Title
de Broglie's wave hypothesis from Fisher information

Journal
Physica A-Statistical Mechanics and its Applications, A 388

Authors
Frieden, B.Roy
Soffer, Bernard H

Publication Date
2009

Peer reviewed

De Broglie's wave hypothesis from Fisher information

B. Roy Frieden

College of Optical Sciences, Univ. of Arizona, Tucson, AZ 85721

Bernard H. Soffer

University of California, Los Angeles, CA 90024

(Date textdate; Received textdate; Revised textdate; Accepted textdate; Published textdate)

Abstract

Seeking the unknown dynamics obeyed by a particle gives rise to the de Broglie wave representation, without the need for physical assumptions specific to quantum mechanics. The only required physical assumption is conservation of momentum μ. The particle, of mass m, moves through free space from an unknown source-plane position a to an unknown coordinate x in an aperture plane of unknown probability density $p_X(x)$, and then to an output plane of observed position $y = a + z$. There is no prior knowledge of the probability laws $p(a, \mathbf{M}), p(a)$ or $p(\mathbf{M})$, with $\mathbf{M}$ the particle momentum at the source. It is desired to (i) optimally estimate a, in the sense of a maximum likelihood (ML) estimate. The estimate is further optimized, by minimizing its error through (ii) maximizing the Fisher information about a that is received at y. Forming the ML estimate requires (iii) estimation of the likelihood law $p_Z(z)$, which (iv) must obey positivity. The relation $p_Z(z) \equiv |u(z)|^2 \geq 0$ satisfies this. The same $u(z)$ conveniently defines the Fisher channel capacity, a concept central to the principle of Extreme physical information (EPI). Its output $u(z)$ achieves aims (i)-(iv). The output is parametrized by a free parameter K. For a choice $K = 0$, the result is $u(z) = \delta(z)$, indicating classical motion. Or, for a finite, empirical choice $K = \hbar$ (Planck's constant), $u(z)$ obeys the familiar de Broglie representation as the Fourier transform of the particle's probability amplitude function $P(\mu)$ on momentum μ. For a definite momentum μ, $u(z)$ becomes a sinusoid of wavelength $\lambda = h/\mu$, the de Broglie result.

The famous De Broglie-Fourier (D-F) representation (32) [1] of a particle expresses the particle's *probability amplitude* function $u(x)$ as a Fourier superposition [2] of waves in momentum-component space μ, as weighted by amplitudes $P(\mu)$. These have respective $x-$component wavelengths $\lambda = h/\mu$, h =Planck's constant. This D-F representation (or "transformation") expresses a basic equivalence of the two spaces x and μ in describing quantum phenomena.

The D-F is often used as a given *starting point* of a development of quantum mechanics [2],[3],[4],[5]. Indeed, to our knowledge, in the 80 yrs since de Broglie, no one has derived the D-F representation. Perhaps this was because it is the expression *par excellence* of the wave nature of matter, and therefore so fundamental as to be incapable of implication by *any prior physical* principle. Clearly a new tack is needed.

During the past two decades many physical processes have been found to be derivable by uses of principles of *information.* Here we utilize such a principle, with the aim of *deriving* the D-F transformation and momentum-dependent wavelength. This is to be without the prior assumption of any physical property that is unique to wave phenomena, including diffraction theory and (more to the point) quantum theory. Instead, the derivation uses a principle of optimum information transmission, called EPI (see below), under an assumption of conservation of momentum (Sec. I.D). The information in question is Fisher information [6]. The derivation develops within the framework of *an estimation problem* for a specific physical scenario.

General comment 1. But, is such a derivation possible? Does it not already presume what it intends to derive? That is, the argument can be made that any attempt to derive the D-F representation – a relation between probability amplitudes – must *presume* the existence of meaningful probability amplitudes, which in turn presumes quantum mechanics, since quantum mechanics (or the closely related field of diffraction optics) is the only scientific field that uses the concept. However, this argument is moot, because (i) we define and use each probability amplitude as merely a *mathematical* quantity, whose squared modulus conveniently defines a probability density function (PDF) obeying positivity PDF≥ 0. Also, in fact (ii) R.A. Fisher used probability amplitudes to describe the frequencies of occurrence of *biological* species [7], definitely outside the realm of quantum mechanics. Another convenient mathematical property of an amplitude is that its gradient defines the Fisher channel capacity (3), the very information that is used in the EPI principle.

Regarding *physical properties* of the amplitudes, these arise only as *a result* of the derivation, after use of the EPI principle. For example, they are found to obey the D-F representation (32) and Heisenberg principle (38). Note also that, since we do not a priori assume the Schrodinger wave equation (SWE), but, rather, derive it (subject to an additional assumption), we are also not a priori assuming the *Born interpretation* that amplitudes are probability amplitudes *obeying the SWE.* Rather, they are a priori merely mathematical constructions that are later *shown* to obey the SWE. In essense, we derive the Born interpretation.

General comment 2. It is useful to mention in one place the *types* of assumptions that are made in deriving the D-F. The sole *physical assumption* is that of conservation of momentum (Sec. I.D). There are also many *mathematical assumptions* made – of the particular geometrical constraints (1) imposed by the experiment, lack of prior statistical knowledge (Sec. II), a weak aperture amplitude (Sec. I.A), and analyticity (Secs. II.D, III.A and IV.C). Although mathematical, these of course have strong physical ramifications.

A synopsis of the overall approach is as follows.

I.A particle motion experiment is performed. The particle, of arbitrary mass m, travels through the general system of Fig. 1, from object- to aperture- to image plane over some trajectory. All indicated coordinates and momenta are *vertical* components, in the plane of the paper. The experiment operates under special *geometrical conditions* usually called 'far-field' (1). The particle obeys *unknown dynamics* and is not tracked, so that neither its (vertical) coordinate a nor its position x in the aperture plane are known. *Its dynamics are to be found.* The particle's final position y in the image plane is observed. Its net motion $a \to x \to y$ is affected in some unknown way by an (also) unknown probability density function PDF $p_X(x) \equiv |U(x)|^2$ in the aperture, with $|U(x)| << 1$ (see *General comment 2*) and termed 'weak.' Just beyond the aperture, the particle has a momentum value μ_0 with vertical component μ. As the sole explicitly physical assumption, the particle is assumed to obey *conservation of momentum* at each point along path AB. By this property, a given coordinate x in the aperture is found to be *linear* in the corresponding momentum value μ of the particle in the image plane.

II. There is a severe lack of prior information about the particle. Neither the particle's initial conditions a, $\mathbf{M}$, nor any prior PDF $p(a, \mathbf{M})$, $p_A(a)$ or $p(\mathbf{M})$, are known. With such a lack of knowledge, it is usual to seek a maximum likelihood (ML) solution. Based on an assumed analyticity property and the far-field conditions (1), the particle statistics in the image plane *are found to obey invariance to shift* a. Tools of the ML approach are this shift invariance property and the concept of 'Fisher channel capacity' (3) of the system. Conveniently, this *information* quantity is expressed in terms of the same complex amplitude function $A(x)$ (including the phase) that determines *its PDF* $|A(x)|^2$. Depending upon EPI application, the phase part can have, or not have, physical significance (in our case, it will have it, for a parameter value $K \neq 0$).

III. In general, ML assumes knowledge of the likelihood PDF $p_Y(y|a) = p_Z(y-a) = p_Z(z)$, $z = y-a$, by the shift invariance proven in II. ML requires this PDF to be maximized through choice of a. Hence, $p_Z(z)$ must first be estimated. This PDF is, as usual, represented as a squared modulus, in particular $p_Z(z) \equiv |u(z)|^2 \geq 0$ in terms of its amplitude law $u(z)$. Hence, $u(z)$ must be found. By use of the small-amplitude approximation $|U(x)| << 1$ mentioned in I above, $u(z)$ is found to relate linearly to $U(x)$ via an integral transformation $u(z) \equiv T[U(x)]$. The transformation has an unknown kernel $k(z,x)$ which, therefore, must be found. It turns out to be most convenient to first estimate $u(z|x)$, the image amplitude due to a slit particle source located at position x. The EPI approach is chosen for this task, because its estimates contain maximal information about the unknown parameter a and, by implication, the unknown amplitude $u(z|x)$. Prior knowledge that is used consists of far-field conditions (1), and the small-amplitude approximation $|U(x)| << 1$ defined in I. Also, linearity $u(z|x) \propto U(x)$ is derived at Eq. (8).

IV. The information $I_X(x)$ about the unknown a at the slit position x is regarded as fixed, but of unknown value (found in V below). The use of EPI finds the unknown kernel $k(z,x)$ to obey (23). This, in turn, represents the transformation $T[U(x)]$ as the first Born approximation (24), if the weak amplitude $U(x)$ is re-interpreted as a weak *potential* function. This is our first derived quantum property of the probability amplitudes heretofore only defined mathematically as forming PDFs that obey positivity.

