacobi equation for the free particle. The generalization to the Schrödinger equation for n particles in a potential is straightforward, in that it only requires modifying Φ by the addition of a potential term $V(x_i)$, which leads to

$$\Phi_S=\int P\left(\frac{\partial S}{\partial t}+\frac{1}{2}\sum_{i,k=1}^{3n} g^{ik}\,\frac{\partial S}{\partial x^i}\,\frac{\partial S}{\partial x^k}+V\right)d^{3n}x\;dt$$
$$+\lambda\sum_{i,k=1}^{3n} g^{ik}\int\,\frac{1}{P}\,\frac{\partial P}{\partial x^i}\,\frac{\partial P}{\partial x^k}\,d^{3n}x\;dt \tag{16}$$

and extending the inverse metric g^{ik} to the case of the kinematical line element in a configuration space of dimension $3n$. Variations of Φ_S with respect to S and P leads to the two equations

$$\frac{\partial P}{\partial t}+\sum_{i,k=1}^{3n} g^{ik}\,\frac{\partial}{\partial x^i}\left(P\,\frac{\partial S}{\partial x^k}\right)=0, \tag{17}$$

$$\frac{\partial S}{\partial t}+\frac{1}{2}\sum_{i,k=1}^{3n} g^{ik}\,\frac{\partial S}{\partial x^i}\,\frac{\partial S}{\partial x^k}+V(x)$$
$$-\lambda\sum_{i,k=1}^{3n} g^{ik}\left(\frac{2}{P}\,\frac{\partial^2 P}{\partial x^i\partial x^k}-\frac{1}{P^2}\,\frac{\partial P}{\partial x^i}\,\frac{\partial P}{\partial x^k}\right)=0. \tag{18}$$

The Schrödinger equation appears in the form of two coupled nonlinear differential equations, which are identical to the equations used by Bohm in his formulation of quantum mechanics [7]. Bohm interprets Eq. (18) as a Hamilton-Jacobi equation, and assumes that the last term on the left-hand side (which he calls the quantum potential Q) is part of the potential that acts on the particles. This is strange, in that it forces the potential acting on the particles to depend on the probability assignment used to infer their positions, with the result that in his formulation the potential is a mixture of both ontological elements (the potential function V) and epistemological elements (the probability assignment P). This suggests that Bohm's quantum potential is not a property of the system (and therefore not part of the ontology of quantum mechanics in the sense of Jaynes [8,9]), but a consequence of the process of inference used here. It is perhaps remarkable that there is a connection between the average of the quantum potential Q and the Fisher information:

$$\int\;PQ\;d^{3n}x\;dt=-\frac{\hbar^2}{8}\sum_{i,k=1}^{3n} g^{ik}\int P\left(\frac{2}{P}\,\frac{\partial^2 P}{\partial x^i\partial x^k}-\frac{1}{P^2}\,\frac{\partial P}{\partial x^i}\,\frac{\partial P}{\partial x^k}\right)d^{3n}x\;dt=\frac{\hbar^2}{8}\sum_{i,k=1}^{3n} g^{ik}\int\;\frac{1}{P}\,\frac{\partial P}{\partial x^i}\,\frac{\partial P}{\partial x^k}\,d^{3n}x\;dt. \tag{19}$$

The average value of the quantum potential introduced by Bohm is proportional to the Fisher information.

CONCLUDING REMARKS

The approach used here to derive the equations of nonrelativistic quantum mechanics parallels in some ways the information-theoretic approach to statistical mechanics initiated by Jaynes [10], which considers statistical mechanics to be a form of inference. In his formulation, the probability distribution function describing an ensemble of systems subject to constraints (such as mean energy, mean number of particles, etc.) is derived using the maximum entropy principle, which asserts that the probability distribution that has the maximum statistical entropy subject to whatever is known provides the most unbiased representation of our knowledge of the state of the system. For example, given a constraint on the mean energy of the system, the principle of maximum entropy leads to the well-known canonical distribution function. The information-theoretic approach to statistical mechanics leads to a formulation that maintains a sharp distinction between the ontological and epistemological contents of the theory. From this point of view, the maximization of entropy is not an application of a law of physics, but a method of inference. The physical content is in the constraints that are enforced.

