International Journal of Theoretical Physics, Vol. 30, No. 7, 1991

Nonlinear Wave Mechanics, Information Theory, and Thermodynamics

James D. Brasher[1]

Received October 18, 1990

A logarithmic nonlinear term is introduced in the Schrödinger wave equation, and a physical justification and interpretation are provided within the context of information theory and thermodynamics. From the resulting nonlinear Schrödinger equation for a system at absolute temperature $T > 0$, the energy equivalence, $kT \ln 2$, of a bit of information is derived.

1. INTRODUCTION

The Lagrangian formulation of quantum field theory and the Hamiltonian formulation of quantum mechanics in general use today are both largely phenomenological descriptions of the dynamics of particles and fields. Interactions are represented by adding the corresponding terms to the Lagrangian or Hamiltonian operator, as appropriate. In the Hamiltonian of the Schrödinger wave equation, these are generally independent of the wave function; thus, the familiar linear Schrödinger equation of nonrelativistic quantum mechanics. However, various model potentials used in the Schrödinger equation, such as the Thomas–Fermi approximation (Slater, 1968), Hartree-type potentials (Efinger, 1973, 1984), and certain hydrodynamic analogies (Gridnev *et al.*, 1983), lead to nonlinear wave equations.

Nonlinear generalizations (Weinberg, 1989*a*) and extensions (Bialynicki-Birula and Mycielski, 1975*a*,*b*, 1976; Delion *et al.*, 1978; Efinger, 1984; Hefter, 1985*a*) of nonrelativistic quantum mechanics and relativistic quantum field theory (Kibble, 1978; Morris, 1978; Ventura and Marques, 1978) have been proffered. Several forms of nonlinearity have been examined in these extensions and a thorough study of the logarithmic nonlinearity in particular has been made by Bialynicki-Birula and Mycielski

---

[1] Teledyne Brown Engineering, Optical Systems Department, Cummings Research Park, Huntsville, Alabama 35807-7007.

0020-7748/91/0700-0979$06.50/0

(1975*b*, 1976). In this paper, it is demonstrated that such a term can be ascribed a physical significance. Its introduction into the Schrödinger wave equation is thus justified on physical grounds and it arises quite naturally, not merely as an ad hoc mathematical facility. It is interpreted within the context of information theory and thermodynamics. As a consequence, the energy equivalence of a bit of information is derived.

2. THEORY

Imagine a subsystem of N particles, which is part of, and in contact with, a large closed system at a fixed temperature, $T > 0$. A good example is a computer in its operating environment. It is assumed that information and energy may be exchanged between the subsystem and its surroundings, e.g., by thermal contact or other means.

The time-independent Schrödinger equation for the subsystem in a quantum state $|\Psi\rangle$, whose coordinate representation is the wavefunction Ψ and whose energy eigenvalue is E_0, is

$$H_0|\Psi\rangle = E_0|\Psi\rangle \tag{1}$$

in which the Hamiltonian operator H_0 has the coordinate representation

$$H_0 = -\sum_{j=1}^{N} \frac{1}{2m_j} \nabla_j^2 + U \tag{2}$$

in atomic units. It consists of the usual kinetic energy operator (for particles of masses m_j) and a potential function U. One includes in this as many terms as are required to account for all the interactions between the particles and fields comprising the subsystem. Such terms are generally wavefunction-independent, resulting in a linear Schrödinger equation.

However, the interactions do depend on the distribution of particles in configuration space. This distribution, described by $|\Psi|^2$, contains information and, consequently, requires the expenditure or, at least, the transfer of energy to be established. Thus, there is some potential energy stored as information in $|\Psi\rangle$ or in the measurement of $|\Psi\rangle$. The probability density of a given configuration is proportional, within a normalization factor, to $|\Psi|^2$. The information (Shannon and Weaver, 1949) I acquired upon measurement of the state $|\Psi\rangle$ is proportional to the logarithm of the probability that a measurement results in the outcome $|\Psi\rangle$, or

$$I = -\log_2(a|\Psi|^2) \quad \text{bits} \tag{3}$$

The parameter $a > 0$ can be tentatively regarded as a normalization factor, though it will subsequently become apparent that it has a physical significance. Suppose that the measurement of the state $|\Psi\rangle$ produces I bits of information according to equation (3) and that an amount ε of energy per bit is expended, transferred, stored, or otherwise associated with this information or with the particular organization of the matter in which the information is encoded. Then, accordingly, there should be an additional term εI added to the Hamiltonian in equation (2). Thus,

$$H = H_0 + \varepsilon I = -\sum_{j=1}^{N} \frac{1}{2m_j} \nabla_j^2 + U - \varepsilon \log_2(a|\Psi|^2) \tag{4}$$

is the system Hamiltonian, which now includes the potential energy εI associated with the information encoded or stored in the distribution of matter described by the probability density $|\Psi|^2$.