V. The EPI output $T[U(x)] \equiv u(z)$ is parametrized by an arbitrary constant K. If $K \equiv 0$, then the particle behaves completely classically, with $u(z) \propto \delta(z)$ implying no spread in image positions. This option then predicts deterministic motion. Or, if K is kept finite, $u(z)$ has finite spread, and the particle behaves probabilistically. In the latter case, empirical evidence demands that K have the particular value of Planck's constant $\hbar$. Using the known (via step I) proportionality between aperture coordinate x and particle momentum

μ leads to an effective 'pupil function' $P(\mu)$ proportional to $U(x)$. Pupil function $P(\mu)$ and an amplitude function $u(z)$ are found to be Fourier transform mates (32). These are parameterized by an arbitrary constant b_1. Defining b_1 appropriately allows one to cancel all particular experimental parameters from Fourier relations (32), casting them into the universal D-F form. This is the second derived *physical* property of amplitude functions that were defined (see *General comment 1* above) as purely mathematical abstractions. Knowledge of the D-F relation also allows the unknown information $I_X(x)$ to be found, as quadratic in aperture position x [Eq. (35)]. A free parameter K of the solution is shown to imply classical particle dynamics, and thus a classical trajectory, if $K \equiv 0$, or quantum dynamics if $K \equiv \hbar$, an empirical choice. The derived D-F relation also predicts a particle wavelength obeying $\lambda = h/\mu_0$. The D-F relation also allows the SWE (37) to be derived after assuming, as well, conservation of energy (Sec. V.I). This is the third derived physical property of our mathematical amplitude functions, and is equivalent to deriving the well known Born interpretation (note: not *scattering*) of amplitude functions. Examples of resulting ML estimates a are given in Sec. V.G.

VI. The D-F answer gives the Heisenberg uncertainty principle (38). This is the fourth derived physical property found for the postulated probability amplitudes. The Heisenberg principle is consistent with our assumed lack of prior knowledge of the particle's input position and momentum, since it implies that they *couldn't have* been known. The classical, $K = 0$ solution is found to be ill-defined if $m \to 0$, or well-defined for large m.

The derivation depends upon two properties of the Fisher *channel capacity* I: (A) I is an upper bound to the classical Fisher information; and (B) I is physically realizable. These properties are crucial to allowing meaningful phase functions in EPI output amplitude laws. The properties are proven, respectively, in Appendices A and B. These connections with Fisher information constitute a fifth physically useful property (through physical applications of EPI) of the postulated amplitude functions.

Details follow in corresponding sections I-VI and Appendices A,B.

I. PARTICLE EXPERIMENT

In Fig. 1 a particle obeying *unknown dynamics* moves through fixed apparatus, over some unknown path SAB from an object or source plane to an intermediary aperture plane and on to an image or output plane. The paths SA and AB are first allowed to have any *generally curved,* Feynman-type shapes, not necessarily the straight paths shown (used in a later trigonometric analysis). A typically curving path is shown dashed in the figure.The particle has small mass m (on the atomic level or less), except in special cases taken up later. For simplicity, the system in Fig. 1 operates at equal conjugates L in the analogous sense of optical imaging (taken up below). At input point S the particle has a total vector momentum $\mathbf{M}$, which becomes some vector momentum $\mu_0 \neq \mathbf{M}$ generally at point A in the image. The vertical component of μ_0 is μ, as indicated. For simplicity, only the vertical components of the motion are analysed, with all coordinates in the plane of the page. A 3D, and even (3+1)D, generalization of the experiment is probably straightforward. A 2D version is sketched in Sec. V.5.

The particle is not tracked, so that the coordinates a,x and z, identifying positions S,A and B, are unknown, as are the shapes of the paths SA and AB. As will become evident, the system can more productively be regarded as a *communication channel* for transducing individual coordinate values from plane to plane as $a \rightarrow x \rightarrow y \equiv a + z$. That is, it transduces informations, specifically *information about* the source position a. This suggests R.A. Fisher's form [6] of information in particular, whose use later turns out to be key to the derivation.

The source coordinate a is to be estimated. Pursuing this estimate will automatically lead to an estimate of the particle's dynamics defined by $u(z)$ and, then, to the required de Broglie wave representation. For the purpose of estimating a, the image position y is observed (see eye in Fig. 1).

A. Unknown amplitudes; potentials

The particle obeys a *PDF* $p_X(x) \equiv |U(x)|^2$ in the aperture of Fig. 1, where $U(x)$ is an unknown, generally complex, *aperture amplitude* (see *General comment 1* at outset of paper). The PDF is, ideally, perfectly compressed longitudinally, so as to influence the particle trajectory only within the aperture plane AC. In practice, $p_X(x)$ inevitably has some longitudinal extensions beyond the aperture on either side. These could be accomodated by the approach, but, for simplicity, are assumed to quickly and continuously go to zero (see also related effect in Appendix B). Also, $p_X(x)$ is assumed small, $p_X(x) << 1$ (see *General* *comment 2*). An example is a Gaussian with very large σ. Then the amplitude $|U(x)|$ is small as well, and termed 'weak' as convenient terminology. The analysis will show that, *for finite* m, $U(x)$ enforces both the familiar Born quantum scattering [8] upon the particle and the D-F transformation (32); or, *for mass* $m \to \infty$, a *deterministic* deflection to the image point B. These familiar alternatives are automatic consequences of the approach, through the action of a free parameter K.

Between the planes of Fig. 1, the particle travels in free space, with potential $\Phi \equiv 0$.

B. Far-field approximation

The system obeys special *geometrical* length inequalities typically called the 'far field approximation.' Here the (vertical) image coordinates z, a are small compared to a typical aperture position x, and all coordinates are small compared to the conjugate distance L:

$$z, a << x \text{ and } z, a, x << L. \tag{1}$$

These inequalities are also sometimes called the 'small angle' approximation, since the angles θ, θ' in Fig. 1 are then small (as will be shown). *Although conditions (1) are purely geometrical in nature, they will prove to be vital to many of the derivations to follow.*

C. Analogous optical experiment

The geometry of Fig. 1 also follows that of a famous gedanken 'optical diffraction experiment' [9] that is often used to introduce quantum ideas. In the optical experiment the 'particle' is a photon. We shall often use the optical aspect of this experiment as a guide to suggest corresponding particle properties. For example, the unknown (dashed) paths SA and AB are defined, by diffraction theory [10], to obey a condition of 'stationary phase.' However, *neither this fact nor any other that is unique to either diffraction theory or quantum theory will be used in establishing the particle's trajectory.* Indeed, the single physical assumption about the particle that is made is that, because potential $\Phi = 0$ in the space between A and B, momentum is conserved in this space. Note that this property is obeyed by either classical or quantum particles.

Also, the analysis will be purely computational, with the aim of optimally estimating

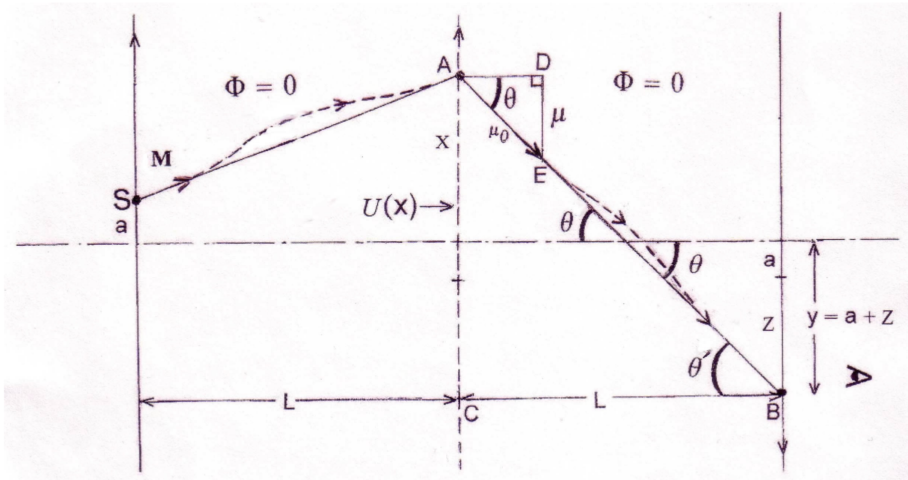

FIG. 1: Particle experiment. A particle travels on an unknown trajectory (shown as dashed where it possibly departs from given straight paths) from an unknown source point S to an aperture plane A, and then to an image plane B. With free fields $\Phi = 0$ en route, what is the relation between the unknown particle amplitude laws $U(x)$ in plane A and $u(z)$ in plane B?

unknowns of the experiment. Information theory will be used in, by now, a standard way, to determine the required probability amplitude function of the estimation approach. The computed result will be of interest in having, as well, the *physical* significance of establishing the D-F transform and wavelength.

D. Linear relation between z and momentum

A key result of diffraction optics is that, in Fig. 1, the vertical component μ of the photon momentum is approximately linearly related to its aperture coordinate position x. This is shown next to follow, as well, for our material particle experiment.

As in [3],[9], assume that some fixed, but necessarily *approximate* $x(\mu)$ relation connects a given coordinate x with the vertical momentum defined as $\mu \equiv mv$, v the velocity at x. Note that $x(\mu)$ has to be approximate since, in principle, all possible momenta could emenate from the point A. In the *optical* experiment, the method of stationary phase is used to find the dominant such momentum, which lies along the *straight* path AB. In our particle case of unknown dynamics, stationary phase cannot be used. The approximate nature is also required by the Heisenberg uncertainty principle (38). If x were exactly a known function of μ, then both could be exactly known simultaneously, violating the principle. We will ultimately derive (38), and not assume its form here (but, of course, cannot violate it either!) What is the approximate relation between x and μ?

As was mentioned, the generally curved (dashed) path shape AB in Fig. 1 is some unknown path. The path, and angles θ' and θ, are next approximately found in steps (i) and (ii). The key consideration will be the zero potentials $\Phi = 0$ in the intervening spaces.

