Interpretation of the extreme physical information principle in terms of shift information

G. V. Vstovsky
VST Laboratories, Podbel'skogo Street, 1-1-39, Moscow 107150, Russia
(Received 8 April 1993; revised manuscript received 15 August 1994)

It is shown that Fisher information (FI) can be considered as a limiting case of a related form of Kullback information—a shift information (SI). The compatibility of the use of SI with a basic physical principle of uncertainty is demonstrated. The scope of FI based theory is extended to the nonlinear Klein-Gordon equation.

PACS number(s): 05.40.+j, 02.30.+g, 03.40.−t, 11.10.Lm

I. INTRODUCTION

An extreme physical information principle is a generalization of the minimum Fisher information principle (MFIP), which was found to be a unifying principle of physics. As was shown by Frieden it gives a universal motivation as well as the way for derivation of a variety of equations and distributions in theoretical physics: the Maxwell-Boltzmann and Boltzmann laws [1,2], Schrödinger wave equation [1–3], Helmholtz wave equation [1,2], diffusion equation [1], Klein-Gordon equation [1,2], Dirac equation [2,4], Maxwell equations [5], and some uncertainty principles [2].

The conceptual basis for such derivations was a gedanken experiment on seeking an expression for a minimum mean-square error of estimation of the mean parameter ζ characterizing the system of particles under investigation. The particles may be material particles or quasiparticles, photons, etc., and the physical parameter ζ may be a position coordinate, velocity, etc. As was shown [1–5], the requirement that an efficient estimation error is to be maximum results in

$$I\{\psi\}=4\int d\zeta \sum_{ks}(\partial\psi_s/\partial\zeta_k)^2=\text{extremum}\ , \tag{1}$$

where $I\{\psi\}$ is a so-called Fisher information (FI), $\zeta=\{\zeta_1,\zeta_2,\ldots,\zeta_N\}$ is a vector parameter. ψ_s are the well-behaved probability density amplitudes,

$$\sum_{s=1}^{M}[\psi_s(\zeta)]^2=p(\zeta)\ , \tag{2}$$

so that $p(\zeta)$ is a probability density. (In this and subsequent cases the integration limits are from $-\infty$ to $+\infty$ unless otherwise specified.)

Frieden had discovered that all the above results are the consequences of the principle (1) when an additional constraint is imposed by, for example, the mean kinetic energy. Operationally, MFIP is a variational problem:

$$I\{\psi\}-\lambda\langle E_{\text{kin}}\rangle=\text{extremum}\ . \tag{3}$$

Other constraints (e.g., the normalization condition) may be imposed as well. Sometimes instead of mean kinetic energy, the constraint term should be a mean-squared energy, as in relativistic quantum mechanics. $I\{\psi\}$ is assumed to be a disorder measure in the system, so the principle (3) is akin to the second law of thermodynamics. The principle (3) may be also thought of as a "balance" between the kinetic energy and disorder.

Frieden has proposed an axiomatic formulation of MFIP and developed an agenda for derivations [2,4], where the constraint term has been considered as an equivalent piece of information obtained from FI through some transform, e.g., the Fourier transform. This has been named the extreme physical information principle. However, for the purposes of this paper it is sufficient to use MFIP in the simple form (3).

The derivation of the time-independent Schrödinger equation can serve as a simple example of the use of MFIP (3). Consider a one-dimensional probability distribution $p(x)$, so that $p(x)dx$ is the probability of finding some particle in the neighborhood dx of the point x, and $p(x)=[q(x)]^2$, i.e., $q(x)$ is an absolute value of a normalized ψ function of the particle. The kinetic energy of the particle is given by $E_{\text{kin}}=W-U(x)$; W and $U(x)$ are the total energy of the particle and potential at the point x. And now principle (3) reads ($q'=dq/dx$)

