Aust. J. Phys., 1977, **30**, 533–75

Energy–Momentum Tensors for Dispersive Electromagnetic Waves

R. L. Dewar

Department of Theoretical Physics, Research School of Physical Sciences, Australian National University; present address: Plasma Physics Laboratory, Princeton University, P.O. Box 451, Princeton, N.J. 08540, U.S.A.

Abstract

Classical relativistic field theory is used as a basis for a general discussion of the problem of splitting up the total energy–momentum tensor of a system into contributions from its component subsystems. Both the Minkowski and Abraham forms (including electrostriction) arise naturally in alternative split-up procedures applied to a nondispersive dielectric fluid. The case of an electromagnetic wave in a (spatially and temporally) dispersive medium in arbitrary but slowly varying motion is then treated. In the dispersive case the results cannot be found by replacing the dielectric constant ε with $\varepsilon(k, \omega)$ but include derivatives with respect to the wave vector k and the frequency ω. Ponderomotive force expressions are obtained and the perturbation in the total energy–momentum tensor due to a one-dimensional wavepacket is found. A nonlinear Schrödinger equation is obtained for the evolution of a three-dimensional wavepacket. Both hot and cold plasmas are treated.

1. Introduction

The ancient Abraham–Minkowski dispute regarding the correct form of the energy–momentum tensor for nondispersive electromagnetic waves in material media has recently received renewed attention in the literature. For a review of the older literature the reader is referred to the English edition of Pauli's book on relativity (1958, pp. 109, 110), while for the current status of the controversy the review by Robinson (1975) is recommended.

The latter author, while emphasizing the difficulty of obtaining a general microscopic derivation, points out that the problem of obtaining a macroscopic solution in arbitrarily moving media has been solved by Penfield and Haus (1967) using their method of virtual power. As Robinson points out, the result is in fact a generalization of Helmholtz's (1882) solution for the force density acting on a medium subject to static fields, a result which predates the controversy itself by almost 30 years! Stated succinctly, the conclusion is that the force density acting on the medium (henceforth called the ponderomotive force density) is that expected from the Abraham form of the energy–momentum tensor, plus a part described macroscopically as electrostrictive and magnetostrictive effects. This result can be found in such text books as Landau and Lifshitz (1960) or Panofsky and Phillips (1962) for the special case of quasistatic nondispersive media. It is the calculation from first principles of the electrostrictive and magnetostrictive coefficients which makes a microscopic treatment difficult (Robinson 1975; Peierls 1976).

There is a case, however, in which a microscopic treatment is possible, namely that of the collisionless plasma. The problem of ponderomotive forces of electromagnetic waves in inhomogeneous plasmas is of great interest in laser fusion research (Hora

1969) and also has application in magnetic containment devices in RF confinement and microwave heating. The reason for the tractability of the problem in the plasma case is that the particles of a plasma are weakly interacting and may be adequately described using a self-consistent field model: the Vlasov equation or some fluid approximation to it. Klima and Petrzilka (1968) have shown that the ponderomotive force in a cold plasma is that expected from the Abraham tensor with the electrostrictive correction. This is actually quite surprising since a plasma is a highly dispersive medium and it is not clear that the conventional treatment holds. Landau and Lifshitz (1960, p. 256) give a derivation of the time-averaged internal energy density in a medium exhibiting temporal dispersion, but they explicitly state that ponderomotive force expressions have not been derived for such a medium. We shall see in Section 4*d* here that the reason why the result holds is that there is no *spatial* dispersion in a cold plasma. This is no longer true in a warm plasma, and it is one of the principal aims of this paper to derive the ponderomotive force for a system exhibiting spatial dispersion. Our overall aim is to provide a unified macroscopic (continuum) description in which ponderomotive effects in all states of matter (solids, fluids and plasmas) can be discussed.

A simple application of the ponderomotive force expression combined with the electromagnetic energy–momentum tensor is to calculate the total perturbed energy–momentum tensor convected with a one-dimensional wavepacket. This has been discussed by Haus (1969) and by Robinson (1975) in the nondispersive case, and by Hora (1974) and Klima and Petrzilka (1973, 1975) in the cold plasma case. As the equation of motion for the background medium must be solved to find the amount of momentum and energy carried by the background, the result corresponds neither to the Abraham nor to the Minkowski result in general—the medium 'dresses' the wavepacket and modifies the energy–momentum tensor. A three-dimensional wavepacket leaves a sonic wake behind it (Peierls 1976), which is related to induced Brillouin scattering (Kroll 1965). There are also self-focusing effects and modulational instabilities (Karpman and Krushkal' 1969) which tend to break an initially one-dimensional wavetrain into three-dimensional wavepackets.

Although the preceding remarks would appear to imply that Minkowski was 'wrong' and Abraham (and Helmholtz) were 'right', the situation is not as simple as this since there is no unique way of splitting up a system into interacting subsystems. This point has been made clearly by Penfield and Haus (1967). The Minkowski form is wrong only if one demands that the energy–momentum tensor for the background subsystem be unaffected in form by the introduction of interacting fields. But surely this is the only 'natural' assumption? In this paper we argue that there is at least one other equally natural form for the background energy–momentum tensor. By basing the treatment on Hamilton's principle and the methods of relativistic field theory (Pauli 1941; Hill 1951) it becomes apparent that there is a canonical procedure which, from a Hamiltonian viewpoint, is also very natural. Just as the canonical momentum for a particle in general differs from its physical momentum, so does the canonical energy–momentum tensor for a subsystem differ from what we shall call its physical energy–momentum tensor. This distinction is different from that between the canonical and the symmetrized energy–momentum tensor (Pauli 1941) for the system as a whole. The canonical and physical split-up procedures could be applied to either the canonical or symmetrized tensor, although we will not find it useful to talk about the physical split-up of the canonical energy–momentum tensor.

The approach we adopt is a natural extension of earlier work (Dewar 1970) on hydromagnetic waves, in which the idea of canonical and physical split-up procedures was introduced, and Whitham's (1965) averaged Lagrangian principle was used to effect a general treatment of dispersive waves within the WKB approximation. Some of the techniques have also been used to discuss the analogy between electrostatic plasma waves and galactic density waves (Dewar 1972*a*), and to treat modulational instability of electrostatic plasma waves (Dewar 1972*b*). The concept of canonical background momentum in the presence of waves has also been approached from the point of view of canonical transformation theory (Dewar 1973, 1976).

In generalizing the previous work to fully electromagnetic waves the major obstacle has been that the standard treatments of field theory do not include an arbitrarily deformable background medium. It is essential to vary the background coordinate field in Hamilton's principle (taking into account such constraints as mass conservation) in order to obtain the correct ponderomotive force; and so a relativistic variational technique has been developed which includes the constraints explicitly, unlike Penfield and Haus (1967) who use Lagrange multipliers.

Although for most practical purposes a relativistic theory for the background is quite unnecessary, it is required for the electromagnetic field. It has been found that the most efficient technique with any claim to generality is first to do the calculations fully (special) relativistically, exploiting the compactness of 4-vector notation, and then to translate the results into 3-vector form, making nonrelativistic approximations as desired.

A few other authors (e.g. Toupin 1960; Schöpf 1964) have also used Hamilton's principle to treat the electrodynamics of continuous media, but have not treated dispersive waves. On the other hand, Furutsu (1969) has treated dispersive waves relativistically but has omitted to vary the background. Dougherty (1970, 1974) has reviewed the averaged Lagrangian method in the context of the cold plasma model, and has discussed two covariant methods for varying the background. The problem of waves in an arbitrary dielectric was not discussed. Jones (1971) has reviewed the use of Hamilton's principle for waves occurring in geophysics and has also discussed the use of classical field theory techniques.

Hamilton's principle is open to the objection that it requires one to postulate the form of the Lagrangian density, but it should be remembered that *any* macroscopic theory involves a number of postulates, and Hamilton's principle may be deeper than many of these. When one bears in mind that the macroscopic Lagrangian density must be an average of the microscopic density, which is known, and imposes Lorentz invariance, much of the arbitrariness goes out. We also know some of the Euler–Lagrange equations *a priori*, such as Maxwell's equations, and we find that we are unambiguously led to a definite form for the total Lagrangian density. Rules for forming Lagrangian densities are further discussed by Penfield and Haus (1967), who defend Hamilton's principle with the remark that the systematic bookkeeping and standardized set of rules for applying the variational principle allow one to derive equations of motion in a way that is likely to be free of errors.

A more serious defect of Hamilton's principle, when applied to a system of nonlinearly interacting fields, is that it cannot handle dissipation. For treatments which allow entropy flow between subsystems the reader is referred to the books by Penfield and Haus (1967) and de Groot and Suttorp (1972). There has also been considerable work on this subject from the standpoint of continuum mechanics (Lianis 1974).

In Section 2 below we adapt Noether's theorem and the general symmetrization procedure of Belinfante to continuous media and introduce the canonical and physical split-up procedures.

In Section 3 the theory is applied to isotropic dielectrics (as the simplest example), and the connection with the Abraham–Minkowski controversy is made. We also treat longitudinal and transverse waves in an isotropic dispersive medium through a polarization tensor approach, and derive the energy–momentum tensors (and hence ponderomotive forces) for these cases.

The connection with 3-vector formalism is made in Section 4 where we spell out the full 3-vector expressions. Although many of these terms disappear in the non-relativistic limit it is one of the advantages of our general approach that we can see just what is being omitted. The connection with the frequency- and wavenumber-dependent dielectric constant formulation of dispersive electromagnetic waves is also made.

In Section 5 we find the 'dressing' transformation of the physical electromagnetic energy–momentum tensor due to the excitation of background motion by a one-dimensional wavepacket. An evolution equation for a three-dimensional wavepacket is also found which takes into account self-focusing and stimulated Brillouin scattering.

Section 6 contains the adaptation of the previous formalism to the case of a hot collisionless plasma. Because we use the Vlasov description, we are still in a sense dealing with a continuum description, except that the plasma is now regarded as a fluid in phase space. In this case the meaning of the canonical energy–momentum tensor is rather clearer, as a Hamiltonian theory for single-particle motion can be developed; this we do in Appendix 2. Appendix 1 is devoted to a discussion of averaging, in order to clarify the meaning of 'background' in the presence of waves.

The SI system of electromagnetic units (equivalent to MKS) is used throughout this paper.

2. Relativistic Field Theory for Continua

(a) *Notation and Terminology*

We make the theory manifestly Lorentz covariant by working only with 4-scalars (as is the Lagrangian density $\mathscr{L}$), 4-vectors (such as the 4-position $x^\mu = (ct, \boldsymbol{x})$ and 4-gradient $\partial^\mu \equiv \partial/\partial x_\mu$) or 4-tensors. Greek indices run from 0 to 3, Roman from 1 to 3, the summation convention is assumed and the metric tensor $g^{\mu\nu}$ is such that $g^{00} = -g^{11} = -g^{22} = -g^{33} = 1$ (see e.g. Landau and Lifshitz (1971) for an introduction to this notation). The scalar products $a^\mu b_\mu = a_0 b_0 - \boldsymbol{a}.\boldsymbol{b}$, and $a^\mu a_\mu = a_0^2 - \boldsymbol{a}^2$ will be abbreviated $a.b$ and a^2 respectively, and the argument list x_0, x_1, x_2, x_3 will be abbreviated by x.

Although we are dealing here with a continuum description of some material (which can be either a solid or a fluid), it is convenient to use the term *particle* to denote an infinitesimal element of the material. Such a particle forms a *microsystem*, which we assume to be characterized completely by its position, velocity and strain tensor, the vector whose elements consist of these parameters, together with the time at which they were measured, representing a *microstate*. The set of all microstates whose time component equals t is the *state of the system* at time t. We implicitly assume that density, temperature, etc. are related by holonomic constraints to the microstate.

The path traced out in 4-space by a particle as t goes from $-\infty$ to $+\infty$ is its *world line*, and the unit 4-vector tangent to a world line at x is the local 4-velocity $u^{\mu}(x)$.

(*b*) *Reference States*

Unlike the case in quantum field theory (Pauli 1941), we cannot vary the field describing the material arbitrarily at all times. This is because there must be at least one time at which the state of the system (the *reference state*) is held fixed in order that variations in the strain tensor of the material may be defined relative to this reference state. However, it is unsatisfactory for a covariant theory to single out a special reference frame in which 'time' is to be defined. One could instead define a generalized reference state as the set of all the microstates of the system whose 4-position lies on a space-like hypersurface, but this is found to be inconvenient. This is because we wish to describe the system in 4-space, but the map from a general region of 4-space to this hypersurface is a projection, and therefore has no inverse function.

In order to retain the convenient feature of 3-space continuum mechanics that the map from the current state to the reference state is invertible, we introduce the concept of an expanded reference state, which allows microstates with a range of time values. For instance, we could designate the union of all the states of the system at times $-\infty < t_r \leqslant t_0$ as the expanded reference state. That is, all variations vanish prior to t_0. Since Hamilton's principle really only requires variations which can be localized around the current time t (although for a dispersive system we shall assume the variations to be slow with respect to the characteristic memory of the system), holding the system fixed over a range of times not containing t is perfectly compatible with Hamilton's principle.

To avoid specifying a special frame to define t_0 we introduce the concept of a *reference region* of 4-space, denoted by $\mathscr{R}_0$. The *expanded reference state associated with* $\mathscr{R}_0$ is the set of all the microstates whose 4-position lies within $\mathscr{R}_0$, and designating it as a reference state implies that *the allowable variations in Hamilton's principle vanish within* $\mathscr{R}_0$.

(*c*) *Mappings*

In 3-space continuum mechanics (Eringen 1967) the strain tensor is defined in terms of the map from the reference state at $t = 0$ to the state of the system at the current time t (i.e. the time in whose neighbourhood variations are to be taken), generated by the motion of the particles during the time interval $[0, t]$. Because the map is one to one we can equivalently use the inverse map from the current state to the reference state. This is more convenient because it allows 'Eulerian variations' (Dewar 1970) to be used. We shall call this mapping a *reference map*.

To obtain a covariant formulation we introduce a 'pseudotime' parameter τ and seek a continuous one-parameter family of *expanded* reference maps r_τ from the 4-space region $\mathscr{R}$, in which variations are to be taken, onto a family of expanded reference regions $\mathscr{R}_0(\tau)$ (disjoint from $\mathscr{R}$). If the point x is mapped on the point X we can write

$$X^{\mu} = X^{\mu}(x, \tau). \tag{1}$$

We require the map to be one to one and differentiable, and we require that *any point in* $\mathscr{R}$ *be connected to its image in* $\mathscr{R}_0(\tau)$ *by a world line.* Such a mapping is depicted in Fig. 1. We also require that $X^0(x,\tau)$ be a monotonically decreasing function of τ in all frames.

An example of such a map is that provided by sliding every point in $\mathscr{R}$ back in time along the appropriate world line through a distance τ measured along that world line, but there are infinitely many other possibilities. The lack of uniqueness of the map need not be a worry because we find that $X^\mu(x,\tau)$ can always be eliminated in favour of physically observable variables. For a general discussion, however, it is formally much more convenient to express all observables in terms of $X^\mu(x,\tau)$, because this field can be varied without constraint. The resulting Euler–Lagrange equations and conservation relations can always be re-expressed in terms of observables, and then become unique.

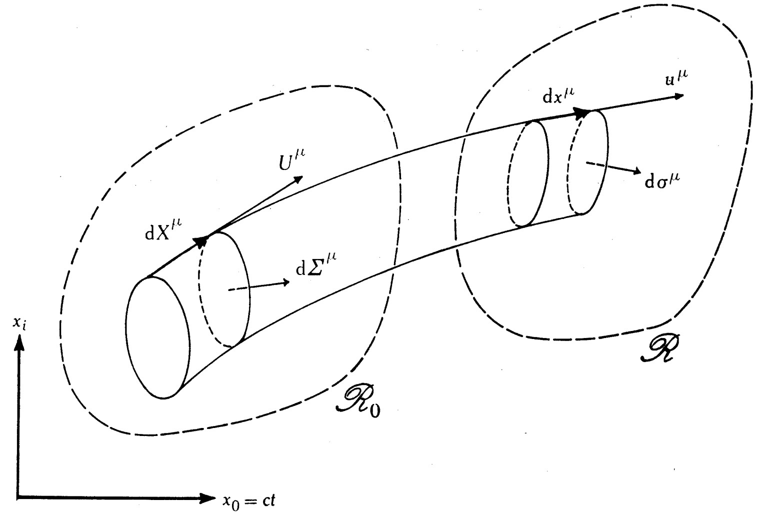

Fig. 1. Schematic diagram of the mapping between the reference world region $\mathscr{R}_0$ and the current region $\mathscr{R}$. The cylindrical volume elements used in Section 3*a* are also shown.

(*d*) *Hamilton's Principle*

For the time being we need only remark that the 4-velocity, proper density and strain tensor can all be expressed as functions of X^μ, $\partial^\mu X^\nu$ and u^μ. Thus the most general Lagrangian density we can encounter is

$$\mathscr{L} = \mathscr{L}(X, \partial X, u, \eta, \partial\eta)\,, \tag{2}$$

where the $\eta^i(x)$ are the other fields entering into the problem. Hamilton's principle is stated in the form

$$\int_{\mathscr{R}} \delta\mathscr{L}\, d^4x = 0\,, \tag{3}$$

where $X^\mu(x)$ and $\eta_i(x)$ are to be varied with τ held constant and the world lines in $\mathscr{R}_0$ held fixed, X^μ changing *as a function of x* because the world lines within $\mathscr{R}$ are varied. Changes in functional dependence on x will be denoted by the Eulerian variation symbol δ, while changes evaluated at the varying 4-position determined by X^μ = const., τ = const. will be denoted by the Lagrangian variation symbol Δ, the relation between the two types of variation being

$$\Delta = \delta + \Delta X . \partial , \tag{4}$$

so

$$\Delta X^{\mu} = 0, \qquad \delta x^{\mu} = 0 \qquad \text{and} \qquad \delta X^{\mu} = -\Delta x . \partial X^{\mu} . \tag{5}$$

(*e*) *Eulerian Variations*

The variation in $\partial^{\mu} X^{\nu}$ is the obvious result

$$\delta\, \partial^{\mu} X^{\nu} = \partial^{\mu}\, \delta X^{\nu} = -\partial^{\mu}(\Delta x . \partial X^{\nu}) . \tag{6}$$

To find the variation in u^{μ}, we first note that a world line is traced out by x^{μ} as τ is varied with X^{μ} fixed. Thus

$$u^{\mu} = -X_{\tau}{}^{\nu}(\partial X^{-1})_{\nu}{}^{\mu}\{(X_{\tau} . \partial X^{-1})^{2}\}^{-\frac{1}{2}} , \tag{7}$$

where $(\partial X^{-1})_{\mu}{}^{\nu}$ is the inverse of the matrix $\partial_{\mu} X^{\nu}$ and the subscript τ denotes partial differentiation with respect to τ. Using the facts that $\Delta X_{\tau}{}^{\mu} = 0$ and

$$\Delta(\partial X^{-1})_{\nu}{}^{\mu} = (\partial X^{-1} . \partial\, \Delta x)_{\nu}{}^{\mu} , \tag{8}$$

we can calculate Δu^{μ} and hence find

$$\delta u^{\mu} = (g^{\mu\nu} - u^{\mu} u^{\nu}) u . \partial\, \Delta x_{\nu} - \Delta x . \partial u^{\mu} . \tag{9}$$

(*f*) *Euler–Lagrange Equations*

Substitution of equations (6) and (9) into the variational principle (3) yields the Lagrange equations of motion for the background material

$$\partial^{\mu} X^{\nu} \partial^{\rho}\left(\frac{\partial \mathscr{L}}{\partial\, \partial^{\rho} X^{\nu}}\right) - \partial^{\mu} X^{\nu} \frac{\partial \mathscr{L}}{\partial X^{\nu}} - \partial_{\rho}\left(u^{\rho}(g^{\mu\nu} - u^{\mu} u^{\nu}) \frac{\partial \mathscr{L}}{\partial u^{\nu}}\right) - \partial^{\mu} u^{\nu} \frac{\partial \mathscr{L}}{\partial u^{\nu}} = 0 , \tag{10}$$

which can also be written as a canonical energy–momentum balance equation for the background material in the form

$$\partial_{\mu} T_{\mathrm{b}}{}^{\mu\nu} = f_{\mathrm{b}}{}^{\nu} ,$$

where

$$T_{\mathrm{b}}{}^{\mu\nu} \equiv u . \partial X_{\sigma} \frac{\partial \mathscr{L}}{\partial\, \partial_{\mu} X_{\sigma}} u^{\nu} + \left(\partial_{\rho} X_{\sigma} \frac{\partial \mathscr{L}}{\partial\, \partial_{\mu} X_{\sigma}} - u^{\mu} \frac{\partial \mathscr{L}}{\partial u^{\rho}}\right)\left(g^{\rho\nu} - u^{\rho} u^{\nu}\right) - \mathscr{L}_{\mathrm{b}}\, g^{\mu\nu} \tag{11}$$

is the canonical energy–momentum tensor for the background subsystem and

$$f_{\mathrm{b}}{}^{\nu} = \partial_{\mathrm{b}}{}^{\nu} \mathscr{L} - \partial^{\nu} \mathscr{L}_{\mathrm{b}} , \tag{12}$$

is the canonical force density acting on the background subsystem, the symbol ∂_{b} denoting the total derivative with respect to the background variables,

$$\partial_{\mathrm{b}}{}^{\nu} \equiv (\partial^{\nu} X_{\rho}) \frac{\partial}{\partial X_{\rho}} + \partial^{\nu}(\partial_{\rho} X_{\sigma}) \frac{\partial}{\partial\, \partial_{\rho} X_{\sigma}} + (\partial^{\nu} u_{\rho}) \frac{\partial}{\partial u_{\rho}} , \tag{13}$$

and the symbol $\mathscr{L}_{\mathrm{b}}$ denoting that part of $\mathscr{L}$ depending on the background variables alone.

