The diffraction and spreading of a wavepacket

Guy Vandegrift[a)]
Department of Physics, Purdue University North Central, Westville, Indiana 46391

(Received 5 June 2002; accepted 23 May 2003)

The spreading of a one-dimensional wavepacket solution of Schrödinger's equation is related to the diffraction of light, as can be verified by considering the three-dimensional spreading of a wavepacket for an arbitrary dispersion relation. This investigation uncovers a special property of Schrödinger's equation for a free particle: A wavepacket with initial spherical symmetry will preserve this symmetry in all Galilean reference frames. This property leads to a derivation of de Broglie's postulate that wave number is proportional to momentum (or velocity). The application to non-Gaussian wavepackets and to Fraunhoffer diffraction also is discussed. © *2004 American Association of Physics Teachers.*
[DOI: 10.1119/1.1591763]

I. INTRODUCTION

Many students encounter Heisenberg's uncertainty principle in the context of the Fraunhoffer diffraction of light by a single slit taking de Broglie's hypothesis, $p=\hbar k=2\pi\hbar/\lambda$, as a starting point.[1] Passage through the slit yields information about the particle's position in the direction perpendicular to motion. The subsequent diffraction of the wave implies uncertainty in that component of the momentum. Although this derivation assumes that both light and particle waves diffract in essentially the same fashion, students should also know that pulses of light do not spread in the direction parallel to motion. This paper explores how arbitrary linear waves spread and diffract. I will show that both properties are governed by a symmetric tensor $\partial^2\omega/\partial k_i k_j$, which is obtained directly from the dispersion relation $\omega=\omega(\mathbf{k})=\omega(k_x,k_y,k_z)$.

We follow the familiar Gaussian wavepacket solution of Schrödinger's equation,[1] and introduce a coherence time, τ. Let $a=\langle(x-\langle x\rangle)^2\rangle^{1/2}$ represent the initial standard deviation Δx of the wavepacket, and assume that $\Delta x\Delta k$ takes on its minimum possible value at time $t=0$. As time evolves, the uncertainty (standard deviation) will grow as,[1,2]

$$\Delta x=a\left[1+\left(\frac{t}{\tau}\right)^2\right]^{1/2}. \tag{1}$$

Under certain circumstances, Eq. (1) can be generalized for an arbitrary dispersion relation, and for higher dimensions,

$$\tau=\frac{a/\Delta k}{\|\partial^2\omega/\partial k^2\|}=\frac{a/\Delta k}{\|\partial v/\partial k\|}, \tag{2}$$

where $v=\partial\omega/\partial k$ is the group velocity, and Δk is the standard deviation of the wave number. If the wavepacket has a Gaussian profile, then $a/\Delta k=2a^2$. If the dispersion relation is $\hbar\omega=\hbar^2k^2/2m+V$, we obtain the well-known result, $\tau=2ma^2/\hbar$. It can be understood as follows: The spread in group velocity is estimated by $\Delta v=(\partial v/\partial k)\Delta k$. For typical or ideal circumstances, a wavepacket of original size, a, will retain this size for a time, τ, where $\tau\Delta v\approx a$. Equation (2) follows directly, though only as an order of magnitude estimate. Sections II and III derive Eqs. (1) and (2) for non-Gaussian wavepackets in three dimensions.

In the following we assume the wavepacket to be of the form,

$$\psi(\mathbf{r},t=0)=\int d^3k\ \eta(\mathbf{k})\exp[i\mathbf{k}_0\cdot\mathbf{r}+i\phi_0], \tag{3}$$

where $\eta=\eta(\mathbf{k})$ is a real valued function, and $\mathbf{k}_0=\langle\psi|\mathbf{k}_{\rm op}|\psi\rangle$ is the expectation value of the wave number. This expectation value may be found either in $\mathbf{r}$ space, $\int d^3r\ \psi^*(-i\boldsymbol{\nabla})\psi$, or in $\mathbf{k}$ space, $\int d^3k\ \varphi^*\mathbf{k}\varphi$, where φ is the Fourier transform of ψ. The generalization to higher dimensions is achieved by replacing $\partial^2\omega/\partial k^2$ in Eq. (2) by the eigenvalues of the 3×3 tensor, $\boldsymbol{\mathcal{D}}=\partial^2\omega/\partial\mathbf{k}\partial\mathbf{k}=\partial^2\omega/\partial k_i\partial k_j$. [See the discussion after Eq. (15) below.] What might not seem intuitive is that $\boldsymbol{\mathcal{D}}$ has nonzero elements even when the dispersion relation has the nondispersive form, $\omega=Ck$, where C is a constant (which need not be the speed of light).

The three-dimensional spreading of a wavepacket unveils a nonrelativistic symmetry argument that motivates Schrödinger's dispersion relation for a free particle. Although the rigorous discussion of Sec. II requires tensor calculus, this symmetry argument can be intuitively understood as follows. Suppose that we begin with the understanding that particles are the limiting form of wavepackets, but without a knowledge of the dispersion relation or wave equation. De Broglie's relations ($E=\hbar\omega$ and $p=\hbar k$) have yet to be established. Nor do we know that the classical Hamiltonian for a free particle is $H(p,q)=p^2/(2m)$. We seek a nonrelativistic dispersion relation. Consider first the familiar dispersion relation, $\omega(k)=Ck$. Because the group velocity, $v=\partial\omega/\partial k=C$, never vanishes, this form is obviously unsuitable as a model for Newtonian particles. The dispersion relation, $\omega(k)=Ck$, is also unsuitable due to the asymmetric manner in which wavepackets spread. Recall that in vacuo light has the peculiar property of diffracting in directions perpendicular to its motion, while exhibiting no spreading (dispersion) in the direction parallel to its motion. If the wavepacket is to represent a particle, an initially spherically symmetric Gauss-

ian wavepacket must remain spherically symmetric for any (nonrelativistic) velocity. The absence of such a symmetry would suggest the existence of a preferred reference frame.

We therefore seek a dispersion relation for which the spreading is equal in all directions, and temporarily restrict ourselves to those of the form, $\omega = Ck^n$. The diffraction of a beam of light arises from the fact it is a superposition of waves traveling in different directions. The spreading of a one-dimensional wavepacket arises from the various components of a wavepacket travelling at different speeds. Equation (2) implies that large n is associated with more spreading along the direction of propagation. Therefore it is plausible that for some unique n, the dispersion relation, $\omega = Ck^n$, might possess the required balance of dispersion and diffraction necessary to maintain spherical symmetry. Although this plausiblility argument fails to suggest a specific value for n, we show in Sec. II how Schrödinger's dispersion relation is uniquely suited for this purpose.

This symmetry argument seems to be unique among plausibility arguments leading to de Broglie's postulates. De Broglie's arguments are based largely on special relativity, a topic not yet fully grasped by students as they begin to learn about wave-particle duality.[3,4] An entirely different class of plausibility arguments (Refs. 5 and 7) link the classical Hamiltonian to the dispersion relation via substitutions such as $p \to \hbar k \to -i\partial/\partial x$. Although this argument leads directly to de Broglie's relations, it requires a presumption that the classical free-particle Hamiltonian is $H(p,q) = p^2/2m$. Few people would object to this assumption, of course. Nevertheless it is fascinating to see the classical free-particle Hamiltonian derived solely from the fact that it represents the limiting case of a wavepacket.

Although most textbooks introduce the coherence time for Schrödinger's equation as $\tau = 2ma^2/\hbar$,[1,2] one advantage of the more general form of Eq. (2) is that numerical evaluation in terms of fundamental parameters is not always required, especially if one considers the coherence length, defined as $v\tau = (\partial\omega/\partial k)\tau$. For any dispersion relation of Schrödinger's form, $\omega = A + Bk^2$, the coherence length for a Gaussian wavepacket is, $v\tau = 2a^2k = 4\pi Na$, where N is the number of wavelengths contained within the wavepacket's initial size (standard deviation). If this concept can be applied (at least qualitatively) to an atom, the difficulty with using wavepackets to model low order atomic states is easily recognized: The wavepacket would lose coherence and cease to be a wavepacket as it makes approximately one orbit.