$$\int dx\{4(q')^2-\lambda[W-U(x)]q^2\}=\text{extremum}\ , \tag{4a}$$

and the solution to the problem (4a) (when q vanishes at infinity for simplicity) is given by

$$\int dx\,\delta q\left[\frac{\partial}{\partial q}-\frac{\partial}{\partial x}\frac{\partial}{\partial q'}+\frac{\partial^2}{\partial x^2}\frac{\partial}{\partial q''}-\cdots\right]\times\{4(q')^2-\lambda[W-U(x)]q^2\}=0$$

or

$$\left[\frac{\partial}{\partial q}-\frac{\partial}{\partial x}\frac{\partial}{\partial q'}+\frac{\partial^2}{\partial x^2}\frac{\partial}{\partial q''}-\cdots\right]\times\{4(q')^2-\lambda[W-U(x)]q^2\}=0\ , \tag{4b}$$

which gives

$$q''+(\lambda/4)[W-U(x)]q=0\ , \tag{4c}$$

and the choice $\lambda=8m/\hbar^2$ results in the time-independent Schrödinger equation (m is a mass of the particle and

$\hbar = h/2\pi$, where h is the Planck constant). The derivation of other equations can be more tedious and demands a sophisticated interpretation of ζ, ψ (or q).

According to a physical scenario, MFIP may imply the minimum or the maximum (an abbreviation EFIP, or extremum Fisher information principle, is more correct, but this is a question of words only). A rather subtle investigation [5] of a variational problem that results in Maxwell's equations shows that it can be rewritten as MFIP, FI being the measure of nonuniformity (roughness, irregularity) of the nonstatistical distribution. For that, ψ_s must be the components of a four-vector potential depending on the coordinates ζ of a Minkowskian space, and the kinetic energy is an interaction energy of the electromagnetic field. So, FI (1) is a measure of a nonuniformity of a field "intensity measure."

That case is also exclusive in that the kinetic energy depends on the field rather than directly on the coordinates ζ, and, therefore, the total kinetic energy is used as a constraint rather than the mean one. It should be mentioned that Eq. (1) implies an additional integration over the time coordinate as compared to Frieden's formulation [5], but this is not significant for finding the extremum.

In this paper an interpretation of Fisher information based principles is proposed in terms of shift information (SI) introduced in Sec. II in its simplest form. The use of SI is shown to remove some arbitrariness of the choice of Lagrange multipliers in MFIP. A consideration of multidimensional SI in Sec. III enables one to interpret FI as a limiting case of SI when the shift is infinitesimal. Section IV demonstrates the use of one possible hitherto uninvestigated constraint and an application of SI based theory to the case of a relativistic particle without spin.

II. SHIFT INFORMATION

Consider a one-dimensional well-behaved distribution $p(\zeta)$ describing some physical system. One can evaluate its nonuniformity (roughness) comparing it with its shifted image $p(\zeta+\Delta)$, and using as a disorder measure a quantity

$$I(\Delta)=\int d\zeta\, p(\zeta)\ln[p(\zeta)/p(\zeta+\Delta)]\ , \tag{5}$$

which can be called shift information. Its mathematical expression is similar to the Kullback information [6,7]. Let the shift be global, $\Delta=$const. When Δ is sufficiently small one can use the expansion of $\ln[p(\zeta+\Delta)]$ in a Taylor series,

$$\begin{aligned}I(\Delta)&=\int d\zeta\left\{p\ln p-p\left[\ln p+\frac{d}{d\zeta}[\ln p]\Delta+\frac{d^2}{d\zeta^2}[\ln p]\Delta^2/2+\dots\right]\right\}\\&=-\int d\zeta\left[p'\Delta-\left[\frac{(p')^2}{p}-p''\right]\frac{\Delta^2}{2}+\dots\right]\end{aligned} \tag{6}$$

