# Momentum, pseudomomentum, and wave momentum: Toward resolving the Minkowski-Abraham controversy

D. F. Nelson
*Department of Physics, Worcester Polytechnic Institute, Worcester, Massachusetts 01609*
(Received 7 February 1991)

We derive the momentum and pseudomomentum conservation laws from a very general nonrelativistic Lagrangian theory of the interaction of the electromagnetic field with a deforming, dispersive dielectric. From the former of these laws, we obtain the momentum density of an electromagnetic wave in matter to be $\epsilon_0\mathbf{E}\times\mathbf{B}$, not the Abraham form of $\epsilon_0\mathbf{E}\times\mu_0\mathbf{H}$. From the latter of these laws, we obtain the electromagnetic pseudomomentum density in the absence of deformation of the matter to be $\mathbf{P}\times\mathbf{B}$ plus a dispersive term, not the Minkowski form of $\mathbf{D}\times\mathbf{B}$ as proposed by Blount (unpublished). We show by quantizing the energy of the wave that the sum of momentum and pseudomomentum, which we name *wave momentum,* corresponds to $N\hbar\mathbf{k}$ ($N$ an integer), the quantity that enters wave-vector conservation or phase-matching relations in wave interactions and that is consistent with the Jones-Richards experiment.

## INTRODUCTION

It is a striking fact that for over eighty years there has been a controversy over a seemingly simple classical quantity, the momentum density of an electromagnetic wave in a material medium. The controversy began with Minkowski's Lorentz-covariant formulation of Maxwell's electromagnetism in which he proposed a relativistic $4\times4$ stress-energy-momentum (SEM) tensor of the electromagnetic field [1]. His momentum density part of the SEM tensor, in its ordinary three-vector SI units, is $\mathbf{D}\times\mathbf{B}$, where $\mathbf{D}$ is the electric displacement and $\mathbf{B}$ is the magnetic induction. Abraham [2] soon criticized the Minkowski tensor for its lack of symmetry. Because of the close connection in relativity between momentum and the flux of energy $\mathbf{E}\times\mathbf{H}$, he proposed an alternate SEM tensor that is symmetric and which contains a momentum density, again in its ordinary three-vector SI units, of $\mathbf{E}\times\mathbf{H}/c^2$ where $\mathbf{E}$ is the electric field, $\mathbf{H}$ the magnetic field, and $c$ the speed of light.

Since neither the Minkowski nor the Abraham tensor was truly derived, the issue has been reconsidered in scores of papers from scores of viewpoints. Interestingly, the "vote" has split about equally over the years between the two forms. Reviews that consider the controversy to about 1970 are Pauli [3], Møller [4], Brevik [5], Penfield and Haus [6], and de Groot and Suttorp [7].

In 1971 Blount [8], in a seminal, but unpublished, work introduced the concept of pseudomomentum or crystal momentum, well known in solid-state physics, into the controversy. He identified by relativistic, macroscopic arguments the Abraham form $\mathbf{E}\times\mathbf{H}/c^2$ as the momentum density and the Minkowski form $\mathbf{D}\times\mathbf{B}$ as the pseudomomentum density. Momentum, of course, is a conserved quantity by virtue of the homogeneity of space, that is, as a result of the invariance of the laws of physics to displacements of the spatial coordinates. Pseudomomentum can also be a conserved quantity provided the material medium is homogeneous, that is, as a result of the invariance of the laws of physics to translations of the material coordinates (those which label a point of matter). This important distinction has reoriented work on the Minkowski-Abraham controversy from a simple issue of right or wrong to one of proper identification and experimental applicability.

More recent papers of Skobel'tsyn [9], Gordon [10], and Robinson [11] have agreed that the Abraham form is the momentum density of the electromagnetic wave while agreeing that the Minkowski form applies to many experiments, in particular, to the Jones and Richards experiment [12] and its improved version [13]. Gordon [10] and Peierls [14–16] accept the identification of the Minkowski form as the pseudomomentum density. Peierls differs, however, from other work by concluding [14] that "Abraham's formula gives correctly the part of the momentum of the medium which resides in the electromagnetic field, but not the mechanical momentum of the medium which travels with the light pulse." A particularly important work is the derivation by de Groot and Suttorp [7] of the Abraham result for the momentum density from a relativistic treatment. Interestingly, they obtain a different result, $\epsilon_0\mathbf{E}\times\mathbf{B}$, from a nonrelativistic treatment. We shall return to a discussion of this later.

The results we derive in this paper differ on the form of both the momentum density and the pseudomomentum density from the Abraham form and the Minkowski form, respectively. We find the momentum density of an electromagnetic wave in a material medium to be $\epsilon_0\mathbf{E}\times\mathbf{B}$ in SI units and the pseudomomentum density (in the absence of deformation) to be $\mathbf{P}\times\mathbf{B}$ plus a dispersive term. Besides presenting a very general deductive mathematical derivation for each of these quantities, we present what we believe to be a strong physical argument in support of each. We also present a new interpretation of the Minkowski momentum density. It represents, in the absence of deformation, the nondispersive part of the sum of momentum and pseudomomentum densities. With the

 

dispersive part included, we define this sum of momentum and pseudomomentum as *wave momentum* and show that it is the quantity which plays the dominant role in wave interactions.

A review of the techniques used in this long controversy produces a number of lessons useful to a present-day worker on the controversy. First, though the controversy arose out of relativistic formulations of electromagnetism, the main issue is not relativistic. The Minkowski momentum density differs from that of Abraham by a factor of the refractive index squared, which can range in value from 2 to 10. Thus a nonrelativistic treatment should be capable of resolving the main issue. The relativistic issues are important, but can be dealt with more easily when the main issue is resolved. Second, the issue cannot be resolved by reliance on Maxwell electromagnetism and relativity alone. The matter equations must be included and handled on an equally fundamental basis, with the action and reaction between field and matter correctly included. Third, because the controversy has always treated the matter as a continuum, the matter equations can be treated in the long-wavelength limit (wavelengths much greater than the atomic or cellular size). Fourth, imposing symmetry in an unfamiliar area from a macroscopic viewpoint may be unjustified; it should rather be derived from a more fundamental viewpoint. For example, imposing symmetry on the $4\times4$ SEM tensor forces the $3\times3$ stress tensor to be symmetric. However, this is untrue as shown for an arbitrary dielectric crystal for dynamic excitations [17,18]. For materials with a soft optic mode the asymmetry can be sizeable. There the torque from the antisymmetric part of the stress tensor is balanced by a change in the internal angular momentum density of the optic mode. Fifth, as is evident from the fourth point, it is important to include all long-wavelength mechanical modes of motion, three of which are acoustic modes and $3N-3$ of which are optic modes ($N$ is the number of particles per primitive unit cell in a crystal). The optic modes not only allow asymmetric stresses but produce the frequency dispersion in the dielectric tensor that accounts for the temporal response of the medium. Sixth, in order to handle these many modes of motion correctly, it is important to introduce appropriate coordinates. These should include the continuum center-of-mass position, which is closely related to the displacement vector that describes the three acoustic modes and which carries all the momentum of the material medium. They also should include $N-1$ vector internal coordinates, linear combinations of which are the $3N-3$ scalar optic mode normal coordinates, which describe the $3N-3$ optic modes. These must be defined so that they cannot carry any momentum of the material medium. Seventh, since a key concept has proven to be the distinction between momentum and pseudomomentum, which have their natural expressions in two different coordinate systems, the spatial (Eulerian) and material (Lagrangian) coordinate systems, respectively, it is important to use a formulation of the problem that employs both of these systems. Eighth, the issue of whether a resolution of the controversy must arise from microscopic [7] or macroscopic [8,11] origins is, to this author's way of thinking, mostly a matter of aesthetics. The important aspect of a resolution is that it be deductive from well-accepted first principles. If it is first done so from macroscopic origins, it is doubtless that those origins will later be justified from microscopic origins, but, by that time, it will not be a resolution of this controversy but rather an expansion of the interconnections that physics must have. Alternatively, if a resolution is first obtained from microscopic origins, the development will produce a macroscopic intermediate formulation in the process of resolution and all viewpoints should be satisfied. Ninth, since the flow of momentum and pseudomomentum hinges on the nature of the forces involved, it is important to distinguish between long-range and short-range forces. Short-range forces arise from bonding forces in the solid and produce mechanical stresses (or contact forces) active only across a surface. They arise as derivatives of a stored energy of the solid which is a function only of the $3N$ mechanical coordinates of the solid. Long-range or volume forces arise from derivatives of a multipole expansion of the interaction energy of the solid with the macroscopic electromagnetic fields. Tenth, unambiguous identification of a density and a flow of a conserved quantity requires, as a minimum, obtaining the actual conservation law, not just a continuity relation of the quantity for some subsystem, such as the electromagnetic field. Otherwise, the density, the flow, and the interaction with other subsystems can be rearranged into many forms with consequent ambiguity of interpretation. Also, the derivation of the general conservation law is preferable in our way of thinking to the evaluation of a thought experiment. While constructing the theory of a thought experiment is often a fruitful approach, it has proven to have many pitfalls in this field because of the rather subtle distinction between momentum and pseudomomentum in experimental situations. Applications to thought experiments are better handled after the general concepts are established. Eleventh, for a conservation law as unfamiliar as that of pseudomomentum in continuum physics, it should be derived from, or verified by, an application of Noether's theorem [17] in order to be certain that the quantity involved is actually pseudomomentum. Twelfth, since the truth of a conservation law never depends on specialized constitutive relations, such as linearized or nondispersive ones, their use should be avoided. At best, their use leads to a loss of generality, and, at worst, they may mislead interpretation. Thirteenth, the constitutive relations for the material response fields (polarization $\mathbf{P}$, magnetization $\mathbf{M}$, quadrupolarization $\mathbf{Q}$, etc.) are given basically in terms of *mechanical* variables of the medium. These response fields combine with the fundamental "vacuum" electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ to produce the combination fields $\mathbf{D}$ and $\mathbf{H}$. These distinctions between the fields are central to the physical interpretation of the momentum and pseudomomentum densities that we derive. Also, the inclusion of magnetization and quadrupolarization in the present derivation is a substantial generalization of the earlier work [17] at the electric dipole (polarization) level. We make use of all of these lessons in the development that follows.