Variation of the fields η_i yields the well-known Euler–Lagrange equations

$$\partial_\mu(\partial\mathscr{L}/\partial\,\partial_\mu\eta_i) - \partial\mathscr{L}/\partial\eta_i = 0. \tag{14}$$

(g) Canonical Energy–Momentum Tensors for Fields

If we associate the η_i in the subspace $i_k \leqslant i < i_{k+1}$ with the kth subsystem, and also associate part of $\mathscr{L}$ with the kth subsystem so that

$$\mathscr{L} = \mathscr{L}_\mathrm{b} + \sum_k \mathscr{L}_k,$$

then we can define the canonical balance equation for the kth subsystem to be

$$\partial_\mu T_k^{\mu\nu} = f_k^{\nu}, \tag{15}$$

where the canonical energy–momentum tensor for the kth subsystem is defined to be

$$T_k^{\mu\nu} = \sum_{i=i_k}^{i_{k+1}-1} \frac{\partial\mathscr{L}}{\partial\,\partial_\mu\eta_i}\partial^\nu\eta_i - \mathscr{L}_k g^{\mu\nu}. \tag{16}$$

Using equation (14), we find the canonical force density to be

$$f_k^{\nu} \equiv \partial_k^{\nu}\mathscr{L} - \partial^\nu\mathscr{L}_k. \tag{17}$$

The symbol ∂_k^{ν} is defined analogously to $\partial_\mathrm{b}^{\nu}$ as

$$\partial_k^{\nu} = \sum_{i=i_k}^{i_{k+1}-1}\left((\partial^\nu\eta_i)\frac{\partial}{\partial\eta_i} + \partial^\nu(\partial^\mu\eta_i)\frac{\partial}{\partial\,\partial^\mu\eta_i}\right). \tag{18}$$

Since the canonical equations for the background subsystem were given in the previous subsection, equation (18) completes the definition of the canonical energy–momentum tensors of all subsystems.

If $\mathscr{L}_k$ depends only on the fields associated with the kth subsystem, and has no explicit dependence on x, then the force density f_k acting on the kth subsystem vanishes and we call this a closed subsystem, i.e. one which has no interaction with any other subsystem. In the more typical and interesting case of interacting subsystems, the $\mathscr{L}_k$ will depend on fields associated with other subsystems and in fact there will be no unique way of defining $\mathscr{L}_k$, although there is usually a most 'natural' way of splitting up $\mathscr{L}$ into contributions from different subsystems.

(h) Translation Invariance

We now consider the conservation equations which the system as a whole must obey. We know from Noether's theorem (Hill 1951) that these are associated with the invariance of the equations of motion under symmetry transformations. In fact, in both classical and quantum field theories, $\mathscr{L}$ itself is form invariant under time and space translations and Lorentz transformations. In our problem we must recognize the fact that a constraint has been applied on allowable variations, namely the requirement that $\mathscr{R}_0$ be held fixed. Thus our first symmetry postulate is that $\mathscr{L}$ (and in fact $\mathscr{L}_\mathrm{b}$ and $\mathscr{L}_k$) is invariant under space–time translations of $\mathscr{R}$ and the world lines and fields within $\mathscr{R}$. That is,

$$\Delta\mathscr{L} = 0 \tag{19}$$

when

$$\Delta x^{\mu} = \varepsilon^{\mu} \qquad \text{and} \qquad \Delta\eta_i = 0\,.$$

The condition that equation (19) hold for all ε^{μ} is

$$\partial^{\mu}\mathscr{L} - \partial_{\mathrm{b}}{}^{\mu}\mathscr{L} - \sum_k \partial_k{}^{\mu}\mathscr{L} = 0\,. \tag{20}$$

From equations (12), (15) and (17) we see that (20) implies that the total 4-force density acting on the system is zero, thus implying the conservation of total canonical energy and momentum

$$\partial_{\mu} T^{\mu\nu} = 0\,, \tag{21}$$

where the total canonical energy–momentum tensor is defined by

$$T^{\mu\nu} \equiv T_{\mathrm{b}}{}^{\mu\nu} + \sum_k T_k{}^{\mu\nu}\,. \tag{22}$$

(i) Local Lorentz Invariance

Our second symmetry postulate is that $\mathscr{L}$ is invariant under 'rigid rotation' (in 4-space) of $\mathscr{R}$ about the origin, the world lines and fields within $\mathscr{R}$ also being 'rotated'. This operation is a Lorentz transformation of $\mathscr{R}$, but $\mathscr{R}_0$ is, as always, held fixed. We shall call this invariance *local Lorentz invariance*, this being a stronger assumption than the global Lorentz invariance implicit in the 4-vector formulation. We further assume that $\mathscr{L}_{\mathrm{b}}$ and $\mathscr{L}_k$ are locally Lorentz invariant and translation invariant. Stated succinctly, we require

$$\Delta\mathscr{L} = 0 \tag{23}$$

when

$$\Delta x^{\mu} = \varepsilon^{\mu\nu} x_{\nu}\,, \qquad \Delta\eta_i = \tfrac{1}{2}\varepsilon^{\mu\nu} I_{\mu\nu ij}\eta_j\,, \tag{24}$$

where $\varepsilon^{\mu\nu}$ is an arbitrary antisymmetric infinitesimal 4-tensor, the matrices $I_{\mu\nu ij}$ being representations of the infinitesimal operators of the Lorentz group (Pauli 1941).

Without loss of generality, $I_{\mu\nu ij}$ can be assumed antisymmetric in μ and ν. The condition that equation (23) be satisfied for all $\varepsilon_{\mu\nu}$ is, on using (20),

$$\left(\partial^{\nu}\eta_i \frac{\partial\mathscr{L}}{\partial\,\partial_{\mu}\eta_i} + \partial^{\nu}X_{\rho}\frac{\partial\mathscr{L}}{\partial\,\partial_{\mu}X_{\rho}} - u^{\mu}\frac{\partial\mathscr{L}}{\partial u_{\nu}}\right)_{\mathrm{a.s.}} + \tfrac{1}{2}I_{ij}^{\mu\nu}\eta_j\frac{\partial\mathscr{L}}{\partial\eta_i} + \tfrac{1}{2}\partial_{\rho}(I_{ij}^{\mu\nu}\eta_j)\frac{\partial\mathscr{L}}{\partial\,\partial_{\rho}\eta_i} = 0\,, \tag{25}$$

where the subscript a.s. denotes the antisymmetric part of a tensor. That is, if $t^{\mu\nu}$ is an arbitrary tensor,

$$t^{\mu\nu}_{\mathrm{a.s.}} \equiv \tfrac{1}{2}(t^{\mu\nu} - t^{\nu\mu})\,.$$

Equation (25) is to hold for $\mathscr{L}_{\mathrm{b}}$ and $\mathscr{L}_k$ as well, and is a restriction on allowable constitutive relations. This aspect will be discussed further in Section 3*a*. In this section we demonstrate that equation (25) has the consequences (i) that angular momentum is conserved and (ii) that the energy–momentum tensor is symmetrizable.

The first consequence (i) follows directly from Noether's theorem (Hill 1951) which yields

$$\partial_\rho M^{\rho\mu\nu} = 0, \tag{26}$$

where

$$M^{\rho\mu\nu} \equiv -x^\mu T^{\rho\nu} + x^\nu T^{\rho\mu} - (\partial\mathscr{L}/\partial\,\partial_\rho\eta_i)I_{ij}^{\mu\nu}\eta_j. \tag{27}$$

It is easily verified that equation (26) follows from equations (11), (16), (21), (22), (25) and (27). Jones (1971) has interpreted the last term in equation (27) as the 'spin' of the fields η_i, but for the purposes of the present paper we limit ourselves to interpreting the energy–momentum tensor.

(j) Symmetrization

The second consequence (ii) above of equation (25) follows directly from application of the method of Belinfante (see Pauli 1941). We define a modified energy–momentum tensor

$$\theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho f^{\rho\mu\nu}, \tag{28}$$

where

$$f^{\rho\mu\nu} \equiv -\tfrac{1}{2}\left(\frac{\partial\mathscr{L}}{\partial\,\partial_\mu\eta_i}I_{ij}^{\rho\nu}\eta_j - \frac{\partial\mathscr{L}}{\partial\,\partial_\rho\eta_i}I_{ij}^{\mu\nu}\eta_j - \frac{\partial\mathscr{L}}{\partial\,\partial_\nu\eta_i}I_{ij}^{\mu\rho}\eta_j\right). \tag{29}$$

In view of the antisymmetry of $f^{\rho\mu\nu}$ with respect to ρ and μ, $\theta^{\mu\nu}$ obeys the same conservation equation as $T^{\mu\nu}$, namely

$$\partial_\mu\theta^{\mu\nu} = 0. \tag{30}$$

After some algebra it can be shown that the modified angular momentum tensor

$$m^{\rho\mu\nu} = -x^\mu\theta^{\rho\nu} + x^\nu\theta^{\rho\mu} \tag{31}$$

obeys the conservation equation

$$\partial_\rho m^{\rho\mu\nu} = 0, \tag{32}$$

whence it follows that $\theta^{\mu\nu}$ is a symmetric tensor. It can be shown that $\theta^{\mu\nu}$ is uniquely determined by requiring symmetry (Pauli 1941).

(k) Physical Split-up

As with the total canonical energy–momentum $T^{\mu\nu}$, we can split $\theta^{\mu\nu}$ into contributions from the various subsystems. There seem to be two natural conventions for effecting this split-up. The first we call, following Dewar (1970), the *physical split-up*. We define the physical energy–momentum tensor for the kth subsystem by

$$\theta_k^{\mu\nu} \equiv u.\partial X_\sigma\frac{\partial\mathscr{L}_k}{\partial\,\partial_\mu X_\sigma}u^\nu + \left(\partial_\rho X_\sigma\frac{\partial\mathscr{L}_k}{\partial\,\partial_\mu X_\sigma} - u^\mu\frac{\partial\mathscr{L}_k}{\partial u^\rho}\right)\left(g^{\rho\nu} - u^\rho u^\nu\right) + \frac{\mathscr{L}_k}{\partial\,\partial_\mu\eta_i}\partial^\nu\eta_i$$
$$-\mathscr{L}_k g^{\mu\nu} + \partial_\rho f_k^{\rho\mu\nu}, \tag{33}$$

where

$$f_k^{\rho\mu\nu} \equiv -\tfrac{1}{2}\left(\frac{\partial\mathscr{L}_k}{\partial\,\partial_\mu\eta_i}I_{ij}^{\rho\nu}\eta_j - \frac{\partial\mathscr{L}_k}{\partial\,\partial_\rho\eta_i}I_{ij}^{\mu\nu}\eta_j - \frac{\partial\mathscr{L}_k}{\partial\,\partial_\nu\eta_i}I_{ij}^{\mu\rho}\eta_j\right), \tag{34}$$

and similarly for $\theta_b^{\mu\nu}$ and $f_b^{\rho\mu\nu}$, simply by replacing the subscript k with b. The subsystems exchange energy and momentum according to the equations

$$\partial_\mu \theta_b^{\mu\nu} = \phi_b^{\ \nu}, \qquad \partial_\mu \theta_k^{\mu\nu} = \phi_k^{\ \nu}, \tag{35}$$

with the physical force density $\phi_k^{\ \nu}$ acting on the kth subsystem given by

$$\begin{aligned}\phi_k^{\ \nu} \equiv{}& \partial^\nu X_\rho \partial_\mu\left(\frac{\partial \mathscr{L}_k}{\partial \partial_\mu X_\rho}\right) - \partial^\nu X_\rho \frac{\partial \mathscr{L}_k}{\partial X_\rho} - \partial_\mu\left(u^\mu(g^{\nu\rho}-u^\nu u^\rho)\frac{\partial \mathscr{L}_k}{\partial u^\rho}\right) - \partial^\nu u_\rho \frac{\partial \mathscr{L}_k}{\partial u_\rho}\\ &+\partial^\nu \eta_i \partial_\mu\left(\frac{\partial \mathscr{L}_k}{\partial \partial_\mu \eta_i}\right) - \partial^\nu \eta_i \frac{\partial \mathscr{L}_k}{\partial \eta_i},\end{aligned} \tag{36}$$

where we have used equation (20). From equations (10) and (14) it is readily verified that the physical forces acting on the system as a whole sum to zero.

We term $\phi_b^{\ \nu}$ the *ponderomotive 4-force density,* as it gives the force acting on the average motion of the material. Note that $\theta_b^{\mu\nu}$ and $\theta^{\mu\nu}$ will always be symmetric, but that $\theta_k^{\mu\nu}$ need not necessarily be so if there are several interacting subsystems. The advantage of the physical breakup is that the background energy–momentum tensor has a very natural form since it is unmodified by the existence of other fields—all interactions are contained in the ponderomotive force. On the other hand, the *canonical* forms for the other subsystems seem more 'natural', except for the absence of the symmetrizing term.

(*l*) *Modified Canonical Split-up*

The above reasoning leads us to introduce a second way of splitting up the symmetric energy–momentum tensor $\theta^{\mu\nu}$, which we call the *modified canonical split-up.* Suppose the matrices $I_{ij}^{\mu\nu}$ have a block diagonal form corresponding to the fact that the fields η_i in the kth subspace $i_k \leqslant i < i_{k+1}$ transform only amongst themselves. (That is, the subspace k corresponds to one or more irreducible representations of the Lorentz group.) Then we can decompose $f^{\rho\mu\nu}$ into a sum of tensors $g_k^{\rho\mu\nu}$ defined by

$$g_k^{\rho\mu\nu} \equiv -\tfrac{1}{2}\sum_{i=i_k}^{i_{k+1}-1}\sum_{j=i_k}^{i_{k+1}-1}\left(\frac{\partial \mathscr{L}}{\partial \partial_\mu \eta_i} I_{ij}^{\rho\nu}\eta_j - \frac{\partial \mathscr{L}}{\partial \partial_\rho \eta_i} I_{ij}^{\mu\nu}\eta_j - \frac{\partial \mathscr{L}}{\partial \partial_\nu \eta_i} I_{ij}^{\mu\rho}\eta_j\right). \tag{37}$$

This decomposition is distinct from that defined by equations (24).

As with the strict canonical split-up we associate the fields in the kth subspace with the kth subsystem. Thus we define the modified canonical energy–momentum tensor for the kth subsystem by

$$S_k^{\mu\nu} \equiv \sum_{i=i_k}^{i_{k+1}-1}\frac{\partial \mathscr{L}}{\partial \partial_\mu \eta_i}\partial^\nu \eta_i - \mathscr{L}_k g^{\mu\nu} + \partial_\rho g_k^{\rho\mu\nu}. \tag{38}$$

We define the modified canonical energy–momentum tensor for the background to be the same as the canonical energy–momentum tensor:

$$S_b^{\mu\nu} \equiv T_b^{\mu\nu} = u.\partial X_\sigma \frac{\partial \mathscr{L}}{\partial \partial_\mu X_\sigma}u^\nu + \left(\partial_\rho X_\sigma \frac{\partial \mathscr{L}}{\partial \partial_\mu X_\sigma} - u^\mu \frac{\partial \mathscr{L}}{\partial u^\rho}\right)\left(g^{\rho\nu} - u^\rho u^\nu\right) - \mathscr{L}_b g^{\mu\nu}. \tag{39}$$

Although none of the subsystem tensors are in general symmetric, they sum to the symmetric tensor

$$S_{\mathrm{b}}{}^{\mu\nu} + \sum_k S_k{}^{\mu\nu} = \theta^{\mu\nu}. \tag{40}$$

Since $S_k{}^{\mu\nu}$ differs from $T_k{}^{\mu\nu}$ by a 4-divergenceless term, the balance equation

$$\partial_\mu S_k{}^{\mu\nu} = f_k{}^\nu \tag{41}$$

applies, with $f_k{}^\nu$ given by equation (17).

3. Constitutive Relations and their Consequences

(a) Dependence on Background Variables

All constitutive equations must obey the two symmetry postulates expressed by equations (20) and (25). The first is trivially satisfied simply by demanding that $\mathscr{L}$ have no explicit dependence on x^μ. The second, however, restricts the allowable dependence of $\mathscr{L}$ on the deformation 4-tensor $\partial^\mu X^\nu$, because X^ν is a scalar, not a vector, under the local Lorentz transformation (24) (i.e. its components are invariant). Also,

$$\Delta\,\partial^\mu X^\nu = \varepsilon^{\mu\sigma}\,\partial_\sigma X^\nu, \tag{42}$$

so that $\partial^\mu X^\nu$ is a vector rather than a tensor under local Lorentz transformation, and similarly for higher derivatives. Furthermore, a vector $a^\mu(X)$ depending only on the initial state is also a scalar under local Lorentz transformations. Thus the only way the deformation 4-'tensor' can appear in $\mathscr{L}$ is through the combinations

$$\partial^\mu X_\sigma\, a^\sigma(X) \qquad \text{and} \qquad \partial^\mu X_\sigma\, \partial_\nu X^\sigma. \tag{43}$$

We shall not enumerate all the ways a scalar $\mathscr{L}$ can be formed from these elements, but shall instead consider as a simple example the construction of the background Lagrangian density for an isotropic fluid.