($p'=dp/d\zeta$, $p''=d^2p/d\zeta^2$). One can recognize the Fisher's "intrinsic accuracy" $(p')^2/p$ [8] in the second-order term. The integral of $(p')^2p$ is the conventional Fisher information [1,2]. The variation of $I(\Delta)$ gives

$$\begin{aligned}\delta I(\Delta)&=-\int d\zeta\,\delta p\left[\frac{\partial}{\partial p}-\frac{\partial}{\partial\zeta}\frac{\partial}{\partial p'}+\frac{\partial^2}{\partial\zeta^2}\frac{\partial}{\partial p''}-\dots\right]\left[p'\Delta-\left[\frac{(p')^2}{p}-p''\right]\frac{\Delta^2}{2}+\dots\right]\\&\cong\frac{\Delta^2}{2}\int d\zeta\,\delta p[(p'/p)^2-2p''/p]\ ,\end{aligned} \tag{7}$$

to second order in powers of Δ. So, the FI (multiplied by some factor) is a part of SI that is significant for finding the extremum, since

$$\begin{aligned}\frac{\Delta^2}{2}\delta\int d\zeta[(p')^2/p]&=\frac{\Delta^2}{2}\int d\zeta q\,\delta p\left[\frac{\partial}{\partial p}-\frac{\partial}{\partial\zeta}\frac{\partial}{\partial p'}+\frac{\partial^2}{\partial\zeta^2}\frac{\partial}{\partial p''}-\dots\right]\left[\frac{(p')^2}{p}\right]\\&=\frac{\Delta^2}{2}\int d\zeta\,\delta p[(p'/p)^2-2p''/p]\cong\delta I(\Delta)\ .\end{aligned}$$

It should be mentioned that instead of (5) one could use $I(\Delta)=\int d\zeta p(\zeta)\ln[p(\zeta)/\sqrt{p(\zeta+\Delta)p(\zeta-\Delta)}]$, which is equivalent to the elimination of the odd powers from the Taylor series of $\ln[p(\zeta+\Delta)]$ in (6), and the result of the variation of $I(\Delta)$ would give the same approximation (7). This is equivalent also to the averaging over two possible directions of the shift $\pm\Delta$ (see also Sec. III below) and makes the approximation more accurate. As is seen from Eq. (7) the first-order term in Eq. (6) can be omitted when defining the variational problems, as it does not influence the solution. The same can be said about the second-order term with p''.

The usefulness of this introduction of SI can be demonstrated by the derivation of the same time-independent one-dimensional Schrödinger equation. As $p(x)$ is the probability density, $\zeta=x$, so the quantity $I(\Delta)$ is undimensional, and so must be the kinetic energy. The intrinsic energy measure for the particle of mass m is mc^2, c being the light velocity. An energy-time uncertainty implies that the particle with energy mc^2 cannot be considered as existing for the time period less than $\hbar/2mc^2$, and a minimum shift value that can be used is $\Delta = c\hbar/2mc^2 = \hbar/2mc$. Solving a problem

$$I(\Delta) - \frac{\lambda}{mc^2}\int dx[W - U(x)]p(x) = \text{extremum} , \quad (8a)$$

instead of (4a), where Δ is the minimum physical shift $\hbar/2mc$, one has $(p' = dp/dx, p'' = dp'/dx)$

$$\int d\xi\, \delta p \left[\frac{\partial}{\partial p} - \frac{\partial}{\partial x}\frac{\partial}{\partial p'} + \frac{\partial^2}{\partial x^2}\frac{\partial}{\partial p''} - \cdots\right] \times \left[\left[\frac{(p')^2}{p} - 2p''\right]\frac{\Delta^2}{2} - \frac{\lambda}{mc^2}(W-U)p\right] = 0 ,$$

or

$$\frac{1}{2}\left[\frac{\hbar}{2mc}\right]^2\left[\left[\frac{p'}{p}\right]^2 - \frac{2p''}{p}\right] - \frac{\lambda}{mc^2}[W - U(x)] = 0 , \quad (8b)$$

and the substitution $p = q^2$ gives finally

$$q'' + \lambda\frac{2m}{\hbar^2}[W - U(x)]q = 0 . \quad (8c)$$

Hence, the Lagrange multiplier $\lambda = 1$ can be omitted, instead of being chosen, although the insertion of a divisor mc^2 in (8a) was arbitrary. The interchanging of the substitution and variation gives the same result, but it is not true for the general case (2).