## LAGRANGIAN FORMULATION

Our approach to the Minkowski-Abraham controversy is to derive the two basic conservation laws relevant to the controversy, as redefined by Blount. These are the conservation laws of momentum and pseudomomentum. To this end we use a very general Lagrangian formulation. The Lagrangian approach has the great merits of economy of construction, guaranteed self-consistency of deductions including equality of action and reaction between subsystems, and the *derivation* of all constitutive relations along with the dynamical equations of the system. The reader is referred to Ref. [17] for a full exposition of this Lagrangian method.

We treat a closed system of an arbitrary dielectric crystal, homogeneous in its natural (unperturbed) state, in interaction with the electromagnetic field in the nonrelativistic regime. The crystal may have arbitrary symmetry and anisotropy, arbitrary structural complexity, and arbitrary order of nonlinearity in the interaction of the various modes of excitation—acoustic, optic (both ionic and electronic), and electromagnetic. The theory is based on a macroscopic Lagrangian which, however, is derived from microscopic physics by a long-wavelength limit [17]. Thus it is hoped that the present treatment of the Minkowski-Abraham controversy will satisfy both those desiring a macroscopic resolution and those desiring a microscopic resolution. While a statistical average is preferred by some [7] for passage from microscopic to macroscopic equations, a long-wavelength limit must also produce the same macroscopic equations or there would be another controversy equal to the one presently discussed.

The Lagrangian consists of three parts: the matter Lagrangian $L_M$, the electromagnetic field Lagrangian $L_F$, and the matter-field interaction Lagrangian $L_I$,

$$L=L_M+L_F+L_I \ . \tag{1}$$

Since we deal with the system in the long-wavelength or continuum limit, the Lagrangian is an integral over a Lagrangian density $\mathcal{L}$, which is the quantity entering the Lagrange equations. The density may be taken with respect to the ordinary spatial coordinates $\mathbf{z}$ or with respect to the material coordinates $\mathbf{X}$. The choice depends on which position coordinates one wishes to use as the independent variables. Of course, transformation between the two choices can be done in the Lagrangian or in the equations of motion with use of the deformation transformation.

The matter Lagrangian consists of the kinetic energy of the matter minus the stored energy of the matter. It is important to choose appropriate fields in which to express these two quantities and thus the mechanical motion of the crystal. The vector field coordinate $\mathbf{x}$ representing the continuum formed from the center of mass of all the primitive unit cells of the crystal should be chosen as one such field because (i) all the momentum carried by the matter can be expressed with this coordinate and (ii) the three acoustic modes of the matter arise from the equation of motion of the center-of-mass variable [19] or the closely related displacement vector. The latter relates the spatial or deformed position $\mathbf{x}$ of a mass point ($\mathbf{z}=\mathbf{x}$ inside matter) with its material or undeformed position $\mathbf{X}$ by

$$u_i=x_i-\delta_{iA}X_A \ . \tag{2}$$

Note that spatial frame components are denoted by lowercase subscripts and material frame components are denoted by uppercase subscripts even though the choice of the Kronecker delta as the shifter tensor in Eq. (2) shows that both types of coordinates are referred to the same rectangular Cartesian axes.

If there are $N$ particles per primitive unit cell, then $N-1$ additional vector coordinates $\mathbf{y}^{T\nu}$ ($\nu=1,2,\ldots,N-1$), called internal coordinates, are needed (the superscript $T$ for total is explained below). They can and should be defined as displacement-invariant coordinates. As a result of this they cannot carry momentum, which is consistent with the center-of-mass field carrying all of it. As a convenience the internal coordinates can be defined so as to diagonalize the kinetic energy. In a linearized theory the coordinates that diagonalize both the kinetic and stored energies are the normal coordinates of the optic modes of the crystal. Thus they are linear combinations of the internal coordinates. The material-frame form of the matter Lagrangian is then

$$L_M=\int \mathcal{L}_{MM}\,dV \ , \tag{3}$$

$$\mathcal{L}_{MM}=\tfrac{1}{2}\rho^0(\dot{x})^2+\sum_{\nu}\tfrac{1}{2}m^{\nu}(\dot{y}^{T\nu})^2-\rho^0\Sigma \ . \tag{4}$$

Here $dV\equiv dX_1dX_2dX_3$, $\rho^0$ is the mass density (per undeformed volume) of the crystal, $m^{\nu}$ is the mass density (per undeformed volume) associated with the $\nu$-internal coordinate, the dot represents the material time derivative ($\mathbf{X}$ held fixed), and $\Sigma$ is the stored energy per unit mass.

The stored energy depends only on the configuration of the particles in the primitive unit cell and the bonding forces between them. As such, it can depend only on the mechanical coordinates of the crystal. The stored energy must permit the theory to conserve energy, momentum, and angular momentum. This requires it to be independent of time, to be independent of $\mathbf{z}$ and $\mathbf{x}$, and to be a rotational invariant (in the spatial frame). To satisfy the last condition it is most convenient to express the stored energy as a function of a minimal but complete set of rotationally invariant measures of the mechanical coordinates. Since perturbations of the medium are typically small, a series expansion in terms of the rotational invariants is useful. Such an expansion can be truncated at a finite number of terms for application to some particular interaction only if the rotational invariants vanish in the natural state of the crystal. These several requirements are met with the choice

$$\Sigma=\Sigma(E_{AB},\Pi_C^{\nu}) \ , \tag{5}$$

where

$$E_{AB}=(x_{i,A}x_{i,B}-\delta_{AB})/2 \ , \tag{6}$$

$$\Pi_C^{\nu}=X_{C,j}y_j^{T\nu}-Y_C^{\nu} \ . \tag{7}$$

Here $E_{AB}$ is the Green finite strain tensor, $\mathbf{Y}^{\nu}$ is the value of $\mathbf{y}^{T\nu}$ in the natural state ($\mathbf{y}^{T\nu}\equiv\mathbf{Y}^{\nu}+\mathbf{y}^{\nu}$), and gradient notation is defined by $x_{i,A}\equiv\partial x_i/\partial X_A$ and $X_{C,j}\equiv\partial X_C/\partial x_j$. Since we do not make use of the series expansion of Eq. (5) until the Appendix, we do not write it down here.

The electromagnetic field Lagrangian is given by

$$L_F=\int\mathcal{L}_{FS}dv\ , \tag{8}$$

$$\mathcal{L}_{FS}=\frac{\epsilon_0}{2}(\mathbf{E})^2-\frac{1}{2\mu_0}(\mathbf{B})^2\ , \tag{9}$$

where

$$\mathbf{E}=-\frac{\partial\mathbf{A}}{\partial t}-\nabla\Phi\ , \tag{10}$$

$$\mathbf{B}=\nabla\times\mathbf{A}\ . \tag{11}$$

Here $\Phi$, the scalar potential, and $\mathbf{A}$, the vector potential, are the Lagrangian variables and the integral with respect to $dv\equiv dz_1dz_2dz_3$ indicates that the field Lagrangian density $\mathcal{L}_{FS}$ is a spatial frame density. Note that $\mathbf{E}$ and $\mathbf{B}$ are the fundamental electromagnetic fields because they are derivable from the potentials and so exist both inside and outside the matter.

The interaction Lagrangian is given by

$$L_I=\int\mathcal{L}_{IM}dV=\int\mathcal{L}_{IS}dv\ , \tag{12}$$

$$\mathcal{L}_{IS}=J^{-1}\mathcal{L}_{IM}=\mathbf{j}\cdot\mathbf{A}-q\Phi\ , \tag{13}$$

where $J=\det(\partial x/\partial X)$ is the Jacobian of the deformation transformation and $q$ and $\mathbf{j}$ are the charge and current densities, respectively, in the spatial frame. Note that the interaction Lagrangian density is given in two forms. We find it convenient (but not necessary) to find the matter equations of motion in the material frame and the electromagnetic equations in the spatial frame. Since the interaction Lagrangian affects both, it is given here as two densities appropriate for the two uses.