(i) We now show that, for the given problem, the actual path AB must be the indicated straightline path. *Assume that the particle's vector momentum* μ_0 *is conserved at each point of this path*, owing to the free field $\Phi = 0$ in that space. (This is the sole explicitly physical knowledge that is used.) Hence its vertical component μ is conserved as well.

Since the preceding argument holds at each position along the point along AB, at each point the vertical component of momentum has the same value μ. Let θ'' denote a local slope value anywhere along curve AB. By the conservation of momentum, the vertical component $\mu = \mu_0 \sin\theta''$. This has two consequences: (a) The local slope angle θ'' is constant over the curve AB, so that AB is a straight line. (b) The particular values θ (at A) and θ' (at B) of θ'' obey $\mu = \mu_0 \sin\theta = \mu_0 \sin\theta'$. Consequently $\theta = \theta' = \sin^{-1}(\mu/\mu_0)$.

(ii) In Fig. 1, simple trigonometry gives $\tan\theta' = (z+a+x)/L \approx x/L$ by the two inqualities (1). Also, by the second inequality (1), $\tan\theta'$ is small, so that θ' is small and consesquently $\tan\theta' \approx \sin\theta' \approx \theta' \approx x/L$. Also, from (i), μ/μ_0 must be small so that $\theta = \theta' \approx (\mu/\mu_0)$. Then $\mu/\mu_0 \approx x/L$ or $x \approx (\mu/\mu_0)L$. In summary, owing to postulated conservation of momentum at each path position between A and B, and because of the geometrical far-field approximations (1), the *aperture coordinate* x of the particle approximates proportionality to its *momentum component* μ.

As an aside, in Sec. II.D below we prove that the image amplitude $u(z)$ obeys shift invariance. It is well known that such shift invariance implies, by Noether's theorem, conservation of momentum, suggesting that we did not have to *assume* conservation of momentum here. However, Noether's theorem holds in a scenario where the Lagrangian of the particle dynamics is known, and we cannot know the Lagrangian at this elemental stage of the development. See also the end of Sec. VI.A on this.

E. Data equation

Path AB, with coordinate $x \to y$, is to a final point of position

$$y = a + z \tag{2}$$

in the image plane B. Position y is observed (see eye in Fig. 1) as a datum. The detector of y is taken to be ideal, not adding noise to the reading. However, the nature of z is of interest. It is taken to be a random vertical-component jiggle that is induced by the interaction between the aperture amplitude and the particle. Thus, in general y departs from a by some randomly unknown amount z.

Our initial aim is to find a, given the imperfect datum y. The form of Eq. (2) indicates that, intuitively, knowledge of y contains 'information' about the value of a. Thus, *given* y, one should be able to estimate a to some extent. This is usually called a 'reverse problem', as distinguished from the 'direct problem' of predicting y *from* a known value a.

F. Trial analytical estimate of a

If the particle were known to be moving classically, we could first try a direct analytical approach to finding a. That is, the datum could be modeled as obeying a relation $y = F(a, L) + z$ where function F is known, and z is as above. Then, *for small enough* z, one could in principle invert F to obtain an approximate, direct-inverse solution $a \approx F^{-1}(y, L)$. Obviously, however, the larger the random component z is the worse an approximation this would be (even excluding error caused by the assumption of classical motion).

To judge how bad an effect this is, consider just the final particle trajectory, $x \to y$. In the corresponding optical analogy experiment of a given thin lens, in the geometrical approximation a photon striking position x in the lens travels to a *known* position y in the image plane via basically Snell's law. However, the reverse problem of inferring x from a known y is more challenging. Indeed the very aim of a good lens design $U(x)$ is to achieve near-stigmatic imagery, for which *many different* aperture positions x give rise to the observed y. Also, of course, the random nature of the component z of y adds in further uncertainty.

G. ML estimate

Because the particle dynamics are unknown, we drop the trial assumption of classical dynamics in Sec. I.F. In fact, *we regard the dynamics u(z) as something to be ultimately estimated*, along with a. Conveniently, it turns out that the two can be estimated simultaneously. Because of its important properties of optimality [11],[12], we first seek an ML estimate of a given the datum y. This is defined as the solution a to the problem $p_Y(y|a) = \max.$, where $p_Y(y|a)$ is the PDF on y in the presence of fixed a. This PDF is often called the 'likelihood law.' It must be known in order to seek the ML answer. Although ML solutions are generally statistical in nature, they can give simple analytic inverses, such as $a = F^{-1}(y, L)$ mentioned above. For example, given one datum, an ML solution is often (e.g., in case of additive noise) just the datum itself [11], irrespective of any other statistical quantity. We carry through on the ML approach. First, some factors affecting the approach are discussed.

II. UNKNOWN INITIAL CONDITIONS

Let the particle's *initial conditions be totally unknown*, i.e., neither the values $(a, \mathbf{M})$ nor their joint probability law $p(a, \mathbf{M})$ nor marginal laws $p(a), p(\mathbf{M})$ are known. Recall that using ML requires us to form the likelihood law $p_Y(y|a)$. Ordinarily this could be found using known rules of statistics: first randomly sampling a and $\mathbf{M}$ values from a probability law $p(a, \mathbf{M})$, then randomly sampling a representative x from these and, finally, a representative y from the x. However, $p(a, \mathbf{M})$ is unknown, and *we do not know the particle dynamics connecting a and x.* Hence, this procedure is *ill-defined*. As a further complication, from another viewpoint (Sec. VI.A), choosing particular laws $p_A(a), p(\mu)$ could in fact lead to inconsistencies. Is there an alternative approach that is, instead, well-defined?

A. Central question

Thus, the central question posed by this paper is:

Given image position y, what mathematically can be said about the unknown a and $u(z)$ in the complete absence of knowledge of initial conditions and particle dynamics?

In summary, our ML approach will have *to estimate* the likelihood law $p_Y(y|a)$ under extremely nebulous conditions: the complete absence of knowledge of initial conditions; *and* no knowledge of the dynamics obeyed by the particle.

B. Use of complex amplitude functions

The likelihood law $p_Y(y|a)$ can be found by the use of the EPI principle (Secs. III,IV). EPI generally gives an output that is a complex amplitude function $A(x) \equiv |A(x)| \exp[i\phi(x)]$, with corresponding probability $p \equiv |A(x)|^2$. (*General comment 1* discussed the dual motivation behind using amplitude functions.) It accomplishes this through use of a concept of classical estimation, called *Fisher information* [6],[11],[12]. In particular, EPI uses the Fisher *channel capacity*

$$I \equiv 4 \int dx\, |A'(x)|^2\,, \quad p \equiv |A(x)|^2\,,\ A'(x) \equiv dA/dx. \tag{3}$$

Thus, I and p are expressed in terms of the same unknown function, $A(x)$. Conveniently, this allows the EPI principle, which depends upon I and p, to be solved straightforwardly for $A(x)$. From this point on, each I (and J below) represents a channel capacity. It has the following properties.

The channel capacity is (i) defined to be an upper bound to the actual Fisher information: See Appendix A; (ii) *equals* the Fisher information in one-dimensional problems [5]; (iii) is a physically realizeable upper bound (Appendix B); and, most important for our purposes, (iv) is sensitive to the phase profile $\phi(x)$ of the particle. For example, use in (3) of an all-phase amplitude function $A(x) \equiv \exp(i\phi(x))$ directly gives $I = 4\int dx \phi'^2(x)$. This explicitly shows strong sensitivity to the local phase gradients $\phi'(x)$.

However, note that *allowing an amplitude* $A(x)$ to have a phase function $\phi(x)$ *does not force* the resulting EPI outputs $A(x)$ to have meaningful phase. The outputs predict the particle's dynamics, and *these can turn out to be classical (e.g., obeying Einstein gravitation)* *or quantum, depending upon case* [4],[5],[13]. The latter has well-defined phase, the former does not.

Our final EPI output will be the amplitude function $u(z)$ in the image plane. The latter will be shown to obey the SWE, Eq. (37), which of course has a well-defined phase function for sufficiently small particle mass m. Under this condition, solving the SWE for the phase function is generally a well-posed mathematical problem.

Alternatively, in the opposite limit of $m \to \infty$, the problem of solving the SWE for the phase is known to be an ill-posed mathematical problem (Sec. V.H.3). That is, as $m \to \infty$ we can expect the dynamics obeyed by $u(x)$ to approach a classical limit. This will be verified by the EPI approach, since it will give both (i) a classical particle solution and (ii) a quantum solution as alternatives. Interestingly, even if quantum solution (ii) is chosen, it goes continuously into the classical solution (i) in the limit of *large particle mass* (Sec. V.H.3), since then the phase can be ignored, i.e. regarded as irrelevent or unphysical – of course the viewpoint of classical mechanics.

C. Unknown amplitudes of problem

As we found, the PDF $p_Z(z)$ is needed in order to form the ML estimate we seek. Random fluctuations z in the image plane obey an unknown PDF $p_Z(z) \equiv |u(z)|^2 \geq 0$ in terms of a complex amplitude law $u(z)$. Possible fluctuations z even exist when the particle can only leave the aperture at a given, *fixed* position x. These follow an unknown PDF $p_Z(z|x) \equiv |u(z|x)|^2$, whose corresponding amplitude law $u(z|x)$ is defined by (11). Amplitude $U(x)$ is likewise regarded as unknown. For a sufficiently weak $U(x)$, the resulting amplitude $u(z)$ will turn out to obey a first Born approximation (24) [8]. An *optical* analogy to $U(x)$ is that of the amplitude in the exit aperture of an unknown lens system [8],[10],[11].