The kinship of form (5) and Eqs. (8) to the second law of thermodynamics can be illustrated by the following consideration. The maximum entropy principle is usually written in the form of a Lagrangian problem, for example, [9]

$$-\sum_i p_i \ln p_i - (\lambda - 1)\sum_i p_i - \sum_i p_i \sum_k \lambda_k f_i^k = \text{extremum} . \quad (9a)$$

Here $\{p_i\}$ is a probability distribution over a state space and $S = -\sum_i p_i \ln p_i$ is its entropy in units of the Boltzmann constant $k_B/\ln 2$. The second and third terms represent the constraints (the normalized distribution, given coordinates, kinetic energy, correlations, or whatever else). Equation (9a) can be written in a form involving a Kullback information (KI)

$$1 - \sum_i p_i \ln\left[\frac{p_i}{\exp\left[-\lambda - \sum_k \lambda_k f_i^k\right]}\right] = \text{extremum} , \quad (9b)$$

which can be interpreted as follows. The constraints always introduce a decrease in information [10,11] and then one always deals with the "distance" between some disorder measures. As the extremum is a maximum here, the solution is, by sight, $p_i = \exp(-\lambda - \sum_k \lambda_k f_i^k)$. This is exactly "the coincidence of the average and most probable distributions" mentioned in [10], p. 44.

Finally, comparison with Eq. (5) shows that the solution p_i here is, in some sense, the shifted probability law $p(\xi + \Delta)$. See also Appendix A of [2] for a closely related approach to that taken in this section.

III. FISHER INFORMATION AS AN INFINITESIMAL SHIFT INFORMATION

We can generalize now the consideration of Sec. II to the case of "multiamplitude" probability density [Eq. (2)] and many-dimensional parameter space $\{\xi = (\xi_1, \ldots, \xi_N)\}$. One can evaluate the nonuniformity of $p(\xi)$ using the information densities

$$p(\xi)\ln[p(\xi)/p(\xi+\Delta)] ,$$
$$p(\xi)\ln[p(\xi)/\sqrt{p(\xi+\Delta)p(\xi-\Delta)}] ,$$

etc. The second one is more appropriate for eliminating the sense of Δ direction. The quantity

$$I(\Delta) = \int d\xi\, p(\xi)\ln[p(\xi)/G_p] \quad (10)$$

[where G_p represents $p(\xi+\Delta)$, $\sqrt{p(\xi+\Delta)p(\xi-\Delta)}$, or other proper functions of the shifted image of $p(\xi)$] can be called a shift information. The shift Δ may be global, $\Delta = \text{const}$, or local $\Delta = \Delta(\xi)$. For the simplest considerations, it is sufficient to use an infinitesimal shift and get rid of the sense of Δ direction, i.e., to make some "deparametrization" of G_p. Let $G_p = \sqrt{p(\xi+\Delta)p(\xi-\Delta)}$, so that

$$i(\Delta) = p\ln(p/G_p) = p\ln p - \frac{p}{2}[\ln p(\xi+\Delta) - \ln p(\xi-\Delta)]$$
$$= -\frac{p}{2}\sum_{jk}(\partial_j\partial_k \ln p)\Delta_j\Delta_k - \cdots . \quad (11)$$