Since we treat a dielectric crystal, the free-charge density and the free-charge current density vanish. Because these terms are the monopole terms of a multipole expansion of $q$ and $\mathbf{j}$, it is convenient to employ a multipole expansion. This is also convenient because it can be truncated at some appropriate order, commonly the electric dipole order. Here we wish to retain magnetic dipole terms so as to include magnetization terms and thus unambiguously distinguish between the magnetic induction $\mathbf{B}$ and the magnetic field $\mathbf{H}$, a distinction crucial to interpreting the form of the momentum density. Because the electric quadrupole terms are of the same order as the magnetic dipole terms and can mix together with the latter in the internal motion equations, we also must include them. The multipole expansions yield the current and charge densities of bound charge

$$\mathbf{j}=\frac{\partial\mathbf{P}}{\partial t}+\nabla\times(\mathbf{P}\times\dot{\mathbf{x}})+\nabla\times\mathbf{M}-\frac{\partial}{\partial t}(\nabla\cdot\overleftrightarrow{\mathbf{Q}})-\nabla\times[\nabla\cdot(\overleftrightarrow{\mathbf{Q}}\times\dot{\mathbf{x}})]\ , \tag{14}$$

$$q=-\nabla\cdot\mathbf{P}+\nabla\nabla:\overleftrightarrow{\mathbf{Q}}\ . \tag{15}$$

Here the polarization $\mathbf{P}$, the magnetization $\mathbf{M}$, and the quadrupolarization $\overleftrightarrow{\mathbf{Q}}$ are given by

$$\mathbf{P}=\frac{1}{J}\sum_{\nu}q^{\nu}\mathbf{y}^{T\nu}\ , \tag{16}$$

$$\mathbf{M}=\frac{1}{2J}\sum_{\mu,\nu}q^{\mu\nu}\mathbf{y}^{T\mu}\times\dot{\mathbf{y}}^{T\nu}\ , \tag{17}$$

$$\overleftrightarrow{\mathbf{Q}}=\frac{1}{2J}\sum_{\mu,\nu}q^{\mu\nu}\mathbf{y}^{T\mu}\mathbf{y}^{T\nu}\ , \tag{18}$$

where $q^{\nu}$ and $q^{\nu\mu}$ are charges associated with the internal coordinates. Equations (14) and (15) contain spatial-frame time derivatives $\partial/\partial t$ ($\mathbf{z}$ held fixed) and gradients $\nabla_i\equiv\partial/\partial z_i$. The magnetization arising from intrinsic spin is not included.

## EQUATIONS OF MOTION

The Lagrangian density developed in the preceding section is a function of the Lagrangian variables $\mathbf{x}$, $\mathbf{y}^{T\nu}$ ($\nu=1,2,\ldots,N-1$), $A$, and $\Phi$, and an equation of motion for each can be found. The material frame Lagrange equation for $\mathbf{x}$ is obtained from

$$\frac{d}{dt}\frac{\partial\mathcal{L}_M}{\partial\dot{x}_i}=\frac{\partial\mathcal{L}_M}{\partial x_i}-\frac{d}{dX_A}\frac{\partial\mathcal{L}_M}{\partial x_{i,A}}\ , \tag{19}$$

where only the matter and interaction parts of the material-frame Lagrangian density $\mathcal{L}_M$ contribute. The result is

$$\rho^0\ddot{x}_i=\frac{d}{dX_A}\frac{\partial\rho^0\Sigma}{\partial x_{i,A}}+p_jE_{i,j}+\epsilon_{ijk}\dot{p}_jB_k+\epsilon_{ijk}\dot{x}_jB_{k,l}p_l+m_kB_{k,i}+q_{kl}E_{i,kl}+\epsilon_{ijk}\dot{x}_jB_{k,lm}q_{lm}+\epsilon_{ijk}\dot{q}_{jl}B_{k,l}\ , \tag{20}$$

where $\mathbf{p}\equiv J\mathbf{P}$, $\mathbf{m}\equiv J\mathbf{M}$, $\overleftrightarrow{\mathbf{q}}\equiv J\overleftrightarrow{\mathbf{Q}}$, and the electromagnetic fields are evaluated at $\mathbf{z}=\mathbf{x}(\mathbf{X},t)$, $\mathbf{X}$ and $t$ being the independent variables. A similar procedure yields the equations of motion for the internal coordinates

$$m^{\nu}\ddot{y}_i^{\nu}=-\frac{\partial\rho^0\Sigma}{\partial y_i^{T\nu}}+q^{\nu}(E_i+\epsilon_{ijk}\dot{x}_jB_k)+\sum_{\mu}q^{\nu\mu}y_j^{T\mu}E_{ij}+\epsilon_{ijk}\dot{x}_j\sum_{\mu}q^{\nu\mu}y_l^{T\mu}B_{k,l}+\epsilon_{ijk}\sum_{\mu}q^{\nu\mu}\dot{y}_j^{\mu}B_k\ . \tag{21}$$

When the last three terms in Eq. (21) are traced to their origins in the Lagrangian, it is seen that they each contain contributions from both the magnetization and the quadrupolarization. It is for this reason that quadrupolarization is included in the treatment even though our main interest is in the inclusion of magnetization in order to distinguish $\mathbf{H}$ from $\mathbf{B}$.

The spatial-frame Lagrange equation for the vector potential $\mathbf{A}$ is obtained from

$$\frac{d}{dt}\frac{\partial\mathcal{L}_s}{\partial(\partial A_i/\partial t)}=\frac{\partial\mathcal{L}_s}{\partial A_i}-\frac{d}{dz_j}\frac{\partial\mathcal{L}_s}{\partial A_{ij}} \tag{22}$$

to be

$$\frac{1}{\mu_0}\nabla\times\mathbf{B}-\epsilon_0\frac{\partial\mathbf{E}}{\partial t}=\mathbf{j}\ , \tag{23}$$

where $\mathbf{j}$ is given by Eq. (14) and the independent variables are $\mathbf{z},t$ with $\mathbf{z}=\mathbf{x}$ inside matter. A similar procedure yields the equation for the scalar potential $\Phi$

$$\epsilon_0\nabla\cdot\mathbf{E}=q\ , \tag{24}$$

where $q$ is given by Eq. (15).

The electromagnetic equations (23) and (24) are in the Maxwell-Lorentz form in which $\mathbf{E}$ and $\mathbf{B}$ are regarded as the fundamental electromagnetic fields and the response of the matter is determined entirely by $\mathbf{j}$ and $q$, which by Eqs. (14) and (15) are expressed in terms of the matter response fields $\mathbf{P}$, $\mathbf{M}$, and $\overleftrightarrow{\mathbf{Q}}$. Note also that the response fields are defined in Eqs. (16)–(18) in terms of the *mechanical* coordinates of the matter. The Maxwell-Lorentz equations (23) and (24) in conjunction with Eqs. (14) and (15) can be reexpressed in the Maxwell form by defining the magnetic field $\mathbf{H}$ and the electric displacement field $\mathbf{D}$ by

$$\mathbf{H}\equiv\frac{1}{\mu_0}\mathbf{B}-\mathbf{P}\times\dot{\mathbf{x}}-\mathbf{M}+\nabla\cdot(\overleftrightarrow{\mathbf{Q}}\times\dot{\mathbf{x}})\ , \tag{25}$$

$$\mathbf{D}\equiv\epsilon_0\mathbf{E}+\mathbf{P}-\nabla\cdot\overleftrightarrow{\mathbf{Q}}\ . \tag{26}$$

Thus $\mathbf{H}$ and $\mathbf{D}$ are combinations of the fundamental vacuum electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ and the matter response fields $\mathbf{P}$, $\mathbf{M}$, and $\overleftrightarrow{\mathbf{Q}}$. The two remaining Maxwell equations,

$$\nabla\times\mathbf{E}+\frac{\partial}{\partial t}\mathbf{B}=\mathbf{0}\ , \tag{27}$$

$$\nabla\cdot\mathbf{B}=0\ , \tag{28}$$

are direct consequences of the definitions of $\mathbf{E}$ and $\mathbf{B}$ in terms of the potentials and so arise as ancillary conditions in the Lagrangian method.

## MOMENTUM CONSERVATION

Momentum conservation is one of the basic laws of nature and results from the homogeneity of free space. This is proven by Noether's theorem, which shows that the momentum conservation law results from requiring the laws of physics to be invariant to an arbitrary infinitesimal displacement of the spatial coordinates $\mathbf{z}\rightarrow\mathbf{z}+\delta\mathbf{d}$. Besides being able to obtain the momentum conservation law directly from Noether's theorem, it may also be obtained from the equations of motion. Useful and different insights are obtained from the two methods. We will present here the derivation from the equations of motion.