Strictly speaking, Eqs. (1) and (2) do not describe the spreading of an electron wavepacket near an atomic nucleus because the spatial variation in the potential energy, $V(\mathbf{r})$, violates the assumption of spatial homogeneity.[2] By Ehrenfest's theorem, wavepackets of Schrödinger's equation obey Hamiltonian equations of motion, $\dot{\mathbf{r}} = \partial\omega/\partial\mathbf{k}$, $\dot{\mathbf{k}} = -\partial\omega/\partial\mathbf{r}$, $\dot{\omega} = \partial\omega/\partial t$.[1] Here the wavepacket's coordinates in $(\omega,\mathbf{k},\mathbf{r},t)$-space represent expectation values, and $\omega = \omega(\mathbf{k},\mathbf{r},t)$ is a classical Hamiltonian.[5] These same canonical equations of wavepacket motion also describe other situations in the eikonal limit[5] that include solid-state physics,[6] plasma waves,[7] and general relativity.[8] The consideration of the spreading of a wavepacket when the dispersion relation exhibits such spatial inhomogeniety is beyond the scope of this paper.[2]

II. SPREADING AND DIFFRACTION IN THREE DIMENSIONS

We Taylor expand an arbitrary dispersion relation, assuming that the wavepacket is localized in $\mathbf{k}$ space, centered at some wave number k_0, and frequency $\omega_0 = \omega(\mathbf{k}_0)$:

$$\omega(\mathbf{k}) = \omega_0 + \mathbf{v}\cdot(\mathbf{k}-\mathbf{k}_0) + \frac{1}{2}(\mathbf{k}-\mathbf{k}_0)\cdot\frac{\partial^2\omega}{\partial\mathbf{k}\partial\mathbf{k}}\cdot(\mathbf{k}-\mathbf{k}_0)+\cdots$$

$$\equiv \omega_0 + \mathbf{v}\cdot\boldsymbol{\kappa} + \tfrac{1}{2}\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\kappa}+\cdots, \tag{4}$$

where $\boldsymbol{\kappa} = (\mathbf{k}-\mathbf{k}_0)$. The symmetric 3×3 tensor $\boldsymbol{\mathcal{D}}$ may be expressed in a number of forms:

$$\mathcal{D}^{ij} = \frac{\partial^2\omega}{\partial k_i \partial k_j} = \frac{\partial^2\omega}{\partial\mathbf{k}\partial\mathbf{k}} = \partial\mathbf{v}/\partial\mathbf{k}. \tag{5}$$

The symmetry of this matrix ($\mathcal{D}^{ij} = \mathcal{D}^{ji}$) permits the use of the compact but sometimes vague dyadic notation (for example, $\partial\mathbf{v}/\partial\mathbf{k} \equiv \partial v^i/\partial k_j$). For example, the symmetry of $\mathcal{D}^{ij}$ implies that $(\partial/\partial\boldsymbol{\kappa})(\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\kappa}) = (\partial/\partial\kappa_i)(\kappa_j\mathcal{D}^{jk}\kappa_k) = 2\mathcal{D}^{ij}\kappa_j = 2\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}} = 2\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\kappa}$ (written by hand as $2\underline{\underline{D}}\cdot\underline{\kappa}$). We sum over repeated indices unless otherwise stated. The use of both superscripts and subscripts is optional, but facilitates a transition to nonorthogonal coordinate systems.

The symmetry of $\mathcal{D}^{ij}$ also permits the coordinate system to be rotated so that the matrix is diagonal. Henceforth we shall assume that this rotation has taken place, and that $(\mathcal{D}^1,\mathcal{D}^2,\mathcal{D}^3)$ represent the eigenvalues of $\partial\omega^2/\partial\mathbf{k}\partial\mathbf{k}$ in the (x,y,z) directions, respectively. Further simplification occurs when the dispersion relation is such that the angular frequency depends on the magnitude of the wave number but not its direction, $\omega = \omega(|\mathbf{k}|) = \omega(k)$. The chain rule, $\partial\omega/\partial\mathbf{k} = (\partial\omega/\partial k)(\partial k/\partial\mathbf{k})$, implies

$$v^i = \frac{\partial\omega}{\partial k}\frac{\partial}{\partial k_i}(k_1^2+k_2^2+k_3^2)^{1/2} = \frac{\partial\omega}{\partial k}\frac{k_i}{k}, \tag{6}$$

$$\mathcal{D}^{ij} = \frac{\partial v^i}{\partial k_j} = \frac{\partial}{\partial k_i}(k^{-1}k^j v) = \left(\delta^{ij} - \frac{k^i k^j}{k^2}\right)\frac{v}{k} + \frac{k^i k^j}{k^2}\frac{\partial v}{\partial k}, \tag{7}$$

where δ^{ij} is the Kronecker (identity) matrix. If the medium supporting the wave is isotropic, we lose no generality by assuming that $\mathbf{k}$ and hence $\mathbf{v}$ both point in the x direction. Hence, $k_i = 0$, unless $i = 1$. All terms in Eq. (7) proportional to $k_i k_j$ vanish, except the xx term corresponding to $i = j = 1$. However, nonzero yy and zz terms arise from $(v/k)\delta^{ij}$ in Eq. (7). Thus $\partial^2\omega/\partial\mathbf{k}\partial\mathbf{k}$ is given by

$$\mathcal{D}^{ij} = \begin{bmatrix} \partial v/\partial k & 0 & 0 \\ 0 & v/k & 0 \\ 0 & 0 & v/k \end{bmatrix}. \tag{8}$$

The three diagonal elements are the eigenvalues $\mathcal{D}^j$.

A wavepacket is formed by multiplying the plane wave, $\exp[i\mathbf{k}\cdot\mathbf{r} - i\omega(\mathbf{k})t]$, by a wave number amplitude factor, $\phi(\mathbf{k})$, and then integrating over wave number. The integration is greatly facilitated by the change of variables, $\boldsymbol{\kappa} = \mathbf{k} - \mathbf{k}_0$, $\boldsymbol{\xi} = \mathbf{r} - \mathbf{v}t$, and $\eta(\boldsymbol{\kappa}) = \phi(\mathbf{k})$. This change of variables also facilitates the analysis of a one-dimensional Gaussian wavepacket of Schrödinger's equation often encountered in introductory textbooks,[1] because it converts the problem of a moving wavepacket into that of a stationary one:

$$\psi(\mathbf{r},t)=\frac{e^{i\mathbf{k}_0\cdot\mathbf{r}-i\omega_0 t}}{(2\pi)^{3/2}}\int d^3\boldsymbol{\kappa}\,\eta(\boldsymbol{\kappa})\exp\left[i\kappa_j\xi^j-i\frac{\mathcal{D}^j\kappa_j^2}{2}t\right]. \tag{9}$$

If the amplitude $\eta(\boldsymbol{\kappa})$ represents a Gaussian wavepacket, $\eta\propto\exp[-(\mathbf{a}\cdot\boldsymbol{\kappa})^2]$, then the components, a^j, represent standard deviations (in ξ space) along each of the three principle axes defined by the eigenvalues of $\boldsymbol{\mathcal{D}}$. As one might guess, an initially spherically symmetrical wavepacket ($a_1=a_2=a_3$) will spread in a spherically symmetrical fashion only if all three eigenvalues are equal ($\mathcal{D}^1=\mathcal{D}^2=\mathcal{D}^3$). This is easily verified because the three-dimensional integration in Eq. (9) for a spherically symmetrical Gaussian wavepacket separates into a product of three integrals, each equivalent to the familiar one-dimensional Gaussian wavepacket. An initially spherical Gaussian wavepacket will retain its spherical symmetry only if $v/k=\partial v/\partial k$ in Eq. (8). Because $v=\partial\omega/\partial k$, this condition can be integrated to show that the dispersion relation must be of the form, $\omega=A+Bk^2$. This completes the argument that Schrodinger's equation for free particles (and hence de Broglie's relations) can be deduced from non-relativistic symmetry considerations.