Since $\mathscr{L}_{\mathrm{b}}$ is to be a scalar, it suffices to evaluate it in the local rest frame of the medium. In this frame the kinetic energy vanishes and $\mathscr{L}_{\mathrm{b}}$ is just the negative of the total internal energy density, including the rest energy, as

$$\mathscr{L}_{\mathrm{b}} = -\rho'(x)\,c^2 - \mathscr{E}'(\rho'), \tag{44}$$

where to be consistent with the assumption of scalar pressure we have assumed $\mathscr{E}'$, the proper internal energy density with rest energy subtracted, to be a function only of ρ', the proper density. By 'proper' we mean evaluated in the local rest frame, and this we indicate with a prime. We evaluate $\rho'(x)$ by a geometric argument similar to that used for flux conservation in hydromagnetics (Newcomb 1962). Suppose $\mathrm{d}\sigma_\mu$ is an element of area on a spacelike hypersurface cutting $\mathscr{R}$ and $\mathrm{d}\Sigma_\mu$ is its image under the mapping equation (1). Then mass conservation requires

$$\rho'(x)\,u.\mathrm{d}\sigma = \rho'(X)\,U.\mathrm{d}\Sigma, \tag{45}$$

where U^μ, the 4-velocity at X^μ, is given by

$$U^\mu = u.\partial X^\mu\{(u.\partial X)^2\}^{-\frac{1}{2}}. \tag{46}$$

A cylindrical volume element of side $\mathrm{d}X^\mu$ maps into a cylindrical volume element of side $\mathrm{d}x^\mu$, the ratio of the volumes $\mathrm{d}X.\mathrm{d}\Sigma/\mathrm{d}x.\mathrm{d}\sigma$ being the Jacobian, $\det(\partial_\alpha X^\beta)$ (see Fig. 1). Since this holds for all $\mathrm{d}x^\mu$, we have

$$\mathrm{d}\sigma_\mu = \partial_\mu X^\nu \,\mathrm{d}\Sigma_\nu/\det(\partial_\alpha X^\beta). \tag{47}$$

Substituting equation (47) into (45) we find

$$\rho'(x) = \det(\partial_\alpha X^\beta)\,\{(u.\partial X)^2\}^{-\frac{1}{2}}\rho'(X). \tag{48}$$

If $\mathscr{L}$ depends on $\partial^\mu X^\nu$ only through ρ' (as in an isotropic fluid) then we can use equation (48) to simplify some of the equations of the previous section, since

$$\frac{\partial\rho'}{\partial\,\partial_\mu X_\rho}\partial^\nu X_\rho - u^\mu\frac{\partial\rho'}{\partial u_\nu} = \rho' g^{\mu\nu}, \qquad u.\partial X_\sigma\frac{\partial\rho'}{\partial\,\partial_\mu X_\sigma} = 0. \tag{49}$$

For example, the physical energy–momentum tensor corresponding to equation (44) is from (33)

$$\theta_{\mathrm{b}}{}^{\mu\nu} = (\rho' c^2 + \mathscr{E}')u^\mu u^\nu - P'(g^{\mu\nu} - u^\mu u^\nu), \tag{50}$$

where the proper pressure P' is defined by

$$P' \equiv \rho'\,\partial\mathscr{E}'/\partial\rho' - \mathscr{E}'. \tag{51}$$

Equation (50) is the expected form for the energy–momentum tensor of a continuous system (Landau and Lifshitz 1971).

(*b*) *Low Frequency Electromagnetic Field*

The slowly varying parts of the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are contained in the antisymmetric tensor

$$B^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu, \tag{52}$$

where A^μ is the 4-vector potential. In the absence of dispersion, $\mathscr{L}$ depends on A^μ through A^μ and $B^{\mu\nu}$ only, the Euler–Lagrange equations (14) resulting from taking $\eta_i = A^\mu$ being the covariant Maxwell equation

$$\partial_\mu H^{\mu\nu} = J^\nu, \tag{53}$$

where

$$H^{\mu\nu} \equiv -2\,\partial\mathscr{L}/\partial B_{\mu\nu}, \qquad J^\nu \equiv -\partial\mathscr{L}/\partial A_\nu, \tag{54}$$

J^ν being the 4-current carried by free charges. The canonical energy–momentum density (16) and force density (17) are given by

$$T^{\mu\nu}_{\mathrm{em}} = -H^{\mu\sigma}\partial^\nu A_\sigma - \mathscr{L}_{\mathrm{em}}g^{\mu\nu}, \qquad f^{\nu}_{\mathrm{em}} = -\tfrac{1}{2}H_{\rho\sigma}\partial^\nu B^{\rho\sigma} - \partial^\nu\mathscr{L}_{\mathrm{em}}. \tag{55}$$

Provided we work in Lorentz gauge, or some other relativistically invariant gauge, the 4-potential A^μ is a 4-vector. Thus the infinitesimal operators for the electromagnetic subspace are represented by

$$I_{\mu\nu\rho\sigma} = g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma}. \tag{56}$$

From equation (25) the constraint of local Lorentz invariance requires that

$$\left(H^{\mu\sigma}_{\text{em}}B_{\sigma}{}^{\nu}+\partial^{\nu}X_{\sigma}\frac{\partial\mathscr{L}_{\text{em}}}{\partial\,\partial_{\mu}X_{\sigma}}-u^{\mu}\frac{\partial\mathscr{L}_{\text{em}}}{\partial u_{\nu}}-J^{\mu}_{\text{em}}A^{\nu}\right)_{\text{a.s.}}=0\,, \tag{57}$$

where we have taken $\mathscr{L}_{\text{em}}$ to be that part of $\mathscr{L}$ depending only on A^{μ} and the background variables, and $H^{\mu\nu}_{\text{em}}(J^{\mu}_{\text{em}})$ to be that part of $H^{\mu\nu}(J^{\mu})$ contributed by $\mathscr{L}_{\text{em}}$.

Using equations (34), (54) and (56) we find

$$f^{\rho\mu\nu}_{\text{em}}=H^{\mu\rho}_{\text{em}}A^{\nu}. \tag{58}$$

Thus from equation (33) the physical energy–momentum tensor for the electromagnetic subsystem is

$$\theta^{\mu\nu}_{\text{em}}=u\,.\,\partial X_{\sigma}\frac{\partial\mathscr{L}_{\text{em}}}{\partial\,\partial_{\mu}X_{\sigma}}u^{\nu}+\left(\partial_{\rho}X_{\sigma}\frac{\partial\mathscr{L}_{\text{em}}}{\partial\,\partial_{\mu}X_{\sigma}}-u^{\mu}\frac{\partial\mathscr{L}_{\text{em}}}{\partial u^{\rho}}\right)\left(g^{\rho\nu}-u^{\rho}u^{\nu}\right)$$

$$+H^{\mu\sigma}_{\text{em}}B_{\sigma}{}^{\nu}-\mathscr{L}_{\text{em}}g^{\mu\nu}-(\partial_{\rho}H^{\rho\mu}_{\text{em}})A^{\nu}. \tag{59}$$

From equation (57) it is easily seen that $\theta^{\mu\nu}_{\text{em}}$ is a symmetric tensor when there is no free current. Because of this it is tempting to identify $\theta^{\mu\nu}_{\text{em}}$ as the general form of the *Abraham* electromagnetic energy–momentum tensor. As we shall see, however, this is not quite correct. Nevertheless, the modified canonical energy–momentum tensor, defined by equation (38), is given by

$$S^{\mu\nu}_{\text{em}}\equiv H^{\mu\rho}B_{\rho}{}^{\nu}-\mathscr{L}_{\text{em}}g^{\mu\nu}, \tag{60}$$

and we shall now show that this is identical with the *Minkowski* electromagnetic energy–momentum tensor.

(*c*) *Linear Isotropic Case*

Piezoelectric effects are represented by a term in $\mathscr{L}_{\text{em}}$ linear in $B^{\mu\nu}$ while the linear dielectric, magnetic and magnetoelectric response is represented by a quadratic term and the nonlinear response is represented by higher order terms. We shall consider only a linear uncharged, insulating isotropic fluid, for which (see Section 4*c* below)

$$\mathscr{L}_{\text{em}}=\tfrac{1}{2}\mu_0^{-1}\{((\mu')^{-1}-\varepsilon')B^{\rho}B_{\rho}+\tfrac{1}{2}(\mu')^{-1}B^{\rho}{}_{\sigma}B^{\sigma}{}_{\rho}\}, \tag{61}$$

where $\varepsilon'(\rho')$ and $\mu'(\rho')$ are the proper dielectric permittivity and proper magnetic permeability respectively, relative to the vacuum values ε_0 and μ_0, and

$$B^{\mu}\equiv B^{\mu}{}_{\nu}u^{\nu}. \tag{62}$$

Consider the case where only the background and low frequency electromagnetic subsystems are present, i.e.

$$\mathscr{L}=\mathscr{L}_{\text{b}}+\mathscr{L}_{\text{em}}. \tag{63}$$

Then the canonical background energy–momentum tensor is, from equations (11) and (49),

$$T_b{}^{\mu\nu} = (\rho' c^2 + \mathscr{E}')u^\mu u^\nu$$
$$-[P' - \tfrac{1}{2}\mu_0^{-1}\{((\mu')^{-1} - \varepsilon')_{\ln\rho'} B^\rho B_\rho - \tfrac{1}{2}((\mu')^{-1})_{\ln\rho'} B^\rho{}_\sigma B^\sigma{}_\rho\}](g^{\mu\nu} - u^\mu u^\nu)$$
$$+\tfrac{1}{2}\mu_0^{-1}(\varepsilon' - (\mu')^{-1})u^\mu B_\rho B^\rho{}_\tau (g^{\tau\nu} - u^\tau u^\nu), \tag{64}$$

where the subscript $\ln\rho'$ denotes $\rho'\,\partial/\partial\rho'$. From equations (54) we find

$$H^{\mu\nu} = \mu_0^{-1}\{((\mu')^{-1} - \varepsilon')(u^\mu B^\nu - u^\nu B^\mu) + (\mu')^{-1}B^{\mu\nu}\}, \tag{65}$$

where $\mathscr{L}_{\mathrm{em}}$ can be written

$$\mathscr{L}_{\mathrm{em}} = \tfrac{1}{4}B^\rho{}_\sigma H^\sigma{}_\rho. \tag{66}$$

The modified canonical electromagnetic energy–momentum tensor is, by equations (60) and (66),

$$S_{\mathrm{em}}^{\mu\nu} = H^\mu{}_\rho B^{\rho\nu} - \tfrac{1}{4}B^\rho{}_\sigma H^\sigma{}_\rho g^{\mu\nu}. \tag{67}$$

Comparison with equation (301) of Pauli (1941) confirms that the present equation (67) is indeed (to within a sign convention) the Minkowski form of the electromagnetic energy–momentum tensor. The interaction 4-force density acting on the canonical background subsystem is most easily obtained from the conservation equation (21),

$$f_b{}^\nu = -f_{\mathrm{em}}^\nu = -\partial_\mu S_{\mathrm{em}}^{\mu\nu}. \tag{68}$$

The two subsystems are clearly coupled by any inhomogeneity in the background, thus illustrating the futility of discussing the 'true' form of the electromagnetic energy–momentum tensor in isolation from the background. Even worse, the canonical energy–momentum tensor for the *background*, equation (64), contains terms quadratic in the electromagnetic field. At first sight this appears unphysical (hence the designation 'physical' for $\theta_b{}^{\mu\nu}$), especially as it leads to an asymmetric tensor, but it is really no more unphysical than the fact that a term qA appears in the canonical momentum of a particle in an electromagnetic field. One can carry this analogy further using 'oscillation centre' canonical transformation theory (Dewar 1973, 1976).

The physical energy–momentum tensor for the background, $\theta_b{}^{\mu\nu}$, is unchanged, and given by equation (50). From equations (49), (59), (65) and (66) we find the physical energy–momentum tensor for the electromagnetic subsystem to be

$$\theta_{\mathrm{em}}^{\mu\nu} = \theta_A{}^{\mu\nu} + \tfrac{1}{4}\rho'(\partial H^\rho{}_\sigma/\partial\rho')B^\sigma{}_\rho(g^{\mu\nu} - u^\mu u^\nu), \tag{69}$$

where $\theta_A{}^{\mu\nu}$ is the Abraham energy–momentum tensor (equation (303) of Pauli 1958) for an isotropic medium, given by

$$\theta_A{}^{\mu\nu} \equiv H^\mu{}_\rho B^{\rho\nu} - \tfrac{1}{4}H^\rho{}_\sigma B^\sigma{}_\rho g^{\mu\nu} - (\varepsilon'\mu' - 1)u^\mu \Omega^\nu, \tag{70}$$

Ω^ν being the 'Ruhstrahlvektor'

$$\Omega^\nu \equiv (u^\nu H^{\sigma\tau} + u^\sigma H^{\tau\nu} + u^\tau H^{\nu\sigma})B_\sigma u_\tau. \tag{71}$$

Since $\theta_{\rm em}^{\mu\nu}$ is symmetric in the rest frame, it is symmetric in all frames. The ponderomotive force is easily obtained from the conservation equation (30) as

$$\phi_{\rm b}{}^{\nu} = -\phi_{\rm em}^{\nu} = -\partial_{\mu}\theta_{\rm em}^{\mu\nu}. \tag{72}$$

Thus our resolution of the famous controversy is as follows: The Minkowski form is correct provided the canonical energy–momentum tensor is used for the background subsystem; the Abraham form is not quite correct when the physical energy–momentum tensor for the background subsystem is used, but may be corrected by the addition of a tensor which accounts for electrostrictive and magnetostrictive effects. It will be seen in Section 4*c* below that this correction term corresponds to the Helmholtz form of the ponderomotive force (Robinson 1975).

(*d*) *High Frequency Electromagnetic Field*

Consider the high frequency electromagnetic field to be due to the passage of a nearly monochromatic wave train, described by

$$A_{\rm hf}^{\mu} = \sum_{n=-\infty}^{\infty}{}' \, a_n^{\mu}\exp(in\theta), \tag{73}$$

where $a_n^{\mu}(x)$ is the slowly varying complex amplitude of the nth harmonic of the wave 4-potential and $\theta(x)$ is the phase of the wave. At this stage we make no assumption regarding the linearity of the response, so harmonics will in general be present. However, we assume that the amplitudes of the higher harmonics can be expressed in terms of that of the fundamental $a_1^{\mu} \equiv a^{\mu}$. Also associated with the wave is the slowly varying wave 4-vector $k^{\mu}(x)$ defined (Dougherty 1970) as the 4-gradient of θ, that is

$$k^{\mu} \equiv -\partial^{\mu}\theta. \tag{74}$$

Within the WKB approximation, $\mathscr{L}$ is a function only of X^{μ}, $\partial^{\mu}X^{\nu}$, a^{μ}, $a^{\mu *}$ and k^{μ}. Following Whitham (1965) we assume local averaging to have been applied to $\mathscr{L}$ (which has negligible effect on the action integral), so that $\mathscr{L}$ is independent of θ.

The Euler–Lagrange equations (14) corresponding to $\eta_i = a^{\mu}, a^{\mu *}$ are the 'wave equations'

$$\partial\mathscr{L}/\partial a^{\mu *} = \partial\mathscr{L}/\partial a^{\mu} = 0, \tag{75}$$

which besides giving the dispersion relation for the wave also determine its polarization. The Euler–Lagrange equation from variation of θ is the continuity equation for wave action

$$\partial_{\mu}N^{\mu} = 0, \tag{76}$$

where

$$N^{\mu} \equiv \partial\mathscr{L}/\partial k_{\mu} \tag{77}$$

is the wave action current (see Section 4*d*). From equations (16) and (17) we find the canonical energy–momentum tensor and force density for the wave subsystem to be

$$T_{\rm w}{}^{\mu\nu} = N^{\mu}k^{\nu} - \mathscr{L}_{\rm w}g^{\mu\nu}, \qquad f_{\rm w}{}^{\mu} = N^{\nu}\partial^{\mu}k_{\nu} - \partial^{\mu}\mathscr{L}_{\rm w}, \tag{78a, b}$$

where $\mathscr{L}_{\rm w}$ is that part of $\mathscr{L}$ depending on k^{μ} and a^{μ}. Since $g_{\rm w}{}^{\rho\mu\nu}$ vanishes, equation (78a) also gives the modified canonical tensor $S_{\rm w}{}^{\mu\nu}$.

We now show that equations (78) are consistent with equations (55) for non-dispersive waves. In this case the only k^μ dependence in $\mathscr{L}_w$ comes from

$$B^{\mu\nu} = \sum_n in(k^\nu a^\mu - k^\mu a^\nu)\exp(in\theta). \tag{79}$$

Then, since $\mathscr{L}_w$ is the averaged electromagnetic Lagrangian,

$$\mathscr{L}_w = \langle \mathscr{L}_{em} \rangle, \tag{80}$$

we have, to lowest order in the WKB approximation,

$$N^\mu = \left\langle \frac{\partial \mathscr{L}_{em}}{\partial B_{\rho\sigma}} \frac{\partial B_{\rho\sigma}}{\partial k_\mu} \right\rangle = \left\langle H^{\mu\sigma} \frac{\partial A}{\partial \theta^\sigma} \right\rangle, \tag{81}$$

where the angle brackets denote local time and space averaging. It is then easily seen that

$$T_w{}^{\mu\nu} = \langle T_{em}^{\mu\nu} \rangle. \tag{82}$$

However, we saw that $T_{em}^{\mu\nu}$ differs from the Minkowski tensor $S_{em}^{\mu\nu}$ only by the addition of $\partial_\rho(H^{\mu\rho}A^\nu)$. To lowest order in the WKB approximation this averages to zero, and hence

$$T_w{}^{\mu\nu} = \langle S_{em}^{\mu\nu} \rangle. \tag{83}$$

We have thus established that, for nondispersive waves, the wave energy–momentum tensor (78a) is equal to the averaged Minkowski energy–momentum tensor. For dispersive waves we adopt equation (78a) as the *definition* of the Minkowski tensor. It is interesting to note that this equation is consistent with the remark by Peierls (1976) that the Minkowski tensor corresponds to assigning pseudomomentum $\hbar \boldsymbol{k}$ to the wave.

Since a^μ is a 4-vector and θ a scalar, local Lorentz invariance from equations (25) and (56) implies

$$\left(H_w{}^{\mu\sigma} B_\sigma{}^\nu + N^\mu k^\nu + \partial^\nu X_\sigma \frac{\partial \mathscr{L}_w}{\partial\, \partial_\mu X_\sigma} - u^\mu \frac{\partial \mathscr{L}_w}{\partial u_\nu}\right)_{\text{a.s.}} - \tfrac{1}{2}\left(a^\mu \frac{\partial \mathscr{L}_w}{\partial a_\nu} - a^\nu \frac{\partial \mathscr{L}_w}{\partial a_\mu}\right) = 0, \tag{84}$$

where $H_w{}^{\mu\sigma}$ is that part of $H^{\mu\sigma}$ contributed by $\mathscr{L}_w$ (assuming the low frequency electromagnetic field to influence the dispersive properties of the high frequency field). The physical energy–momentum tensor for the wave subsystem is, from equations (33) and (34),

$$\theta_w{}^{\mu\nu} = u.\partial X_\sigma \frac{\partial \mathscr{L}_w}{\partial\, \partial_\mu X_\sigma} u^\nu + \left(\partial_\rho X_\sigma \frac{\partial \mathscr{L}_w}{\partial\, \partial_\mu X_\sigma} - u^\mu \frac{\partial \mathscr{L}_w}{\partial u^\rho}\right)\left(g^{\rho\nu} - u^\rho u^\nu\right) + H_w{}^{\mu\sigma} B_\sigma{}^\nu + N^\mu k^\nu - \mathscr{L}_w g^{\mu\nu}. \tag{85}$$

From equations (75) and (84) this is seen to be a symmetric tensor.