Since momentum is inherently a spatial-frame quantity, it is natural to express its conservation law in the spatial description, that is, with $\mathbf{z},t$ as the independent variables. If Eq. (20) is multiplied by the Jacobian $J$, it can be reexpressed after considerable manipulation as

$$\rho\ddot{x}_i=t^E_{il,l}+qE_i+(\mathbf{j}\times\mathbf{B})_i\ , \tag{29}$$

where $\rho\equiv\rho^0/J$, $\mathbf{j}$ and $q$ are given in Eqs. (14) and (15), and

$$\begin{aligned}t^E_{il}\equiv&\frac{x_{l,A}}{J}\frac{\partial\rho^0\Sigma}{\partial x_{i,A}}+\mathcal{E}_iP_l\\&+\epsilon_{ijk}\left[\frac{\partial Q_{lj}}{\partial t}+(Q_{lj}\dot{x}_m)_{,m}+\epsilon_{ljm}M_m\right]B_k\\&+Q_{lm}E_{i,m}-Q_{lm,m}E_i\\&+\epsilon_{ijk}Q_{lm}\dot{x}_jB_{k,m}-\epsilon_{ijk}(Q_{lm}\dot{x}_j)_{,m}B_k\\=&\frac{x_{i,A}x_{l,B}}{J}\frac{\partial\rho^0\Sigma}{\partial E_{AB}}+\sum_\nu\rho^\nu\ddot{y}^\nu_iy^{T\nu}_l-(\mathcal{E}_iQ_{lm})_{,m}\ ,\end{aligned} \tag{30}$$

where $\mathcal{E}_i\equiv E_i+\epsilon_{ijk}\dot{x}_jB_k$, $\rho^\nu\equiv m^\nu/J$, and the last form of $t^E_{il}$ is obtained with the use of the internal motion equation [20]. Equation (29) can be put into a momentum continuity form by converting a material time derivative in the inertial term to spatial time and space derivatives:

$$\frac{\partial}{\partial t}(\rho\dot{x}_i)-\frac{\partial}{\partial z_l}(t^E_{il}-\rho\dot{x}_i\dot{x}_l)=qE_i+(\mathbf{j}\times\mathbf{B})_i\ . \tag{31}$$

The momentum continuity equation for the electromagnetic field is obtained by forming a vector product of Eq. (23) with $-\mathbf{B}$, a vector product of Eq. (27) with $-\mathbf{E}$, a product of Eq. (24) with $-\mathbf{E}$, and a product of Eq. (28) with $-\mathbf{B}/\mu_0$, and adding them to obtain

$$\frac{\partial}{\partial t}(\epsilon_0\mathbf{E}\times\mathbf{B})_i-\frac{\partial m_{il}}{\partial z_l}=-qE_i-(\mathbf{j}\times\mathbf{B})_i\ , \tag{32}$$

where the electromagnetic stress tensor is defined by

$$m_{il}\equiv\epsilon_0E_iE_l+\frac{1}{\mu_0}(B_iB_l)-\tfrac{1}{2}\left[\epsilon_0E_kE_k+\frac{1}{\mu_0}B_kB_k\right]\delta_{il}\ . \tag{33}$$

The addition of Eqs. (31) and (32) gives the conservation law of momentum in the spatial frame

$$\frac{\partial}{\partial t}(\rho\dot{x}+\epsilon_0\mathbf{E}\times\mathbf{B})_i-\frac{\partial}{\partial z_l}(t^E_{il}+m_{il}-\rho\dot{x}_i\dot{x}_l)=0\ . \tag{34}$$

There are several points to be made about this. First, that this is the conservation law of momentum is verified by obtaining it also from the Lagrangian by applying Noether's theorem for invariance to displacements in the spatial coordinate $\mathbf{z}$. Second, the action and reaction in Eqs. (31) and (32) between the electromagnetic field and the matter are the familiar Lorentz force. Third, the internal motion equations are *not* used to obtain the momentum conservation law, a fact revealed only by the equations-of-motion-type derivation. This is to be expected because the internal coordinates are displacement invariant and so cannot carry momentum. Fourth, there cannot be any confusion between the momentum density of the matter $\rho\dot{x}$ and the momentum density of the electromagnetic field $\epsilon_0\mathbf{E}\times\mathbf{B}$ because the matter speed $\dot{x}$ of an acoustic disturbance is about $10^{-5}c$ and a rigid body speed would likely be even smaller. We thus find no material momentum traveling at the speed of the electromagnetic momentum. Thus the electromagnetic momentum density $\epsilon_0\mathbf{E}\times\mathbf{B}$ that we derive is different

from the Abraham form $\mathbf{E}\times\mathbf{H}/c^2=\epsilon_0\mu_0\mathbf{E}\times\mathbf{H}$. The difference between these is readily interpretable: as seen from Eq. (25) $\mathbf{H}$ differs from $\mathbf{B}$ by terms which involve the response fields $\mathbf{P}$, $\mathbf{M}$, and $\overleftrightarrow{Q}$ which, in turn, are given in terms of the internal coordinates, as seen in Eqs. (16)–(18). Since the internal coordinates cannot contribute to momentum, their absence from the momentum density is expected, and so our derived form of the electromagnetic momentum density in matter is supported on physical grounds.

We have chosen to express the momentum conservation law, Eq. (34), in the spatial description, that is, with $\mathbf{z},t$ ($\mathbf{z}=\mathbf{x}$ inside matter) as independent variables. This is natural because the invariance $\mathbf{z}\rightarrow\mathbf{z}+\delta\mathbf{d}$ that gives rise to momentum conservation is a spatial-frame transformation with the result that momentum is inherently a spatial-frame vector (lowercase subscripts). Nonetheless, the momentum density and the stress tensor can be regarded as functions of $\mathbf{X},t$ as independent variables (the material description) by using the deformation transformation $\mathbf{x}=\mathbf{x}(\mathbf{X},t)$. The momentum conservation law can then be transformed to be

$$\frac{d}{dt}(\rho^0\dot{\mathbf{x}}+J\epsilon_0\mathbf{E}\times\mathbf{B})_i - \frac{\partial}{\partial X_J}\{JX_{J,l}[t^E_{il}+m_{il}+\epsilon_0(\mathbf{E}\times\mathbf{B})_i\dot{x}_l]\}=0\ . \quad (35)$$

Note that the momentum density remains a spatial-frame vector even though it now is in the material description. This description is the natural choice to examine the momentum density of a linear acoustic wave since it is naturally expressed with a material description phase $\mathbf{K}\cdot\mathbf{X}-\omega t$. We see from Eq. (35) that the momentum density of such a wave is simply $\rho^0\dot{x}$, a strictly linear expression, which thus integrates to zero over a wavelength in space or a period in time. In this sense a linear acoustic wave possesses no momentum. It follows that an acoustic phonon possesses no momentum either.

Our derivation differs with those works [6–8,11] which advocate the Abraham result. Other recent works [9,10,14] do not include magnetic effects and so do not distinguish between our result and Abraham's result. Interestingly, Livens [21] and Tiersten and Tsai [22] both obtain $\epsilon_0\mathbf{E}\times\mathbf{B}$, though neither related their expression to the Minkowski-Abraham controversy. Livens obtains the result by reexpressing the Lorentz force [see Eq. (32)], which he takes as fundamental. The Tiersten-Tsai result arises from much the same origin, but through a much more elaborate theory. Since we find the Lorentz force to be the form of the interaction between the matter and the electromagnetic field subsystems, it is not surprising that we agree with them.

Peierls, in a series of papers [14–16], has proposed that the material momentum term $\rho\dot{x}$ of Eq. (34) can be driven by bilinear electromagnetic terms and so be reexpressed in terms of the electromagnetic momentum term $\epsilon_0\mathbf{E}\times\mathbf{B}$ (as he considers a nonmagnetic medium, he does not distinguish between this and the Abraham form). He then concludes that the material momentum $\rho\dot{x}$ travels at the optical velocity. We find this puzzling because the term, by its nature, represents the momentum of the mass of a volume element traveling at the material velocity $\dot{x}$. We do not disagree that a light wave could impart momentum to the center-of-mass continuum, but we do not believe that material momentum can travel at an optical velocity (except for relativistic motion of whole bodies).

de Groot and Suttorp derive a momentum conservation law from both nonrelativistic [23] and relativistic [24] points of view with use of a statistical average to obtain the macroscopic equations from microscopic beginnings. Their nonrelativistic derivation obtains the electromagnetic momentum density $\epsilon_0\mathbf{E}\times\mathbf{B}$ in agreement with our derivation. However, their relativistic treatment obtains $\mathbf{E}\times\mathbf{H}/c^2$, the Abraham result. They believe that this is in agreement with their nonrelativistic result because the magnetization part of $\mathbf{H}$ disappears in the nonrelativistic limit [25]. We find this puzzling based on our argument that the magnetization, being a function of the internal coordinates, cannot carry momentum.

## PSEUDOMOMENTUM CONSERVATION

We pointed out above that a conservation law can be obtained by two methods, from the equations of motion and from Noether's theorem, and that different insights are often obtained from the two methods. Since the pseudomomentum conservation law is uncommon in classical field theories and the prescription for finding it from the equations of motion is not commonly known, we find it useful to derive it first from Noether's theorem.