III. NON-GAUSSIAN WAVEPACKETS

We now show how Eqs. (1) and (2) can be generalized to include a certain class of non-Gaussian wavepackets. This class is defined by the constraint that η in Eqs. (3) and (12) be real-valued. It is of course no consequence if this real-valued η is multiplied by a complex constant. For wavepackets of light or Schrödinger's equation, this implies that at time $t=0$, the wavepacket's phase is constant throughout any plane oriented so that its normal is parallel to the wavepacket's motion.

Consider the integral Fourier transform pair inspired by Eq. (9):

$$\Psi(\boldsymbol{\xi},t)=(2\pi)^{-3/2}\int d^3\boldsymbol{\kappa}\,\Phi(\boldsymbol{\kappa},t)\exp[i\boldsymbol{\kappa}\cdot\boldsymbol{\xi}], \tag{10}$$

$$\Phi(\boldsymbol{\kappa},t)=(2\pi)^{-3/2}\int d^3\boldsymbol{\xi}\,\Psi(\boldsymbol{\xi},t)\exp[-i\boldsymbol{\kappa}\cdot\boldsymbol{\xi}], \tag{11}$$

$$\Phi(\boldsymbol{\kappa},t)=\eta(\boldsymbol{\kappa})\exp\left[-i\frac{\mathcal{D}^j\kappa_j^2}{2}t\right]\equiv\eta(\boldsymbol{\kappa})e^{iS(\boldsymbol{\kappa},t)}, \tag{12}$$

where we have identified the phase associated with wavepacket dispersion as $S=S(\boldsymbol{\kappa},t)=-(t/2)(\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\kappa})$.

Following quantum mechanics, we treat $\boldsymbol{\kappa}$ and $\boldsymbol{\xi}$ as observables representing position and momentum, respectively. Here the expectation values are most conveniently taken in momentum space, where the components of $\boldsymbol{\kappa}$ act as a numbers, while $\boldsymbol{\xi}\rightarrow -i\partial/\partial\boldsymbol{\kappa}$ becomes an operator.[1] As the wavepacket has been tranformed to the origin in both $\mathbf{k}$ and $\mathbf{r}$ space, the expectation values vanish: $\langle\boldsymbol{\kappa}\rangle=0=\langle\boldsymbol{\xi}\rangle$. We seek the expectation value of the standard deviation tensor, $\langle\boldsymbol{\xi}\,\boldsymbol{\xi}\rangle=-\langle\partial/\partial\boldsymbol{\kappa}\,\partial/\partial\boldsymbol{\kappa}\rangle$. The diagonal components represent the size (variance) of the wavepacket, while the off-diagonal elements are correlation coefficients.

The careful reader will observe that the condition that $\eta(\boldsymbol{\kappa})$ must be real-valued in Eq. (12) is imposed as an extra constraint. Some justification for this constraint can be understood by considering three simple modifications whereby $\eta(\boldsymbol{\kappa})$ is not real: (i) Multiplication of η by a complex number has no consequence. (ii) Multiplication of η by $\exp[i\boldsymbol{\kappa}\cdot\mathbf{r}_0]$ is not permitted because it shifts the wavepacket away from the origin. (iii) In one dimension, multiplication by $\exp[i\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{M}}\cdot\boldsymbol{\kappa}]$ shifts the time when a Gaussian wavepacket converges to a minimum value of uncertainty, $\Delta x\Delta k=1/2$. But in three dimensions, this simple interpretation holds only if strong restrictions are placed on $\boldsymbol{\mathcal{M}}$. In the special case that, $\boldsymbol{\mathcal{M}}=\frac{1}{2}\boldsymbol{\mathcal{D}}\,t_0$, the time when the wavepacket takes on its minimum size is shifted from $t=0$ to $t=t_0$.[2] (Before that time, the wavepacket had been converging.) If $\boldsymbol{\mathcal{M}}$ is diagonal, but not proportional to $\boldsymbol{\mathcal{D}}$, then the wavepacket takes on minimum size at three different times for the three principle directions. In this case, the Gaussian wavepacket is never realized. We therefore see that non-real values of $\eta(\boldsymbol{\kappa})$ in Eq. (12) can greatly complicate the meaning of the wavepacket as a state of maximum compactness (minimum $\Delta x\Delta k$). To avoid such complications, we henceforth take $\eta(\boldsymbol{\kappa})$ to be real.

To evaluate the expectation values, $\langle\boldsymbol{\xi}\,\boldsymbol{\xi}\rangle$, we let the tensor operator $(\partial/\partial\boldsymbol{\kappa}\,\partial/\partial\boldsymbol{\kappa})$ act on Ψ in Eq. (12):

$$\frac{\partial^2(\eta e^{iS})}{\partial\boldsymbol{\kappa}\partial\boldsymbol{\kappa}}=\frac{\partial}{\partial\boldsymbol{\kappa}}\left(\frac{\partial\eta}{\partial\boldsymbol{\kappa}}e^{iS}+i\eta\frac{\partial S}{\partial\boldsymbol{\kappa}}e^{iS}\right)$$
$$=\frac{\partial^2\eta}{\partial\boldsymbol{\kappa}\partial\boldsymbol{\kappa}}-\frac{\partial S}{\partial\boldsymbol{\kappa}}\frac{\partial S}{\partial\boldsymbol{\kappa}}\eta+i\left\{2\frac{\partial S}{\partial\boldsymbol{\kappa}}\frac{\partial\eta}{\partial\boldsymbol{\kappa}}+\frac{\partial^2 S}{\partial\boldsymbol{\kappa}\partial\boldsymbol{\kappa}}\eta\right\}. \tag{13}$$

If we multiply by η^* and integrate over $\boldsymbol{\kappa}$ space, the term in curly brackets $\{\cdots\}$ vanishes if η is real. This is most quickly verified from the requirement that the Hermitian operators $\boldsymbol{\xi}\,\boldsymbol{\xi}$ must have real expectation values, but also can be proven using integration by parts, after using the reality of η to substitute η^2 for $\eta^*\eta$. Because $\boldsymbol{\mathcal{D}}$ is diagonal, the jth component of $\partial S/\partial\boldsymbol{\kappa}=-(\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}})t$ is $\kappa_j\mathcal{D}^j t$ (not summed over j). Therefore,

$$\langle\boldsymbol{\xi}\;\boldsymbol{\xi}\rangle=\int d^3\boldsymbol{\kappa}\left\{-\eta^*\frac{\partial^2\eta}{\partial\boldsymbol{\kappa}\partial\boldsymbol{\kappa}}+|\eta|^2(\boldsymbol{\kappa}\cdot\boldsymbol{\mathcal{D}})(\boldsymbol{\mathcal{D}}\cdot\boldsymbol{\kappa})t^2\right\}$$
$$=\{1+(t/\tau_{ij})^2\}\langle\boldsymbol{\xi}\,\boldsymbol{\xi}\rangle_0, \tag{14}$$

where the 0-subscript on $\langle\boldsymbol{\xi}\,\boldsymbol{\xi}\rangle_0=\langle\xi^i\xi^j\rangle_0$ represents the expectation value evaluated at $t=0$, and

$$\tau_{ij}^2=\left(\frac{1}{\mathcal{D}^i\mathcal{D}^j}\right)\frac{\langle\xi^i\xi^j\rangle_0}{\langle\kappa_i\kappa_j\rangle_0}. \tag{15}$$

Equation (14) is the generalization of Eq. (1). The diagonal elements of Eq. (15) represent the generalization of Eq. (2) because $\mathcal{D}^i$ is an eigenvalue of $\partial^2\omega/\partial k_i\partial k_j$, while $\langle\xi^2\rangle$ and $\langle\kappa^2\rangle$ represent variances in position and wave number, respectively. For a one-dimensional Gaussian distribution, $\Delta\xi\equiv\langle\xi^2\rangle^{1/2}=a$, and $\Delta\kappa\equiv\langle\kappa^2\rangle^{1/2}=(2a)^{-1}$.