(e) General Linear Case

To treat the linear response of the system in a general covariant fashion we adopt the rank-2 polarization tensor description of quantum electrodynamics. This has been expounded by Melrose (1973) and is far simpler than the rank-4 susceptibility tensor description used by O'Dell (1970) and O'Sullivan and Derfler (1973). The polarization tensor $\alpha^{\mu\nu}(k)$ is defined as the linear response function for the high frequency current J^{μ}_{hf} taking the 4-vector potential A^{μ}_{hf} as the driving term,

$$J^{\mu}_{\mathrm{hf}} \equiv \alpha^{\mu}{}_{\nu} A^{\nu}_{\mathrm{hf}}\,. \tag{86}$$

Thus the wave equation is

$$(k^2 g^{\mu}{}_{\nu} - k^{\mu}k_{\nu} + \mu_0 \alpha^{\mu}{}_{\nu})a^{\nu} = 0\,. \tag{87}$$

Now we know that the vacuum contribution to $\mathscr{L}_{\mathrm{w}}$ from equations (61) and (79) is

$$\tfrac{1}{4}\mu_0^{-1}\langle B^{\rho}_{\mathrm{hf}\,\sigma} B^{\sigma}_{\mathrm{hf}\,\rho}\rangle = \mu_0^{-1}(|k.a|^2 - k^2 a^*.a)\,. \tag{88}$$

When it is considered that the wave Lagrangian density must be derivable, at least in principle, from the exact microscopic Lagrangian density, it is clear that the effect of polarization of the material must be to provide an additional term to be added to equation (88). This extra term must yield equation (87), and its complex conjugate, on use of (75). The following Lagrangian density fulfills these requirements, provided $\alpha^{\mu\nu}$ is a hermitian matrix (nondissipative case):

$$\mathscr{L}_{\mathrm{w}} = \mu_0^{-1}(|k.a|^2 - k^2 a^*.a) - a^*_{\rho}\alpha^{\rho}{}_{\sigma}a^{\sigma}\,. \tag{89}$$

The requirement that $\mathscr{L}_{\mathrm{w}}$ be gauge invariant implies that the conditions

$$k_{\rho}\alpha^{\rho}{}_{\sigma} = \alpha^{\rho}{}_{\sigma}k^{\sigma} = 0 \tag{90}$$

be satisfied by $\alpha^{\mu\nu}$. These conditions are met automatically in the representation

$$\alpha^{\mu}{}_{\nu} = \varepsilon_0\left(g^{\mu}{}_{\rho} - \frac{u^{\mu}k_{\rho}}{k.u}\right)\Pi^{\rho}{}_{\sigma}\left(g^{\sigma}{}_{\nu} - \frac{k^{\sigma}u_{\nu}}{k.u}\right). \tag{91}$$

By virtue of the conditions (90), the wave equations (87) are not linearly independent and the determinant of the coefficient matrix vanishes identically. This trivial singularity can be removed, without affecting the component of a^{μ} orthogonal to k^{μ}, by adding $k^{\mu}k_{\nu}$ to the matrix. The general covariant dispersion relation is therefore

$$\det(k^2 g^{\rho}{}_{\sigma} + \mu_0 \alpha^{\rho}{}_{\sigma}) = 0\,. \tag{92}$$

Note that equations (87) and (89) imply $\mathscr{L}_{\mathrm{w}} = 0$. Thus the Minkowski tensor for a linear dispersive wave is

$$T_{\mathrm{w}}{}^{\mu\nu} = N^{\mu}k^{\nu}\,, \tag{93}$$

where

$$N^{\mu} = \mu_0^{-1}(k.a\,a^{\mu *} + k.a^*\,a^{\mu} - 2k^{\mu}a^*.a) - a^*_{\rho}(\partial\alpha^{\rho}{}_{\sigma}/\partial k_{\mu})a^{\sigma}\,. \tag{94}$$

(f) Linear Isotropic Case

In an isotropic dielectric fluid with no DC fields, $\Pi^{\mu}{}_{\nu}$ is completely determined (up to terms proportional to u^{μ} or u^{ν}, which do not contribute to $\alpha^{\mu}{}_{\nu}$) by two scalar functions $\Pi_l(k.u, k^2; \rho')$ and $\Pi_t(k.u, k^2; \rho')$, measuring the longitudinal and transverse responses respectively:

$$\Pi^\mu{}_\nu = \frac{\Pi_l k^\mu k_\nu}{(k.u)^2-k^2} - \Pi_t\left(g^\mu{}_\nu + \frac{k^\mu k_\nu}{(k.u)^2-k^2}\right). \tag{95}$$

Substitution of equation (95) into (91) yields

$$c^2\mu_0\alpha^\mu{}_\nu = \frac{(\Pi_l-\Pi_t)k^\mu k.u}{(k.u)^2-k^2} - \frac{\{(k.u)^2\Pi_t - k^2\Pi_l\}k^2u^\mu u_\nu}{\{(k.u)^2-k^2\}(k.u)^2}$$
$$-\Pi_t g^\mu{}_\nu + \frac{(k.u)^2\Pi_t - k^2\Pi_l}{(k.u)^2-k^2}\,\frac{k^\mu u_\nu + u^\mu k_\nu}{k.u}. \tag{96}$$

We now seek the eigenvectors of the wave equation (87). First we observe that

$$(k^2g^\mu{}_\nu - k^\mu k_\nu + \mu_0\alpha^\mu{}_\nu)u^\nu$$
$$= -\{1 - \Pi_l/c^2(k.u)^2\}(k^2g^\mu{}_\nu - k^\mu k_\nu)u^\nu. \tag{97}$$

Thus longitudinal waves obey the dispersion relation

$$1 - \Pi_l(k.u, k^2)/c^2(k.u)^2 = 0 \tag{98}$$

and have the polarization vectors

$$u^\mu + \lambda_l k^\mu, \tag{99}$$

where λ_l is arbitrary, depending on the gauge.

To find the transverse wave solutions, define two vectors $\tau_{1,2}{}^\mu$ by

$$k.\tau_i = u.\tau_i = 0, \qquad \tau_i^*.\tau_j = -\delta_{ij}. \tag{100}$$

Then

$$(k^2g^\mu{}_\nu - k^\mu k_\nu + \mu_0\alpha^\mu{}_\nu)\tau^\nu = (k^2 - \Pi_t/c^2)\tau^\mu. \tag{101}$$

Thus the transverse waves have the dispersion relation

$$1 - \Pi_t(k.u, k^2)/c^2k^2 = 0 \tag{102}$$

and the polarization vectors

$$\tau_{1,2}{}^\mu + \lambda_{1,2}k^\mu, \tag{103}$$

where λ_1 and λ_2 are arbitrary.

Expand a^μ in terms of these normal modes as

$$a^\mu = a_l u^\mu + a_1\tau_1{}^\mu + a_2\tau_2{}^\mu + \lambda k^\mu. \tag{104}$$

Then the wave Lagrangian density reduces to the sum of

$$\mathscr{L}_l \equiv \mu_0^{-1}\{(k.u)^2-k^2\}\{1 - \Pi_l/c^2(k.u)^2\}|a_l|^2 \tag{105}$$

and

$$\mathscr{L}_t \equiv \mu_0^{-1}k^2(1 - \Pi_t/c^2k^2)|a_t|^2, \tag{106}$$

where

$$|a_t|^2 \equiv |a_1|^2 + |a_2|^2.$$

The longitudinal and transverse wave action currents, as defined by equation (77), are therefore

$$N_l^{\mu} = \frac{(k.u)^2-k^2}{(k.u)^2}\left(2k.u\,u^{\mu} - \frac{2}{c^2}\frac{\partial \Pi_l}{\partial k^2}k^{\mu} - \frac{1}{c^2}\frac{\partial \Pi_l}{\partial (k.u)}u^{\mu}\right)\frac{|a_l|^2}{\mu_0}, \tag{107}$$

$$N_t^{\mu} = \left(2k^{\mu} - \frac{2\,\partial \Pi_t\, k^{\mu}}{c^2\,\partial k^2} - \frac{\partial \Pi_t}{c^2\,\partial (k.u)}u^{\mu}\right)\frac{|a_t|^2}{\mu_0}, \tag{108}$$

where use has been made of the dispersion relations (98) and (102). The Minkowski tensors then follow directly from equation (93). In general neither is symmetric, although that for the transverse wave is symmetric if Π_t is not a function of $k.u$.

From equations (49) and (85), the physical energy–momentum tensors for longitudinal and transverse waves are found to be

$$\theta_l^{\mu\nu} = -\varepsilon_0|a_l|^2\frac{(k.u)^2-k^2}{(k.u)^2}\left\{\rho'\frac{\partial \Pi_l}{\partial \rho'}\left(g^{\mu\nu}-u^{\mu}u^{\nu}\right)+2\frac{\partial \Pi_l}{\partial k^2}k^{\mu}k^{\nu}\right.$$
$$\left. -k.u\,u^{\mu}u^{\nu}\left(2c^2\,k.u - \frac{\partial \Pi_l}{\partial (k.u)}\right)\right\}, \tag{109}$$

$$\theta_t^{\mu\nu} = \varepsilon_0|a_t|^2\left\{2\left(c^2 - \frac{\partial \Pi_t}{\partial k^2}\right)k^{\mu}k^{\nu} - \rho'\frac{\partial \Pi_t}{\partial \rho'}\left(g^{\mu\nu}-u^{\mu}u^{\nu}\right) - k.u\frac{\partial \Pi_t}{\partial (k.u)}u^{\mu}u^{\nu}\right\}. \tag{110}$$

The ponderomotive force is therefore

$$\phi_b^{\ \nu} = -\partial_{\mu}\theta_l^{\mu\nu} \quad \text{or} \quad -\partial_{\mu}\theta_t^{\mu\nu}. \tag{111}$$

(g) Cold Plasma

The preceding formalism is readily adaptable to a plasma made up of beams of particles of various species (labelled by the subscript s). Each of the beams may be regarded as a continuum with reference position X_s, 4-velocity u_s^{μ}, proper charge density $q_s n_s'$ and proper mass density $m_s n_s'$, where n_s' is the proper number density of species s. The Lagrangian is a linear superposition of the contributions from the various species. The background Lagrangian density is (cf. equation 44)

$$\mathscr{L}_b = -\sum_s m_s n_s' c^2, \tag{112}$$

and the electromagnetic Lagrangian density for a plasma with no strong low frequency (or DC) field is

$$\mathscr{L}_{em} = -c\sum_s q_s n_s' u_s.A. \tag{113}$$

The electromagnetic contribution differs from equation (61) in that we have neglected terms quadratic in the DC field but have included a term in A^{μ}, which, by equation (53), correctly gives the 4-current density.

The wave contribution can easily be derived from first principles by perturbation expansion, but in the spirit of the phenomenological approach of this section we simply observe that the isotropic polarization tensor must apply for each species. Comparison of the known dispersion relations for longitudinal and transverse waves in a cold plasma (Stix 1962) with equations (98) and (102) reveals that

$$\Pi_{l,s} = \Pi_{t,s} = \omega^2_{\mathrm{p},s}, \tag{114}$$

where $\omega^2_{\mathrm{p},s}$ is the proper plasma frequency of species s, given by

$$\omega^2_{\mathrm{p},s} \equiv n'_s q_s^2 / \varepsilon_0 m_s. \tag{115}$$

These relations are verified in Section 4*d*. The wave Lagrangian density has therefore the particularly simple form

$$\mathscr{L}_{\mathrm{w}} = \sum_s \varepsilon_0 \omega^2_{\mathrm{p},s}\left(g^\mu{}_\rho - \frac{u_s^\mu k_\rho}{k.u_s}\right)\left(g^\rho{}_\nu - \frac{k^\rho u_{s\nu}}{k.u_s}\right) a^*_\mu a^\nu. \tag{116}$$

4. Three-vector Expressions

In the preceding sections we have used 4-vector notation in the interest of elegance and economy. However, most practical calculations are done in 3-vector notation and so in this section we provide a bridge between the two formalisms in order to facilitate the utilization of our general results, and to assist in their physical interpretation.

(*a*) *Matrix Notation*

The relation between 3-vector notation and 4-vector notation is most easily visualized through a partitioned matrix notation in which an arbitrary contravariant 4-vector a^μ is represented by a column vector

$$a^\mu = \begin{bmatrix} a_0 \\ \\ \boldsymbol{a} \end{bmatrix} \tag{117}$$

and a covariant 4-vector a_μ by a row vector

$$a_\mu = [a_0, \quad -\boldsymbol{a}], \tag{118}$$

where $\boldsymbol{a} = a_i \boldsymbol{e}_i$, the $\boldsymbol{e}_i$ being the orthonormal basis vectors of the 3-space reference frame. In equation (117) $\boldsymbol{a}$ is a column vector and in equation (118) $\boldsymbol{a}$ is a row vector. As is usual in vector notation we do not distinguish notationally between 3-row and 3-column quantities but allow the context to determine which is meant. Examples are

$$x^\mu = \begin{bmatrix} ct \\ \\ \boldsymbol{x} \end{bmatrix}, \qquad \partial^\mu = \begin{bmatrix} c^{-1}\,\partial/\partial t \\ \\ -\nabla \end{bmatrix} \tag{119a}$$

and

$$x_\mu = [ct, \quad -\boldsymbol{x}], \qquad \partial_\mu = [c^{-1}\,\partial/\partial t, \quad \nabla]. \tag{119b}$$

The velocity is given by

$$u^\mu = \begin{bmatrix} \gamma \\ \\ \gamma \boldsymbol{v}/c \end{bmatrix}, \qquad u_\mu = [\gamma, \quad -\gamma \boldsymbol{v}/c], \tag{120}$$

where

$$\gamma \equiv (1 - v^2/c^2)^{-\frac{1}{2}}.$$

Mixed tensors will be represented by a 4×4 matrix partitioned into a 1×1 block in the upper left-hand corner, a 3×1 block in the lower left-hand corner, a 3×3 block in the lower right-hand corner, and a 1×3 block in the upper right-hand corner. Thus, if $a^{\mu}{}_{\nu}$ is an arbitrary 4-tensor

$$a^{\mu}{}_{\nu} = \begin{bmatrix} a_{00} & : & -a_{0j}\boldsymbol{e}_j \\ \cdots & : & \cdots \\ a_{i0}\boldsymbol{e}_i & : & -a_{ij}\boldsymbol{e}_i\boldsymbol{e}_j \end{bmatrix}. \tag{121}$$

In dyadic notation $\boldsymbol{e}_i$ is a column vector when on the left and a row vector when on the right. In this representation the metric tensor $g^{\mu}{}_{\nu}$ is the unit 4×4 matrix.

Consider the equation obeyed by the physical energy–momentum tensor of some subsystem

$$\partial_{\mu}\theta^{\mu}{}_{\nu} = \phi_{\nu}. \tag{122}$$

If we denote the components of $\theta^{\mu}{}_{\nu}$ and ϕ^{μ} as

$$\theta^{\mu}{}_{\nu} = \begin{bmatrix} W & : & -c\boldsymbol{G} \\ \cdots & : & \cdots \\ c^{-1}\boldsymbol{S} & : & -\mathbf{T} \end{bmatrix}, \qquad \phi_{\nu} = \begin{bmatrix} c^{-1}p \\ \\ -\boldsymbol{f} \end{bmatrix}, \tag{123a, b}$$

then equation (122) can be written

$$\begin{bmatrix} c^{-1}(\partial W/\partial t + \nabla\boldsymbol{.}\boldsymbol{S}) \\ -(\partial \boldsymbol{G}/\partial t + \nabla\boldsymbol{.}\mathbf{T}) \end{bmatrix} = \begin{bmatrix} p/c \\ -\boldsymbol{f} \end{bmatrix}. \tag{124}$$

The interpretations of the symbols are now clear: W is the physical energy density of the subsystem, $\boldsymbol{S}$ the physical energy flux, $\boldsymbol{G}$ the physical momentum density, $\mathbf{T}$ the physical stress tensor, $\boldsymbol{f}$ the physical force density acting on the subsystem and p the physical power input to the subsystem.

(b) Background Subsystem

Substitution of equation (120) in (50) yields

$$\theta_{\mathrm{b}}{}^{\mu}{}_{\nu} = \begin{bmatrix} (\rho'c^2 + \mathscr{E}' + v^2P'/c^2)\gamma^2 & : & -(\rho'c^2 + \mathscr{E}' + P')\gamma^2\boldsymbol{v}/c \\ \cdots & : & \cdots \\ (\rho'c^2 + \mathscr{E}' + P')\gamma^2\boldsymbol{v}/c & : & -P'\mathbf{I} - (\rho'c^2 + \mathscr{E}' + P')\gamma^2\boldsymbol{v}\boldsymbol{v}/c^2 \end{bmatrix}, \tag{125}$$

where $\mathbf{I}$ is the unit dyadic. By comparing with equation (123a) we can easily read off the expressions for W_{b}, $\boldsymbol{S}_{\mathrm{b}}$, $\boldsymbol{G}_{\mathrm{b}}$ and $\mathbf{T}_{\mathrm{b}}$. The proper density ρ' obeys the continuity equation

$$\partial_{\mu}(\rho' u^{\mu}) = 0, \tag{126}$$

which becomes

$$\partial(\rho'\gamma)/\partial t + \nabla\boldsymbol{.}(\rho'\gamma\boldsymbol{v}) = 0. \tag{127}$$

In nonrelativistic work it is undesirable to include the rest mass energy in W_{b} and $\boldsymbol{S}_{\mathrm{b}}$, and this may be removed, in any given frame, by subtracting the tensor $\rho'c^2u^{\mu}g^{0}{}_{\nu}$ from $\theta_{\mathrm{b}}{}^{\mu}{}_{\nu}$. By virtue of equation (126) the 4-divergence of this term vanishes, so the force balance equation (122) remains unaffected, although the modified tensor is no longer symmetric. The effect on W_{b} and $\boldsymbol{S}_{\mathrm{b}}$ is to change the expression $\rho'c^2\gamma^2$ to $\rho'c^2\gamma(\gamma-1)$.