D. Property of shift invariance

We next show that the amplitude u_Y at any image coordinate value y is invariant to off-axis shift a of the source, obeying

$$u_Y(y|a) \equiv u_Z(y-a|a) \equiv u(z|a) = u(z|0) \equiv u(z), \text{ and} \tag{4}$$
$$p_Y(y|a) \equiv p_Z(y-a|a) \equiv p_Z(z|a) = p_Z(z|0) \equiv p_Z(z) = |u(z)|^2$$

for corresponding density functions. This will be for suffiiciently small values of the relative shift a/L. In each line, the first equality defines new amplitudes and densities u_Z, p_Z by the use of (2); the second equality is again by (2); the third is by the assumed invariance to shift a; and the fourth is a matter of notation. The fifth, on the 2nd line, defines $p_Z(z)$ as an unknown PDF. The optical analogy to $u(z)$ is that of the point amplitude function, which likewise does not change appreciably in modulus or phase within its isoplanatic region [10] of shifts a.

Before presenting the quantitative proof, we argue from intuition. By Fig. 1, the source S is off axis by an *angular* amount a/L. Hence, since $a/L << 1$, even for a finite shift a the source S is *angularly* off axis by a negligable amount. Moreover, as shown in the Derivation to follow, the *longitudinal shift* due to finite a is, by comparison, negligable (second-order in a/L). Therefore a 'wave' from S effectively sees *the same* aperture function $U(x)$ for any shift a, and its output wave $u(z)$ is independent of a.

1. *Derivation*

An increased shift a off axis causes both an angular and a longitudinal shift about the aperture. The angular is by amount $a/L \equiv a_\alpha$, and the longitudinal is by the relative amount $(\sqrt{L^2+a^2}-L)/L \approx 0.5(a/L)^2 = 0.5a_\alpha^2$. This is of second order in angle a_α which, by (1), is already a very small number and, hence, negligable compared to the first-order dependence a_α just found for the angular shift. Hence the longitudinal shift may be ignored.

To facilitate the angular analysis we define angular coordinates $y_\alpha \equiv y/L$, $z_\alpha \equiv z/L$, and an angular amplitude function $u_\alpha(y_\alpha|a_\alpha)$. This relates to $u_Y(y|a)$ as $u_Y(y|a) = L^{-1/2}u_\alpha(y_\alpha|a_\alpha)$, where factor $L^{-1/2}$ is needed to satisfy normalization of the corresponding PDF $|u_\alpha(y_\alpha|a_\alpha)|^2 \equiv p_\alpha(y_\alpha|a_\alpha)$. Let $u_\alpha(y_\alpha|a_\alpha) \equiv u_{\Delta\alpha}(y_\alpha - a_\alpha|a_\alpha)$ define a new, shifted, angular amplitude law $u_{\Delta\alpha}$. Therefore, combining, $u_Y(y|a) = L^{-1/2}u_{\Delta\alpha}(y_\alpha - a_\alpha|a_\alpha)$, or

$$u_Y(y|a) = L^{-1/2}u_{\Delta\alpha}(z_\alpha|a_\alpha)$$

by Eq. (2).

Consider, for fixed z_α,the dependence of $u_{\Delta\alpha}(z_\alpha|a_\alpha)$ upon shift a_α. By definition (4), amplitude $u_{\Delta\alpha}(z_\alpha|0)$ does not depend upon angular shift. It is reasonable, then, that for a *small* such shift a_α, $u_{\Delta\alpha}(z_\alpha|a_\alpha)$ should only depart from $u_{\Delta\alpha}(z_\alpha|0)$ by a small amount, and larger for larger shifts. Or, $u_{\Delta\alpha}(z_\alpha|a_\alpha)$ should be *analytic* over small shifts about the shift $a_\alpha = 0$. Then we may expand it as a Taylor series of powers of a_α. Consequently, the last set-out equation becomes $u_Y(y|a) = L^{-1/2}\left[u_{\Delta\alpha}(z_\alpha|0) + u'_{\Delta\alpha}(z_\alpha|0)a_\alpha + ...\right]$, where the prime denotes a derivative $\partial/\partial a_\alpha$ and dots stand for terms in quadratic and higher powers of a_α.

By far-field conditions (1), although *linear* shift a may be finite a_α is very small. Therefore we may drop all terms beyond the first, giving simply $u_Y(y|a) = L^{-1/2}u_{\Delta\alpha}(z_\alpha|0)$. But $u_Y(y|a) \equiv u(z|a)$ by the top Eq. (4). Also, by normalization of corresponding PDFs, $L^{-1/2}u_{\Delta\alpha}(z_\alpha|0) = u(z|0)$. Consequently $u(z|a) = u(z|0) \equiv u(z)$. The amplitude obeys shift invariance, confirming the top Eq. (4). Then, by modulus-squaring, so does the PDF $p_Z(z|a)$, giving the bottom Eqs. (4). In summary, the shift invariance follows from an argument of analyticity and the far-field conditions (1). Shift invariance gives the following immediate benefit.

E. Effectively $a_\alpha = 0$

Since the unknown amplitude law $u(z)$ obeys shift-invariance relations (4), we can model $u(z)$ as being formed in effectively the presence of an on-axis system $a_\alpha = 0$.This simplifies the analysis. Even though a is still finite (and to be found), *the size* of information $I_Z(a)$ – formed by (3) with z and $u(z)$ replacing the generic x and $A(x)$ – is *independent of both* a_α *and* a. Hence, we can denote $I_Z(a)$ as $I_Z(0)$. Also, $U(x)$ is by definition independent of the size of a, so that by (3) again we may denote $I_X(a) \equiv I_X(0)$. In summary, the informations at the positions of the transition $a \to x \to y$ (or z) of the particle are now denoted as

$$J(\underset{\delta(a)}{a}) \to \underset{U(x)}{I_X(0)} \to \underset{u(z)}{I_Z(0)}. \tag{5}$$

Note that information $I_Z(0)$ *about* a is still finite, even though its level does not depend upon a. It is the observable y that provides this information to the observer, via $y = a + z$.

In (5), the amplitude law defining each level of information about a is denoted below it. Thus, by (3), the information $J(a)$ for an observer *in the object plane is thereby infinite.*

However, its values after transitions to x and then y *are finite*, owing to the fact that Fisher information decreases or remains constant [4],[5] after any transition. In fact the two informations will turn out to be equal.

F. ML estimate

What approach should be used for estimating a, given the unknown nature of its prior probability law $p(a)$? We use the standard maximum-likelihood (ML) value [6],[11],[12]. As in the above, this satisfies $p_Y(y|a) =$ maximum, and so we need to find the likelihood law $p_Y(y|a)$. Now, as a probability density, $p_Y(y|a)$ must obey positivity $p_Y(y|a) \geq 0$. Hence, we seek an estimate that obeys this constraint. One important way of accomplishing this is to use (4), $p_Y(y|a) = |u(y-a)|^2 \geq 0$ by construction. This is an arbitrary but reasonable choice, and it requires that we first estimate the amplitude law $u(z)$.

III. EPI APPROACH

As discussed in Sec. I.G, the ML estimate a is to be the solution to a problem $p_Y(y|a) = \max$. Then by (4) we must know the PDF $p_Z(z)$ and, consequently, its amplitude function $u(z)$.This is addressed here and in Sec. IV.

We choose to estimate $u(z)$ by use of a principle of Extreme physical information or EPI [4],[5],[13],[14],[15],[16],[17],[18],[19]. This is a variational principle $I - J \equiv$ *extremum* through variation of the amplitude $u(z)$. In general, $I \leq J$ is the acquired information about an unknown parameter a in received data (see Sec. II.B). Hence J is its maximum possible level at the data source. In real-coordinate z cases, as here, the extremum is a *minimum.* Therefore, for a given level of source information J, our EPI output $u(z)$ will convey data containing *maximum received information.* In fact, it will attain the maximum possible value $I = J$ here. By the Cramer-Rao inequality $e^2 \geq 1/I$, where e^2 is the mean-square error in the estimated a. Then with our I maximal, the mean-squared error e^2 tends to be minimal, reflecting a high level of accuracy in both the ML estimate of a and (by implication) the amplitude $u(z)$ forming the estimate. In general, *EPI derives* out of a premise that nature is stable, and tends to convey information in a maximally coherent manner [19].

An important property of the Fisher channel capacity is that it is sensitive to *the phase* of the amplitude function involved (see below (3)). This allows it to serve as a bridge between classical and quantum physics. *Depending upon application, the EPI solution can follow either classical or quantum physics* (see, e.g., Sec. V.H). In fact, in our case both choices will hold, depending upon the chosen size of a free parameter K and the mass m of the particle.

A. Prior knowledge of type A, and weak-field approximation

Any approach to quantifying a physical effect requires at least some prior knowledge about the effect. In general, EPI classifies prior knowledge into three definite types [13] A,B and C. These define a descending order of accuracy in EPI outputs. We use a type A approach, the 'A' standing for an 'abduction' in the sense of a universal truth. Here the EPI outputs are correct to within some specified range of application, such as inequalities (1). Thus, we can expect EPI to give an accurate output $u(z)$ within the confines of assumptions (1). The type A prior knowledge is that of conservation of momentum (Sec. I.D) over path AB.