IV. FRAUNHOFFER DIFFRACTION

To model the diffraction of a beam of light, we take Eq. (8) as the dispersion matrix with x representing the direction of propagation, and $\mathcal{D}=v/k$ representing the eigenvalues along the other two (perpendicular) directions. Assume also that the (parallel) x-direction separates as: $\eta(\boldsymbol{\kappa})=\eta_\parallel(\kappa_x)\eta_\perp(\kappa_y,\kappa_z)$. It is sufficient to begin with only one of the perpendicular directions. If we let ξ represent either y or z, we can define a phase, F, in Eqs. (10) and (12):

$$\Psi(\xi,t)=(2\pi)^{-1/2}\int d\kappa\ \ \eta(\kappa)e^{iF(\kappa)}, \tag{16}$$

$$F(\kappa)=-\frac{1}{2}\mathcal{D}t\,\kappa^2+\kappa\xi=-\frac{1}{2}\mathcal{D}t\left(\kappa-\frac{\xi}{\mathcal{D}t}\right)^2+\frac{\xi^2}{2\mathcal{D}t}. \tag{17}$$

At large times, $F(\kappa)$ varies so rapidly that $\exp(iF)$ acts as a Dirac delta function, $\delta(\kappa-\kappa_0)$. In Eq. (17) this occurs at $\kappa_0=\xi/\mathcal{D}t$. The integrals over all K of $\int\cos(bK^2)dK$ and $\int\cos(bK^2)dK$ both equal $(\pi/2b)^{1/2}$. To the extent that extremely rapid oscillations causes a function to be effectively zero, we have

$$\lim_{b\to\infty} e^{ibK^2}=\left(\frac{\pi}{2b}\right)^{1/2}\delta(K). \tag{18}$$

Because the real part of the left-hand side is positive for any sign of b, we should choose the branch for which the real part on the right-hand side is positive. This choice yields an expression that can also be obtained using the method of steepest descent:[9]

$$\lim_{t\to\infty} e^{iF(\kappa)}=\exp\left(\frac{i\xi^2}{2\mathcal{D}t}\right)\ \sqrt{\frac{2i\pi}{\mathcal{D}t}}\,\delta\left(\kappa-\frac{\xi}{\mathcal{D}t}\right). \tag{19}$$

Consider diffraction by a slit of width, W, with Ψ_0 representing the uniform intensity at the slit. Hence, $\Psi(\xi,t=0)=\Psi_0$, and Eq. (12) suggests that $\eta(\kappa)=\Phi(\kappa,0)$. Therefore Eq. (11) becomes an expression for η, which yields upon integration (at the slit):

$$\begin{aligned}\eta(\kappa)&=\Phi(\kappa,\ \ t=0)\\&=\frac{1}{\sqrt{2\pi}}\int_{-W/2}^{W/2}d\xi'\,\Psi(\xi',0)\exp[-i\kappa\xi']\\&=\frac{\Psi_0 W}{\sqrt{2\pi}}\,\frac{\sin(\kappa W/2)}{\kappa W/2}.\end{aligned} \tag{20}$$

To recover the single-slit diffraction pattern, we combine Eqs. (16), (19), and (20). We then make the substitutions, $\mathcal{D}=v/k$, and $x=vt$. We also make the approximation that $\theta\approx\xi/x\ll 1$ is the angle with respect to the optical axis:

$$\begin{aligned}\Psi(\xi,t)&=\sqrt{\frac{i}{2\pi\mathcal{D}t}}\,\Psi_0 W\,\frac{\sin(\xi W/2\mathcal{D}t)}{\xi W/2\mathcal{D}t}\exp\left(\frac{i\xi^2}{2\mathcal{D}t}\right)\\&\approx\sqrt{\frac{ik}{2\pi x}}\,\Psi_0 W\,\frac{\sin(kW\theta/2)}{kW\theta/2}\exp(ik\Delta\ell),\end{aligned} \tag{21}$$

where $\Delta\ell$ is a pathlength correction to the distance along the optical axis: $x+\Delta\ell\approx(x^2+\xi^2)^{1/2}$. The small-angle approximation, $\theta\ll 1$, is not required in traditional treatments of Fraunhoffer diffraction, where $\sin(\theta)$ would replace θ in Eq. (21).[5] The need for this approximation can be seen from the truncated Taylor expansion at Eq. (4). Light that has been diffracted by a large amount represents wave number far from the center at $\mathbf{k}_0$, which is aligned along the x direction. For light, the higher order terms in Eq. (4) do not vanish.

Diffraction by a rectangular aperture is obtained using the same methods outlined above. The result is the product of two terms such as given by Eq. (21). The same factor of $x^{-1/2}$ appears in both terms. In this small-angle approximation, x can be replaced by the distance to the slit, r. Thus the intensity of light far from a rectangular slit obeys the expected $1/r$ radiation law.

[a]Electronic mail: gvandegr@purduenc.edu

[1]Richard L. Liboff, *Introductory Quantum Mechanics* (Addison-Wesley, Reading, MA, 1992), pp. 56–57, 157–161, 208–209.
[2]Albert Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, The Netherlands, 1965), Vol. I, pp. 218–222.
[3]E. MacKinnon, "De Broglie's thesis: A critical retrospective," Am. J. Phys. **44**, 1047–1055 (1976).
[4]F. Crawford, "Derivation of the de Broglie relation from the Doppler effect," Am. J. Phys. **50**, 269–270 (1982).
[5]Max Born and Emil Wolf, *Principles of Optics*, 3rd ed. (Pergamon, New York, 1965), pp. 392–395, 738–746.
[6]Charles Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976), pp. 212–219.
[7]Steven Weinberg, "Eikonal method in magnetohydrodynamics," Phys. Rev. **126**, 1899–1909 (1962).
[8]Ronald Adler, Maurice Bazin, and Menahem Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1965), pp. 220–224.
[9]George B. Arfken and Hans J. Weber, *Mathematical Methods for Physicists*, 4th ed. (Academic, San Diego, CA, 1995), pp. 439–453.

A new appraisal of old formulations of mechanics

Antonino Drago[a]
Department of Physical Sciences, University "Federico II," Napoli 80125, Italy

(Received 17 December 2002; accepted 23 September 2003)

Lazare Carnot's formulation of mechanics (1783) is re-evaluated in relation to other formulations of classical mechanics and modern theories.
[DOI: 10.1119/1.1625924]

In a recent paper on the teaching of quantum mechanics, D. F. Styer *et al.* discussed nine formulations of quantum mechanics in order to gain some perspective for teaching quantum mechanics.[1] In Appendix A the authors list the different formulations of classical mechanics known to them. However, this list omits a formulation of a very different nature given by Lazare Carnot.[2] Historians have recognized this formulation as the only one based on principles that are entirely supported by experiments.[3] Furthermore, it is the only formulation able to suggest the basic ideas of thermodynamics, whose main author was the son of Lazare, Sadi Carnot.[4] Remarkably, the theories of Lazare and Sadi Carnot

both share the following characteristic features: (a) they disregard absolute space and trajectories; (b) they make use of discrete time (time evaluated before and after a process); (c) they consider extended bodies rather than point-masses; (d) their theories are based on the physical notion of work, while neglecting the notion of force; (e) in agreement with D'Alembert, they disregard action at a distance; (f) their theories are designed to solve the problem of evaluating the optimum efficiency in machines performing energy conversions; and (g) their theories make use of elementary mathematics—nothing beyond exponential functions—instead of differential equations.