To do a Noether's theorem derivation one must know the relevant symmetry transformation that leaves the equations of physics invariant. For pseudomomentum conservation the symmetry transformation must represent the homogeneity of the material body (which ideally must extend throughout all space). This can be readily determined in analogy to the symmetry transformation that expresses the homogeneity of space and leads to momentum conservation. That transformation is $\mathbf{x}\rightarrow\mathbf{x}+\delta\mathbf{d}$, where $\mathbf{x}$ is the spatial position coordinate vector and $\delta\mathbf{d}$ is an arbitrary infinitesimal displacement. Thus, by complete analogy the transformation expressing the homogeneity of a material body (in the continuum limit) is $\mathbf{X}\rightarrow\mathbf{X}+\delta\mathbf{D}$, where $\mathbf{X}$ is the material position vector and $\delta\mathbf{D}$ is an arbitrary infinitesimal displacement in the material coordinate system. The conservation law that results from Noether's theorem with the use of the latter symmetry transformation is defined as the pseudomomentum conservation law. We believe that our use of the term "pseudomomentum" to refer to the quantity conserved as a result of material homogeneity is conventional. Our work departs from others by our insistence that the transformation $\mathbf{X}\rightarrow\mathbf{X}+\delta\mathbf{X}$ is the proper one to characterize material homogeneity.

This conservation law has been obtained from Noether's theorem before [17] in the form

$$0=\frac{d}{dt}\left[-\frac{\partial\mathcal{L}_M}{\partial\dot{x}_i}x_{i,A}-\sum_\nu\frac{\partial\mathcal{L}_M}{\partial\dot{y}_i^{T\nu}}y_{i,A}^{T\nu}-\frac{\partial\mathcal{L}_M}{\partial\dot{A}_i}A_{i,A}\right]$$
$$+\frac{d}{dX_B}\left[-\frac{\partial\mathcal{L}_M}{\partial x_{i,B}}x_{i,A}-\sum_\nu\frac{\partial\mathcal{L}_M}{\partial y_{i,B}^{T\nu}}y_{i,A}^{T\nu}-\frac{\partial\mathcal{L}_M}{\partial\Phi_{,B}}\Phi_{,A}\right.$$
$$\left.-\frac{\partial\mathcal{L}_M}{\partial A_{i,B}}A_{i,A}+\mathcal{L}_M\delta_{AB}\right], \tag{36}$$

though no further exploration of its meaning was pursued there. This law is inherently a material-frame law with all fields regarded as functions of $\mathbf{X},t$ and with the pseudomomentum density vector and the pseudostress tensor having components referred to that system (uppercase subscripts). Note that this does not prevent $\mathbf{x}$, $\mathbf{y}^{T\nu}$, and $\mathbf{A}$ from being spatial-frame vectors. The material-frame Lagrangian density $\mathcal{L}_M$ is used in Eq. (36) and consists of Eqs. (4), (13), and $\mathcal{L}_{FM}=J\mathcal{L}_{FS}$ with $\mathcal{L}_{FS}$ given by Eq. (9) where the fields are regarded as having been converted to dependence on $\mathbf{X},t$ by substitution of the deformation transformation $\mathbf{x}=\mathbf{x}(\mathbf{X},t)$. Substitution of the Lagrangian density into Eq. (36) leads to terms in both the pseudomomentum density and pseudostress (or flow of pseudomomentum) that are gauge variant. The combination of terms can be shown to vanish by a lengthy algebraic procedure that is quite analogous to that [26] used for the momentum conservation law when it is found from Noether's theorem. The final result is

$$\frac{d}{dt}\left[-\rho^0\dot{x}_i x_{i,C}-\sum_\nu m^\nu\dot{y}_i^\nu y_{i,C}^\nu+x_{i,C}\epsilon_{ijk}p_jB_k+x_{i,C}\epsilon_{ijk}q_{jl}B_{k,l}+\tfrac{1}{2}\epsilon_{ijk}\sum_{\mu\nu}q^{\mu\nu}y_{i,C}^\mu y_j^\nu B_k\right]$$
$$-\frac{d}{dX_A}\left[-\delta_{CA}\left[\rho^0\dot{x}^2/2+\sum_\nu m^\nu(\dot{y}^\nu)^2/2-\rho^0\Sigma+\mathbf{p}\cdot(\mathbf{E}+\dot{\mathbf{x}}\times\mathbf{B})+\mathbf{m}\cdot\mathbf{B}\right.\right.$$
$$\left.\left.+q_{kl}(E_{k,l}+\epsilon_{kjl}\dot{x}_jB_{i,l})\right]-x_{i,C}\frac{\partial\rho^0\Sigma}{\partial x_{i,A}}\right]=0\ . \tag{37}$$

Several things should be said about this pseudomomentum conservation law. First, now that the form of the general pseudomomentum conservation law is known, it is a simple matter to find a prescription for obtaining it from the equations of motion. That prescription is to form a scalar product of Eq. (20) with $-x_{i,C}$ over $i$, a scalar product of Eq. (21) with $-y_{i,C}^\nu$ over $i$, sum the latter equation over $\nu$ from 1 to $N-1$, add the two results, and rearrange the terms into perfect derivatives. Equation (37) again results. Thus this general conservation law exists independently of Noether's theorem argument; the latter is useful, however, to interpret the origin of it. Second, an interesting insight results from this equations-of-motion derivation: the electromagnetic equations (23), (24), (27), and (28) are not used in the derivation. This fact remains true even if the material form of electromagnetism is used. We interpret this fact by realizing that the pseudomomentum is a quantity that is inherently referenced with respect to the matter, that is, the material coordinate system. The matter equations of motion are thus clearly relevant to it. However, the electromagnetic equations describe the $\mathbf{E}$ and $\mathbf{B}$ fields that can exist in a vacuum and are inherently spatial-frame quantities and so do not contribute. This distinction also expresses itself in the pseudostress tensor where the quantity in the large parentheses in Eq. (37) is just the material-frame Lagrangian density of the matter and matter-field interaction. The remaining part of the Lagrangian, the Lagrangian of the electromagnetic field, which is inherently a spatial-frame density, disappears from there in spite of the entire Lagrangian entering there in Eq. (36). Third, it should be emphasized that in a homogeneous body there are two distinct, separately conserved quantities, the momentum in Eq. (34) and the pseudomomentum in Eq. (37), which have the same physical dimensions. In spite of this similarity they are distinctly different quantities. Pseudomomentum is inherently a vector having material-frame components and momentum is inherently a vector having spatial-frame components. Each is expressed in the coordinate system in which the symmetry transformation that produces its conservation operates. Fourth, because pseudomomentum and momentum are vectors in different frames, they cannot in general be combined. Furthermore, since there are only two distinctly different coordinate systems, the spatial and material coordinate systems, in which to invoke homogeneity, there are only two momentumlike quantities conserved. Thus we are certain that there cannot be any further *general* conservation law beyond the two presented involving a quantity having the physical dimensions of momentum. Fifth, if a situation in which there is no deformation of the material is considered ($x_{i,C}=\delta_{iC}$, $J=1$), then the distinction between the spatial and material coordinate systems vanishes. In this special case, momentum and pseudomomentum can be combined, added for example. Since this special case applies to an optical frequency wave in a material medium, which is the situation traditionally considered in the Minkowski-Abraham controversy, it is considered in the following section. We also discuss there other special circumstances where the two conservation laws can be usefully combined.

It is important to show that the classical expression for the pseudomomentum density found in Eq. (37) agrees with the expected quantum expression for a linear acoustic wave. For this application the magnetization and electric quadrupolarization terms should be dropped. Also, the term representing the interaction of the polar-

ization and the magnetic induction has nothing to do with a linear acoustic wave since even in a piezoelectric crystal the linear electric-field wave that accompanies the linear displacement wave is completely longitudinal and so has no accompanying magnetic-field wave; hence this term should also be dropped. The pseudomomentum density for a linear acoustic wave is thus

$$G_C=-\rho^0\dot{x}_i x_{i,C}-\sum_\nu m^\nu \dot{y}^\nu_i y^\nu_{i,C} \ . \tag{38}$$

We prove in the Appendix that the second term vanishes for a linear acoustic wave in the long wavelength or continuum limit. We reexpress the first term with the use of the displacement **u**

$$u_i(X,t)=U_0 b_i \cos(\mathbf{K}\cdot\mathbf{X}-\omega t) \ , \tag{39}$$

where $U_0$ is a constant amplitude, **b** is a unit displacement eigenvector, and **K** is a material-frame wave vector. If we substitute into Eq. (38) and take a time average, then

$$\overline{\mathbf{G}}=\frac{\rho^0\omega U_0^2}{2}\mathbf{K} \ . \tag{40}$$

The time-averaged energy density in such a wave is

$$\overline{H}=\frac{\rho^0\omega^2 U_0^2}{2} \ . \tag{41}$$

When quantized, we take this wave to have $N$ phonons per unit volume and so an energy density of

$$\overline{H}=N\hbar\omega \ . \tag{42}$$

Combining the three preceding equations yields

$$\overline{\mathbf{G}}=N\hbar\mathbf{K} \ , \tag{43}$$

the expected amount of pseudomomentum for $N$ phonons. Thus the general expression for pseudomomentum density in Eq. (37) is consistent with conventional thinking on pseudomomentum in linear acoustic waves.