As preliminary to the use of EPI, we have to quantify, as generally as possible, the image amplitude $u(z)$ that will result from a given source amplitude $U(x)$. The amplitude $u(z)$ at a fixed point z must depend upon $U(x)$ over its entire range of points x, i.e. as some functional $u(z) = \int dx' \mathcal{F}[z, x', U(x')]$. The ultimate aim of EPI will be, then, to estimate the function $\mathcal{F}$. Note that the left side of the preceding equation has the unit of amplitude, and hence to *balance units* on the right side $\mathcal{F}$ could be linear in $U(x')$. This is quantified next.

For brevity, use the notation $U(x') \equiv U$. Assume that, at any fixed x' and z, $\mathcal{F}$ is an analytical function of U about the amplitude value $U = 0$. (See *General comment 2* at outset.) This will allow the above balancing of units argument to hold. By the analyticity, over a sufficiently small range of values U, $\mathcal{F} = \mathcal{F}(z, x', 0) + (\partial\mathcal{F}/\partial U)U$ + ...+ higher-order terms in U. Of these terms, the first $\mathcal{F}(z, x', 0)$ must not contribute since it would give a nonzero $u(z)$ even in *the absence* of a source $U(x)$. Also, the coefficient $(\partial\mathcal{F}/\partial U)$ and those of higher powers are evaluated at $U = 0$, and hence are only functions of z, x'. For example, we can denote $(\partial\mathcal{F}/\partial U) \equiv k(z, x')$, some unknown 'kernel' function of its arguments.

By the weak-field approximation $|U| << 1$ (Sec. I.A) the higher-order terms of $\mathcal{F}$ are negligable. Then, from all these considerations, only the linear term $k(z,x')U$ in $\mathcal{F}$ contributes, giving as its integral

$$u(z) = \int dx' U(x') k(z,x') \equiv T[U(x)]. \tag{6}$$

The integration x' is over the aperture. Hence, the problem to be attacked by EPI is now reduced *to finding functional* T, or equivalently, the general kernel function $k(z,x')$.This is a decisive step since T will ultimately become the de Broglie Fourier relation we seek.

Note that *no prior assumption of unitarity has been used to establish (6).* In fact, by the use of EPI the kernel $k(z,x')$ in (6) *will be derived* by EPI to have the unitary form (23), with (24) expressing the unitary transform.

B. Aperture slit

Let the aperture have finite extension x_0 on the symmetric interval $(-x_0/2, x_0/2)$. Temporarily consider the case where *a slit is placed at a position* x *in the aperture*,

$$U(x') = x_0 U(x) \delta(x'-x) \text{ and } u(z) \equiv u(z|x), \tag{7}$$

the latter chosen notation for "if x". (We could equally well have used a comma, as $u(z,x)$.) Also, note that the factor x_0 is needed in (7) to balance its units, since, by $\int_{-x_0/2}^{x_0/2} dx' \delta(x') = 1$, $\delta(x') \sim l^{-1}$, i.e., it has units of $length^{-1}$. (However, x_0 will drop out of consideration later.) The given particle passes through the arbitrary, unknown point x en route to image space, where the resulting amplitude u at any z is contingent upon the fixed x, thus the notation $u(z|x)$.

Eqs. (6) and (7) together give

$$u(z|x) = x_0 \int dx' U(x') \delta(x'-x) k(z,x') = x_0 U(x) k(z,x) \tag{8}$$

by the sifting property of the delta function. This relation will be used below in order to find $k(z,x)$.

IV. EPI PROBLEM

We continue with the special case of a slit at a fixed position x. Here, there is a transition $x \rightarrow z$ from a fixed slit position x to a random image position z. In this "if x" scenario, the EPI output due to transition (9) will be the amplitude function $u(z|x)$. Also, informations (5) are now appropriately designated as $I_X(0) \equiv I_X(x)$ and $I_Z(0) \equiv I_Z(x)$,emphasizing the contingency upon x. The information transition is

$$I_X(x) \rightarrow I_Z(x), \tag{9}$$

where $I_X(x)$ is a fixed but unknown number formed by plugging the unknown $U(x)$ into (3). We find next how $U(x)$ and $I_X(x)$ relate, respectively, to $u(z)$ and $I_Z(x)$.

A. Variational problem

In general, an EPI problem describes an information transition $J \rightarrow I$, via a principle $I - J = extrem.$ Quantities J and I are generally *functionals* [4],[5] of an unknown system amplitude function. Here we are tracking the information from *the single slit position* x to a general point z along the trajectory, with effective source and data informations denoted as $I_X(x)$ and $I_Z(x)$, where $I =< I_Z(x) >$ and $J =< I_X(x) >$. The problem is then

$$I_Z(x) - I_X(x) = extrem., \; I =< I_Z(x) >, \; J =< I_X(x) > \tag{10}$$

for a general *fixed* value of x, with $u(z|x)$ to be varied to obtain the extremum. The two average informations I and J will later be found to be equal.

B. Shift-invariance revisited

The shift-invariance condition (4) holds for any aperture function $U(x)$, and therefore for our *particular* case of a slit aperture centered on x. Here the 2nd line of (4) becomes

$$p_Y(y|a, x) \equiv p_Z(y - a|a, x) = p_Z(y - a|x) \equiv p_Z(z|x) \equiv u^*(z|x)u(z|x). \tag{11}$$

In this x-conditional case, the information $I_Z(x)$ directly becomes the general form (3) with $A(x) \equiv u(z|x)$,

$$I_Z(x) = 4 \int dz u^{*\prime}(z|x) u'(z|x). \tag{12}$$

Primes $'$ denote a derivative $\partial/\partial z$. The information (12) defines the left-most input into principle (10). The other information, $I_X(x)$, is found next.

C. Information in aperture

Informations are generally positive if the physical coordinates a, x, y, z are real [4],[5], as here. Hence, the unknown but fixed information $I_X(x)$ in the aperture obeys $I_X(x) \geq 0$. This allows us to represent $I_X(x) \equiv 4f^2(x)$, with f some real function. But, since slit parameters a and L, are also fixed, more generally $I_X(x) = 4f^2(x, a, L)$, with f to be found. However, as was discussed following (5), the *level* of Fisher information is independent of the shift a. Then $f \to f(x, L)$ so that

$$I_X(x) = 4f^2(x, L) = 4f^2(x, L) \int dz u^*(z|x) u(z|x), \tag{13}$$

where additionally we used the normalization property of the PDF $p_Z(z|x)$. (We could have instead tried the normalization of $p_X(x)$ over $x-$space, but this would not work in implementing the variational problem (10), which must contain functionals $I_Z(x)$, $I_X(x)$ over *the same* space, and $I_Z(x)$ in (12) is already over $z-$space.) Function $f(x, L)$ is to be found from the EPI solution.

It is interesting that the solution $u(z|x)$ to extremum principle (10) will actually give an extremum value of zero (Sec. V.F), i.e. the minimum possible, so that $I_Z(x) - I_X(x) = 0$. Information is conserved at each end of the trajectory AB from aperture to image plane.

Finally, *we require function $I_X(x)$ to be analytic* about positions x centered on $x = 0$. Thus, $I_X(x)$ and its derivatives must be finite at $x = 0$. This makes intuitive sense, since $I_X(x)$ represents the contribution of the aperture point x to the information about a, and there is nothing about position $x = 0$ that would make it convey infinite information, i.e., be a pole. (To the contrary, it will be found that $I_X(0) = 0$.) We may note in this respect that Fisher information is always an analytic function of the parameters [20] (specifically, quadratic in x [Eqs. (35)]). But most importantly, since $I_X(x) = 4f^2(x, L)$ then $f(x, L)$ is itself *analytic* about point $x = 0$. This will be key to subsequently finding $f(x, L)$.

D. Euler-Lagrange output

Using Eqs. (12) and (13) in principle (10) then defines $u(z|x)$ as the solution to the variational problem

$$I_Z(x) - I_X(x) = \int dz u^{*\prime}(z|x)u'(z|x) - f^2(x,L)\int dz u^*(z|x)u(z|x) = extrem. \tag{14}$$

(after division by 4). As usual, the solution is provided by an Euler-Lagrange equation

$$\frac{\partial}{\partial z}\left(\frac{\partial \mathcal{L}}{\partial u^{*\prime}}\right) = \frac{\partial \mathcal{L}}{\partial u^*}\ , \quad u \equiv u(z|x),\ \text{where}\ \mathcal{L} = u^{*\prime}u' - f^2 u^* u \tag{15}$$

is the integrand of (14). See also [16] for a related problem using Fisher information.

Combining Eqs. (15) directly gives the differential equation

$$u''(z|x) = -f^2(x,L)u(z|x). \tag{16}$$

The primes denote derivatives $\partial/\partial z$. This is a stationary-state wave equation governing the amplitude $u(z|x)$. Note that the fully x-dependent shape of the *function* $f^2(x,L)$ will later come into play.

E. Linearity condition, and solution

Linearity $u(z|x) \propto U(x)$ was derived at (8). The solution to Eq. (16) that satisfies this is directly

$$u(z|x) = CU(x)\exp\left[if(x,L)z\right], \tag{17}$$

for some constant C. This may be easily verified by trial substitution into (16).

Note that (17) has resulted *without assuming* prior knowledge of quantum properties. Thus, its phase part $f(x,L)z$ is *not presumed* to have physical significance. Any such significance will have to *result* from the dynamics that are implied by (17).