L. Carnot's mechanics is essentially a theory of interacting bodies by means of collisions. A collision is regarded as the basic phenomenon. In particular, continuously accelerated motion is obtained as a limiting case of a system driven by a series of pulses. The notion of Newtonian force is dismissed and masterly criticized. Newton's second law is replaced by L. Carnot's second fundamental equation for a system of n bodies

$$\sum m_i \mathbf{U}_i \cdot \mathbf{u}_i = 0, \tag{1}$$

where m_i is the mass of the ith body, $\mathbf{U}_i$ is the velocity lost by that body during the collision, and $\mathbf{u}_i$ is a velocity called by him "geometrical motion," that is, a motion whose reverse motion is not obstructed by the geometrical configuration of the system, owing to the impenetrability of the bodies. By disregarding L. Carnot's obscure arguments for supporting Eq. (1), it is easy to recognize it as an extension of the principle of virtual velocities to the collision of several bodies.

Even more interesting is the development of his theory from Eq. (1). Let us consider the case $\mathbf{u}_i = \text{constant}$, that is, the same translation for all bodies; because it is a collective motion, it can be reversed and hence constitutes a geometrical motion. From Eq. (1) we have

$$\sum_i m_i \mathbf{u} \cdot \mathbf{U}_i = 0 \quad \text{or} \quad \mathbf{u} \cdot \sum_i m_i \mathbf{U}_i = 0. \tag{2}$$

Due to the arbitrariness of $\mathbf{u}$, it follows that

$$\sum_i m_i \mathbf{U}_i = 0. \tag{3}$$

Because

$$\mathbf{U}_i = \mathbf{W}_i - \mathbf{V}_i, \tag{4}$$

where $\mathbf{W}_i$ is the initial velocity of the ith body and $\mathbf{V}_i$ is the final velocity, we finally obtain

$$\sum_i m_i \mathbf{W}_i = \sum_i m_i \mathbf{V}_i, \tag{5}$$

which represents the conservation of the total momentum of the system.

Next, consider another geometrical motion, $\mathbf{u}_i = \boldsymbol{\omega} \times \mathbf{r}_i$, that is, a rotation of the system with angular velocity $\boldsymbol{\omega}$ around a fixed axis. In this case we have

$$\sum_i m_i \mathbf{U}_i \cdot \boldsymbol{\omega}_i \times \mathbf{r}_i = 0, \tag{6}$$

or, using the properties of the triple product:

$$\sum_i m_i \boldsymbol{\omega} \cdot \mathbf{r}_i \times \mathbf{U}_i = 0. \tag{7}$$

Due to the arbitrariness of $\boldsymbol{\omega}$, we obtain

$$\sum_i m_i \mathbf{r}_i \times \mathbf{U}_i = 0. \tag{8}$$

and from Eq. (4) we have

$$\sum_i m_i \mathbf{r}_i \times \mathbf{W}_i = \sum_i m_i \mathbf{r}_i \times \mathbf{V}_i, \tag{9}$$

which is the conservation of the total angular momentum. It is interesting to observe that L. Carnot was proud to propose a "new theory," nowadays recognized as a theory of symmetries.[5] (Note that such a theory was proposed 50 years before the celebrated work of Galois, which was not appreciated until twenty years later.) Incidentally, because the solution of the collision problem is expressed in terms of velocities, we have to solve the conservation laws as algebraic–trigonometric equations; these equations are an example of point (g) in the above.

We now compare Carnot's formulation of mechanics with special relativity. Both theories can make use of the hyperbolic geometry. Indeed, the main equation in L. Carnot's formulation, that is, Eq. (1), deals with the product of two vectors $(\mathbf{U}_i, \mathbf{u}_i)$ applied at the same point. As a consequence, this formula does not depend on the Euclidean axiom of parallel lines and the underlying geometry can be either Euclidean or hyperbolic geometry.[6] On the other hand, special relativity can be formulated, as first done by Sommerfeld in 1909, in velocity space,[7] whose geometry turns out to be hyperbolic.[8] Hence, there is no substantial difference in the underlying geometry of special relativity and classical mechanics when the latter is considered in L. Carnot's formulation in velocity space; both may be linked together by the same geometry.

Some recent suggestions have introduced special relativity in a straightforward way, by essentially generalizing the classical conservation laws.[9] One can view these suggestions as extensions of Carnot's formulation. Finally, in Carnot's formulation, as well as in special relativity, the concept of action at a distance is absent. We recall that in 1905 the conceptual difference between classical mechanics—in Newton's formulation—and special relativity was a dramatic one; this difference is drastically reduced when L. Carnot's rather than Newton's formulation of classical mechanics is taken into account.

In addition, in Carnot's formulation of mechanics, a superposition principle for geometrical motion holds in analogy with the superposition principle for waves, including Schrödinger's wave mechanics.

The various formulations of classical mechanics can be grouped according to two basic criteria. One criterion concerns the kinds of mathematics that they use: while L. Carnot's formulation makes use of algebraic mathematics, all other ones make use of calculus. The other criterion concerns the different ways of organizing a scientific theory. Accord-

ing to L. Carnot, a theory can be internally organized according to two models: the purely deductive model (''where one has to take out all from one's own bag'' of a few axioms) and the ''merely empirical'' model.[10] These two models can be compared with those suggested by Einstein, namely the model of ''constructive theories'' (like statistical mechanics), and the model of the ''theories of principle'' (like thermodynamics).[11] These two criteria distinguish the differences between the various formulations of classical mechanics. Whereas Newton's mechanics is a deductive theory and makes use of the calculus of infinitesimals, both L. Carnot and S. Carnot's theories are empirical theories and make use of purely algebraic mathematics.

The analysis of the basic differences in the formulations of classical mechanics can be useful in characterizing the wide spectrum of the possible formulations of quantum mechanics (for example, matrix and wave mechanics),[12] in order to suggest a more suitable attitude for teaching this theory.

[a]Electronic mail: drago@unina.it

[1]D. F. Styer *et al.*, ''Nine formulations of quantum mechanics,'' Am. J. Phys. **70** (3), 288–297 (2002).

[2]L. Carnot, *Essai sur les Machines en Général* (Defay, Dijon, 1783) and *Principes Fondamentaux de l'Equilibre et du Mouvement* (Crapelet, Paris, 1803). For a sketch of his life and his wide scientific work, see the corresponding issue in *Dictionary of Scientific Biography*, edited by C. C. Gillispie (Scribner's, New York, 1971). A more comprehensive presentation of his scientific work is in C. C. Gillispie, *Lazare Carnot Savant* (Princeton U.P., Princeton, NJ, 1971). An anonymous referee suggested a reference to Christopher Tong's letter, ''Various formulations of classical mechanics,'' Am. J. Phys. **70** (7), 664 (2002).

[3]R. Dugas, *Histoire de la Mécanique* (Griffon, Neuchâtel, 1956), pp. 309–317. See also A. Drago and S. D. Manno, ''I principi della meccanica secondo Lazare Carnot,'' Epistemologia **12**, 305–330 (1989) and A. Drago, ''Le lien entre mathématique et physique dans la mécanique de Lazare Carnot,'' *Lazare Carnot ou le Savant-Citoyen*, edited by J.-P. Charnay (Presse de l'Université de la Sorbonne, Sorbonne, 1990), pp. 501–515.

[4]C. C. Gillispie, in Ref. 2, Chap. iii D.

[5]L. Carnot, *Principes*, in Ref. 2, p. x; A. Drago: ''The birth of symmetries in theoretical physics: Lazare Carnot's mechanics,'' in *Symmetry of Structure*, edited by G. Darvas and D. Nagy (Hung. Acad. Sci., Budapest, 1989), pp. 98–101.