Here $\partial_k = \partial/\partial\xi_k$, $\{\Delta_k\}$ are the coordinates of Δ. One can eliminate Δ by a limit $\lim_{\Delta\to 0} i(\Delta)/|\Delta|^2$ and average it over all possible directions by an operation $A = \prod_{n=1}^N (1/\pi)\int_0^\pi d\phi_n$, so that

$$A \lim_{\Delta\to 0} i(\Delta)/|\Delta|^2 = -(p/2)A\sum_{jk}(\partial_j\partial_k \ln p)\cos\phi_j\cos\phi_k$$
$$= -(p/4)\sum_k(\partial_k^2 \ln p) ,$$

$\cos\phi_k = \Delta_k/|\Delta|$. Thus one can simplify the consideration, investigating, instead of $I(\Delta)$, the quantity

$$I' = \int d\xi\, A \lim_{\Delta\to 0} i(\Delta)/|\Delta|^2 = -\int d\xi \frac{p}{4}\sum_k(\partial_k^2 \ln p) , \quad (12)$$

which can be considered as some type of derivative.

Substitution for $p(\xi)$ its expression via Eq. (2) gives

$$-\frac{p}{4}\sum_k(\partial_k^2 \ln p)=\frac{1}{4}\sum_k[(\partial_k p)^2/p-(\partial_k^2 p)]$$

$$=\frac{1}{4}\sum_k\left[\left[2\sum_s\psi_s\partial_k\psi_s\right]^2/p-2\sum_s\psi_s\partial_k^2\psi_s-2\sum_s(\partial_k\psi_s)^2\right]$$

$$\le\frac{1}{4}\sum_k\left[4\sum_s(\partial_k\psi_s)^2-2\sum_s\psi_s\partial_k^2\psi_s-2\sum_s(\partial_k\psi_s)^2\right]$$

$$=\tfrac{1}{2}\sum_k\left[\sum_s(\partial_k\psi_s)^2-\sum_s\psi_s\partial_k^2\psi_s\right]$$

(the Cauchi-Bunjakovsky-Schwarz inequality is used). Thus

$$I'\le I=\tfrac{1}{2}\int d\zeta\sum_k\left[\sum_s(\partial_k\psi_s)^2-\sum_s\psi_s\partial_k^2\psi_s\right], \tag{13}$$

the equality being achieved for $M=1$. Both terms in the right hand side of (13) are the inner products of two vectors.

It can be easily verified that

$$\delta(\psi_s)\sum_k\left[\sum_s(\partial_k\psi_s)^2-\sum_s\psi_s\partial_k^2\psi_s\right]\equiv-4\sum_k\partial_k^2\psi_s\ ,$$

$$s=1,\ldots,M,\quad \delta(\psi_s)=\left[\frac{\partial}{\partial\psi_s}-\sum_i\frac{\partial}{\partial\zeta_i}\frac{\partial}{\partial(\partial_i\psi_s)}+\sum_{ij}\frac{\partial^2}{\partial\zeta_i\partial\zeta_j}\frac{\partial}{\partial(\partial_i\partial_j\psi_s)}\right].$$

[$\delta(\psi_s)$ is an operator with explicitly specified quantity to be varied.] So the extremum of the estimation (13) is provided by the functions that simply obey the Laplace equation (when no constraints are imposed). Instead of using I' one can use its estimation via Eq. (13). The further investigation is a question of interpretation.

In the case when ζ is a four-dimensional space time, ψ is an electromagnetic four potential and $(\sum_k\partial_k^2)\psi=J$ is the four-vector current density; the requirement that I be extremal is equivalent to the minimum Fisher information principle (but with an additional integration over the time coordinate) resulting in Maxwell's equations [2,5].

When the quantities ψ_s or all their first-order derivatives vanish on some closed hypersurface (or at infinity), one has [2]

$$I=\int d\zeta\sum_{ks}(\partial_k\psi_s)^2\ . \tag{14}$$

Equation (14) fits into an axiomatic MFIP formulation by Frieden [2], who used it to derive a variety of equations and distributions in theoretical physics [1–5] from the same minimum Fisher information principle. Thus the term with second derivative can be excluded also in this case, as was indicated in [6,7], because of the regularity conditions imposed on a (generalized) probability density. The similar relation between the Fisher and "infinitesimal" Kullback information was described in [6,7] when the support space and parameter one were different. Thus Fisher information may be thought of as a measure of totality of distinctions between close shifted images of the same distribution, the measure being averaged over all possible directions of shift.