Indeed, energy is the medium of generation, transport, and storage of information. Energy is required to measure, arrange, and maintain any distribution of matter that encodes information. Equation (4) is a reflection of this fact, which is the physical origin of the logarithmic nonlinearity in the Schrödinger equation.

The expectation value of the Hamiltonian of equation (4) in the state $|\Psi\rangle$, normalized so that $\langle\Psi|\Psi\rangle = 1$, is the total energy

$$\begin{aligned} E &= \langle\Psi|H|\Psi\rangle = \langle\Psi|H_0|\Psi\rangle + \varepsilon\langle\Psi|I|\Psi\rangle \\ &= E_0 - \varepsilon\langle\Psi|\log_2(a|\Psi|^2)|\Psi\rangle \end{aligned} \tag{5}$$

Of course, taking the expectation value includes the appropriate statistical ensemble averaging. Let us define the entropy S, associated with the subsystem by

$$S = -k\langle\Psi|\ln(a|\Psi|^2)|\Psi\rangle \tag{6}$$

the expectation value of the information (Shannon and Weaver, 1949),

$$I = -\ln(a|\Psi|^2) \quad \text{nats} \tag{7}$$

expressed in natural units. The factor k is the Boltzmann constant and converts the units of entropy to the familiar thermodynamic units. Peres (1989) and Weinberg (1989*b*) have discussed alternative definitions of entropy in quantum systems. From equations (5) and (6), it follows that

$$E = E_0 + \frac{\varepsilon}{k \ln 2} S \tag{8}$$

Equation (8) involves the entropy S expressed in terms of the elementary entropy change $k \ln 2$, accompanying the measurement of a bit of information (Landauer, 1961).

Now, let us reconsider the parameter a introduced in equation (3). From equation (6), it is apparent that

$$S = -k\langle\Psi|\ln|\Psi|^2|\Psi\rangle - k\langle\Psi|\ln a|\Psi\rangle \tag{9}$$

Thus, the significance of a is that it establishes a reference for the measurement of the entropy. For continuous systems, the entropy is measured relative to a given coordinate system or configuration. There is no absolute entropy as in discrete systems (Shannon and Weaver, 1949). One could, for instance, establish the reference entropy as that for a uniform distribution, $|\Psi_0|^2$. Thus,

$$S_0 = -k\langle\Psi|\ln(a|\Psi_0|^2)|\Psi\rangle = 0 \tag{10}$$

yields

$$a = |\Psi_0|^{-2} \tag{11}$$

and

$$S = -k\langle\Psi|\ln|\Psi/\Psi_0|^2|\Psi\rangle \tag{12}$$

Equation (6), (9), or (12) gives the entropy associated with the information contained in $|\Psi\rangle$ relative to that in $|\Psi_0\rangle$. If the subsystem has fixed volume Ω, and the states are box-normalized over Ω, then $a = |\Psi_0|^{-2} = \Omega$ sets the scale of the coordinate system, and thus fixes the reference entropy.

The temperature T of the subsystem is defined (Landau and Lifshitz, 1977) as

$$T = (\partial E/\partial S)_\Omega \tag{13}$$

Equations (8) and (13) give

$$T = \frac{\varepsilon}{k \ln 2} \tag{14}$$

or

$$\varepsilon = kT \ln 2 \tag{15}$$

This gives the parameter ε a direct physical interpretation as the energy per bit associated with the information acquired in the measurement of the state $|\Psi\rangle$ or, equivalently, of the information stored in $|\Psi\rangle$. Equation (15) is a familiar result, although it was originally derived in a different context (Landauer, 1961) and has been the subject of much discussion (Bennett, 1982; Porod *et al.*, 1984).[2]

From equations (8) and (14), we obtain the free energy

$$E_0 = E - TS \tag{16}$$

---

[2] See also the comments and replies thereto in *Physical Review Letters*, **53**, 1202, 1203, 1204, 1205, 1206 (1984) and references therein.

available to do work (compute, measure, rearrange matter, maintain a given configuration, etc.), and the energy TS bound up in information is unavailable for this purpose.

The Hamiltonian in equation (4) is appropriate for a subsystem in which information is being exchanged with its environment. In other words, information is not being conserved. This is precisely analogous to a system having a variable number of particles, which is accommodated by adding the term $-\mu N$ to the Hamiltonian H_0, resulting in

$$H = H_0 - \mu N \tag{17}$$

The chemical potential μ measures the rate of change of the energy as the particle number N varies. In the case at hand, the (number of bits of) information is variable, so the Hamiltonian (4),

$$H = H_0 + \varepsilon I$$

contains a parameter ε, which measures the variation of the energy E with respect to information I as the number of bits varies.