(*c*) *Low Frequency Electromagnetic Field*

The 4-vector potential and 4-current are given by

$$A^{\mu} = \begin{bmatrix} c^{-1}\Phi \\ \\ A \end{bmatrix}, \qquad J^{\mu} = \begin{bmatrix} c\sigma \\ \\ J \end{bmatrix}, \tag{128a, b}$$

where Φ and A are the scalar and vector potentials and σ and J the charge and current densities. From equations (52) and (128), $B^{\mu}{}_{\nu}$ is found to be given in terms of the electric field E and magnetic induction B by

$$B^{\mu}{}_{\nu} = \begin{bmatrix} 0 & : & c^{-1}E \\ \cdots & : & \cdots \\ c^{-1}E & : & -\mathbf{I}\times B \end{bmatrix}. \tag{129}$$

Following Minkowski (see e.g. Penfield and Haus 1967) we *define* D and H in a material medium by

$$H^{\mu}{}_{\nu} \equiv \begin{bmatrix} 0 & : & cD \\ \cdots & : & \cdots \\ cD & : & -\mathbf{I}\times H \end{bmatrix}. \tag{130}$$

Thus the modified canonical energy–momentum tensor (60) can be written

$$S_{\mathrm{em}}{}^{\mu}{}_{\nu} = \begin{bmatrix} E.D-\mathscr{L}_{\mathrm{em}} & : & -cD\times B \\ \cdots & : & \cdots \\ c^{-1}E\times H & : & DE+BH-(B.H+\mathscr{L}_{\mathrm{em}})\mathbf{I} \end{bmatrix}. \tag{131}$$

We arrived at equation (61) in Section 3*c* for the linear isotropic case by requiring that in the local rest frame it have the form

$$\mathscr{L}'_{\mathrm{em}} = \tfrac{1}{2}\mu_0^{-1}(\varepsilon' E'^2/c^2 - B'^2/\mu'). \tag{132}$$

In the local rest frame, $u'_{\mu} = [1,\ \mathbf{0}]$ and therefore

$$c^{-2}E'^2 = u'_{\mu}B'^{\mu}{}_{\rho}B'^{\rho}{}_{\nu}u'^{\nu}, \qquad c^{-2}E'^2 - B'^2 = \tfrac{1}{2}B'^{\mu}{}_{\rho}B'^{\rho}{}_{\mu}. \tag{133}$$

Since $\mathscr{L}_{\mathrm{em}}$ is to be a scalar we generalize equation (132) to an arbitrary frame simply by deleting the primes from $\mathscr{L}'_{\mathrm{em}}$, u'^{μ} and $B'^{\mu}{}_{\nu}$. In order to verify (132) we calculate $H^{\mu}{}_{\nu}$ from equation (65) and compare it with equation (130) to find the expressions for D and H in a linear isotropic medium moving with velocity v,

$$D = \varepsilon_0\,\varepsilon' E + \gamma^2\varepsilon_0\big(\varepsilon' - (\mu')^{-1}\big)v\times\big(B - c^{-2}v\times E\big), \tag{134a}$$

$$H = \big(\mu_0\,\mu'\big)^{-1}B + \gamma^2\varepsilon_0\big(\varepsilon' - (\mu')^{-1}\big)v\times\big(E + v\times B\big). \tag{134b}$$

In the rest frame, when $v = \mathbf{0}$, these expressions reduce to $\varepsilon_0\,\varepsilon' E$ and $B/\mu_0\,\mu'$ as expected, thus verifying equation (132). Note that equation (66) can be written

$$\mathscr{L}_{\mathrm{em}} = \tfrac{1}{2}(E.D - B.H). \tag{135}$$

The 'Ruhstrahlvektor' Ω^{ν} is given by

$$\Omega^{\nu} = (\gamma^3/c\mu_0\mu')\begin{bmatrix} \boldsymbol{g}\,.\,\boldsymbol{v}/c \\ \boldsymbol{g} \end{bmatrix}, \tag{136}$$

where

$$\boldsymbol{g} \equiv (\gamma^{-2}\mathrm{I} + c^{-2}\boldsymbol{vv})\,.\,\{\boldsymbol{E}\times\boldsymbol{B} + (\boldsymbol{v}\times\boldsymbol{B})\times\boldsymbol{B} + c^{-2}(\boldsymbol{v}\times\boldsymbol{E})\times\boldsymbol{E}\} + c^{-2}\boldsymbol{vv}\,.\,\boldsymbol{E}\times\boldsymbol{B}. \tag{137}$$

Thus the components of the Abraham tensor (equation 70) are

$$W_{\mathrm{A}} = \tfrac{1}{2}(\varepsilon_0\varepsilon' E^2 + (\mu_0\mu')^{-1}B^2) - \gamma^4\varepsilon_0(\varepsilon' - (\mu')^{-1})[(1 + c^{-2}v^2)\boldsymbol{v}\,.\,\boldsymbol{E}\times\boldsymbol{B}$$
$$- \tfrac{1}{2}(1 + c^{-2}v^2)(\boldsymbol{v}\times\boldsymbol{B})^2 - \tfrac{1}{2}(3 - c^{-2}v^2)(c^{-1}\boldsymbol{v}\times\boldsymbol{E})^2], \tag{138a}$$

$$\boldsymbol{S}_{\mathrm{A}} = c^2\boldsymbol{G}_{\mathrm{A}} = (\mu_0\mu')^{-1}\boldsymbol{E}\times\boldsymbol{B} + \gamma^4\varepsilon_0(\varepsilon' - (\mu')^{-1})[(1 - c^{-2}v^2)\boldsymbol{E}\times\{\boldsymbol{v}\times(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B})\}$$
$$- \boldsymbol{v}\{(1 + c^{-2}v^2)\boldsymbol{v}\,.\,\boldsymbol{E}\times\boldsymbol{B} - (\boldsymbol{v}\times\boldsymbol{B})^2 - c^{-2}(\boldsymbol{v}\times\boldsymbol{E})^2\}], \tag{138b}$$

$$\mathbf{T}_{\mathrm{A}} = \tfrac{1}{2}(\varepsilon_0\varepsilon' E^2 + (\mu_0\mu')^{-1}B^2)\mathrm{I} - \varepsilon_0\varepsilon'\boldsymbol{EE} - (\mu_0\mu')^{-1}\boldsymbol{BB}$$
$$- \gamma^2\varepsilon_0(\varepsilon' - (\mu')^{-1})[\{\boldsymbol{v}\,.\,\boldsymbol{E}\times\boldsymbol{B} + \tfrac{1}{2}(\boldsymbol{v}\times\boldsymbol{B})^2 - \tfrac{1}{2}c^{-2}(\boldsymbol{v}\times\boldsymbol{E})^2\}\mathrm{I}$$
$$+ (\boldsymbol{v}\times\boldsymbol{B})\boldsymbol{E} + \boldsymbol{E}(\boldsymbol{v}\times\boldsymbol{B}) + v^2(c^{-2}\boldsymbol{EE} - \boldsymbol{BB}) - (c^{-2}E^2 + B^2)\boldsymbol{vv}$$
$$+ \boldsymbol{v}\,.\,\boldsymbol{B}(\boldsymbol{Bv} + \boldsymbol{vB}) + \gamma^2c^{-2}\boldsymbol{vv}\{2\boldsymbol{v}\,.\,\boldsymbol{E}\times\boldsymbol{B} - (\boldsymbol{v}\times\boldsymbol{B})^2 - c^{-2}(\boldsymbol{v}\times\boldsymbol{E})^2\}]. \tag{138c}$$

The symmetry is now manifest, and also we can verify that ${\theta_{\mathrm{A}}}^{\mu}{}_{\nu}$ is traceless since

$$W_{\mathrm{A}} = \mathrm{Tr}(\mathbf{T}_{\mathrm{A}}). \tag{139}$$

Evaluating the final term of the physical energy–momentum tensor for the electromagnetic subsystem, equation (69), we find

$$W_{\mathrm{em}} = W_{\mathrm{A}} - \tfrac{1}{2}\rho'\left(\boldsymbol{E}\,.\frac{\partial\boldsymbol{D}}{\partial\rho'} - \boldsymbol{B}\,.\frac{\partial\boldsymbol{H}}{\partial\rho'}\right)\frac{\gamma^2v^2}{c^2}, \tag{140a}$$

$$\boldsymbol{S}_{\mathrm{em}} = c^2\boldsymbol{G}_{\mathrm{em}} = \boldsymbol{S}_{\mathrm{A}} - \tfrac{1}{2}\rho'\left(\boldsymbol{E}\,.\frac{\partial\boldsymbol{D}}{\partial\rho'} - \boldsymbol{B}\,.\frac{\partial\boldsymbol{H}}{\partial\rho'}\right)\gamma^2\boldsymbol{v}, \tag{140b}$$

$$\mathbf{T}_{\mathrm{em}} = \mathbf{T}_{\mathrm{A}} - \tfrac{1}{2}\rho'\left(\boldsymbol{E}\,.\frac{\partial\boldsymbol{D}}{\partial\rho'} - \boldsymbol{B}\,.\frac{\partial\boldsymbol{H}}{\partial\rho'}\right)\left(\mathrm{I} + \frac{\gamma^2\boldsymbol{vv}}{c^2}\right). \tag{140c}$$

The ponderomotive force density

$$\boldsymbol{f}_{\mathrm{b}} = -\boldsymbol{f}_{\mathrm{em}} = -\nabla\,.\,\mathbf{T}_{\mathrm{em}} - \partial\boldsymbol{G}_{\mathrm{em}}/\partial t \tag{141}$$

is seen, in a medium in which $\boldsymbol{v}$ is small, to reduce to the well-known expression (Robinson 1975)

$$\boldsymbol{f}_{\mathrm{b}} = -\tfrac{1}{2}\varepsilon_0E^2\nabla\varepsilon' - \tfrac{1}{2}\mu_0H^2\nabla\mu' - \frac{\partial(c^{-2}\boldsymbol{E}\times\boldsymbol{H})}{\partial t} + \frac{\partial(\boldsymbol{D}\times\boldsymbol{B})}{\partial t}$$
$$+ \nabla\left(\tfrac{1}{2}\varepsilon_0E^2\rho'\frac{\partial\varepsilon'}{\partial\rho'} + \tfrac{1}{2}\mu_0H^2\rho'\frac{\partial\mu'}{\partial\rho'}\right). \tag{142}$$

(d) High Frequency Electromagnetic Field

The wave 4-vector (74) expressed in terms of the frequency ω and wave vector $\boldsymbol{k}$ defined by

$$\omega \equiv -\partial\theta/\partial t, \qquad \boldsymbol{k} \equiv \nabla\theta, \tag{143}$$

is

$$k^\mu = \begin{bmatrix} \omega/c \\ \\ \boldsymbol{k} \end{bmatrix}, \tag{144}$$

whence we have $k^2 = (\omega/c)^2 - \boldsymbol{k}^2$. The nth harmonic has the complex amplitudes $\boldsymbol{e}_n$ and $\boldsymbol{b}_n$ of the electric and magnetic fields respectively given by

$$\boldsymbol{e}_n = n\mathrm{i}(\omega \boldsymbol{a}_n - \boldsymbol{k}\phi_n), \qquad \boldsymbol{b}_n = n\mathrm{i}\,\boldsymbol{k}\times\boldsymbol{a}_n, \tag{145a, b}$$

where ϕ_n and $\boldsymbol{a}_n$ are defined by

$$a_n{}^\mu = \begin{bmatrix} c^{-1}\phi_n \\ \\ \boldsymbol{a}_n \end{bmatrix}. \tag{146}$$

The action density N (Whitham 1965) and group velocity $\boldsymbol{v}_\mathrm{g}$, defined by

$$N \equiv \frac{\partial\mathscr{L}}{\partial\omega}, \qquad \boldsymbol{v}_\mathrm{g} \equiv -\frac{\partial\mathscr{L}}{\partial\boldsymbol{k}}\Big/\frac{\partial\mathscr{L}}{\partial\omega}, \tag{147a, b}$$

are related to the action current 4-vector (77) by

$$N^\mu = \begin{bmatrix} Nc \\ \\ N\boldsymbol{v}_\mathrm{g} \end{bmatrix}. \tag{148}$$

Thus equation (76) can be written

$$\partial N/\partial t + \nabla\cdot(N\boldsymbol{v}_\mathrm{g}) = 0. \tag{149}$$

The canonical wave energy–momentum tensor (Minkowski tensor) is thus

$$T_\mathrm{w}{}^\mu{}_\nu = \begin{bmatrix} N\omega - \mathscr{L}_\mathrm{w} & : & -cN\boldsymbol{k} \\ \cdots & : & \cdots \\ c^{-1}N\omega\boldsymbol{v}_\mathrm{g} & : & -(N\boldsymbol{v}_\mathrm{g}\boldsymbol{k} + \mathscr{L}_\mathrm{w}\mathbf{I}) \end{bmatrix}. \tag{150}$$

We recall that $\mathscr{L}_\mathrm{w}$ vanishes for linear waves.

In relating the polarization tensor $\alpha^\mu{}_\nu$ to the dielectric constant and magnetic permeability we run into the problem that $\boldsymbol{D}_\mathrm{hf}$ and $\boldsymbol{H}_\mathrm{hf}$ are not uniquely defined in a dispersive medium. (In fact they are not uniquely defined in any moving medium.) We shall adopt the standard convention used in plasma physics, and define $\boldsymbol{D}$ and $\boldsymbol{H}$ by the relations

$$\boldsymbol{H}_\mathrm{hf} = \boldsymbol{B}_\mathrm{hf}/\mu_0, \qquad \nabla\times\boldsymbol{H}_\mathrm{hf} = \partial\boldsymbol{D}_\mathrm{hf}/\partial t. \tag{151a, b}$$

Then

$$\boldsymbol{D}_\mathrm{hf} = \varepsilon_0\,\boldsymbol{\varepsilon}(\boldsymbol{k},\omega)\cdot\boldsymbol{E}_\mathrm{hf}, \tag{152}$$

where the dielectric tensor $\boldsymbol{\varepsilon}$ is defined by

$$\boldsymbol{\varepsilon} \equiv \mathbf{I} + \mathrm{i}(\varepsilon_0\,\omega)^{-1}\boldsymbol{\sigma}, \tag{153}$$

$\boldsymbol{\sigma}(\boldsymbol{k}, \omega)$ being the high frequency conductivity tensor. Note, however, that if equation (151a) is to hold in all frames of reference, the Minkowski transformation laws for $\boldsymbol{D}$ and $\boldsymbol{H}$ do not apply, since they mix $\boldsymbol{E}$ and $\boldsymbol{B}$. In other words $H^{\mu}{}_{\nu}$ defined by equations (130) and (151) is not a true 4-tensor. This is one reason we prefer to work with the 4-vector potential and the polarization tensor $\alpha^{\mu}{}_{\nu}$.

From equations (86) and (153) we can show

$$\alpha^{\mu}{}_{\nu} = \mathrm{i}\omega \begin{bmatrix} -\boldsymbol{n}.\boldsymbol{\sigma}.\boldsymbol{n} & : & \boldsymbol{n}.\boldsymbol{\sigma} \\ \cdots & : & \cdots \\ -\boldsymbol{\sigma}.\boldsymbol{n} & : & \boldsymbol{\sigma} \end{bmatrix}, \tag{154}$$

where $\boldsymbol{n} \equiv c\boldsymbol{k}/\omega$. Note that equation (154) satisfies the requirements (90). We can rewrite equation (154) in the form

$$\alpha^{\mu}{}_{\nu} = \mathrm{i}\omega \begin{bmatrix} 0 & : & \boldsymbol{n} \\ \cdot\cdot & : & \cdot\cdot \\ \boldsymbol{0} & : & \mathbf{I} \end{bmatrix} \begin{bmatrix} 0 & : & \boldsymbol{0} \\ \cdot\cdot & : & \cdot\cdot \\ \boldsymbol{0} & : & \boldsymbol{\sigma} \end{bmatrix} \begin{bmatrix} 0 & : & \boldsymbol{0} \\ \cdot\cdot & : & \cdot\cdot \\ -\boldsymbol{n} & : & \mathbf{I} \end{bmatrix}. \tag{155}$$

In the rest frame, the factorization (155) is identical with that in equation (91). Thus, *in the rest frame,*

$$\Pi^{\mu}{}_{\nu} = \mathrm{i}\omega \begin{bmatrix} \lambda & : & \boldsymbol{\mu} \\ \cdot\cdot & : & \cdot\cdot \\ -\boldsymbol{\mu} & : & \boldsymbol{\sigma} \end{bmatrix}, \tag{156}$$

where λ and $\boldsymbol{\mu}$ are arbitrary.

If the medium is isotropic in the rest frame then

$$\boldsymbol{\sigma} = \sigma_l \boldsymbol{kk}/\boldsymbol{k}.\boldsymbol{k} + \sigma_t(\mathbf{I} - \boldsymbol{kk}/\boldsymbol{k}.\boldsymbol{k}). \tag{157}$$

Comparison of equation (156) with (95) reveals that

$$\Pi_l = -\mathrm{i}\omega\sigma_l/\varepsilon_0, \qquad \Pi_t = -\mathrm{i}\omega\sigma_t/\varepsilon_0. \tag{158}$$

For instance, in a cold plasma with no external field

$$\sigma_l = \sigma_t = \mathrm{i}\varepsilon_0\,\omega_{\mathrm{p}}^2/\omega \tag{159}$$

and therefore

$$\Pi_l = \Pi_t = \omega_{\mathrm{p}}^2, \tag{160}$$

thus verifying equation (114).

Substitution of equation (104) into (145a) shows that the electric field $\boldsymbol{e}_l$ of a longitudinal wave and the field $\boldsymbol{e}_t$ of a transverse wave are given by

$$\boldsymbol{e}_l = -\mathrm{i}\omega a_l(\boldsymbol{n} - \gamma\boldsymbol{v}/c), \qquad \boldsymbol{e}_t = \mathrm{i}\omega a_t(\mathbf{I} - \boldsymbol{nv}/c).\boldsymbol{\tau}, \tag{161}$$

where $\boldsymbol{n} \equiv c\boldsymbol{k}/\omega$, as above, and $\boldsymbol{\tau}$ is a vector orthogonal to $\boldsymbol{n} - \boldsymbol{v}/c$ and normalized so that

$$(\boldsymbol{\tau})^2 - (\boldsymbol{\tau}.\boldsymbol{v}/c)^2 = 1. \tag{162}$$

Thus

$$\boldsymbol{e}_l^*.\boldsymbol{e}_l = \omega^2|a_l|^2(n^2 - \gamma c^{-1}\boldsymbol{n}.\boldsymbol{v} + \gamma^2 c^{-2} v^2), \tag{163a}$$

$$\boldsymbol{e}_t^*.\boldsymbol{e}_t = \omega^2|a_t|^2\{1 - 2c^{-1}\boldsymbol{v}.\boldsymbol{\tau}\,\boldsymbol{\tau}.\boldsymbol{n} + (1+n^2)(c^{-1}\boldsymbol{v}.\boldsymbol{\tau})^2\}. \tag{163b}$$

Since $|a_l|^2$ and $|a_t|^2$ are invariants and $\boldsymbol{e}_l^*.\boldsymbol{e}_l$ and $\boldsymbol{e}_t^*.\boldsymbol{e}_t$ are not, it is useful to define new invariants equal to the average electric field energy in the rest frame,

$$W_{\mathrm{e}} \equiv \varepsilon_0 \boldsymbol{e}'^*.\boldsymbol{e}', \tag{164}$$

where $\boldsymbol{e}'$ is the proper electric field. Thus

$$|a_l|^2 = W_{\mathrm{e},l}/\varepsilon_0 c^2|\boldsymbol{k}'|^2, \qquad |a_t|^2 = W_{\mathrm{e},t}/\varepsilon_0 \omega'^2, \tag{165}$$

where

$$\omega' \equiv ck.u = \gamma(\omega - \boldsymbol{k}.\boldsymbol{v}), \qquad |\boldsymbol{k}'| \equiv \{(k.u)^2 - k^2\}^{\frac{1}{2}} \tag{166}$$

are the proper frequency and wave vector magnitude respectively.

The proper longitudinal and transverse dielectric constants are, by equations (153), (157) and (158), given by

$$\varepsilon_l'(|\boldsymbol{k}'|, \omega') = 1 - \Pi_l/\omega'^2, \qquad \varepsilon_t'(|\boldsymbol{k}'|, \omega') = 1 - \Pi_t/\omega'^2. \tag{167}$$

Thus the action densities defined by equation (147a) are, from equations (107) and (108),

$$N_l = \frac{\gamma W_{\mathrm{e},l}}{\omega'^2}\frac{\partial(\omega'^2\varepsilon_l')}{\partial\omega'} + \frac{(\gamma\omega' - \omega)W_{\mathrm{e},l}}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_l'}{\partial|\boldsymbol{k}'|}, \tag{168}$$

$$N_t = \frac{\gamma W_{\mathrm{e},t}}{\omega'^2} + \frac{\partial(\omega'^2\varepsilon_t')}{\partial\omega'} + \left(\gamma\omega' - \omega\right)\left(\frac{1}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_t'}{\partial|\boldsymbol{k}'|} - \frac{2}{\omega'^2}\right). \tag{169}$$

In the rest frame these reduce to known expressions (Tsytovich 1970).