Several words about the Fisher information itself are needed. In 1925, Fisher [8] wrote about the measure of "intrinsic accuracy" $\langle-\partial^2\ln(y)/\partial\theta^2\rangle$ (where θ is a parameter of a statistical distribution density y), that "it may be equally conceived as the amount of information . . ." exactly in the sense that Brillouin [10,11] (and many other authors) discussed the amount of information obtained from the experimental measurement. Taking into account the shift information, which is some sort of KI, one makes the kinship between Fisher's and Brillouin's views on experimental accuracy more explicit [12].

IV. NONLINEAR KLEIN-GORDON EQUATION

Apply now the disorder measure $I(\Delta)$ to some vector field, $\psi(\zeta)=\{\psi_1,\ldots,\psi_M\}$, assuming that there exists its statistical interpretation, such that its intensity measure defined via Eq. (2) is a probability density. The kinetic energy can be calculated as in the case of derivation of the time-independent Schrödinger equation, but the potential depends on the field $U=U(\psi)$ rather than on the coordinates. Such a constraint has not been investigated yet. Making use of Eqs. (10)–(14) giving the estimation of $I(\Delta)/\Delta^2$, and assuming $\zeta=\{\zeta_k\}=\{ict,x,y,z\}$ to be the Minkowskian space, one can set a problem

$$\Delta^2\int d\zeta\sum_{k,s=1}^{4,M}(\partial\psi_s/\partial\zeta_k)^2-\lambda\int d\zeta[W-U(\psi)]=\text{extremum}\ , \tag{15}$$

where the first term is SI averaged over all possible directions of shift in the Minkowskian space time and W is a total energy of the system. Probability p does not weight the second integral [compare with Eq. (8a)].

The solution to the problem (15) is

$$\int d\zeta\,\delta\psi_s\left[\frac{\partial}{\partial\psi_s}-\sum_k\frac{\partial}{\partial\zeta_k}\frac{\partial}{\partial(\partial_k\psi_s)}\right]\times\left[\Delta^2\sum_{jt}\left[\frac{\partial\psi_t}{\partial\zeta_j}\right]^2-\lambda[W-U(\psi)]\right]=0$$

($\partial_k=\partial/\partial\zeta_k$), which gives

$$2\Delta^2\left[\sum_k\partial_k^2\right]\psi_s=\frac{\lambda\partial U}{\partial\psi_s}\ , \tag{16a}$$

which are the nonlinear Klein-Gordon equations ($s=1,\ldots,M$) generalizing the sine-Gordon equation when $M=1$, $U(\psi)\propto[1-\cos(\psi)]$, and the ϕ^4 model when $U\propto(\psi^2-\alpha)^2$ [13]. Equations (16a) can be rewritten in a vector form

$$\Box\psi=(\lambda/2\Delta^2)(\nabla_\psi U)\ , \tag{16b}$$

where $\Box=\sum_k \partial_k^2$ is the D'Alembertian operator, and ∇_ψ is a gradient in the "internal" space $\psi=\{\psi_1,\ldots,\psi_M\}$.