F. Kernel $k(z,x)$

Eqs. (8) and (17) can be equated, giving

$$u(z|x) = x_0 U(x) k(z,x) = CU(x) \exp\left[if(x,L)z\right]. \quad (18)$$

Then the kernel is known as

$$k(z,x) = C \exp\left[if(x,L)z\right], \quad (19)$$

for convenience absorbing a factor x_0^{-1} into C. Using (19) in (6) gives

$$u(z) = C \int dx' U(x') \exp\left[if(x',L)z\right]. \quad (20)$$

The analytic function $f(x,L)$ is to be found. Note that once this is done, the likelihood amplitude function $u(z)$ is known from (20).

G. Expansion for $f(x',L)$

We found in Sec. IV.C that function $f(x,L)$ is analytic about $x=0$. Also, by (1), x/L is small. Hence, as in Sec. II.D, we can take advantage of these properties by re-expressing $f(x,L)$ as some function $b(x/L,L)$. This must likewise be analytic around point $x/L=0$. Thus, over some non-trivial range of values x/L, function $b(x/L,L)$ obeys a power series

$$f(x,L) \equiv b(x/L,L) = b(0,L) + b'(0,L)(x/L) + \frac{1}{2} b''(0,L)(x/L)^2 + ... \quad (21)$$

Primes denote derivatives $\partial/\partial(x/L)$ and the zeros mean as evaluated at $x/L=0$. By (1), x/L is small, so that second- and higher-power terms in (21) are insignificant. The result is

$$f(x,L) = b_0 + b_1 x/L, \quad b_0 \equiv b(0,L) = \text{const.}, \; b_1 \equiv b'(0,L) = \text{const.} \quad (22)$$

The constants b_0, b_1 can generally depend upon the constants μ_0, L and one or more universal physical constants. Also, by (19) and (22) the kernel $k(z,x)$ is

$$k(z,x) = C \exp\left(ib_0 z + ib_1 xz/L\right) = C \exp\left(ib_0 z\right) \exp\left(ib_1 xz/L\right) \quad (23)$$

Then, by (6) and (23),

$$u(z) \exp\left(-ib_0 z\right) \equiv \hat{u}(z) = C \int dx' U(x') \exp\left(ib_1 z x'/L\right) \quad (24)$$

where $\hat{u}(z)$ is simply $u(z)$ offset by the linear phase $-ib_0z$. Also, the constant C may be found by demanding that the PDF $|\hat{u}(z)|^2$ obey normalization, given that $|U(x')|^2$ does as well. Squaring and integrating out (24) in this way gives

$$C = (2\pi L)^{-1/2} b_1^{1/2}. \tag{25}$$

We next evaluate the constant b_0 which, from definition (21), is some fixed function of L. Multiplying both sides of (24) by $\exp(-ib_1zx/L)$ and integrating dz gives a shifted function $U(b_0 + b_1x/L)$, and hence a shifted PDF $p_X(b_0 + b_1x/L)$. This predicts that all PDFs with the same L are *universally shifted* by the same amount b_0.This result is inconsistent with the system property of shift invariance previously found, and also with empirical evidence, so that

$$b_0 = 0, \text{ and } \hat{u}(z) = u(z) \tag{26}$$

by (24).

For our weak amplitude function $|U(x')| << 1$, and if $U(x')$ is now regarded as a scattering potential, (24) determines $u(z)$ as obeying familiar *first Born-approximation scattering* [8]. However, for purposes of analysis, we continue regarding $U(x')$ as a probability amplitude.

V. FINAL DE BROGLIE FOURIER FORM

A. Change of coordinate

By Sec. I.D, coordinate x is linear in the momentum μ, obeying

$$x \approx (\mu/\mu_0)L, \quad dx = (L/\mu_0)\, d\mu, \tag{27}$$

provided L is large. Then by (24)-(27),

$$u(z) = \left(\frac{Lb_1}{2\pi\mu_0^2}\right)^{1/2} \int d\mu U\left(L\mu/\mu_0\right) \exp(ib_1z\mu/\mu_0). \tag{28}$$

Since the exponent in (28) must be unitless, $b_1$has the unit of $length^{-1}$, denoted as $b_1 \sim l^{-1}$. Also, in (28), the argument $(L\mu/\mu_0)$ of U has the unit of position, which checks.

B. Probability amplitude for momentum

We have yet to define the required probability amplitude $P(\mu)$ on the *momentum* (notation P adapted by analogy with the lens "pupil" function in the corresponding optical analogy). This must be proportional to U, obeying $P(\mu) \equiv BU(L\mu/\mu_0)$ for some constant B. Requiring that the PDF $|P(\mu)|^2$ integrate to unity, assuming $|U(L\mu/\mu_0)|^2$ does, fixes $B = \sqrt{L/\mu_0}$. This gives

$$P(\mu) = \sqrt{L/\mu_0}U(L\mu/\mu_0). \tag{29}$$

Then (28) becomes

$$u(z) = \sqrt{\frac{b_1}{2\pi\mu_0}} \int d\mu P(\mu) \exp(ib_1 z\mu/\mu_0). \tag{30}$$

As another units check, each side of (30) has the same unit $l^{-1/2}$ (since PDF $|u(z)|^2 \sim l^{-1}$, and using the above unit for $b_1 \sim l^{-1}$ and that $P \sim \mu^{-1/2}$).

C. Fixing $b_1 \propto \mu_0$

We note that (30) contains the particular experimental momentum value μ_0, and an unidentified parameter b_1. With the usual aim of maximally broadening the scope of application of a theoretical study, we ask whether μ_0 in (30) can be somehow cancelled, thereby recasting it as a *universal Fourier-transform result.* Since μ_0 occurs in (30) only as ratio b_1/μ_0 (in two places), to attain the desired cancellation b_1 must obey

$$b_1 \equiv \mu_0/K, \quad K = \text{const.}, \tag{31}$$

with K a new constant. In addition, (30) does not depend upon *any other* parameters $a, L, \mathbf{M}$ of the experiment. Then this choice of b_1 gives the result (30) universal applicability. Hence, as a working hypothesis, we require such universal applicability, and let b_1 have this value. Will it lead to physically valid predictions?

D. Resulting De Broglie-Fourier representation

Using (31) in (30) gives the main result of the paper,

$$u(z) = \frac{1}{\sqrt{2\pi K}} \int d\mu P(\mu) \exp(iz\mu/K), \text{ with } P(\mu) = \frac{1}{\sqrt{2\pi K}} \int dz u(z) \exp(-iz\mu/K) \quad (32)$$

as the direct inverse. These lack any dependence upon factors a, L or μ_0, thereby achieving our goal of establishing a universal Fourier relation of quantum mechanics. They also define the unitary transformation pair we set out to find in Sec. III.A. Thus, transform relations (32) have been formed as *predictions* of the overall approach (including the working hypothesis of the preceding subsection). As a bonus, unitarity also implies [5,11] the validity of the very EPI approach we used.

Of course relations (32) have also been abundantly verified over the past 80 yrs, verifying the overall calculation, with the empirically chosen value

$$K \equiv \hbar. \quad (33)$$

Eqs. (32), (33) are then the De Broglie-Fourier transform pair that is the main result of the paper.

However, as shown below, a value $K \equiv 0$ is another physically meaningful choice. This choice will yield a particle that obeys *classical mechanics.* Note that the choice (33) has the required units of *momentum* $\times$ *length* mentioned above.

In summary, if we require the result (30) to represent a *universal law* of quantum mechanics, the particular answers (31)-(33) follow. The result (32) will support *both* classical and quantum descriptions of the particle, *depending upon choice of the empirically defined constant* K.

E. de Broglie wavelength

The use of primed coordinates μ' and a pupil function $P(\mu') = \delta(\mu' - \mu)$ in the first Eq. (32) describes the case of a particle with definite x-component momentum μ. The delta function sifts out an amplitude function obeying proportionality $u(z) \propto \exp(iz\mu/\hbar)$; this repeats itself at points $z + 2\pi n\hbar/\mu$, $n = 1, 2, ...$ The minimal $z-$distance for repetition is for $n = 1$. This defines, respectively, an $x-$component wavelength and its usual reciprocal-space [8] wavenumber k_x, as

$$\lambda = h/\mu,\ \ k_x \equiv 2\pi/\lambda = \mu/\hbar. \tag{34}$$

The latter expression derives from the first. It also is one component of the usual vector expression $\mathbf{k} = \mu_0/\hbar$ of the de Broglie hypothesis in wavenumber space. This vector relation may be *derived* in 2D, in analogous steps to the EPI approach beginning in Sec. 3 and ending at (34). This is for the space (x, w) lateral to the imaging axis of Fig. 1, with coordinate w (not shown) at right angles to the page.

In summary, the de Broglie wave result (32),(34) followed from the requirements of universality preceding Eq. (31) and the periodic behavior noted above (34).

F. Resulting informations

With b_1 determined by (31), and by use of (22), (26) and (33) in (13), the information at the aperture point x is

$$I_X(x) = 4f^2(x, L) = (\hbar L)^{-2}\,(2\mu_0 x)^2\,,\ \text{with}\ \rho_J(x) \equiv |U(x)|^2(\hbar L)^{-2}\,(2\mu_0 x)^2 \tag{35}$$

defining an information *density* by definition (10) of total information J. Thus, the information density about a increases quadratically with x, as modulated by the local aperture intensity $|U(x)|^2$.

This has a parallel in the optical analog to Fig. 1 that we discussed: It is long known [10] that the outer (high x) zones of a lens permit *higher resolution* of a source position a than do the inner zones. By (35), $\rho_J(x)$ likewise tends to carry a locally higher level of information about the value of a. Conversely, the innermost zone at $x = 0$ contributes only a real, DC amount to $i(z)$, which has zero phase and hence cannot provide phase shift information proportional to a. This is again confirmed by the prediction (35) that $I_X(0) = \rho_J(0) = 0$.