The physical energy–momentum tensors (109) and (110) have the components

$$W_l = W_{\mathrm{e},l}\left\{\frac{\gamma^2}{\omega'}\frac{\partial(\varepsilon_l'\omega'^2)}{\partial\omega'} - \left(\frac{\omega^2}{\omega'^2} - \gamma^2\right)\frac{\omega'^2}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_l'}{\partial|\boldsymbol{k}'|} - \frac{\gamma^2 v^2\rho'}{c^2}\frac{\partial\varepsilon_l'}{\partial\rho'}\right\}, \tag{170a}$$

$$\boldsymbol{S}_l = c^2\boldsymbol{G}_l = W_{\mathrm{e},l}\left\{\frac{\gamma^2\boldsymbol{v}}{\omega'}\frac{\partial(\varepsilon_l'\omega'^2)}{\partial\omega'} - \left(\frac{c^2\boldsymbol{k}\omega}{\omega'^2} - \gamma^2\boldsymbol{v}\right)\frac{\omega'^2}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_l'}{\partial|\boldsymbol{k}'|} - \gamma^2\boldsymbol{v}\rho'\frac{\partial\varepsilon_l'}{\partial\rho'}\right\}, \tag{170b}$$

$$\mathbf{T}_l = W_{\mathrm{e},l}\left\{\frac{\gamma^2\boldsymbol{vv}}{c^2\omega'}\frac{\partial(\varepsilon_l'\omega'^2)}{\partial\omega'} - \left(\frac{c^2\boldsymbol{kk}}{\omega'^2} - \frac{\gamma^2\boldsymbol{vv}}{c^2}\right)\frac{\omega'^2}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_l'}{\partial|\boldsymbol{k}'|} - \left(\mathbf{I} + \frac{\gamma^2\boldsymbol{vv}}{c^2}\right)\rho'\frac{\partial\varepsilon_l'}{\partial\rho'}\right\}; \tag{170c}$$

$$W_t = W_{\mathrm{e},t}\left\{\left(\frac{\omega^2}{\omega'^2} - \gamma^2\right)\left(2 - \frac{\omega'^2}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_t'}{\partial|\boldsymbol{k}'|}\right) + \frac{\gamma^2}{\omega'}\frac{\partial(\varepsilon_t'\omega'^2)}{\partial\omega'} - \frac{\gamma^2 v^2\rho'}{c^2}\frac{\partial\varepsilon_t'}{\partial\rho'}\right\}, \tag{171a}$$

$$\boldsymbol{S}_t = c^2\boldsymbol{G}_t = W_{\mathrm{e},t}\left\{\left(\frac{c^2\boldsymbol{k}\omega}{\omega'^2} - \gamma^2\boldsymbol{v}\right)\left(2 - \frac{\omega'^2}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_t'}{\partial|\boldsymbol{k}'|}\right) + \frac{\gamma^2\boldsymbol{v}}{\omega'}\frac{\partial(\varepsilon_t'\omega'^2)}{\partial\omega'} - \gamma^2\boldsymbol{v}\rho'\frac{\partial\varepsilon_t'}{\partial\rho'}\right\}, \tag{171b}$$

$$\mathbf{T}_t = W_{\mathrm{e},t}\left\{\left(\frac{c^2\boldsymbol{kk}}{\omega'^2} - \frac{\gamma^2\boldsymbol{vv}}{c^2}\right)\left(2 - \frac{\omega'}{c^2|\boldsymbol{k}'|}\frac{\partial\varepsilon_t'}{\partial|\boldsymbol{k}'|}\right) + \frac{\gamma^2\boldsymbol{vv}}{c^2\omega'}\frac{\partial(\varepsilon_t'\omega'^2)}{\partial\omega'} - \rho'\frac{\partial\varepsilon_t'}{\partial\rho'}\left(\mathbf{I} + \frac{\gamma^2\boldsymbol{vv}}{c^2}\right)\right\}. \tag{171c}$$

It is an interesting exercise to compare equations (170) and (171) with their non-dispersive analogues, equations (140), in the rest frame. One finds that *in the case of no spatial dispersion* (ε independent of $\boldsymbol{k}$) one can derive the momentum density and stress tensor, and thus the ponderomotive force, simply by averaging the nondispersive result and replacing ε' by $\varepsilon_l'(\omega)$ or $\varepsilon_t'(\omega)$ as appropriate. This is *not* valid when there is spatial dispersion.

5. Wavepackets

(*a*) *One-dimensional Wavepackets*

The results of the previous sections do not have immediate experimental significance since they say nothing about how the wave is generated or absorbed. This is a complicated subject, but a simple thought experiment sheds some light on the physics involved. We suppose that the wave arrives in the region of interest, which is initially uniform and stationary, as a wavepacket and propagates through the region without changing the shape of its envelope. For this it is necessary that the pulse be broad enough for dispersive spreading to be negligible during the transit time. Furthermore, we suppose the medium to be unchanged after the wavepacket has departed, and for this it is necessary that the wavepacket be finite in only one direction in order to avoid setting up a Cerenkov-like wake (Peierls 1976).

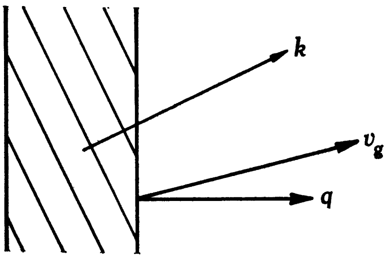

Fig. 2. Three vectors associated with a wavepacket:
$\boldsymbol{k}$, the wave vector;
$\boldsymbol{v}_g$, the group velocity;
$\boldsymbol{q}$, a vector normal to the contours of constant intensity.

We therefore assume the contours of constant $a^*.a$ to be parallel planes with unit normal $\hat{q} = \boldsymbol{q}/|\boldsymbol{q}|$, as depicted in Fig. 2. The analysis is simple in 3-dimensional notation only if we work in the rest frame of the unperturbed medium. However, the analysis is simple in an arbitrary frame if we employ the 4-dimensional formalism of Section 3, and this we choose to do.

Because $\hat{q}$ and $\boldsymbol{v}_g$ cannot be parallel in all frames we allow them to make an arbitrary angle with each other, and with $\boldsymbol{k}$. The one-dimensional propagation of the wavepacket can be summarized by forming a 4-vector q^μ such that $a^*.a$ (and hence N^μ) depends only on $\psi \equiv q.x$. By integrating the action conservation equation (76) from $\psi = \pm\infty$ to $q.x$ we see that q^μ is orthogonal to the action current N^μ,

$$q.N = 0. \tag{172}$$

From equation (148) this implies

$$q_0 = \boldsymbol{q}.\boldsymbol{v}_g/c. \tag{173}$$

The normalization of $\boldsymbol{q}$ is arbitrary.

As the average energy–momentum tensor must also be a function of ψ alone, integration of the conservation equation (30) implies

$$q_\mu\,\delta\theta^{\mu\nu} = 0, \tag{174}$$

where

$$\delta\theta^{\mu\nu}(q.x) \equiv \theta^{\mu\nu}(q.x) - \theta^{\mu\nu}(\pm\infty). \tag{175}$$

Equations (123a), (173) and (174) imply

$$\boldsymbol{q}\cdot\delta\boldsymbol{S} = \boldsymbol{q}\cdot\boldsymbol{v}_{\mathrm{g}}\,\delta W. \tag{176}$$

In a frame in which $\boldsymbol{q}$, $\boldsymbol{v}_{\mathrm{g}}$ and $\delta\boldsymbol{S}$ are parallel, equation (176) and the symmetry of $\theta^{\mu\nu}$ led Burt and Peierls (1973) to the inescapable conclusion that the perturbed momentum density and the perturbed energy density of the system obey the relation

$$|\delta\boldsymbol{G}| = v_{\mathrm{g}}\,\delta W/c^2. \tag{177}$$

However, δW is not simply given by equation (171a) (Haus 1969; Klima and Petrzilka 1973; Peierls 1976), but includes the contribution δW_{b} from the perturbation in the background due to the ponderomotive force. As the assumption of symmetric $\theta_{\mathrm{b}}{}^{\mu\nu}$ implies that the rest energy is to be retained, as in equation (125), δW_{b} is not negligible and we must therefore integrate the equation of motion for the background.

(b) Linear Isotropic Dispersive Medium

From the continuity equation (126) the perturbation $\delta\rho'$ in the background proper density due to the passage of the wave is given by

$$\delta\rho'/\rho' = -(q\cdot\delta u)/(q\cdot u), \tag{178}$$

neglecting second-order terms. Linearizing equation (50) and using (174) and (178) we find

$$q_\mu\,\delta\theta_{\mathrm{b}}{}^{\mu\nu} = \rho_0\,c^2 q\cdot u\{\delta u^\nu + C_{\mathrm{s}}^2\,q_\perp^\nu\,q\cdot\delta u/c^2(q\cdot u)^2\} = -q_\mu\,\theta_{\mathrm{w}}{}^{\mu\nu}, \tag{179}$$

where

$$q_\perp^\mu \equiv (g^\mu{}_\nu - u^\mu u_\nu)q^\nu,$$

and the sound speed C_{s} and effective mass density ρ_0 are defined by

$$C_{\mathrm{s}}^2 \equiv (\rho'/\rho_0)\,\partial P'/\partial\rho', \qquad \rho_0 \equiv \rho' + (\mathscr{E}' + P')/c^2. \tag{180}$$

Since δu^μ is clearly proportional to W_{e}, which is assumed small for the linear theory to apply, we are amply justified in the linearizations leading to equation (179). Furthermore, we can neglect perturbations in the background quantities in evaluating the right-hand side of (179).

Contracting equation (179) with q_ν we find

$$\rho_0\,c^2 q\cdot\delta u = -q_\mu\,\theta_{\mathrm{w}}{}^{\mu\nu}\,q_\nu/q\cdot u\,D_{\mathrm{s}}, \tag{181}$$

where

$$D_{\mathrm{s}} \equiv 1 + C_{\mathrm{s}}^2\,q_\perp^2/c^2(q\cdot u)^2 \tag{182}$$

is the sound wave dispersion function (that is, $D_{\mathrm{s}} = 0$ means that q^μ is the wave 4-vector for a sound wave). Equation (181) can now be fed back into (179) to find δu^ν, which in turn is fed back into $\delta\theta_{\mathrm{b}}{}^{\mu\nu}$. Adding $\theta_{\mathrm{w}}{}^{\mu\nu}$ we can then manipulate the total perturbed energy–momentum tensor into the form

$$\delta\theta^\mu{}_\nu = Q^\mu{}_\rho\{\theta_{\mathrm{w}}{}^\rho{}_\sigma - C_{\mathrm{s}}^2\,q_\alpha\,\theta_{\mathrm{w}}{}^{\alpha\beta}\,q_\beta\,g^\rho{}_\sigma/c^2(q\cdot u)^2 D_{\mathrm{s}}\}Q_\nu{}^\sigma, \tag{183}$$

where

$$Q^\mu{}_\nu \equiv g^\mu{}_\nu - u^\mu q_\nu/q\cdot u.$$

The form (183) makes manifest how the background medium 'dresses' the wave energy–momentum tensor in a wavepacket.

In the case of a transverse wave (e.g. a laser pulse), substitution of equation (110) into (183) yields

$$\delta\theta^{\mu}{}_{\nu} = \varepsilon_0 |a_t|^2 Q^{\mu}{}_{\rho}\left\{2\left(c^2 - \frac{\partial \Pi_t}{\partial k^2}\right)k^{\rho}k_{\sigma} - \frac{\rho'}{D_s}\frac{\partial \Pi_t}{\partial \rho'}g^{\rho}{}_{\sigma} - \frac{C_s^2}{c^2 q.u}\frac{k_\perp . q_\perp}{D_s}\frac{\partial \Pi_t}{\partial (k.u)}g^{\rho}{}_{\sigma}\right\}Q_{\nu}{}^{\sigma}, \tag{184}$$

where we have used equations (108) and (172). Evaluating the result (184) in the rest frame of the unperturbed medium, replacing Π_t with the transverse dielectric constant from equations (167) and taking the $0i$ components of $\delta\theta^{\mu\nu}/c$, we obtain the momentum density $\boldsymbol{G}$ in the wavepacket as

$$\boldsymbol{G} = W_e\left[\left(\frac{2c^2}{\omega^2} - \frac{1}{|\boldsymbol{k}|}\frac{\partial \varepsilon_t}{\partial |\boldsymbol{k}|}\right)\frac{\boldsymbol{k}|\boldsymbol{k}|}{|\boldsymbol{v}_g|} - \frac{\boldsymbol{q}\rho}{\boldsymbol{q}.\boldsymbol{v}_g D_s}\frac{\partial \varepsilon_t}{\partial \rho} + \frac{\boldsymbol{q}C_s^2|\boldsymbol{k}|}{\boldsymbol{q}.\boldsymbol{v}_g\omega|\boldsymbol{v}_g|D_s} - \left\{\frac{1}{\omega}\frac{\partial(\varepsilon_t\omega^2)}{\partial\omega} + \frac{\omega^2}{c^2|\boldsymbol{k}|}\frac{\partial \varepsilon_t}{\partial|\boldsymbol{k}|} - 2\right\}\right], \tag{185}$$

with

$$D_s = 1 - C_s^2|\boldsymbol{q}^2|/(\boldsymbol{q}.\boldsymbol{v}_g)^2 .$$

Using equations (108) and (148) we can also write (185) in the form

$$\boldsymbol{G} = N\boldsymbol{k} - \frac{W_e\boldsymbol{q}\rho}{\boldsymbol{q}.\boldsymbol{v}_g D_s}\frac{\partial \varepsilon_t}{\partial \rho} + \frac{\boldsymbol{q}C_s^2 N\omega}{\boldsymbol{q}.\boldsymbol{v}_g c^2 D_s}\left(\frac{c^2|\boldsymbol{k}|}{\omega|\boldsymbol{v}_g|} - 1\right). \tag{186}$$

If $\boldsymbol{q}$ is parallel to $\boldsymbol{v}_g$ and $|\boldsymbol{v}_g| \lesssim c$, then the terms proportional to C_s^2 are of order C_s^2/c^2 relative to the other terms and may be neglected. In this case $D_s \approx 1$. For instance, in a nondispersive medium we have

$$G \approx \frac{nW_e}{c}\left(2n^2 - \rho\frac{\partial\varepsilon}{\partial\rho}\right) = \frac{nW_t}{c}\left(1 - \frac{\rho}{2n^2}\frac{\partial\varepsilon}{\partial\rho}\right), \tag{187}$$

where $W_t = 2n^2 W_e$ is the wave energy density, equation (171a). Without the electrostrictive contribution, this result would be the so-called Minkowski result $N\boldsymbol{k} = \langle \boldsymbol{D}\times\boldsymbol{B}\rangle$. With the correction term, the result is in agreement with Robinson's (1975) assertion that the momentum travelling with a short wide wavepacket is $(1 + \alpha/2\varepsilon + \beta/2\mu)$ times the Minkowski result, with α and β the electrostrictive and magnetostrictive coefficients respectively (with $\beta = 0$ in our case).

If $\boldsymbol{q}$ is almost perpendicular to $\boldsymbol{v}_g$, however, we can make D_s arbitrarily small and upset the relative ordering of the terms. In this case the electrostrictive term dominates, and is still of order C_s^2/c^2 larger than the final term in equation (185). Thus this term may be neglected in all ranges of $\boldsymbol{q}$ (except in a relativistic gas). The resonance at $D_s = 0$ corresponds to fulfilment of the ω and $\boldsymbol{k}$ matching conditions for decay of the electromagnetic wave into an acoustic wave and another electromagnetic wave close to the original in ω and $\boldsymbol{k}$. This phenomenon, stimulated Brillouin scattering (Kroll 1965), clearly tends to invalidate the assumption that the wave profile is unchanged, and requires a more sophisticated treatment. This will be sketched out in Section 5*d*.

For $\boldsymbol{q}$ even closer to the perpendicular than required for the acoustic resonance, $|D_s|$ begins to increase again and approaches infinity as $\boldsymbol{q}.\boldsymbol{v}_g \to 0$. In this limit the electrostrictive contribution vanishes and the momentum reduces to the Minkowski result Nk. This is also in accord with Robinson's (1975) assertion that the momentum of a long thin wavepacket is the Minkowski result.

(c) Cold Plasma

Variation of A^μ in equation (113) gives the charge neutrality constraint for a plasma composed of particles of charge e_s as

$$\sum_s e_s n_s' u_s{}^\mu = 0. \tag{188}$$

In linearized form this states, using equations (178) and (183), that

$$\sum_s e_s n_s' Q_s{}^{\mu\nu} \delta u_{s\,\nu} = 0. \tag{189}$$

Each individual species obeys the equation of motion (cf. equation 11)

$$\partial_\mu\left\{\left(n_s'\frac{\partial\mathscr{L}}{\partial n_s'}g^\mu{}_\rho - u_s{}^\mu\frac{\partial\mathscr{L}}{\partial u_s{}^\rho}\right)\left(g^{\rho\nu} - u_s{}^\rho u_s{}^\nu\right) - \mathscr{L}_{\mathrm{bs}}g^{\mu\nu}\right\}$$
$$= \partial^\nu n_s'\frac{\partial\mathscr{L}}{\partial n_s'} + \partial^\nu u_s{}^\rho\frac{\partial\mathscr{L}}{\partial u_s{}^\rho} - \partial^\nu\mathscr{L}_{\mathrm{bs}}, \tag{190}$$

where $\mathscr{L}_{\mathrm{bs}}$ is the contribution of species s to the background Lagrangian density. Linearizing and integrating we find

$$m_s n_s' \delta u_s{}^\mu = -\left(e_s n_s A^\mu + n_s'\frac{\partial\mathscr{L}_{\mathrm{ws}}}{\partial n_s'}\frac{q^\mu}{q.u_s} - \frac{\partial\mathscr{L}_{\mathrm{ws}}}{\partial u_{s\mu}}\right)_\perp. \tag{191}$$

The general procedure is to feed equation (191) into (189) and solve for A^μ, thus coupling the wavepacket to the low frequency electromagnetic modes of the plasma. However, we shall be content here with an important special case where this is not necessary.

If we suppose the plasma to have only two species (ions and electrons) which have equal velocities and charge densities in the unperturbed state, then equation (189) becomes

$$e_{\mathrm{i}} n_{\mathrm{i}}' Q^\mu{}_\nu(\delta u_{\mathrm{i}}{}^\nu - \delta u_{\mathrm{e}}{}^\nu) = 0. \tag{192}$$

Since $\delta u_{\mathrm{i}}{}^\mu$ and $\delta u_{\mathrm{e}}{}^\mu$ can have no component parallel to u^μ, the unique solution of equation (192) is

$$\delta u_{\mathrm{i}}{}^\mu = \delta u_{\mathrm{e}}{}^\mu. \tag{193}$$

Thus, *in this special case only*, the plasma behaves as a single fluid and we may apply the results of the previous subsection. Using equation (114) in (184) we have

$$\delta\theta^\mu{}_\nu = \varepsilon_0 c^2k^2|a_t|^2(2k^\mu k_\nu/k^2 - Q^\mu{}_\rho Q_\nu{}^\rho), \tag{194}$$

where we have used the dispersion relation (102) and the fact, following equations

(108) and (172), that

$$q\,.\,k = 0\,. \tag{195}$$

Using the results of Section 4*d* we can show that equation (194) is in complete agreement with the results of Klima and Petrzilka (1973), despite their statement that their results cannot be obtained by the averaged Lagrangian method.