[6]I. Adler, *A New Look at Geometry* (Day, New York, 1966), pp. 253–257.

[7]A. Sommerfeld, ''Ueber die Zusammensetzung der Geschwidigkeiten in der Relativtheorie,'' Phys. Z. **10**, 826–829 (1909).

[8]B. A. Rosenfeld, *A History of Non-Euclidean Geometry* (Springer, Berlin, 1988), pp. 270–272.

[9]See W. C. Davidon, ''Consequences of the inertial equivalence of energy,'' Found. Phys. **5**, 525–541 (1975) and J.-M. Lévy-Leblond, ''What is so 'special' about 'Relativity,'" in *Group Theoretical Methods*, edited by A. Jenner (Springer, Berlin, 1976), pp. 617–627.

[10]L. Carnot, *Essai*, in Ref. 2, pp. 102–105; and *Principes*, in Ref. 2, pp. xiii–xvii.

[11]M. J. Klein, ''Thermodynamics in Einstein's thought,'' Science **57**, 505–516 (1967); A. I. Miller, *Albert Einstein's Special Theory of Relativity* (Addison–Wesley, Reading, MA, 1981), pp. 123–142.

[12]With A. Pirolo, I have suggested a variant of T. F. Jordan, *Quantum Mechanics in Simple Matrix Form* (Wiley, New York, 1986) and have developed an approach to quantum mechanics based on symmetries. See A. Drago and A. Pirolo, ''Quantum mechanics reformulated by means of symmetries,'' in *The Foundations of Quantum Mechanics*, edited by C. Garola and A. Rossi (Kluwer Academic, Dordrecht, 1995), pp. 229–237.

Comment on "Delta functions in spherical coordinates and how to avoid losing them: Fields of point charges and dipoles," by S. M. Blinder [Am. J. Phys. 71 (8), 816–818 (2003)]

Ben Yu-Kuang Hu[a)]
Department of Physics, University of Akron, Akron, Ohio 44325-4001

(Received 8 August 2003; accepted 23 September 2003)

[DOI: 10.1119/1.1625928]

In a recent paper,[1] S. M. Blinder addressed the apparent disappearance of the δ function that is generated by $\nabla^2|r|^{-1}$ when the spherical coordinates expression of the Laplacian ∇^2 is used. Blinder showed that the ''lost'' delta function at the origin can be recovered by a judicious introduction of an ad hoc step function $\mathrm{sgn}(r)$ to the $|r|^{-1}$ potential. The purpose of this comment is to show that the insertion of this step function can be put on a more rigorous footing.

When spherical coordinates are used, it is conventional to let the radial distance r range from 0 to ∞, and the polar angle θ range from 0 to π. However, the occurrence of the origin at the beginning of the range of r is a source of ambiguity [see, for example, Eqs. (7) and (38) in Ref. 1] and contributes to the seeming disappearance of the delta function of $\nabla^2|r|^{-1}$. To avoid these problems at the origin, we can instead let r range from $-\infty$ to $+\infty$, and θ from 0 to $\pi/2$. This choice does not affect the various spherical coordinate expressions of the Laplacian,

$$\nabla^2=\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}+\cdots, \tag{1a}$$

$$\nabla^2=\frac{1}{r}\frac{\partial^2}{\partial r^2}r+\cdots, \tag{1b}$$

where $\cdots$ signifies terms with derivatives with respect to ϕ and θ. However, because r can be negative, the potential of a point charge at the origin must be written as $|r|^{-1}=(r^2)^{-1/2}=\mathrm{sgn}(r)/r$. This is the mathematical justification for the insertion of the $\mathrm{sgn}(r)$ term.

The δ function in $\nabla^2|r|^{-1}$ is recovered as in Ref. 1. For example, the Eq. (1a) form of ∇^2 gives

$$\nabla^2\frac{1}{|r|}=\frac{1}{r^2}\frac{d}{dr}r^2\left[\frac{d}{dr}\frac{1}{(r^2)^{1/2}}\right]=-\frac{1}{r^2}\frac{d}{dr}\frac{r^3}{(r^2)^{3/2}}$$

$$=-\frac{1}{r^2}\frac{d}{dr}\mathrm{sgn}(r)=-\frac{2\delta(r)}{r^2}$$

$$=-4\pi\delta(\mathbf{r}). \tag{2}$$

The last equality in Eq. (2) results from

$$\int\int\int_{\text{all space}} d\mathbf{r}\delta^3(\mathbf{r})=\int_{-\infty}^{\infty}r^2dr\int_0^{2\pi}d\phi\int_0^{\pi/2}\sin\theta d\theta\delta^3(\mathbf{r})$$

$$=2\pi\int_{-\infty}^{\infty}r^2\,dr\,\delta^3(\mathbf{r})$$

$$=2\pi\int_{-\infty}^{\infty}r^2dr\frac{\delta(r)}{2\pi r^2}. \tag{3}$$

Note that the ambiguous integral $\int_0^\infty\delta(r)dr$ that occurs in Ref. 1 is absent here because the r integration runs from $-\infty$ to ∞. Using Eq. (1b) for ∇^2 gives

$$\nabla^2\frac{1}{|r|}=\frac{1}{r}\frac{d^2}{dr^2}\frac{r}{|r|}$$

$$=\frac{1}{r}\frac{d^2}{dr^2}\mathrm{sgn}(r)$$

$$=\frac{2}{r}\frac{d}{dr}\delta(r)$$

$$=-\frac{2\delta(r)}{r^2}$$

$$=-4\pi\delta^3(\mathbf{r}), \tag{4}$$

where we have used the identity $r[d\delta(r)/dr]=-\delta(r)$.[2]

[a]Electronic mail: benhu@physics.uakron.edu

[1]S. M. Blinder, "Delta functions in spherical coordinates and how to avoid losing them: Fields of point charges and dipoles," Am. J. Phys. **71**, 816–818 (2003).

[2]This identity can be shown by inserting $r[d\delta(r)/dr]$ into an integral over r of a well-behaved test function $f(r)$ and integrating by parts: $\int_{-\infty}^{\infty}f(r)r[d\delta(r)/dr]dr=-\int_{-\infty}^{\infty}[f(r)+r(df/dr)]\delta(r)dr=-\int_{-\infty}^{\infty}f(r)\delta(r)dr$.

Comment on "Algebraic approach to the radioactive decay equations," by L. Moral and A. F. Pacheco [Am. J. Phys. 71 (7), 684–686 (2003)]

T. M. Semkow[a)]
Wadsworth Center, New York State Department of Health, and University at Albany, SUNY, Albany, New York 12201-0509

(Received 15 July 2003; accepted 1 October 2003)

[DOI: 10.1119/1.1629092]

The radioactive decay chain equations were originally solved by Bateman[1] using a Laplace transform method. This approach was recently reconsidered by Pressyanov.[2] The matrix (or algebraic) solution, recently described by Moral and Pacheco,[3] was first introduced in Ref. 4. The purposes of this comment are to outline the matrix solution with branching, to give physical applications of the solution, and to give new insight when all the decay constants in the chain are equal.

Consider a branching decay chain described by the set of differential equations

$$\frac{dN_1}{dt}=-\lambda_1N_1, \tag{1a}$$

$$\frac{dN_j}{dt}=\sum_{i=1}^{j-1}\lambda_ib_{ji}N_i-\lambda_jN_j\quad(j=2,\ldots,n), \tag{1b}$$

where $N_j\equiv N_j(t)$ is the average number of radioactive atoms at time t, λ_j is the decay constant, j is the position index in the decay chain consisting of n components, and b_{ji} is a branching fraction from component i to j ($\Sigma_{j=i+1}^{n}b_{ji}=1$).