Eq. (35) gave the solution for the aperture information $I_X(x)$. What is the resulting image information $I_Z(x)$? From the first Eq. (35), and (22), (31) and (33), we obtain $I_Z(x) = I_X(x)$. Then also, by (10), $I = J$, i.e., both locally and (now) overall there is conservation of information from aperture to image plane. This is also intuitively correct, since the intervening space over path AB is free space. Note that the equality of the two informations is the most direct way of stating the truism that *position space and momentum space contain equivalent information about the phenomenon* (here, parameter a).

It is interesting that result (35) also confirms past results where Fig. 1 describes a quantum scenario in particular. There [4,5] the aperture (called the 'source') information had value $J = 4 < \mu^2 > /\hbar^2$. Here, (35) gives, by use of the identity $x/L = \mu/\mu_0$ (Sec. I.D.(ii)), $I_X(x) = (2\mu/\hbar)^2$. Next, by the material above Eq. (10), the full information $J = \langle I_X(x) \rangle$. Thus the last two equations give

$$J = 4 < \mu^2 > /\hbar^2,$$

confirming the previous value for J.

G. ML solutions

The original aim of the calculation was to ML estimate position a, which is by definition the a obeying $p_Z(y-a) \equiv |u(y-a)|^2 = \max$. It is most convenient to work from Eq. (24). This and Eqs. (31),(32) give a condition on a of

$$\left| \int dx U(x) \exp\left[iB(y-a)x\right] \right|^2 = \max, \quad B \equiv \mu_0/\hbar L \tag{36}$$

after a factor $|\exp[ib_0(y-a)]|^2 = 1$ becomes irrelevent. In general, the solution a to (36) depends upon the form of $U(x)$. As examples, if $U(x)$ is Gaussian, the ML solution is simply $a = y$, the data value. Or, if $U(x) = \exp(ix/X)$ over some interval of length X, the ML solution is $a = y + (\hbar/\mu_0)(L/X)$. Notice that the units of this solution balance as those of length l.

H. Limiting cases K

We next show that, depending upon the sizes of K and m, the EPI outputs (32), (35) give either quantum mechanics, with well-defined phase information; or classical mechanics without the phase.

1. *Limit as* $K \to 0$

The first Eq. (32) has an interesting limit. By inspection, for any aperture profile $P(\mu)$, as $K \to 0$ the function $u(z)$ collapses *inward,* toward a delta function $\delta(z)$. Essentially only the value $z = 0$ is approached, a *deterministic limit.* This is confirmed by the use of $A(z) \equiv u(z) \propto \delta(z)$ in definition (3), giving $I \equiv I_Z = \infty$ (e.g., as easily shown in Gaussian case $u(z)$). By the Cramer-Rao inequality [11],[12] $e^2 = 1/I_Z$, the resulting mean-square error e^2 in the estimate of a approaches zero. Then, by (2), the observation $y = a$ perfectly. Also, use of $K = 0$ (replacing $\hbar$) in (35) gives $I_X(x) = \infty$, meaning that the aperture plane likewise contains perfect knowledge about a. Thus the mathematics are saying that in the limit $K \to 0$ the entire trajectory of the particle from object plane to aperture plane to image plane is *deterministic*, following a known trajectory. The particle acts classically, and by $p_Z(z) \equiv |u(z)|^2 \ \to \delta(z) = \delta(y - a)$, the ML estimate a defined as maximizing $|u(y-a)|^2$ becomes simply y, the datum. This solution is discussed further in Sec. VI.A.

2. K taken as finite

For K finite, as discussed above (33) any empirical investigation of the problem would find that $K = \hbar$. Using this in (32) then shows that $u(z)$ and $P(\mu)$ obey the De Broglie-Fourier representation. Here, $u(z)$ generally has meaningful phase. A famous exception is as follows.

3. Limit as mass $m \to \infty$

We found at (33) and (34) that the particle has an effective wavelength $\lambda = 2\pi\hbar/\mu = h/\mu$. Then using $\mu = mv$, v the velocity, gives $\lambda = h/mv$ so that, for given v, as $m \to \infty$ the wavelength $\lambda \to 0$. Hence, the oscillations of the SWE (see (37) below) become ever tighter, and the equation approaches being mathematically *ill-posed.* However, by Ehrenfest's theorem these oscillations can be eliminated out of an appropriate averaging process, yielding classical mechanics. The well-known price paid is that all phase information is lost.

In summary, a particle with very large mass will have large momentum compared to h (or even to any finite K), resulting in a very small wavelength and, therefore, complete loss of phase information. Here the phase function predicted by the EPI output is ill-defined and irrelevent. This is the usual definition of a classical particle. Hence, as $m \to \infty$ the quantum solution resulting from the use of $K = \hbar$ becomes effectively a classical $K = 0$

solution.

I. SWE

At this point we have achieved the intended goals of deriving the Born approximation (24) and the de Broglie wave hypothesis (32)-(34). However, for completeness, we now show how the SWE follows from the preceding. An additional physical assumption, that of energy conservation, is now required.

The preceding EPI calculation was for the special scenario of a slit aperture $U(x)$ at x fixed. We now allow for a *fully open* aperture $U(x)$, all x. Consequently $u(z)$ replaces $u(z|x)$ in (12), with $I \equiv 4\int dzu^{*\prime}(z)u'(z)$, and $J = 4 < \mu^2 > /\hbar^2$ from the preceding. Assume, in addition, conservation of energy $W = \mu^2/2m + V(z)$, where W is a fixed energy value and $V(z)$ is a known potential function (not shown) in the image plane of Fig. 1. Then the new EPI principle $I - J = extremum$ has a Lagrangian $\mathcal{L} = 4|u'(z)|^2 - (8m/\hbar^2)(W - V(z))|u(z)|^2$, and its Euler-Lagrange solution (15) is [4,5]

$$u''(z) + \frac{2m}{\hbar^2}(W - V(z))u(z) = 0. \tag{37}$$

Note that the shift-invariance (4) we assumed for the system is, in fact, reflected in this result, which manifestly does not depend on a. This SWE has well-defined solutions [3] for sufficiently small mass m (atomic level or less).

VI. DISCUSSION

We have shown that the ordinary statistical problem of finding an ML estimate of a source position a requires knowledge of the system likelihood law $p_Y(y|a)$ or, to foster positivity, knowledge of its underlying amplitude law $u(y - a) = u(z)$. The latter is estimated, via EPI, to be the D-F transform (32) of the aperture function $P(\mu)$. This and the SWE (37) have some further important ramifications, discussed next.

A. Could the 'hidden variables' have been known?

The entire calculation assumed that the initial conditions for the particle are unknown, including its position and velocity at $t = 0$. But, in fact, *physically*, *could* the latter have been known? We choose here the more interesting (quantum) case $K = \hbar$.

Standard deviations $\Delta\mu, \Delta z$ are often used to measure the uncertainties in momentum μ and position z of a particle. Fourier relations (32) give rise to uncertainties $\Delta\mu, \Delta z$ that may be shown [21] to obey a Heisenberg uncertainty principle

$$\Delta\mu\Delta z \geq \hbar/2. \tag{38}$$

This of course holds for any conjugate pair of position and momenta, e.g. in the object plane of Fig. 1. Thus, the particle momentum $\mathbf{M}$ and position a there *could not have been known* with arbitrary accuracy. Therefore our problem groundrules that such knowledge is to be missing was not merely hypothetical but, actually, demanded physically. This gives deeper meaning to the assumption of missing initial conditions. See also below.

We temporarily return to hypothetical cases where K is finite and not necessarily of value $\hbar$. The affect of mass $m \to \infty$ on a finite-K solution was taken up in Sec. V.H.3. Here we examine *the opposite limit* $m \to 0$. At the source S, the particle momentum $\mathbf{M}$ is of size $M = mv$, where v is its velocity. Therefore, for an uncertainty $\Delta\mu$ in M, taking differentials gives $\Delta v = \Delta\mu/m$.Then by (38) with $\hbar = K$ as here,

$$\Delta v\Delta z \geq K/2m. \tag{39}$$

Thus, precise knowledge of (now) the velocity v and position z in the object plane could likewise not have been available. In fact the use of classical mechanics requires knowledge of *velocity* (rather than momentum) and position as initial conditions. In the limit $m \to 0$ (39) indicates errors in these initial conditions become ever larger. Therefore a classical approach would have been ill-posed, with the resulting datum y departing from its ideal value a by some randomly large noise amount z. As a result, for particles on the atomic level or smaller, our classical-alternative ML alternative answer taking $K = 0$ in the EPI output (Sec. V.H.1) becomes uselessly inaccurate. Only the quantum ML solution, with $K = \hbar$, remains tenable for such particles. (The opposite case of $m \to \infty$ was discussed in Sec. V.H.3)

Alternatively, the uncertainty principle (39) indicates that in cases of *large* or macroscopic masses the $K = 0$ alternative *would* be useful. As was discussed in Sec. V.H.1, this is the classical-mechanics solution to the problem.

In summary of the preceding, with large mass the classical solution $K = 0$ is the accurate choice, whereas with small mass only the quantum solution, with $K = \hbar$, is the accurate solution.