(*d*) *Three-dimensional Wavepackets*

To treat three-dimensional wavepackets we Fourier analyse the wave envelope. In this case q^μ becomes the wave 4-vector of a Fourier component of the envelope

$$\varepsilon_0\,|\,a_t\,|^2 = \int\frac{\mathrm{d}^4 q}{(2\pi)^4}F_q\exp(-\mathrm{i}q\,.\,x)\,. \tag{196}$$

Note that, by virtue of equation (172), F_q includes a factor $\delta(q\,.\,N)$.

Since the derivation of the perturbed energy–momentum tensor in the preceding subsections was based on linearization in $|\,a_t\,|^2$, the perturbed energy–momentum tensor of a three-dimensional wavepacket is found by linear superposition,

$$\delta\theta^\mu{}_\nu = \int\frac{\mathrm{d}^4 q}{(2\pi)^4}\exp(-\mathrm{i}q\,.\,x)\,F_q\,Q^\mu{}_\rho\left\{2\left(c^2-\frac{\partial\Pi_t}{\partial k^2}\right)k^\rho k_\sigma-\frac{\rho'}{D_\mathrm{s}}\frac{\partial\Pi_t}{\partial\rho'}g^\rho{}_\sigma\right\}Q_\nu{}^\sigma\,, \tag{197}$$

where we have dropped the C_s^2/c^2 term for the reasons previously stated, and D_s is to be given by equation (182) with an infinitesimal positive imaginary part to be added to $q\,.\,u$ for reasons of causality.

We also calculate the perturbation in Π_t due to the presence of the wave:

$$\begin{aligned}\delta\Pi_t &= \delta\rho'\frac{\partial\Pi_t}{\partial\rho'}+k\,.\,\delta u\frac{\partial\Pi_t}{\partial k\,.\,u}\\ &= -\int\frac{\mathrm{d}^4 q}{(2\pi)^4}\,\frac{\exp(-\mathrm{i}q\,.\,x)F_q}{\rho_0\,c^2(q\,.\,u)^2D_\mathrm{s}}\left\{q_\perp^2\left(\rho'\frac{\partial\Pi_t}{\partial\rho'}\right)^2-2q\,.\,u\,k_\perp\,.\,q_\perp\,\rho'\frac{\partial\Pi_t}{\partial\rho'}\,\frac{\partial\Pi_t}{\partial k\,.\,u}\right.\\ &\qquad\left.+\left((q\,.\,u)^2D_\mathrm{s}-\frac{C_\mathrm{s}^2(q_\perp\,.\,k_\perp)^2}{c^2}\right)\left(\frac{\partial\Pi_t}{\partial k\,.\,u}\right)^2\right\}.\end{aligned} \tag{198}$$

This perturbation can lead to self-focusing or modulational instability of the envelope, but since it is proportional to the square of the wave amplitude we must, to be consistent, include the effect of nonlinearity in the dielectric response as well.

In order to limit the growth rate of modulational instabilities at large q it is also necessary to include dispersive spreading of the wavepacket, and to do this we change the meaning of a_t slightly by making it the complex amplitude with respect to a carrier wave with *fixed* $k^\mu \equiv k_0^\mu$. That is, we take $\theta = -k_0\,.\,x$ in equation (73) and allow $a^\mu = a_t\,\tau^\mu$ to take up all the variation in the phase due to inhomogeneities. In fact, in an isotropic medium, we may take τ^μ to be real and quite unambiguously assume a_t to carry all the phase variation.

The carrier wave 4-vector k^μ will be assumed to be the solution to the nonlinear dispersion relation for an unmodulated wave with some reference amplitude a_0 (which may be zero). We assume the nonlinear dispersion relation to be the obvious generalization of equation (102)

$$c^2k_0^2 = \Pi_t(k_0 \cdot u, k_0^2) + |a_0|^2 \Pi_{nl}(k_0 \cdot u, k_0^2). \tag{199}$$

This equation must also apply for a wave with different a and k, so, following Karpman and Krushkal' (1969), we multiply (199) on the right by a ($\equiv a_t$), replace a_0 by a and $k_0{}^\mu$ by $k_0{}^\mu + \mathrm{i}\partial^\mu$ and Taylor expand the linear terms to second order in ∂^μ, to find

$$\mathrm{i}\left(2c^2k^\mu - \frac{\partial \Pi_t}{\partial k_\mu}\right)_0 \partial_\mu a = -\tfrac{1}{2}\left(\frac{\partial^2 \Pi_t}{\partial k_\mu \partial k_\nu}\right)_0 \partial_\mu \partial_\nu a + \Pi_{nl}(k_0)(|a|^2 - |a_0|^2)a + \delta\Pi_t a, \tag{200}$$

$\delta\Pi_t$ being given by equation (198), with F_q determined by

$$F_q = \int \mathrm{d}^4x \exp(\mathrm{i}q \cdot x)\, \varepsilon_0 |a_t|^2, \tag{201}$$

In the frame moving at the linear group velocity of the carrier wave, equation (200) is a nonlinear Schrödinger equation for the wave amplitude. This equation should describe self-focusing, modulational instability and stimulated Brillouin scattering. Although the above derivation is rather heuristic, a similar treatment (Dewar 1972*b*) has been verified by reductive perturbation theory (Ichikawa *et al.* 1973).

6. Collisionless Plasma

(*a*) *Hot Plasma*

Since the plasma case is the simplest one where a microscopic derivation of the energy–momentum tensor is possible, it is of considerable interest in a general discussion. However, one cannot characterize a plasma with a continuous distribution of velocities (hot plasma) in the same way as one does a fluid, so we consider here a microscopic derivation based on the covariant distribution function. We assume there to be no strong DC fields, although the methods of O'Sullivan and Derfler (1973) would probably make this case tractable as well.

(*b*) *Covariant Vlasov Equation*

We work in an 8-dimensional phase space whose position coordinate we denote by $X_i = (x^\mu, p^\mu)$, $i = 1, 2, ..., 8$. In this space the particles appear as an ensemble of world lines, as depicted in Fig. 3. Let us denote the average density vector for world lines by

$$\Phi_i = F_\sigma(X)\, \mathrm{d}X_i/\mathrm{d}s, \tag{202}$$

where $F(X)$ is a scalar quantity and $\mathrm{d}s$ is the distance along a world line in 4-space; $F_\sigma(X)$ is in fact the covariant distribution function introduced by Goto (1958) and Klimontovich (1960). Here σ is the species label, which will henceforth be implicit. The normalization is such that the current vector (128b) is given by

$$J^\mu(x) = \sum_\sigma ec \int \mathrm{d}^4p\, F(x, p)\, u^\mu, \tag{203}$$

so that $F(x,p)$ is related to the 3-space distribution function $f(\boldsymbol{x},\boldsymbol{p},t)$ in a given frame by

$$F(x,p) = f(\boldsymbol{x},\boldsymbol{p},t)\,\delta(p^0 - mc\gamma)/\gamma\,. \tag{204}$$

There is a direct analogy between the world line density Φ_i and the magnetic line density $\boldsymbol{B}$. The analogue of $\nabla . \boldsymbol{B} = 0$ is the covariant Vlasov equation

$$\sum_{i=1}^{8} \frac{\partial \Phi_i}{\partial X_i} = u^\mu \frac{\partial F}{\partial x^\mu} + \frac{\partial}{\partial p^\mu}\left(\frac{\mathrm{d}p^\mu}{\mathrm{d}s} F\right) = 0\,. \tag{205}$$

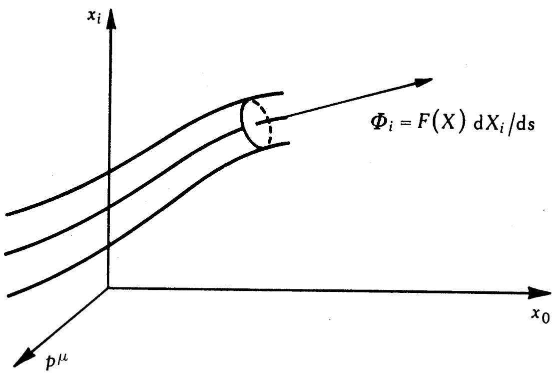

Fig. 3. World lines in 8-dimensional phase space. The 8-vector Φ_i is the line density vector.

(c) *Hamilton's Principle*

We further exploit the analogy between Φ_i and $\boldsymbol{B}$ in order to find the change in F when the world lines are perturbed (or, indeed, when phase space is subjected to any one-to-one mapping). Let the point (x^μ, p^μ) go to $(x^{\mu\dagger}, p^{\mu\dagger})$,

$$x^{\mu\dagger}(X) = x^\mu + \xi^\mu(X)\,, \qquad p^{\mu\dagger}(X) = p^\mu + \pi^\mu(X)\,. \tag{206a, b}$$

Then, just as for $\boldsymbol{B}$ in ideal hydromagnetics (Newcomb 1962), we find the new vector

$$\Phi_i^\dagger = \sum_{j=1}^{8} (\Delta^\dagger)^{-1} (\partial X_i^\dagger / \partial X_j)\,\Phi_j\,, \tag{207a}$$

where

$$\Delta^\dagger \equiv \det(\partial X_i^\dagger / \partial X_j)\,. \tag{207b}$$

Comparing with equation (202) we see that this implies

$$F^\dagger(X^\dagger) = (\Delta^\dagger)^{-1} F(X)\,\mathrm{d}s^\dagger/\mathrm{d}s\,, \tag{208}$$

where

$$\mathrm{d}s^\dagger = (\mathrm{d}x^{\dagger\mu}\,\mathrm{d}x_\mu^\dagger)^{\frac{1}{2}} = \{1 + 2u . \mathrm{D}_s \xi + (\mathrm{D}_s \xi)^2\}^{\frac{1}{2}}\,\mathrm{d}s\,, \tag{209}$$

with

$$\mathrm{D}_s \equiv u^\mu \frac{\partial}{\partial x^\mu} + \frac{\mathrm{d}p^\mu}{\mathrm{d}s}\frac{\partial}{\partial p^\mu}\,.$$

We suppose that the Lagrangian density for a system of fields η_i and particles with distribution function $F(X)$ is given by

$$\mathscr{L} = \sum_{\sigma} \int \mathrm{d}^4 p\, F(X) R(p, \eta, \partial\eta) + \mathscr{L}'(x, \eta, \partial\eta). \tag{210}$$

Comparing with Low's (1958) Lagrangian we see that R is related to the ordinary particle Lagrangian L by $R = \gamma L$. Variation of the particle trajectories may be effected by setting $\xi^\mu = \delta x^\mu$ in equations (206) which implies

$$\delta \int \mathrm{d}^8 X\, FR = \int \mathrm{d}^8 X\, F\{(\mathrm{d}\delta s/\mathrm{d}s)R + \delta R\}. \tag{211}$$

The resulting covariant Lagrange equation is

$$\mathrm{D}_s\{(g^{\mu\nu} - u^\mu u^\nu)\partial R/\partial u^\nu + u^\mu R\} - \partial_\mu R = 0. \tag{212}$$

By setting $R = \gamma L$ one may verify that this equation gives the usual Lagrange equations plus an energy equation. Equation (212) can also be cast in Hamiltonian form as

$$\mathrm{D}_s x^\mu = \partial K/\partial P_\mu, \qquad \mathrm{D}_s P^\mu = -\partial K/\partial x_\mu, \tag{213}$$

where the canonical 4-momentum is defined by

$$P^\mu \equiv -(g^\mu{}_\nu - u^\mu u_\nu)\partial R/\partial u_\nu - u^\mu R \tag{214}$$

and the covariant Hamiltonian is

$$K \equiv \tfrac{1}{2}(P\,.\,u + R). \tag{215}$$

For the derivation of the Hamiltonian see Appendix 2.

(d) Energy–Momentum Tensors

The canonical energy–momentum tensors for the field subsystems k are as defined by equation (16). A similar argument to that employed to derive equation (82) in Section 3 may be used to show that the canonical energy–momentum tensor $T_k{}^{\mu\nu}$ defined with the average Lagrangian density $\mathscr{L}$ is the average of the exact energy–momentum tensor $T_k^{\dagger\mu\nu}$ defined with the exact Lagrangian density $\mathscr{L}^\dagger$.

The energy–momentum tensor for the background plasma defined by the average world lines is

$$T_\mathrm{b}{}^{\mu\nu} = \sum_{\sigma} \int \mathrm{d}^4 p\, F(u^\mu P^\nu + R g^{\mu\nu}) - \mathscr{L}_\mathrm{b} g^{\mu\nu}, \tag{216}$$

where P^ν is the canonical momentum defined by equation (214). Using equations (205) and (212) we can show that the 4-force density acting on the canonical background system is

$$f_\mathrm{b}{}^\nu = \sum_{\sigma} \int \mathrm{d}^4 p\, R\, \partial^\nu F - \partial^\nu \mathscr{L}_\mathrm{b}. \tag{217}$$

Since R is assumed to depend on x^μ only through the fields η_i, it is seen immediately that the force densities sum to zero and the conservation equation (21) is satisfied.

The physical energy–momentum densities are defined by (cf. equation 33)

$$\theta_k^{\mu\nu} = \sum_\sigma \int d^4p\, F(u^\mu P_k^{\ \nu} + R_k g^{\mu\nu}) + \frac{\partial \mathscr{L}_k}{\partial\, \partial_\mu \eta_i}\partial^\nu \eta_i - \mathscr{L}_k g^{\mu\nu} + \partial_\rho f_k^{\ \rho\mu\nu}, \tag{218}$$

where the partial momenta are

$$P_k^{\ \mu} \equiv -(g^\mu{}_\nu - u^\mu u_\nu)\,\partial R_k/\partial u_\nu - u^\mu R_k .$$

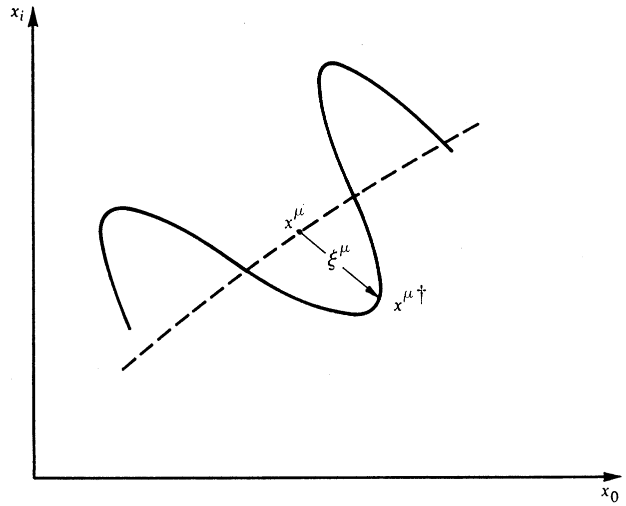

Fig. 4. Representation of the world line $x^{\mu\dagger}(s^\dagger)$ as a motion along the oscillation centre world line $x^\mu(s)$ plus an oscillatory displacement $\xi^\mu(s)$.

(*e*) *Linear Wave Response*

We now distinguish between the fast and slow scales of the wave and its envelope. We identify the displacement $\xi^\mu(x,p)$ in equation (206a) with the oscillatory part of a particle's world line and $x^\mu(s)$ with the smoothed-out oscillation centre (Dewar 1973, 1976) world line, as indicated in Fig. 4. It is natural to require that $\xi^\mu(x,p)$ average to zero, i.e.

$$\langle \xi^\mu(X)\rangle = 0 . \tag{219}$$

For a relativistic plasma we have the exact particle and field Lagrangians (Landau and Lifshitz 1971)

$$R^\dagger = -mc^2 - ceu^\dagger . A^\dagger , \qquad \mathscr{L}'^\dagger = -\tfrac{1}{4}\mu_0^{-1}(\partial_\nu A_\mu^\dagger - \partial_\mu A_\nu^\dagger)(\partial^\nu A^{\dagger\mu} - \partial^\mu A^{\dagger\nu}) . \tag{220}$$

The superscript dagger is used to denote exact quantities, the averaged quantities appearing without the dagger. We now use equations (206) and (207b), expand up to second order in ξ^μ and apply local space–time averaging to obtain

$$\mathscr{L} = \sum_\sigma \int d^4p\, F(X)\langle -mc^2 - ceu . A^\dagger - \tfrac{1}{2}mc^2 D_s \xi_\mu (g^\mu{}_\nu - u^\mu u_\nu) D_s \xi^\nu$$
$$- ce\, D_s \xi^\mu A_\mu^\dagger - ceu^\mu \xi^\nu\, \partial A_\mu^\dagger/\partial x^\nu \rangle + \mathscr{L}' . \tag{221}$$

Now set

$$\xi^\mu = \zeta^\mu \exp(i\theta) + \text{c.c.}, \qquad A^{\dagger\mu} = A^\mu + a^\mu \exp(i\theta) + \text{c.c.}, \tag{222}$$

where $\xi^\mu(X)$ and $a^\mu(x)$ are slowly varying amplitudes and $\theta(x)$ is the wave phase (cf. equation 73). Substituting equations (222) in (221), varying ζ^μ and eliminating it in terms of a^μ, we find

$$\mathscr{L} = \sum_\sigma \int d^4p\, F(X)(-mc^2 - ceu.A + mc^2\tilde{u}^*.\tilde{u})$$

$$- a_\mu^*(k^2 g^\mu{}_\nu - k^\mu k_\nu) a^\nu/\mu_0, \tag{223}$$

where

$$\tilde{u}^\mu \equiv -(e/mc)\{g^\mu{}_\nu - k^\mu u_\nu/(k.u)\} a^\nu. \tag{224}$$

Thus the Lagrangian for the oscillation-centre motion is

$$R = -mc^2 - ceu.A + mc^2 \tilde{u}^*.\tilde{u}. \tag{225}$$

The last term is quadratic in the wave amplitude and gives the radiation force on the particles. Equation (223) may be rearranged in the form (cf. equations 89 and 113)

$$\mathscr{L} = \sum_\sigma d^4p\, F(X)(-mc^2 - ceu.A)$$

$$- a_\mu^*\{k^2 g^\mu{}_\nu - k^\mu k_\nu + \mu_0 \alpha^\mu{}_\nu(k)\} a^\nu/\mu_0, \tag{226}$$

where (cf. equations 91, 95 and 114)

$$\alpha^\mu{}_\nu(k) = -\sum_\sigma \frac{e^2}{m} \int d^4p\, F\left(g^\mu{}_\rho - \frac{u^\mu k_\rho}{k.u}\right)\left(g^\rho{}_\nu - \frac{k^\rho u_\nu}{k.u}\right) \tag{227}$$

is the polarization tensor.

Variation of A^μ yields the constraint (cf. equation 188)

$$\sum_\sigma ec \int d^4p\, F u^\mu = 0. \tag{228}$$

This implies average charge neutrality in all frames.

Variation of the oscillation-centre motion leads to equation (212) with R given by (225). Thus du^μ/ds is implicitly determined and equation (205) may be used to find F.

The physical wave energy–momentum tensor (218) is found to be

$$\theta_w{}^{\mu\nu} = \sum_\sigma mc^2 \int d^4p\, F(\tilde{u}^{\mu*}\tilde{u}^\nu + \tilde{u}^{\nu*}\tilde{u}^\mu - \tilde{u}^*.\tilde{u}\, u^\mu u^\nu)$$

$$+ k^2 \mu_0^{-1}(g^{\mu\nu} a^*.a - a^{\mu*} a^\nu - a^{\nu*} a^\mu - 2a^*.a\, k^\mu k^\nu/k^2). \tag{229}$$

This result may also be obtained directly by averaging the exact energy–momentum tensor.