## WAVE MOMENTUM

We pointed out in the preceding section that in the absence of deformation of the material medium ($x_{i,C}=\delta_{iC}$, $J=1$) the distinction between the spatial and material coordinate systems disappears and thus momentum and pseudomomentum may be combined. Since the last two terms in Eq. (37) in the pseudomomentum density are trilinear, rather than bilinear, they represent nonlinear corrections which have played no role in the Minkowski-Abraham controversy. Thus we drop them here. The momentum density **g** and the pseudomomentum density **G** in the absence of deformation are then given by

$$\mathbf{g}=\rho^0\dot{\mathbf{x}}+\epsilon_0\mathbf{E}\times\mathbf{B} \ , \tag{44}$$

$$\mathbf{G}=-\rho^0\dot{\mathbf{x}}-\sum_\nu m^\nu \dot{y}^\nu_i \nabla y^\nu_i+\mathbf{P}\times\mathbf{B} \ . \tag{45}$$

The forms of these suggest adding them; we denote the sum by

$$\mathcal{G}=\mathbf{D}\times\mathbf{B}-\sum_\nu m^\nu \dot{y}^\nu_i \nabla y^\nu_i \tag{46}$$

and name it *wave momentum* for reasons that will become apparent. Note that if the Abraham momentum $\mathbf{E}\times\mathbf{H}/c^2$ were in Eq. (44), the addition would not give $\mathbf{D}\times\mathbf{B}$ in Eq. (46). Note also the absence of any term representing the momentum of the center-of-mass continuum in the expression for wave momentum.

Let us comment on Eqs. (45) and (46). It is apparent from Eq. (45) that the pseudomomentum of a light wave in the material medium is the combination of the second and third terms on the right side, the first term being that of the (possibly) moving matter. The second term, we will show shortly, accounts for the dispersion of the optical properties of the medium, an effect not included in the work of Blount [8] or Peierls [16]. However, the main difference that we find with them is in the third term: we find the pseudomomentum of a light wave to be $\mathbf{P}\times\mathbf{B}$ while they conclude it is $\mathbf{D}\times\mathbf{B}$. Blount identified the latter quantity as the pseudomomentum by an argument in which he imposed invariance of his equations to a translation of the displacement **u** or what he termed the "disturbance." Since the displacement is the difference of the spatial coordinate and the material coordinate, Eq. (2), invariance to a translation of **u** is equivalent to invariance to translations in both **x** and **X**. Such a procedure would be expected to produce a conservation law that combines momentum and pseudomomentum. From this reasoning it is not surprising that Blount's conserved quantity $\mathbf{D}\times\mathbf{B}$ is just (apart from the dispersive term) the sum of our momentum density Eq. (44) and our pseudomomentum density Eq. (45). Peierl's approach was quite similar. We believe that the invariance $\mathbf{X}\rightarrow\mathbf{X}+\delta\mathbf{D}$ that we used to represent homogeneity of the matter and so to obtain pseudomomentum conservation is so completely analogous to the invariance $\mathbf{x}\rightarrow\mathbf{x}+\delta\mathbf{d}$, which everyone agrees represents homogeneity of space and leads to momentum conservation, that our identification of pseudomomentum must be correct.

Next we evaluate Eqs. (44)–(46) for a linear optical wave in a crystal. The electric field is taken as

$$\mathbf{E}=\mathbf{e}E_0\cos(\mathbf{k}\cdot\mathbf{z}-\omega t) \ , \tag{47}$$

where $\mathbf{k}=\omega n\mathbf{s}/c$ is the wave vector, $n$ the refractive index, **s** a unit propagation vector, and **e** a unit electric-field eigenvector. Also,

$$\mathbf{B}=\frac{n}{c}\mathbf{s}\times\mathbf{e}E_0\cos(\mathbf{k}\cdot\mathbf{z}-\omega t) \ . \tag{48}$$

Since the terms involving $\rho^0\dot{x}$ in Eqs. (44) and (45) represent motion of the center of mass of a volume element, they do not propagate at the optical velocity and can be dropped here. The time-averaged momentum density of a light wave from Eq. (44) is then

$$\overline{\mathbf{g}}=\frac{\epsilon_0 E_0^2 n}{2c}[\mathbf{s}-\mathbf{e}(\mathbf{s}\cdot\mathbf{e})] \ , \tag{49}$$

which has the direction of the group velocity.

To evaluate the pseudomomentum density we must first evaluate the second term of Eq. (45). This is done in

the Appendix. It becomes for the linear optical wave (time-averaged)

$$-\sum_{\nu} m^{\nu}\dot{y}_i^{\nu}\nabla y_i^{\nu}=\frac{\epsilon_0 E_0^2}{4}\left[\mathbf{e}\cdot\frac{\partial\overleftrightarrow{\chi}(\omega)}{\partial\omega}\cdot\mathbf{e}\right]\mathbf{k}\ , \tag{50}$$

where $\overleftrightarrow{\chi}$ is the linear electric susceptibility. The time-averaged pseudomomentum of a light wave then becomes from Eq. (45)

$$\overline{\mathbf{G}}=\frac{\epsilon_0 E_0^2 n}{2c}\left[\mathbf{se}\cdot\left[\overleftrightarrow{\chi}(\omega)+\frac{\omega}{2}\frac{\partial\overleftrightarrow{\chi}(\omega)}{\partial\omega}\right]\cdot\mathbf{e}-\mathbf{es}\cdot\overleftrightarrow{\chi}(\omega)\cdot\mathbf{e}\right]\ . \tag{51}$$

Note that this pseudomomentum density is not in the direction of the propagation vector $\mathbf{s}$ (nor the direction of the group velocity).

The time-averaged value of the wave momentum from Eq. (46) for a linear optical wave now becomes [remembering $\mathbf{s}\cdot\overleftrightarrow{\kappa}(\omega)\cdot\mathbf{e}$ is zero]

$$\overline{\mathcal{G}}=\frac{\epsilon_0 E_0^2 n}{2c}\mathbf{e}\cdot\left[\overleftrightarrow{\kappa}(\omega)+\frac{\omega}{2}\frac{\partial\overleftrightarrow{\kappa}(\omega)}{\partial\omega}\right]\cdot\mathbf{es}\ , \tag{52}$$

where $\overleftrightarrow{\kappa}(\omega)=\overleftrightarrow{1}+\overleftrightarrow{\chi}(\omega)$ is the dielectric tensor. Note that $\overline{\mathcal{G}}$ is in the direction of the propagation vector even though neither $\mathbf{g}$ nor $\mathbf{G}$ is. The quantities in Eq. (52) are closely related to the time-averaged energy density of the light wave given by

$$\overline{H}=\frac{\epsilon_0 E_0^2}{2}\mathbf{e}\cdot\left[\overleftrightarrow{\kappa}(\omega)+\frac{\omega}{2}\frac{\partial\overleftrightarrow{\kappa}(\omega)}{\partial\omega}\right]\cdot\mathbf{e}\ . \tag{53}$$

In fact, the magnitude of the quantity in Eq. (52) can be described as the total energy density of the wave divided by the phase velocity of the wave $c/n$. From the quantum viewpoint the plane optical wave can be regarded as consisting of $N$ photons per unit volume having an energy density of

$$\overline{H}=N\hbar\omega\ . \tag{54}$$

Combining the three preceding equations yields

$$\overline{\mathcal{G}}=N\hbar\mathbf{k}\ . \tag{55}$$

Equation (55) may be said to be the expected result, but our derivation of it requires an alternative, but immediately reasonable, interpretation. The quantity $\hbar\mathbf{k}$ has often been called (but without proof, to our knowledge) the pseudomomentum of a photon in a material medium, but our derivation shows that it is the sum of momentum and pseudomomentum. It is for this reason that we give it a new name, *wave momentum*. This different interpretation of $\hbar\mathbf{k}$ of a photon in matter from $\hbar\mathbf{k}$ of a phonon, as discussed in the preceding section, arises quite naturally from the essentially different natures of light and sound. Light can travel in a vacuum, sound cannot. When light travels through matter, part is carried by the fundamental of vacuum fields $\mathbf{E}$ and $\mathbf{B}$, and part is carried by the response fields $\mathbf{P}$, $\mathbf{M}$, $\overleftrightarrow{\mathbf{Q}}$, etc. of the matter. The former transmit the momentum, the latter create the pseudomomentum. The sound wave, being entirely a material based field, possesses only pseudomomentum. Thus its wave momentum is equal to its pseudomomentum. We conclude that the wave momentum, not the pseudomomentum, is the unifying concept between the quantum interpretation of sound and light, Eqs. (43) and (55).

It is important to realize that the Minkowski momentum $\mathbf{D}\times\mathbf{B}$ is not sufficient to obtain the result of Eq. (52) for a realistic medium; the dispersive term, Eq. (50), is essential in order to obtain the full expression for the energy density, Eq. (52). We emphasize that it would be inconsistent to consider a dispersionless material medium having a refractive index differing from unity. The oscillators in the matter (the optic modes) that couple to the electromagnetic field and make the refractive index differ from one, the vacuum value, inherently have resonant frequencies and so inescapably introduce dispersion. Our derivation handles this dispersion in a general manner. It should be further remarked that the addition of momentum and pseudomomentum to produce Eq. (46), which gives the desired wave momentum of Eq. (55), has depended on the momentum density of the electromagnetic field being $\epsilon_0\mathbf{E}\times\mathbf{B}$, not the Abraham momentum density $\epsilon_0\mathbf{E}\times\mu_0\mathbf{H}$. We believe that this gives further support to the correctness of the $\epsilon_0\mathbf{E}\times\mathbf{B}$ form of momentum density that we derive.