3. SUMMARY AND CONCLUDING REMARKS

A physical foundation for the Schrödinger wave equation with logarithmic nonlinearity was provided, and a direct physical significance was attributed to it. Moreover, the connection between the entropies of thermodynamics and of information theory has been made in a precise way. Using the thermodynamic definition of temperature, the energy equivalence of a bit of information was derived. The concept of free energy, with a precise description of exactly how and where the unavailable energy of a system is bound up, was recovered.

It is an interesting observation that the nonlinear term in the Schrödinger equation can be so naturally associated with the nonlinear process of measurement. The fact that the Hamiltonian H governing the dynamics of the state $|\Psi\rangle$ depends itself on $|\Psi\rangle$, through the nonlinear term associated with its measurement, is reminiscent of the inextricable admixture of state and observer.

Various experiments to test the linearity of quantum mechanics in atomic and nuclear systems have been performed (Bollinger *et al.*, 1989; Chupp and Hoare, 1990; Walsworth *et al.*, 1990). In particular, searches for a logarithmic nonlinearity of the form examined in this paper have been made (Shimony, 1979; Shull *et al.*, 1980; Gaehler *et al.*, 1981; Hefter, 1985*b*). All of these studies have indicated that any nonlinearity in the Schrödinger equation must be of extremely small magnitude. However, they have all dealt with microscopic (atomic, nuclear, free-neutron) systems, for which temperature is an ill-defined concept. The nonlinear term studied in this paper was shown to be directly related to a system temperature, suggesting that the nonlinearity might be relevant in many-particle systems, such as macro- or mesoscopic matter.

ACKNOWLEDGMENT

This work was supported by Independent Research and Development Funds provided by Teledyne Brown Engineering.

REFERENCES

Bennett, C. H. (1982). *International Journal of Theoretical Physics*, **21**, 905-940.

Bialynicki-Birula, I., and Mycielski, J. (1975*a*). *Communications in Mathematical Physics*, **44**, 129-132.

Bialynicki-Birula, I., and Mycielski, J. (1975*b*). *Bulletin de l'Academie Polonaise des Sciences III*, **23**, 461-466.

Bialynicki-Birula, I., and Mycielski, J. (1976). *Annals of Physics*, **100**, 62-93.

Bollinger, J. J., Heinzen, D. J., Itano, W. M., Gilbert, S. L., and Wineland, D. J. (1989). *Physical Review Letters*, **63**, 1031-1034.

Chupp, T. E., and Hoare, R. J. (1990). *Physical Review Letters*, **64**, 2261-2264.

Delion, D. S., Gridnev, K. A., Hefter, E. F., and Semjonov, V. M. (1978). *Journal of Physics G: Nuclear Physics*, **4**, 125-132.

Efinger, H. J. (1973). *Acta Physica Austriaca*, **37**, 343-349.

Efinger, H. J. (1984). *Nuovo Cimento*, **B80**, 260-278.

Gaehler, R., Klein, A. G., and Zeilinger, A. (1981). *Physical Review A*, **23**, 1611-1617.

Gridnev, K. A., Hefter, E. F., Mikulas, K., Semjonov, V. M., and Subbotin, V. B. (1983). *Australian Journal of Physics*, **36**, 155-161.

Hefter, E. F. (1985*a*). *Nuovo Cimento*, **A89**, 217-224.

Hefter, E. F. (1985*b*). *Physical Review A*, **32**, 1201-1204.

Kibble, T. W. B. (1978). *Communications in Mathematical Physics*, **64**, 73-82.

Landau, L. D., and Lifshitz, E. M. (1977). *Statistical Physics*, 2nd rev. ed., Pergamon Press, England.

Landauer, R. (1961). *IBM Journal of Research and Development*, **5**, 183-191.

Morris, T. F. (1978). *Canadian Journal of Physics*, **56**, 1405-1411.

Peres, A. (1989). *Physical Review Letters*, **63**, 1114.

Porod, W., Grondin, R. O., Ferry, D. K., and Porod, G. (1984). *Physical Review Letters*, **52**, 232-235.

Shannon, C. E., and Weaver, W. (1949). *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Illinois.

Shimony, A. (1979). *Physical Review A*, **20**, 394-396.

Shull, C. G., Atwood, D. K., Arthur, J., and Horne, M. A. (1980). *Physical Review Letters*, **44**, 765-768.

Slater, J. C. (1968). *Quantum Theory of Matter*, 2nd ed., McGraw-Hill, New York.

Ventura, I., and Marques, G. C. (1978). *Journal of Mathematical Physics*, **19**, 838-847.

Walsworth, R. L., Silvera, I. F., Mattison, E. M., and Vessot, R. F. C. (1990). *Physical Review Letters*, **64**, 2599-2602.

Weinberg, S. (1989*a*). *Physical Review Letters*, **62**, 485-488.

Weinberg, S. (1989*b*). *Physical Review Letters*, **63**, 1115.