We define a vector of radioactive atoms, $\mathbb{N}^T=(N_1,\ldots,N_n)$, and matrix $\mathbb{A}$ with elements, $A_{ji}=\lambda_ib_{ji}$ for $j>i$, $A_{jj}=-\lambda_j$, $A_{ji}=0$ for $j<i$, and express Eq. (1) in matrix form

$$\frac{d\mathbb{N}}{dt}=\mathbb{A}\mathbb{N}. \tag{2}$$

When all the λ_j are different, the matrix $\mathbb{A}$ is diagonalizable ($\mathbb{P}^{-1}\mathbb{A}\mathbb{P}=\mathrm{diag}(-\lambda_j)$), and the solution of Eq. (2) is[4]

$$\mathbb{N}(t)=\mathbb{P}\,\mathrm{diag}(e^{-\lambda_jt})\mathbb{P}^{-1}\mathbb{N}(0)\,. \tag{3}$$

The methods for calculating the matrix $\mathbb{P}$ and its inverse $\mathbb{P}^{-1}$ were given in Ref. 5.

The matrix solution, Eq. (3), has been used to study the nuclear fission β-decay chains, where branching frequently occurs due to nuclear isomerism.[4,5] Another application is to the decay and growth of radon daughters.[6] Equation (3) can incorporate a decrease of activity due to exponential decay during counting, as well as chemical separations, in a general and compact way.[6] Equations of the type of Eq. (1) are ex-

amples of compartment (or box) models, which have numerous applications. One such general treatment of the matrix solution in health physics is in Ref. 7.

There is a limiting condition of the matrix solution, when some or all λ_j are equal. Then, $\mathbb{A}$ is singular and no longer diagonalizable. Although this case is less likely in nuclear physics because the decay constants are generally different, it may nevertheless be of importance in general compartment modeling. In addition, it provides interesting physical insight.

Consider a case for which all the λ_j are equal, there is only sequential decay in the chain, that is, no branching, and the boundary condition is $N_1(0)\neq 0$, $N_j(0)=0$ for $j>1$. One way of solving this case would be to consider the general solution and take the limit. For example, for $n=2$, Eq. (3) gives $N_2(t)=N_1(0)\lambda_1[(e^{-\lambda_1 t}-e^{-\lambda_2 t})/(\lambda_2-\lambda_1)]$. By resolving the exponential series and taking the limit λ_1, $\lambda_2 \rightarrow\lambda$, we obtain $N_2(t)=N_1(0)\lambda t e^{-\lambda t}$. This procedure is too cumbersome in the general case. It is more convenient to rewrite Eq. (1) as

$$\frac{dN_1}{dt}=-\lambda N_1, \tag{4a}$$

$$\frac{dN_j}{dt}=\lambda(N_{j-1}-N_j). \quad (j=2,\ldots,n). \tag{4b}$$

If we apply the Laplace transform[1,2] to Eq. (4), we obtain

$$\tilde{N}_n(s)=\frac{\lambda^{n-1}}{(s+\lambda)^n}N_1(0). \tag{5}$$

The function $e^{st}/(s+\lambda)^n$ has a pole of the nth order, and the residue theorem yields

$$N_n(t)=N_1(0)\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}. \tag{6}$$

Before interpreting Eq. (6), let us shift the index j in Eq. (4), so that it starts from 0, and let us define $P_j\equiv P_j(t)=N_j(t)/N_0(0)$ as the probability of finding a radioactive atom at position j in the chain. Then, from Eq. (4),

$$\frac{dP_0}{dt}=-\lambda P_0, \tag{7a}$$

$$\frac{dP_j}{dt}=\lambda(P_{j-1}-P_j) \quad (j\geqslant 1). \tag{7b}$$

Equation (7) was first solved by Bateman in another of his seminal papers on statistical fluctuations in radioactive decay.[8] By means of the generating function $G(z,t)=\Sigma_{j\geqslant 0}P_j(t)z^j$, Eq. (7) can be coverted to[9]

$$\frac{\partial G}{\partial t}=-\lambda(1-z)G, \tag{8}$$

whose solution is $G(z,t)=e^{-(1-z)\lambda t}$. It is known from the theory of statistical distributions[10] that the exponential generating function is a Poisson distribution given by Eq. (6) (with an index shift).

It is remarkable, that the *average* number of atoms of the chain members for sequential decay, when all the decay constants are equal, is governed by the Poisson distribution, similar to the *fluctuations* in radioactive decay. The physical origin of this result lies in the renewal process. An atom that has decayed in the chain with equal decay constants is renewed, that is, it continues to decay with the same decay constant, and only its position index in the chain has incremented. It is known that such an ordinary renewal process leads to the Poisson distribution.[11] The mean of this Poisson distribution, Eq. (6), is equal to λt, which can take any value. There is an important difference from the Poisson *fluctuations* in radioactive decay, however.[12] To satisfy the Poisson fluctuations, exemplified by the decay of a long-lived radionuclide, we must have the total number of radioactive atoms $N\gg 1$ and the probability of a single decay $1-e^{-\lambda t}\approx\lambda t\ll 1$. The mean in this case is equal to $N\lambda t$, which can take any value.[13] The small number of atoms that decay in a short time interval t is much less than the total number of atoms and, therefore, the decay does not disturb the distribution.

[a]Electronic mail: tms15@health.state.ny.us

[1]H. Bateman, "The solution of a system of differential equations occurring in the theory of radio-active transformations," Proc. Cambridge Philos. Soc. **15**, 423–427 (1910).

[2]D. S. Pressyanov, "Short solution of the radioactive decay chain equations," Am. J. Phys. **70** (4), 444–445 (2002).

[3]L. Moral and A. F. Pacheco, "Algebraic approach to the radioactive decay equations," Am. J. Phys. **71** (7), 684–686 (2003).

[4]G. P. Ford, K. Wolfsberg, and B. R. Erdal, "Independent yields of the isomers of ^{133}Xe and ^{135}Xe for neutron-induced fission of ^{233}U, ^{235}U, ^{238}U, ^{239}Pu, and ^{242}Am^m," Phys. Rev. C **30**, 195–213 (1984).

[5]T. M. Semkow, A. C. Wahl, and L. Robinson, "Yields of In and Sn products from thermal- and 14-MeV-neutron-induced fission of ^{235}U," Phys. Rev. C **30**, 1966–1975 (1984).

[6]T. M. Semkow, P. P. Parekh, C. D. Schwenker, R. Dansereau, and J. S. Webber, "Efficiency of the Lucas scintillation cell," Nucl. Instrum. Methods Phys. Res. A **353**, 515–518 (1994).

[7]P. A. Assimakopoulos, K. G. Ioannides, and A. A. Pakou, "A general multiple-compartment model for the transport of trace elements through animals," Health Phys. **61** (2), 245–253 (1991).

[8]H. Bateman, "On the probability distribution of α particles," Philos. Mag. Ser. 6 **20** (118), 704–707 (1910)

[9]N. T. J. Bailey, *The Elements of Stochastic Processes with Applications to the Natural Sciences* (Wiley, New York, 1964), p. 69.

[10]N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate Discrete Distributions* (Wiley, New York, 1993), p. 151.

[11]D. R. Cox, *Renewal Theory* (Methuen, London, 1962), p. 29.

[12]The Bateman paper (Ref. 8) is cited in statistical texts (Ref. 10) historically as a second derivation of the Poisson distribution.

[13]Note that N was omitted in Ref. 8.