(f) Wavepackets

As an application of the preceding theory we treat the one-dimensional wavepacket problem of Section 5. Linearizing equation (205), and using equations (212) or (213), we find the perturbation in the oscillation centre distribution function as

$$mc^2\delta F = -\frac{\partial}{\partial u^\mu}\left\{\left((g^{\mu\nu}-u^\mu u^\nu)\frac{\partial\,\delta R}{\partial u^\nu}+u^\mu\,\delta R\right)F_0\right\}$$

$$+\left(u^\mu\frac{\partial\,\delta R}{\partial u^\mu}-\delta R\right)F_0+\frac{q^\mu(\partial F_0/\partial u^\mu)\delta R}{q\,.\,u+\mathrm{i}0}, \qquad (230)$$

where δ denotes the difference between a quantity in the presence of the wave and in its absence. Using $\delta R = mc^2\,\tilde{u}^*.\tilde{u} - ceu.A$ in equation (230), to evaluate the perturbation in the background physical energy–momentum tensor, and adding equation (229) we find

$$\delta\theta^{\mu\nu} = \sum_\sigma mc^2\int \mathrm{d}^4p\; F\left\{\tilde{u}^{\mu *}\,\tilde{u}^\nu - \frac{k\,.\,\tilde{u}}{k\,.\,u}\left(u^\mu\,\tilde{u}^{\nu *}+u^\nu\,\tilde{u}^{\mu *}\right)-\frac{k\,.\,\tilde{u}^*}{k\,.\,u}\left(u^\mu\,\tilde{u}^\nu+u^\nu\,\tilde{u}^\mu\right)\right\}$$

$$+\sum_\sigma\int \mathrm{d}^4p\; u^\mu u^\nu q^\rho\frac{(\partial F/\partial u^\rho)\delta R}{q\,.\,u+\mathrm{i}0}+\theta'^{\mu\nu}, \qquad (231)$$

where $\theta'^{\mu\nu}$ is the last term in equation (229) and A^μ is to be found from equations (228) and (230), which give

$$\alpha^\mu{}_\nu(q)\,A^\nu = \sum_\sigma ec\int \mathrm{d}^4p\;(u^\mu\,u^{\nu *}+u^{\mu *}\,u^\nu)\frac{k_\nu F}{k\,.\,u}$$

$$-\tilde{u}^*.\,\tilde{u}\,u^\mu\frac{q\,.(\partial F/\partial u)}{q\,.\,u+\mathrm{i}0}. \qquad (232)$$

The resonance at $q\,.\,u = 0$ corresponds to particles having a velocity component in resonance with the group velocity, that is, to nonlinear Landau damping (Dewar 1972*b*). The presence of the polarization tensor in equation (232) also means that ion acoustic resonance can occur, in a similar manner to that described in Section 5. It can be verified that taking delta function distribution functions in equations (231) and (232) reproduces the cold plasma results of Section 5.

7. Discussion

We have seen that the present variation model provides a unifying description of dispersive waves in all three states of matter, with minimal specific assumptions about the material in question. The covariant reference map method introduced in Section 2 is believed to be a technical innovation in the calculus of variations which allows one to take over in a very direct manner the techniques of classical relativistic field theory and to apply them to continuum mechanics, thus exposing the essential unity of physics. Equations (33)–(36) and equations (17) and (38)–(41) provide two very general, if rather abstract, bases for assigning energy and momentum to subsystems. Special emphasis has been placed on the non-uniqueness of this splitting-up procedure, as it is particularly striking when we add the essentially novel element of the paper, namely the presence of a dispersive wave. As a 'bonus' this framework has allowed us to cast new light on the old Abraham–Minkowski controversy.

Because this work bears on several branches of physics normally regarded as disparate: field theory, the physics of magnetic and dielectric materials and plasma physics; it has been considered appropriate to include some fairly specific working of representative special cases. Thus we have examined the isotropic fluid case, both with essentially DC fields and with a high frequency field, and have looked at both hot and cold plasmas. This juxtaposition of examples has been pedagogically rewarding, though it has led to this being rather a long paper.

The new results in 3-vector form, are essentially embodied in equations (168)–(171). A large amount of new material in covariant form has also been included. The covariant treatment of linear dispersive waves in Sections 3*e*–3*g* is rather novel and, although this makes the results rather inaccessible to the average reader, it is noted that covariant methods have recently been used in such fields as relativistic electron beam technology. Also Section 4*a* spells out very specifically how to 'translate' the results to 3-vector form.

In Section 5 we have taken a fresh look at the wavepacket problem and have shown how the general covariant result (183) agrees with previous more specialized derivations. We have also derived a new covariant generalization of the nonlinear Schrödinger equation (200) for treating the modulational stability of wavetrains.

The present variational method has allowed us to cover a vast territory fairly superficially. For the treatment of specific problems in depth it must be admitted that the method has some serious limitations, the most serious of which being the restriction to nondissipative systems. Thus, for instance, the treatment of a collisionless plasma in Section 6 is seriously hampered by the inability to treat Landau damping. However, just as Lagrangian theory historically paved the way for the more powerful Hamiltonian mechanics, it is hoped that the canonical transformation methods currently being developed (Dewar 1976) will greatly extend the scope of the general theory of the interaction between waves and matter.

Acknowledgments

It is a pleasure to thank Professor Sir R. Peierls, Dr K. Kumar, Dr D. B. Melrose, Dr M. Siegrist and Dr L. J. Tassie for useful conversations.

References

Burt, M. G., and Peierls, R. (1973). *Proc. R. Soc. London* A **333**, 149.
Dewar, R. L. (1970). *Phys. Fluids* **13**, 2710.
Dewar, R. L. (1972*a*). *Astrophys. J.* **174**, 301.
Dewar, R. L. (1972*b*). *J. Plasma Phys.* **7**, 267.
Dewar, R. L. (1973). *Phys. Fluids* **16**, 1102.
Dewar, R. L. (1976). *J. Phys.* A **9**, 2043.
Dougherty, J. P. (1970). *J. Plasma Phys.* **4**, 761.
Dougherty, J. P. (1974). *J. Plasma Phys.* **11**, 331.
Eringen, A. C. (1967). 'Mechanics of Continua' (Wiley: New York).
Furutsu, K. (1969). *Phys. Rev.* **185**, 257.
Goto, K.-i. (1958). *Prog. Theor. Phys.* **20**, 1.
de Groot, S. R., and Suttorp, L. G. (1972). 'Foundations of Electrodynamics' (North-Holland: Amsterdam).
Haus, H. A. (1969). *Physica* **43**, 77.
von Helmholtz, H. (1882). *Wied. Ann.* **13**, 798.
Hill, E. L. (1951). *Rev. Mod. Phys.* **23**, 253.
Hora, H. (1969). *Phys. Fluids* **12**, 182.
Hora, H. (1974). *Phys. Fluids* **17**, 1042.
Ichikawa, Y. H., Suzuki, T., and Taniuti, T. (1973). *J. Phys. Soc. Jpn* **34**, 1089.

Jones, W. L. (1971). *Rev. Geophys. Space Phys.* **9**, 917.
Karpman, V. I., and Krushkal', E. M. (1969). *Sov. Phys. JETP* **28**, 277.
Klima, R., and Petrzilka, V. A. (1968). *Czech. J. Phys.* B **18**, 1292.
Klima, R., and Petrzilka, V. A. (1973). *Phys. Lett.* A **43**, 151.
Klima, R., and Petrzilka, V. A. (1975). *J. Phys.* A **8**, 829.
Klimontovich, Yu. L. (1960). *Sov. Phys. JETP* **10**, 524.
Kroll, N. A. (1965). *J. Appl. Phys.* **36**, 34.
Landau, L. D., and Lifshitz, E. M. (1960). 'Electrodynamics of Continuous Media', 1st English edn (Pergamon: London).
Landau, L. D., and Lifshitz, E. M. (1971). 'The Classical Theory of Fields', 3rd revised English edn (Pergamon: London).
Lianis, G. (1974). *Arch. Ration. Mech. Anal.* **55**, 300.
Low, F. E. (1958). *Proc. R. Soc. London* A **248**, 282.
Melrose, D. B. (1973). *Plasma Phys.* **15**, 99.
Newcomb, W. A. (1962). *Nucl. Fusion Suppl.* Pt 2, 451.
O'Dell, T. H. (1970). 'The Electrodynamics of Magneto-electric Media' (North-Holland: Amsterdam).
O'Sullivan, R. A., and Derfler, H. (1973). *Phys. Rev.* A **8**, 2645.
Panofsky, W. K. H., and Phillips, M. (1962). 'Classical Electricity and Magnetism', 2nd edn (Addison–Wesley: Reading, Mass.).
Pauli, W. (1941). *Rev. Mod. Phys.* **13**, 203.
Pauli, W. (1958). 'Theory of Relativity' (Pergamon: London).
Peierls, R. (1976). *Proc. R. Soc. London* A **347**, 475.
Penfield, P., Jr, and Haus, H. A. (1967). 'Electrodynamics of Moving Media' (M.I.T. Press).
Robinson, F. N. H. (1975). *Phys. Rep.* **16**, 313.
Schöpf, H. G. (1964). *Ann. Phys. (Leipzig)* **13**, 41.
Stix, T. H. (1962). 'The Theory of Plasma Waves' (McGraw–Hill: New York).
Toupin, R. A. (1960). *Arch. Ration. Mech. Anal.* **5**, 440.
Tsytovich, V. N. (1970). 'Nonlinear Effects in a Plasma', pp. 80–2 (Plenum: New York).
Whitham, G. B. (1965). *J. Fluid Mech.* **22**, 273.

Appendix 1

Except in Section 6 we have adopted a phenomenological approach, and have not been much concerned with the relation between the microscopic fluctuating Lagrangian density (or energy–momentum tensor) and the averaged macroscopic quantity. Also we have worked only to lowest order in a WKB expansion. In the following we consider the rudiments of a formal theory of averaging which represents a considerable extension and simplification over the author's previous attempt in this direction (Dewar 1970), and helps to place the preceding work in a deeper perspective.

Inherent in the discussion is the assumption that there are two time and length scales, which would be represented, in the stretched coordinate method, by a fast oscillatory scale x^μ and a slow scale εx^μ, where ε is the WKB expansion parameter. However, we shall not explicitly exhibit the dependence on these scales; if something depends only on the slow scale we shall indicate this by calling it slowly varying.

As in previous discussion (Dewar 1970), we shall appeal to the Riemann–Lebesgue lemma to justify our assumption that, corresponding to a function $f^\dagger(x)$ varying on both the fast and slow scales, there exists a slowly varying function $f(x) = \langle f^\dagger(x)\rangle$ such that

$$\int f^\dagger(x)\,\phi(x)\,\mathrm{d}^4x \sim \int f(x)\,\phi(x)\,\mathrm{d}^4x \tag{A1}$$

for all slowly varying test functions $\phi(x)$, where the tilde denotes asymptotic equality to arbitrary order in the ε expansion. The function $f(x)$ is called the *local average* of $f^{\dagger}(x)$, and is unique to all orders in ε.

If $\phi(x)$ is slowly varying then so is $\partial_{\mu}\phi(x)$, which is therefore a valid test function in equation (A1). Using this test function and integrating by parts (assuming ϕ sufficiently localized to vanish near the boundaries), we prove the following lemma.

Lemma 1. Averaging and differentiation commute to all orders in ε; that is,

$$\langle \partial_{\mu} f^{\dagger}(x)\rangle \sim \partial_{\mu}\langle f^{\dagger}(x)\rangle = \partial_{\mu} f(x). \tag{A2}$$

Thus, for example, if $T^{\dagger\mu\nu}$ obeys a conservation equation then so does its average $T^{\mu\nu}$.

As a further example suppose we wish to average the mass current $c\rho^{\dagger}u^{\dagger\mu}$ and rest mass energy–momentum tensor $c^2\rho^{\dagger}u^{\dagger\mu}u^{\dagger\nu}$ in the presence of a wavelike disturbance $\xi^{\mu}(x)$ averaging to zero (cf. Fig. 4). We might also want the average of the corresponding Lagrangian density $-\rho^{\dagger}c^2$.

The exact and average world lines are connected by a mapping of the form of equation (1) in Section 2, with X^{μ} corresponding to the exact position $x^{\dagger\mu}$,

$$x^{\dagger\mu} = x^{\mu}+\xi^{\mu}. \tag{A3}$$

Admittedly the interpretation of X^{μ} has changed, and the regions $\mathscr{R}$ and $\mathscr{R}^{\dagger}$ are no longer disjoint, but equations (46) and (48) of Section 3 remain valid, and become

$$u^{\dagger\mu}(x+\xi) = (u^{\mu}+u.\partial\xi^{\mu})\{(u^{\sigma}+u.\partial\xi^{\sigma})(u_{\sigma}+u.\partial\xi_{\sigma})\}^{-\frac{1}{2}}, \tag{A4}$$

$$\rho^{\dagger}(x+\xi) = \{(u^{\sigma}+u.\partial\xi^{\sigma})(u_{\sigma}+u.\partial\xi_{\sigma})\}^{\frac{1}{2}}\rho'(x)/\det(\partial^{\alpha}x_{\beta}), \tag{A5}$$

where all implicit dependences are on x^{μ}. There is no need to solve for $u^{\dagger\mu}(x)$ and $\rho^{\dagger}(x)$, nor to expand the determinant, as we may use the following theorem.

Theorem

$$\langle f^{\dagger}(x)\rangle \sim \langle \Delta(x) f^{\dagger}(x+\xi)\rangle - \partial_{\mu}\langle \xi^{\mu}\Delta(x) f^{\dagger}(x+\xi)\rangle + (1/2!)\partial_{\mu}\partial_{\nu}\langle \xi^{\mu}\xi^{\nu}\Delta(x) f^{\dagger}(x+\xi)\rangle - \ldots, \tag{A6}$$

where

$$\Delta(x) \equiv \det(\partial^{\alpha}x_{\beta}). \tag{A7}$$

Clearly, each successive term is $O(\varepsilon)$ smaller than the preceding one. To prove the theorem, label the dummy integration variable in equation (A1) with a dagger. Now change variables from $x^{\dagger}$ to x, using equations (A3) and (A7), and Taylor expand $\phi(x+\xi)$ to obtain

$$\int d^4x\, f^{\dagger}(x)\,\phi(x) = \int d^4x\, \Delta(x) f^{\dagger}(x+\xi)\{\phi(x)+\xi.\partial\phi(x)+(1/2!)\xi\xi:\partial\partial\phi+\ldots\}. \tag{A8}$$

Integration by parts of the terms on the right-hand side, and use of Lemma 1 above, proves the theorem.

Use of the theorem and equations (A4) and (A8), in that order, shows that to lowest order in the WKB expansion

$$\langle \rho^\dagger \rangle \sim \langle \{(u^\sigma + u \,.\, \partial \xi^\sigma)(u_\sigma + u \,.\, \partial \xi_\sigma)\}^{\frac{1}{2}} \rangle \rho', \tag{A9}$$

$$\langle \rho^\dagger u^{\dagger\mu} \rangle \sim \rho' u^\mu, \tag{A10}$$

$$\langle \rho^\dagger u^{\dagger\mu} u^{\dagger\nu} \rangle \sim \langle (u^\mu + u \,.\, \partial \xi^\mu)(u^\nu + u \,.\, \partial \xi^\nu)\{(u^\sigma + u \,.\, \partial \xi^\sigma)(u_\sigma + u \,.\, \partial \xi_\sigma)\}^{-\frac{1}{2}} \rangle \rho', \tag{A11}$$

where we have used the following lemma.

Lemma 2. If g is slowly varying, then

$$\langle f^\dagger g \rangle \sim \langle f^\dagger \rangle g \tag{A12}$$

to all orders in ε.

Lemma 2 follows by replacing ϕ by $g\phi$ in equation (A1). We can also use the continuity relation (126) of Section 4 and Lemma 1 to show

$$\langle \rho' u \,.\, \partial \xi^\mu \rangle \sim \partial_\sigma \langle \rho' u^\sigma \xi^\mu \rangle . \tag{A13}$$

This term arises in equation (A10) but has been omitted because it is $O(\varepsilon)$ smaller than the leading term.

These results can be used, for instance, in obtaining a microscopic derivation of the cold plasma Lagrangian density, but here we simply use them to point out the physical meaning of the background proper density ρ' and background 4-velocity u^μ. From equation (A9), ρ' is clearly *not* the average of the exact density $\rho^\dagger$. Rather u^μ is the unit vector in the direction of $\langle \rho^\dagger u^{\dagger\mu} \rangle$, and ρ' is the magnitude of this vector.

Appendix 2

As the covariant Lagrange equations of motion (212) in Section 6 are unusual, the derivation of the covariant Hamiltonian (215) is not covered in the standard textbooks, and accordingly the derivation is presented here.

First note that the canonical 4-momentum

$$P^\mu = -(g^\mu{}_\nu - u^\mu u_\nu) \partial R / \partial u_\nu - u^\mu R \tag{A14}$$

is closely related to the standard canonical 4-momentum. Specifically, using

$$R = \gamma L(\boldsymbol{v}, \boldsymbol{x}, t) \tag{A15}$$

with

$$\gamma = u_0, \qquad \boldsymbol{v} = c\boldsymbol{u}/u_0, \tag{A16}$$

we can readily show that

$$P^\mu = \begin{bmatrix} \boldsymbol{v} \,.\, (\partial L / \partial \boldsymbol{v}) - L \\ \\ c\, \partial L / \partial \boldsymbol{v} \end{bmatrix} . \tag{A17}$$

Thus P^μ differs by a factor of c from the standard definition. This is because we use the interval $c\gamma\, \mathrm{d}t$ as our 'time' interval. The equivalence between equations (212), which we rewrite as

$$\mathrm{d}P^{\mu}/\mathrm{d}s = -\partial R/\partial x_{\mu}, \tag{A18}$$

and the standard Lagrange equations is now apparent. Also it can be shown that equation (A18) automatically conserves the normalization

$$u^2 = 1 \tag{A19}$$

for all physical motions. Contracting equation (A14) with u_{μ} we find

$$u.P + R = 0. \tag{A20}$$

Also, by rearranging equation (A14) and using (A20), we have

$$\partial R/\partial u_{\mu} = -(g^{\mu\nu} - u^{\mu}u^{\nu})P_{\nu} + u^{\mu}\,u.(\partial R/\partial u). \tag{A21}$$

From equations (A18), (A20) and (A21) we obtain

$$\mathrm{d}R = -\mathrm{d}x.(\mathrm{d}P/\mathrm{d}s) - P.\mathrm{d}u + \{u.(\partial R/\partial u) - R\}u.\mathrm{d}u \tag{A22}$$

for all differential displacements away from a physical motion. The last term in equation (A22) can be eliminated by using the relation following from (A14) and (A19), namely

$$u.\mathrm{d}P = \{u.(\partial R/\partial u) - R\}u.\mathrm{d}u - \mathrm{d}x.(\partial R/\partial x), \tag{A23}$$

and we find

$$\mathrm{d}R = -2\,\mathrm{d}x.(\mathrm{d}P/\mathrm{d}s) + u.\mathrm{d}P - P.\mathrm{d}u. \tag{A24}$$

By use of equation (A24) it is then apparent that

$$\mathrm{d}\{\tfrac{1}{2}(P.\mathrm{d}u + R)\} = u.\mathrm{d}P - (\mathrm{d}P/\mathrm{d}s).\mathrm{d}x \tag{A25}$$

for all displacements away from a physical motion. Referring now to equation (212), we see that the left-hand side of (A25) is just

$$\mathrm{d}K \equiv (\partial K/\partial P).\mathrm{d}P + (\partial K/\partial x).\mathrm{d}x. \tag{A26}$$

Equating the coefficients of the differentials in equations (A25) and (A26) proves the Hamiltonian equations of motion (213).

Manuscript received 28 February 1977