The derivation of the Schrodinger equation presented here seems to indicate that one can also view quantum mechanics as a form of inference. In this case, the epistemological content of the theory lies in the prescription to minimize the Fisher information, while the physical content of the theory lies in assuming the relationship in Eq. (6), which leads to the constraint expressed in Eq. (7). There are, however, important differences with respect to the procedure used in the information-theoretic approach to statistical mechanics. The constraint on the probability density is not derived from measurements on the system, but is due to assuming *a priori* that Eq. (7) holds. Furthermore, the main concern here is with a probability assignment that is optimal with respect to measurements of the position of particles. Position and momentum are not on the same footing, in the sense that position is considered the fundamental variable, and the probability assignment is one in configuration space, as opposed to a probability assignment in phase space. Possible modifications of quantum mechanics can by arrived at by modifying Φ_S in Eq. (16), by replacing the Fisher information by a different measure of information, or by modifying the nature of the constraint used in the minimization.

ACKNOWLEDGMENTS

I wish to thank Harold Beck, Engelbert Schucking, and Peter Shebell for helpful discussions.

APPENDIX: DERIVATION OF THE KLEIN-GORDON EQUATION

The derivation of the Klein-Gordon equation is similar to the derivation of the Schrodinger equation. It can be derived from minimizing (with respect to both P and S) a function Φ_{KG} given by

$$\Phi_{\mathrm{KG}}=\int P\left(\sum_{i,k=0}^{4}\gamma^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k}+mc^2\right)d^4x$$

$$+\lambda\sum_{i,k=0}^{4}\gamma^{ik}\int\frac{1}{P}\frac{\partial P}{\partial x^i}\frac{\partial P}{\partial x^k}d^4x, \tag{A1}$$

where $\gamma^{ik}=\mathrm{diag}\,(-1/m,1/m,1/m,1/m)$. The equations that result from the minimization of Φ_{KG} take the form

$$\sum_{i,k=0}^{4}\gamma^{ik}\frac{\partial}{\partial x^i}\left(P\frac{\partial S}{\partial x^k}\right)=0, \tag{A2}$$

$$\sum_{i,k=0}^{4}\gamma^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k}+mc^2-\lambda\sum_{i,k=0}^{4}\gamma^{ik}$$

$$\times\left(\frac{2}{P}\frac{\partial^2 P}{\partial x^i\partial x^k}-\frac{1}{P^2}\frac{\partial P}{\partial x^i}\frac{\partial P}{\partial x^k}\right)=0. \tag{A3}$$

Equations (A2) and (A3) are identical to the Klein-Gordon equation,

$$\sum_{i,k=0}^{4}\gamma^{ik}\frac{\partial^2\Psi}{\partial x^i\partial x^k}-\frac{mc^2}{\hbar^2}\Psi=0 \tag{A4}$$

provided that the parameter λ is set equal to

$$\lambda=\frac{\hbar^2}{4} \tag{A5}$$

and the wave function ψ is expressed in terms of P and S as

$$\psi=P^{1/2}\exp(iS/\hbar). \tag{A6}$$

[1] B. R. Frieden and Soffer, Phys. Rev. E **52**, 2274 (1995).
[2] R. A. Fisher, Proc. Cambridge Philos. Soc. **22**, 700 (1925).
[3] B. R. Frieden, J. Mod. Opt. **35**, 1297 (1988).
[4] B. R. Frieden, Am. J. Phys. **57**, 1004 (1989).
[5] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959; corrected and revised edition, Dover Publications, Inc., New York, 1968).
[6] J. L. Synge, *Classical Dynamics*, in *Encyclopedia of Physics*, Vol. III/1, edited by S. Flugge (Springer-Verlag, Berlin, 1960).
[7] D. Bohm, Phys. Rev. **85**, 166 (1952); **85**, 180 (1952).
[8] E. T. Jaynes, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer, Dordrecht, 1989).
[9] E. T. Jaynes, in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Redwood City, CA, 1990).
[10] E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).