Comment on "Frustrated total internal reflection: A simple application and demonstration," F. P. Zanella *et al.* [Am. J. Phys. 71 (5), 494–496 (2003)]

A. A. Stahlhofen[a)]
Universität Koblenz, Institut für Naturwissenschaften, Abteilung Physik, Universitätsstr. 1, D-56075 Koblenz, Germany

(Received 23 May 2003; accepted 17 October 2003)

[DOI: 10.1119/1.1632492]

In a recent article in this journal an interesting application and demonstration of frustrated total internal reflection (FTIR) was presented.[1] Two items should be noted to place this article in a proper context: (i) The physics of FTIR as well as the propagation of the light beam in the piece of glass sketched in Fig. 1 of Ref. 1 is a very simplified picture: the authors did not discuss the Goos–Hänchen shift in FTIR, which is a very important deviation from geometrical optics (predicted by Newton and discussed by him in his book on optics cited as Ref. 3 in Ref. 1). This shift describes a longitudinal shift for polarized beams in the plane of incidence and can be as large as 4 to 5 wavelengths.[2] For unpolarized or circular polarized beams, a second shift perpendicular to the plane of incidence can be observed, which is, however, an order of magnitude smaller than the longitudinal shift.[4] (ii) Many references to applications and demonstration of FTIR can be found in a review article that appeared in this journal some time ago.[5]

[a)]Electronic mail: alfons@uni-koblenz.de

[1]F. P. Zanella, D. V. Magalhaes, M. M. Oliveira, R. F. Bianchi, L. Misoguti, and C. R. Mendoca, "Frustrated total internal reflection: A simple application and demonstration," Am. J. Phys. **71**(5), 494–496 (2003).
[2]A. Haibel, G. Nimtz, and A. A. Stahlhofen, "Frustrated total reflection: The double prism revisited," Phys. Rev. E **63**(1), 047601 (2001).
[3]A. A. Stahlhofen, "The photonic tunneling time in frustrated total internal reflection," Phys. Rev. A **62**(2), 012112 (2000).
[4]References for the transverse Goos–Hänchen shift can be found in Refs. 2 and 3.
[5]S. Zhu, A. W. Yu, D. Hawley, and D. Roy, "Frustrated total internal reflection: A demonstration and review," Am. J. Phys. **54(7)**, 601–607 (1986).

The validity of the Helmholtz theorem

F. Rohrlich[a)]
Department of Physics, Syracuse University, Syracuse, New York 13244-1130

(Received 11 August 2003; accepted 7 November 2003)

[DOI: 10.1119/1.1637041]

The Helmholtz theorem states that a vector field, $\mathbf{V}(\mathbf{x},t)$, in three-dimensional space can be separated uniquely into two components, a transverse (solenoidal) component, $\mathbf{V}_t$, and a longitudinal (irrotational) one, $\mathbf{V}_l$, such that $\mathbf{V}=\mathbf{V}_t+\mathbf{V}_l$, where[1]

$$\nabla\cdot V_t=0, \quad \nabla\times V_l=0. \tag{1}$$

The time dependence of $\mathbf{V}$ is irrelevant. This irrelevance can also be seen from the explicit expressions for $\mathbf{V}_t$ and $\mathbf{V}_l$ in terms of $\mathbf{V}$ as given in Jackson's[1] equations (6.27) and (6.28).

This theorem has recently been challenged by Heras[2] (cited in the following as *H*). He claims that there is a "common misconception that the standard Helmholtz theorem ...can be applied to retarded vector fields." He then proceeds to claim that therefore my paper[3] (cited as *R*) is incorrect. The purpose of this note is to show that Heras's claim is false, and that the Helmholtz theorem also applies to retarded (and advanced) vector fields.

At first, it might seem that retardation may indeed make a difference because it introduces an additional spatial dependence into the integral that expresses the field in terms of its sources. However, the Helmholtz theorem has nothing to do with the source of the vector field that is being separated. Therefore, it holds also for retarded (and advanced) fields.

Because this issue has bearing on causality (the propagation of fields with the speed of light), and is encountered often in electrodynamics, further discussion is desirable. As a preliminary remark, note that in the Coulomb gauge the *scalar potential* ϕ satisfies the Poisson equation as a consequence of Maxwell's field equations and is therefore *not causal*. Nevertheless, the *total fields are causal*, independent of the gauge. That fact has been known for a long time.[4] Its proof does *not* involve the Helmholtz theorem.

The objection to the validity of the Helmholtz theorem refers to retardation, that is, to the relation of the field to its sources. A vector field that is a source, therefore, should not encounter that objection. In electrodynamics, one can therefore define the separation of the field, $\mathbf{E}$, in terms of the separation of its sources. For the equation

$$\Box \mathbf{E} = 4\pi(\partial \mathbf{j}/\partial t + \nabla \rho), \tag{2}$$

a separation of sources into transverse and longitudinal ones, results in the *retarded* electric fields

$$\mathbf{E}_t(x) = -4\pi \int D_R(x-x') \frac{\partial \mathbf{j}_t(x')}{\partial t'} d^4x' \quad \text{(transverse)}, \tag{3a}$$

$$\mathbf{E}_l(x) = -4\pi \int D_R(x-x') \left[\frac{\partial \mathbf{j}_l(x')}{\partial t'} + \nabla \rho(x') \right] d^4x' \quad \text{(longitudinal)}. \tag{3b}$$

It is easy to verify that the conditions (1), applied to these electric fields, hold for Eq. (3). This involves integration by parts with the assumption of asymptotic boundedness of the sources).

On the other hand, Heras[2] claims the decomposition $(H8)$, $(H9)$, and $(H10)$ of his paper. If those equations are added, we find that (a) the **B** dependent terms cancel, (b) the last term in Eq. $(H10)$ separates into two terms according to $\mathbf{j} = \mathbf{j}_t + \mathbf{j}_l$, and (c) the sum of the three equations is exactly the sum of $\mathbf{E}_t$ and $\mathbf{E}_l$ given in Eq. (3) as well as in *R*. It follows that the separation advocated by Heras is entirely arbitrary.

The conclusion is therefore that the Helmholtz theorem applies to three-vector fields of *any* time dependence, including retarded or advanced time dependence on its sources, as long as the sources of these fields are bounded in space.

[a]Electronic mail: rohrlich@syr.edu

[1]For the present purpose, the best reference is J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999), 3rd ed., pp. 241–242.

[2]J. A. Heras, "Comment on 'Causality, the Coulomb field, and Newton's law of gravitation,' " by F. Rohrlich, Am. J. Phys. **71** (7), 729 (2003), cited as *H*.

[3]F. Rohrlich, "Causality, the Coulomb field, and Newton's law of gravitation," Am. J. Phys. **70** (4), 411 (2002), cited as *R*.

[4]C. W. Gardiner and P. D. Drummond, "Causality in the Coulomb gauge: A direct proof," Phys. Rev. A **38**, 4897–4898 (1988).

Erratum: "Propagating and evanescent waves in absorbing media" [Am. J. Phys. 71 (6), 562–567 (2003)]

S. Anantha Ramakrishna[a)]
The Blackett Laboratory, Imperial College, London SW7 2BW, United Kingdom

A. D. Armour
School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom

[DOI: 10.1119/1.1629285]

In Fig. 1, the inequality should read $\varepsilon_2 < \varepsilon_1$, and the caption should state: "A wave incident on a dielectric medium with a smaller dielectric constant at an angle greater than the critical angle, $\theta > \theta_c$, results in an exponentially decaying wave in medium 2." Similarly, the sentence in the third paragraph of the left-hand column of p. 562 should read: "The physics of evanescent waves is readily understood by considering what happens when light is shone on a nonabsorbing medium at an angle greater than the angle of total internal reflection from a medium with higher refractive index."

In the paragraph below Eq. (5), we should have stated that an examination of the Poynting vector shows that no energy flows inside medium 2 normal to the interface.

In the paragraph below Eq. (8), the first sentence should read: "In other words, the normal component of the Poynting vector is only nonzero if $\varepsilon_i \neq 0$."

Also, Eq. (24) should read:

$$T = \frac{2k_z^{(a)}/\varepsilon_a}{[k_z^{(m)}/\varepsilon_m + k_z^{(a)}/\varepsilon_a]}. \tag{24}$$

Although these errors are regrettable, they do not affect any of the main results of our paper.

[a)]Present address: Department of Physics, Indian Institute of Technology, Kanpur 208 016, India.