The inclusion of dispersion is important from another viewpoint. Examination of the energy density in Eq. (53) reveals that there is energy in the light wave contained in the dispersive term that does not contribute to the momentum of the wave as given in Eq. (49). This is again explained by the fact that the internal coordinates (optic modes) can store energy as a light wave passes but cannot contain momentum because of their displacement-invariant nature.

The significance of the wave momentum in nonlinear optics, such as, for example, in sum-frequency generation $\omega_1+\omega_2=\omega_3$, is that its conservation law (with $\hbar$ divided out)

$$\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}_3 \tag{56}$$

expresses the phase-matching condition needed to obtain maximum power transfer in a finite crystal from waves 1 and 2 into wave 3. In this situation it is not a rigorous requirement as is apparent from the solution for such an interaction being proportional to the phase-matching function. In this situation Eq. (56) is not a rigorous conservation law because a finite-sized crystal does not have material homogeneity throughout all space. It does become a rigorously exact conservation law for an infinite crystal.

The wave momentum density of Eq. (55) is proportional to the refractive index of the material medium if the energy density of the wave is held constant ($N=$const) when comparing one medium to another. This wave momentum is clearly the proper quantity to explain momentum transfer from a light wave in a medium to a test object such as the torsion-balance-mounted mirror used in the Jones-Richard experiment [12] and its improved successor [13]. Those experiments accurately measured a deflection proportional to the refractive index, as we would expect. Of course, detailed application

of our result to those experiments cannot be made because they were necessarily performed in dielectric liquids while our theory applies to crystalline media.

From these several considerations we conclude that wave momentum is a more important quantity in wave interactions than either momentum or pseudomomentum alone.

The fact that the sum of momentum and pseudomomentum in the absence of deformation gives such a readily interpretable and important quantity as the wave momentum suggests that the entire conservation laws, Eq. (34) and (37), should be added under this condition also. The result is a combined, specialized conservation law for optical interactions,

$$\frac{\partial \mathcal{G}_i}{\partial t}-\frac{\partial \mathcal{T}_{ij}}{\partial z_j}=0\ , \tag{57}$$

where $\mathcal{G}$ is given by Eq. (46) and

$$\mathcal{T}_{ij}=E_iD_j+H_iB_j-\tfrac{1}{2}(\mathbf{E}\cdot\mathbf{D}+\mathbf{B}\cdot\mathbf{H})\delta_{ij} - \left[\sum_\nu \frac{m^\nu}{2}(\dot{\mathbf{y}}^\nu)^2-\rho^0\Sigma+\tfrac{1}{2}\mathbf{P}\cdot\mathbf{E}+\tfrac{1}{2}\mathbf{M}\cdot\mathbf{B}\right]\delta_{ij}\ . \tag{58}$$

Here, once again, the electric quadrupolarization terms are dropped (but not the magnetization terms) for simplicity, and the moving medium terms are dropped also ($\dot{x}=0$). The first four terms comprise the stress tensor of Minkowski [1] and Abraham [2]; the last four are new but have a zero time average for a linear light wave in a dielectric.

Up to this point we have considered adding the momentum and pseudomomentum conservation laws only in the complete absence of deformation. There is also a way that they can be combined (added) in the presence of deformation for application to a specific interaction such as acousto-optic diffraction. The procedure is to expand the fields in the various terms of each conservation law to a level appropriate to the interaction. For example, the acousto-optic interaction is a three-field interaction, two fields being electric fields and one field being a displacement field. Once this is done, the distinction between the spatial coordinate $\mathbf{x}$ and the material coordinate $\mathbf{X}$ can be dropped because their difference by Eq. (2) would then be of higher order. The conservation laws can then be added and the resultant quantity again called wave momentum. If the sum is integrated over a volume larger than the interacting waves, the divergence terms yield surface integrals that vanish. If the time derivative terms are integrated from a time before the interaction begins to a time after it is complete and the terms evaluated, the equation of conservation of wave momentum for the particular interaction results. It is equivalent to what is called the phase-matching condition of the interaction.

## SUMMARY

We approached the Minkowski-Abraham controversy on the nature of momentum of a light wave in matter by deducing the momentum conservation law from a very general nonrelativistic Lagrangian theory of the interaction of the electromagnetic field with a dispersive, deforming dielectric. Because of the redefinition of the controversy by Blount [8], we also calculated the pseudomomentum conservation law that results from homogeneity of the material medium. Particularly because of the unfamiliarity of pseudomomentum in classical field theories, we found it advantageous to obtain each conservation law from two methods: Noether's theorem and combining the equations of motion. Each method gives different and important information. Noether's-theorem approach guarantees that the nature of the conserved quantity is properly defined by relating it to the symmetry transformation (invariance) from which it arises. The derivations from the equations of motion reveal that certain fields do not contribute to each of the quantities: the internal motion field equations representing the optic modes of the medium do not contribute to momentum and the electromagnetic-field equations do not contribute to pseudomomentum. We found that each of these facts aided physical interpretation.

The momentum and pseudomomentum densities found from the conservation laws contain contributions from both the electromagnetic field and from the matter. In each case, however, the division between the contributions is crisp and obvious.

We find that the electromagnetic momentum density is $\epsilon_0\mathbf{E}\times\mathbf{B}$, not the Abraham form $\epsilon_0\mathbf{E}\times\mu_0\mathbf{H}$. We find the difference between these forms physically interpretable and in support of the former. The $\mathbf{E}$ and $\mathbf{B}$ fields are the fundamental, vacuum electromagnetic fields, that is, those derivable from the potentials that are the more fundamental quantities in the Lagrangian theory. The magnetic field $\mathbf{H}$, on the other hand, contains in addition response fields of the matter such as the magnetization. Since the response fields are all proportional to some power of the internal motion coordinates, which we showed cannot contribute to momentum, their absence from our momentum expression is understandable and to be expected.

We find the electromagnetic pseudomomentum density is easily distinguishable from the matter pseudomomentum when the deformation of the medium vanishes. Then it is $\mathbf{P}\times\mathbf{B}$ plus a dispersive term involving the optic modes. The first term is physically reasonable because it is proportional to a response field of the matter, the polarization. It differs from the form advocated by Blount and Peierls, which is $\mathbf{D}\times\mathbf{B}$.

When we applied the above expressions for momentum and pseudomomentum densities to a linear light wave, we found that a very important quantity was the sum of momentum and pseudomomentum, which we thus named *wave momentum*. It is the momentum that in the quantum regime is expressible as $\hbar\mathbf{k}$ per photon. We also found that an acoustic phonon has pseudomomentum of $\hbar\mathbf{k}$, zero momentum, and thus wave momentum of $\hbar\mathbf{k}$.

We found that the components of the momentum and pseudomomentum vectors are inherently referred to different coordinate systems, momentum to the spatial coordinates, pseudomomentum to the material coordinates. Thus, under general circumstances they cannot be combined. However, we indicated how they can be indi-

vidually expanded for a particular interaction and then added to form wave momentum. After integration over space and time the result is the well-known conservation of wave-vector or phase-matching condition of the wave interaction.

We close with a caveat. We expressed confidence in our derivation of the expression $\epsilon_0\mathbf{E}\times\mathbf{B}$ for the electromagnetic momentum density in contradistinction to the Abraham form $\epsilon_0\mathbf{E}\times\mu_0\mathbf{H}$ based on a physical argument about the natures of **B** and **H**. However, de Groot and Suttorp [7] obtained the Abraham form from a relativistic derivation of the momentum conservation law and they argued that only in the nonrelativistic limit does the magnetization drop out leaving the form we obtain. Their conservation law is exceptionally long, complex, and difficult to interpret. This leaves a puzzling discrepancy between their work and ours. Thus, while we argued that our nonrelativistic treatment should be capable of deciding the main issues in the Minkowski-Abraham controversy, it appears that this one, numerically small but conceptually large, issue can be fully resolved to our satisfaction only with a relativistic generalization of our method.

## ACKNOWLEDGMENTS

We wish to acknowledge helpful conversations and correspondence with E. Blount, M. W. Klein, M. Lax, G. A. Maugin, R. Peierls, L. G. Suttorp, H. F. Tiersten, and A. Walther. This material is based upon work supported by the National Science Foundation under Grant No. DMR-8922578.

## APPENDIX

While the derivation of the general conservation laws does not require specifying the form of the stored energy, Eq. (5), application to linear acoustic and optical waves does. This is done by a series expansion in terms of rotationally invariant measures of the deformation gradient and the internal motion coordinates defined in Eqs. (6) and (7). Each of these measures is chosen to vanish for the natural state of the crystal, that is, in the absence of any static or dynamic perturbation of the perfect crystal. Since linear terms in the stored energy are needed only when there are initial electric fields or strains, as met in ferroelectric and ferroelastic phenomena, they may be dropped here and we have

$$\rho^0\Sigma={}^{02}M_{ABCD}E_{AB}E_{CD}+\sum_{\mu,\nu}{}^{20}M^{\mu\nu}_{AB}\Pi^{\mu}_{A}\Pi^{\mu}_{B}+\sum_{\nu}{}^{11}M^{\nu}_{ABC}\Pi^{\nu}_{A}E_{BC}\ , \tag{A1}$$

where the various ${}^{mn}M$ are called material descriptors and serve as expansion coefficients.