Parenthetically, we note that were the groundrules of the problem *different*, with PDFs $p(a), p(\mu)$ now *known*, they would of course have to satisfy the inequality (38). It is possible to construct such PDFs, e.g. the well-known case [3] where both are Gaussian. However, this defines a new estimation problem, with different prior knowledge than as assumed in this paper. We do not know what the new ML answers for a and $p_Z(z)$ would be.

In Sec. I.D, even despite the knowledge of shift invariance, *Noether's theorem* was *not* used to imply conservation of momentum. Rather, the conservation was *postulated*. In [22] the authors discuss circumstances under which known symmetry implies, via Noether, a conservation law. These can be quite involved and generally require additional physical assumptions. On this basis, we do not use Noether.

VII. CAPSULE SUMMARY

A full summary I.-VI is given at the outset. The following, shorter overview may be useful: We seek to ML-estimate (Sec. I.G) the input source position a of a particle that obeys unknown dynamics as it travels through the apparatus of Fig. 1. There is a complete absence of prior knowledge of initial conditions (Secs. II, II.A). The only prior physical assumption about the particle dynamics is that of conservation of momentum (Sec. I) along path AB. Forming the ML estimate requires that the particle's likelihood amplitude function $u(z)$ (Secs. II.C,D) be known; see also the *General comments 1,2* at the outset of the paper. Because the EPI principle (Secs. III,IV) maximizes the received level I of information about a, and outputs a maximally informative $u(z)$, EPI is chosen for this task. Depending on the value of a free parameter K, the output $u(z)$ is found to obey either the Fourier representation (32) (K finite, Secs. V.D,E,H) or Dirac $\delta(z)$ behavior ($K \equiv 0$, Sec. V.H.1). These, respectively, correspond to either quantum-*like* or classical motion for the particle. The particular empirical choice $K \equiv \hbar$, Planck's constant, in (32) makes it the fully quantum D-F relation we sought, with (34) now the $x-$component de Broglie wavelength. Geometrical far-field assumptions (1) are extensively used, as are assumptions of analyticity and a small-amplitude mathematical approximation (Sec. I.A).

APPENDIX A: FISHER CHANNEL CAPACITY

Consider the shift-invariant Fisher channel-capacity measure

$$I \equiv 4 \int dx u^{*\prime} u', \quad u' \equiv du/dx \tag{A1}$$

in a space x. We show that it is a channel capacity, i.e. an *upper bound*

$$I \geq I_f \tag{A2}$$

to the Fisher information

$$I_f \equiv \int dx p'^2/p, \quad p \equiv u^* u. \tag{A3}$$

First, since p is real, $p' = u^* u' + u^{*\prime} u \equiv 2\,\mathrm{Re}(u^* u')$ is likewise real. From the latter and the far-right (A3)

$$p'^2/p = \frac{[2\,\mathrm{Re}(u^* u')]^2}{u^* u}. \tag{A4}$$

Then, the theorem (A2) will follow if we can show that, at each x, the integrand of (A1), and (A4), obey

$$u^{*\prime} u' \geq \frac{[\mathrm{Re}(u^* u')]^2}{u^* u} \tag{A5}$$

(the factors 4 cancelling out). Cross-multiplying, (A5) is valid if

$$|u^* u'|^2 \geq [\mathrm{Re}(u^* u')]^2. \tag{A6}$$

But, by definition of the modulus-square, $|u^* u'|^2 = [\mathrm{Re}(u^* u')]^2 + [\mathrm{Im}(u^* u')]^2$. Since both these right-hand terms are positive, (A6) follows. QED.

But, can the Fisher channel capacity (A1) be *physically realized*? That is, can form (A1) for the channel capacity equal the I_f for a physically meaningful scenario?

APPENDIX B: REALIZATION OF FISHER CHANNEL CAPACITY

The concept of Fisher channel capacity would have limited real import if it could not be realized by physical systems. We show here that EPI solutions can, in fact, *physically achieve* their channel capacities.

Generally let the particle amplitude function $u(x) \equiv q_1(x) + iq_2(x)$, where q, q_2 are real. Any $u(x)$ achieving channel capacity (A1) for its I_f is one where the real and imaginary parts $q_1(x), q_2(x)$ *do not overlap in* x. (An example is separated triangle functions $q_1(x), q_2(x)$.) The reader can easily verify this, by using this separated $u(x)$ in (A3), giving

$$I_f = 4 \int dx(q_1'^2 + q_2'^2) = 4 \int dx u^{*\prime} u'. \tag{B1}$$

All cross terms have dropped out of the latter, because of the non-overlap property. QED.

However, does such a non-overlapping amplitude distribution satisfy the requirements of (say) quantum mechanics? Consider a generally time-dependent wave function $\Psi(x,t)$ solution of the time-varying SWE. The rules of quantum mechanics [3] require that any such wave function $\Psi(x,t)$ have *continuously varying* amplitude and phase at each x and t. This should then hold as well for its value $\Psi(x,0)$, where $\Psi(x,0) \equiv u(x) \equiv q_1(x) + iq_2(x)$ preceding. Amplitude $\Psi(x,0)$ is, in fact, an input boundary-value condition to be *imposed* upon the SWE [3]. But, for spatially non-overlapping functions $q_1(x), q_2(x)$, the phase $\phi(x)$ in particular, cannot be everywhere continuous. For example, let the domain of the real part $q_1(x)$ be all $x \leq 0$, and that of the imaginary part $q_2(x)$ be $x \geq 0$. Then the phase $\phi(x) \equiv \tan^{-1}[q_2(x)/q_1(x)] = 0$ for $x < 0$, but is $\pi/2$ for $x > 0$, with a discontinuity at $x = 0$ (where the two separated regions touch). The requirement of continuity is therefore violated. Can it somehow be retrieved?

We take recourse in the concept of a 'generalized function' [23]. In particular, slightly blur each of $q_1(x), q_2(x)$ by convolution with a Gaussian (say) of variance σ^2. This has the physical significance of a slight coarse graining operation due to increased disorder. After the convolution, the two amplitude functions slightly overlap, so that *the phase is now well defined* at all x, including $x = 0$. However, since the two amplitude functions are no longer strictly separated, (B1) is no longer strictly obeyed. The overlap gives a small contribution to a cross term. Hence, we simply use cases where σ^2 is nonzero, *but small*, i.e. $\sigma^2 \to 0$. This now defines a pair of generalized amplitude functions $q_1(x), q_2(x)$ with the required properties in the limit. QED

REFERENCES

1. L. de Broglie thesis, *Research on the theory of quanta* (Université de Paris, 1924)

2. L. de Broglie, *An introduction to the study of wave mechanics* (Dutton, New York, 1928); reprinted (Methuen, London, 2007); P. Roman, *Theory of Elementary Particles* (North-Holland, Amsterdam, 1961, p. 103)

3. L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York,1955), pp. 17-21.

4. B.R. Frieden and B.H. Soffer, Phys. Rev. E **52**, 2274-2286 (1995)

5. B.R. Frieden, *Science from Fisher Information, 2nd ed.* (Cambridge Univ. Press, UK, 2004)

6. R.A. Fisher, *Statistical Methods and Scientific Inference, 2nd ed.* (Oliver and Boyd, London, 1959)

7. R.A. Fisher, Ann. Eugenics **12**, 1 (1943)

8. J.M. Cowley, *Diffraction Physics, 3rd ed.* (Elsevier, Amsterdam, Holland, 1995)

9. N. Bohr, "The quantum postulate and the recent development of atomic theory," Nature **121**, 583 (1928). This is the original reference to the famous gedanken microscope experiment (but lacking the usual figure). It was a response to other such gedanken experiments proposed by A. Einstein in the continuing Bohr-Einstein dialogue about early quantum mechanics. See also D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951), pp. 104-105; or Ref. 3, pp. 9-10. Both include the figure.

10. J.W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York,1965); P. Dumontet, Optica Acta **2**, 53-63 (1955); W.T. Welford, *Aberrations of the Symmetrical Optical System* (Academic Press, New York, 1974)

11. B. R. Frieden, *Probability, Statistical Optics and Data Testing, 3rd ed.* (Springer-Verlag, New York, 2001)

12. H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part I* (Wiley, New York,1968)

13. B.R. Frieden and R.A. Gatenby, eds., *Exploratory Data Analysis using Fisher Information* (Springer-Verlag, London, 2007)

14. A precursor of EPI is the recognition that the Von Weiszacker energy term is effec-

tively a Fisher information, in: S.B. Sears, R.G. Parr, U. Dinur, Israel J. Chem. **19**, 165 (1980). We recently learned of this work from A. Nagy [18].

15. R. Carroll, *On the Quantum Potential* (Arima, Suffolk, UK, 2007)

16. S. Luo, Found. Phys. **32**, 1757 (2002)

17. B.R. Frieden and B.H. Soffer, Phys. Rev. A **74**, 052108 (2006)

18. A. Nagy, J. Chem. Phys. **119**, 9401 (2003)

19. B.R. Frieden, "Extreme physical information as a principle of universal stability," in *Information Theory and Statistical Learning*, eds. F. Emmert-Streib and M. Dehmer (Springer-Verlag, N.Y., 2008)

20. R.J. Yanez, P. Sanchez-Moreno, A. Zarzo, J. Math. Physics **49**, 082104 (2008)

21. R.M. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1965)

22. J. Rosen and Y. Freundlich, Am. J. Phys. **46**, 1030-1041 (1978)

23. This approach has famously been used to properly define the Dirac delta function, Feynman path integal, etc. See M. J. Lighthill, *Introduction to Fourier Analysis and Generalised Functions* (Cambridge University Press, UK, 1959); or [21]