One can easily verify that the solution to a problem with a mean-squared kinetic energy, such as a constraint,

$$\Delta^2\int d\xi\sum_{ks}(\partial\psi_s/\partial\xi_k)^2 - \frac{\lambda}{(mc^2)^2}\int d\xi[W-U(\xi)]^2\sum_s(\psi_s)^2=\text{extremum}\ , \tag{17a}$$

which generalizes the problems [(4a),(8)] to the four-dimensional relativistic case (ψ_s are assumed to be the probability amplitudes, $p=\sum_{s=1}^{M}(\psi_s)^2$, $M=2$ is the probability density, W is a particle energy), is an equation

$$\Box\psi=-\frac{\lambda}{(\Delta mc^2)^2}[W-U(\xi)]^2\psi\ , \tag{17b}$$

which is a linear Klein-Fock-Gordon equation (KFG) [13–15]. For the free relativistic particle $W=mc^2$, $m=m_0(1-v^2/c^2)^{-1/2}$ is a relativistic mass, m_0 is a rest mass, and v is a particle velocity. Setting the components of the minimal physical shifts in the Minkowskian space by the Heisenberg uncertainty principle in the relativistic case to be $\Delta x=\Delta y=\Delta z=c\Delta t$, $\Delta t=\hbar/(km_0c^2)$, $|k|<2$, one obtains that $\Delta^2=(ic\Delta t)^2+(\Delta x)^2+(\Delta y)^2+(\Delta z)^2=2(\hbar/m_0c)^2/k^2$, i.e., $\lambda W^2/(mc^2\Delta)^2=\lambda k^2/[2(\hbar/m_0c)^2]$. Denoting a complex wave function by $\phi=\psi_1+i\psi_2$, for $k=\sqrt{2}$, $\lambda=-1$, one obtains

$$\Box\phi=(m_0c/\hbar)^2\phi\ . \tag{17c}$$

The choice of "right" shift (for the Lagrangian multiplier to be unity) can be conceived as a "normalization" of SI. On the other hand, such a normalization may be considered as an argument to set the minimum product of the conjugate uncertainties in the relativistic case to be $\hbar/\sqrt{2}$ rather than $\hbar/2$.

Remarkably, in both cases of Schrödinger and KFG equations, this choice follows from the fundamental uncertainty principles, one being established rigorously using FI [2] for the nonrelativistic case. Although the latter consideration is not completely rigorous, this shows that the use of SI with small but finite shift is consistent with basic physical principles.

V. SUMMARY

The introduction of a concept of shift information enables us to ascribe to Fisher information a sense of an infinitesimally shifted Kullback information, averaged over all possible directions of shift. The shift information is a quantitative measure of distinction between the distribution and its shifted version.

The use of shift information combined with the fundamental Heisenberg uncertainty principle to choose the "right" shift results in the time-independent Schrödinger and linear KFG equations with the right coefficients. The Lagrangian multipliers of the constraints in corresponding Lagrangian problems are equal to unity and can be omitted.

[1] B. R. Frieden, Phys. Rev. A **41**, 4265 (1990).
[2] B. R. Frieden, Physica A **198**, 262 (1993).
[3] B. R.Frieden, Found. Phys. **21**, 757 (1991).
[4] B. R. Frieden, in *Advanced Imaging and Electron Physics,* edited by P. W. Hawkes (Academic Press, Orlando, FL, 1994), Vol. 90, pp. 123–204.
[5] B. R.Frieden, Physica A **180**, 359 (1992).
[6] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
[7] R. L. Stratanovich, *Information Theory* (Soviet Radio, Moscow, 1975) (in Russian).
[8] R. A. Fisher, Proc. Cambridge Philos. Soc. **22**, 700 (1925).
[9] H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems* (Springer-Verlag, Berlin, 1988).
[10] L. Brillouin, *Science and Information Theory* (Academic, New York, 1962).
[11] L. Brillouin, *Scientific Uncertainty and Information* (Academic, New York, 1964).
[12] B. R. Frieden, Am. J. Phys. **57**, 1004 (1989).
[13] R. K. Dodd, J. C. Eilbeck, J. G. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, London, 1984).
[14] A. S. Davydov, *Quantum Mechanics* (Nauka, Moscow, 1973) (in Russian).
[15] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics, Landau and Lifshitz Course of Theoretical Physics,* Vol. IV (Nauka, Moscow, 1989).