The linear terms of the internal motion equations (21) become, with the use Eq. (A1),

$$m^{\nu}\ddot{y}^{\nu}_{i}=q^{\nu}E_i-2\sum_{\mu}{}^{20}M^{\nu\mu}_{ib}(y^{\mu}_{b}+u_{b,c}Y^{\mu}_{c})-{}^{11}M^{\nu}_{ibc}u_{b,c}\ , \tag{A2}$$

where the magnetization and electric quadrupolarization terms are omitted since they represent wave-vector dispersion corrections to linear excitations. Once the terms are expanded to a given order (linear here), the distinction between material and spatial coordinates and components (upper and lower case indices) may be dropped as being of higher order. The internal coordinates can be transformed to optic-mode normal coordinates $\eta^k$ by

$$\eta^k=\sum_{\mu}(m^{\mu})^{1/2}\mathbf{n}^{k\mu}\cdot\mathbf{y}^{\mu}\ , \tag{A3}$$

where the normal-mode eigenvectors $\mathbf{n}^{k\mu}$ satisfy

$$2\sum_{\mu}\frac{n^{k\mu}_{i}\,{}^{20}M^{\mu\nu}_{ij}}{(m^{\mu}m^{\nu})^{1/2}}=\Omega^2_k n^{k\nu}_{j} \tag{A4}$$

and form a complete orthonormal set:

$$\sum_{\mu}\mathbf{n}^{k\mu}\cdot\mathbf{n}^{l\mu}=\delta^{kl}\ , \tag{A5}$$

$$\sum_{k}\mathbf{n}^{k\mu}_{i}n^{k\nu}_{j}=\delta^{\mu\nu}\delta_{ij}\ . \tag{A6}$$

This leads to Eq. (A2) being expressed as

$$\ddot{\eta}^k+\Omega^2_k\eta^k=c^k_iE_i-({}^{11}M^k_{ab}+N^k_{ab})u_{a,b}\ , \tag{A7}$$

where

$$c^k_i\equiv\sum_{\mu}\frac{q^{\mu}n^{k\mu}_{i}}{(m^{\mu})^{1/2}}\ , \tag{A8}$$

$${}^{11}M^k_{ab}\equiv\sum_{\nu}\frac{n^{k\nu}_{i}\,{}^{11}M^{\nu}_{iab}}{(m^{\nu})^{1/2}}\ , \tag{A9}$$

$$N^k_{ab}\equiv\Omega^2_k\sum_{\nu}n^{k\nu}_{a}(m^{\nu})^{1/2}Y^{\nu}_{b}\ . \tag{A10}$$

If the right side of Eq. (A7) is represented as simply $L^k$, then the normal coordinate $\eta^k$ can be written for a plane-wave excitation of angular frequency $\omega$ as

$$\eta^k=\frac{L^k}{\Omega^2_k-\omega^2}\ . \tag{A11}$$

The dispersive term in the pseudomomentum can now be evaluated for a linear acoustic wave to be

$$G^{\rm disp}_{c}=-\sum_{\nu}m^{\nu}\dot{y}^{\nu}_{i}y^{\nu}_{i,c} \tag{A12a}$$

$$=-\sum_{k}\dot{\eta}^k\eta^k_{,c} \tag{A12b}$$

$$=K\frac{c}{2}\frac{\partial}{\partial\omega}\sum_{k}\frac{(L^k)^2}{\Omega^2_k-\omega^2}\ , \tag{A12c}$$

where **K** is the wave vector of the acoustic wave. Though the last form of Eq. (A12) could be elaborated in terms of the linear electric susceptibility, the piezoelectric stress tensor, etc., it is not necessary. The important point is that the frequency dependence of each of these quantities is explicitly exhibited in Eq. (A12c). Since the resonant frequencies $\Omega_k$ $(k=1,2,\ldots,3N-3)$ of optic

modes in normal stable crystals are many terahertz while the frequency of the linear acoustic wave is typically at least one thousand times smaller, we see that these quantities have an exceedingly small frequency dependence and thus the frequency derivative can be taken as zero, as used in the text. Only in the case of a soft-optic-mode frequency nearing zero will this argument need modification, as has been discussed [18]. That case, of course, is not a simple acoustic wave as we wish to treat here.

Next we wish to evaluate Eq. (A12a) for a linear optical wave. Since there is no significant displacement gradient existing at optical frequencies, its terms may be dropped from $L^k$ with the result

$$\mathbf{G}^{\text{disp}}=\frac{\epsilon_0 \mathbf{k}}{2}\frac{\partial \chi_{ij}(\omega)}{\partial \omega}E_i E_j \ , \tag{A13}$$

where

$$\chi_{ij}(\omega)\equiv\frac{1}{\epsilon_0}\sum_k \frac{c_i^k c_j^k}{\Omega_k^2-\omega^2} \tag{A14}$$

is the linear electric susceptibility. Since an optical frequency typically is above the ionic mode resonances and below the electronic mode resonances, the frequency dependence of $\chi_{ij}(\omega)$ is significant and must be retained, as used in the text.


[1] H. Minkowski, Nachr. Ges. Wiss. Göttingen 53 (1908); Math. Ann. **68**, 472 (1910).

[2] M. Abraham, Rend. Pal. **28**, 1 (1909); **30**, 33 (1910).

[3] W. Pauli, *Relativitätstheorie* in *Encyklopädie der Mathematischen Wissenschaften* (Teubner, Leipzig, 1921), Vol. V19 [English translation: *Theory of Relativity* (Pergamon, New York, 1958), see p. 216, note 11].

[4] C. Møller, *The Theory of Relativity* (Clarendon, Oxford, 1952).

[5] J. Brevik, Mat. Fys. Medd. Dan. Vid. Selsk. **37** (11), Pt. 1 (1970); **37** (13), Pt. 2 (1970).

[6] P. Penfield, Jr. and H. A. Haus, *Electrodynamics of Moving Media* (MIT Press, Cambridge, MA, 1967).

[7] S. R. de Groot and L. G. Suttorp, *Foundations of Electrodynamics* (North-Holland, Amsterdam, 1972). For an extended and updated review, see L. G. Suttorp, in *Physics in the Making,* edited by A. Sarlemijn and M. J. Sparnaay (Elsevier, Amsterdam, 1989), p. 167.

[8] E. I. Blount, Bell Telephone Laboratories Technical Memorandum 38139-9 (1971).

[9] D. V. Skobel'tsyn, Usp. Fiz. Nauk **110**, 253 (1973) [Sov. Phys.—Usp. **16**, 381 (1973)].

[10] J. P. Gordon, Phys. Rev. A **8**, 14 (1973).

[11] F. N. H. Robinson, Phys. Rep. C **16**, 314 (1975).

[12] R. V. Jones and J. C. S. Richards, Proc. R. Soc. London Ser. A **221**, 480 (1954).

[13] R. V. Jones and B. Leslie, Proc. R. Soc. London Ser. A **360**, 347 (1978).

[14] R. Peierls, Proc. R. Soc. London Ser. A **347**, 475 (1976).

[15] R. Peierls, Proc. R. Soc. London Ser. A **355**, 141 (1977).

[16] R. Peierls, in *Highlights of Condensed Matter Theory,* edited by F. Bassani, F. Fumi, and M. P. Tosi (North-Holland, Amsterdam, 1985), p. 237.

[17] D. F. Nelson, *Electric, Optic, and Acoustic Interactions in Dielectrics* (Wiley, New York, 1979). (Although this book is no longer in print, a limited number of paperback copies are available from the author.)

[18] D. F. Nelson, Phys. Rev. Lett. **60**, 608 (1988).

[19] Caution should be used in considering these two facts. There is momentum in an acoustic wave in a medium at a particular place and at a particular time. That momentum, however, integrates to zero over one wavelength in space or one period in time. For the same reason the momentum carried by a phonon is zero.

[20] This substitution of the internal motion equations was found useful in proving the inherent asymmetry of the elastic stress tensor $t_{il}^E$ (Ref. 17, p. 132). It was also thought to be significant because at the electric dipole level the substitution eliminates the polarization and the electromagnetic fields from the expression for $t_{il}^E$ as in the second form of Eq. (30). This equation also shows that the substitution eliminates the magnetization and most quadrupolarization terms. Thus it is curious that a single quadrupolarization term remains.

[21] G. H. Livens, *The Theory of Electricity* (University Press, Cambridge, England, 1926), p. 258.

[22] H. F. Tiersten and C. F. Tsai, J. Math. Phys. **13**, 361 (1972). The result of interest here is contained in Eqs. (4.18) and (4.19).

[23] S. R. de Groot and L. G. Suttorp, *Foundations of Electrodynamics* (Ref. 7), p. 49.

[24] S. R. de Groot and L. G. Suttorp, *Foundations of Electrodynamics,* (Ref. 7), p. 272.

[25] S. R. de Groot and L. G. Suttorp, *Foundations of Electrodynamics* (Ref. 7), p. 276.