Pergamon

J. Mech. Phys. Solids, Vol. 43, No. 6, pp. 973–1001, 1995
Copyright © 1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0022–5096/95 $9.50+0.00

0022–5096(95)00009–7

REQUIREMENTS OF THERMODYNAMICS IN THE ANALYSIS OF ELASTIC–PLASTIC SHOCK WAVES

R. M. BRANNON, W. J. DRUGAN† and Y. SHEN

Department of Engineering Mechanics and Astronautics, University of Wisconsin,
1415 Johnson Drive, Madison, WI 53706, U.S.A.

(*Received* 15 *April* 1994)

ABSTRACT

Thermodynamical requirements on elastic–plastic shock waves are investigated to explore the range of validity of, and provide rigorous substantiation for, the previous shock analyses of Drugan and coworkers. These studies assumed (implicitly) that during shock passage, a material particle's stress and deformation history is well-approximated by its history during passage of a smooth wave, and that the material response is purely mechanical. We show precisely the conditions under which these analyses are valid. Courant and Friedrichs' [(1948) *Supersonic Flow and Shock Waves* (Third Printing: 1985). Springer, Berlin] analysis of the order of entropy effects for weak one-dimensional shocks in mechanically conservative fluids is extended to shocks in *general three-dimensional large deformations* in a material of *arbitrary constitution*. Specifically, we prove that the change in thermodynamic state across a suitably-chosen smooth wave coincides with that across a general shock up until third order in material time rates of fundamental field variables, at which point contributions from the shock itself first appear. This result, which is valid even if entropy generation (due to mechanical dissipation) occurs at first or second order, corrects the common misconception that a shock may be approximated by a smooth wave only if the entropy generation is small. We further prove that for the special class of shocks that propagate under steady-state conditions with non-rotating reference configuration images, a smooth wave can be constructed whose change in thermodynamic state coincides with that across the shock through *all* orders of field variable rates. That is, a smooth wave is a potentially *exact* model of a shock in this class. Having legitimized the representation of a shock by a smooth wave, a large deformation statement of the maximum plastic work inequality is integrated across the shock to give general thermomechanical existence conditions for steady shocks. These conditions reduce to those of Drugan and Rice [(1984) Restrictions on quasi-statically moving surfaces of strong discontinuity in elastic–plastic solids. In *Mechanics of Material Behavior* (ed. G. J. Dvorak and R. T. Shield), pp. 59–73. Elsevier Science, Amsterdam] and Drugan and Shen [(1987) Restrictions on dynamically propagating surfaces of strong discontinuity in elastic–plastic solids. *J. Mech. Phys. Solids* **35**, 771–787; (1990) Finite deformation analysis of restrictions on moving strong discontinuity surfaces in elastic–plastic materials: quasi-static and dynamic deformations. *J. Mech. Phys. Solids* **38**, 553–574] whenever thermomechanical coupling is neglected (i.e. when the thermal deformation coefficients or the jump in temperature is neglected, or if the jump in strain has a zero inner product with the thermal deformation coefficient tensor); specific situations where such a simplification is sensible are outlined.

1. INTRODUCTION

This work extends the pioneering one-dimensional inviscid fluid analysis of Courant and Friedrichs (1948) to general three-dimensional deformations and *arbitrary materials* to show that the sequence of states attained in a smooth wave follows the shock Hugoniot until third order in material time rates of fundamental field quantities, *even if entropy generation occurs at first or second order.* Courant and Friedrichs' analysis, which concludes that entropy changes do not occur until third order, is apparently the source of a common, but erroneous, assertion that a shock may be approximated by a smooth wave only if the entropy generation is not too great. In addition to assuming the material is an inviscid fluid, Courant and Friedrichs implicitly assume the material is mechanically conservative (i.e. the dissipation for smooth motions is zero). Given this assumption, it is not surprising that the entropy due to the shock itself does not occur until higher order in a smooth adiabatic approximation of the shock. We prove the following generalization of Courant and Friedrichs' (1948) analysis: *The change in entropy across a shock, in a general three-dimensional deformation of arbitrary material, will agree until third order (i.e. through second order) with the change in entropy for an appropriately constructed smooth motion between the same end states—even if entropy generation occurs at first or second order. Furthermore, the agreement will be through all orders whenever the shock propagates under steady-state conditions with a non-turning reference configuration image.*

---

†To whom correspondence should be addressed.

Having legitimized the representation of a shock by a smooth wave (approximately for low to moderate shock strengths for general non-steady shocks, and possibly exactly for steady-state shocks with non-rotating reference configuration images), we then derive restrictions on thermoelastic–plastic propagating jump discontinuities under steady-state (or nearly so) conditions. Thermoelastic–plastic kinematics and constitutive laws are reviewed in Section 4, and in Section 5 we identify specific thermodynamical mechanisms that may permit the existence of an elastic–plastic shock that would otherwise be ruled out by a purely mechanical analysis. It is well-known that the second law in jump form restricts the class of admissible discontinuities (Lax, 1973). The smooth form of the second law requires that the mechanical dissipation be non-negative, but the maximum plastic work inequality (MPWI) requires that, in general, the dissipation must be *sufficiently* non-negative. That is, the MPWI requires that the dissipation associated with the actual stress state must exceed or equal the dissipation associated with *any* stress state within or on the yield surface. Therefore, in many instances, the MPWI is a stronger restriction on the motion than the second law and may consequently lead to stronger shock admissibility criteria. By enforcing an integrated form of the maximum plastic work inequality, we show that the existence of an elastic–plastic shock often depends critically on thermomechanical coupling. That is, if the jump in temperature is neglected, or if the thermal expansion coefficient (in the anisotropic case, the thermal deformation coefficient tensor) is neglected, or if the jump in strain has a zero inner product with the thermal deformation coefficient tensor, then a shock may be impossible or, at best, may be restricted to propagate only at an elastic wave speed.

One of our primary motivations for the present study is to attempt to provide a *rigorous understanding of when it is legitimate to permit propagating jump discontinuities in solutions of elastic–plastic boundary-value problems when the governing equations employed are purely mechanical.* As explained by Courant and Friedrichs (1948), in actual materials, large gradients in velocity and temperature produce marked effects of friction and heat conduction, causing irreversible thermodynamic processes and entropy production. In such a large gradient region, purely mechanical constitutive equations must be generalized. Such regions are usually very narrow physically; thus, an alternative to generalizing the mechanical constitutive equations is to approximate the narrow, but finite, transition region by a mathematically sharp jump discontinuity.

Entropy production by a shock shows up as a modification to the first law of thermodynamics (energy balance) as compared to the first law's form for a smooth wave in a purely mechanical model. We argue [as did Courant and Friedrichs (1948)] that when conditions are such that the entropy production due to the shock itself is very small, that is, when the change in the first law produced by the shock is negligible in comparison to a smooth wave, then the stress and deformation paths experienced by a material particle during its passage through the shock are the same as when the particle passes through a smooth wave having the same start and end states as the shock. Thus, in such cases we derive restrictions on shock behavior by analyzing permissible stress and deformation paths through a smooth wave characterized by purely mechanical governing differential equations.

Our contention is that a propagating surface of jump discontinuity in the solution of a purely mechanical elastic–plastic boundary-value problem is in accord with the underlying laws of thermodynamics, and hence is physically acceptable, *only* if the propagating discontinuity satisfies the restrictions we derive herein. This concept is analogous to the fact that any physically correct elastic–plastic solution must involve only non-negative plastic work rate, as also required by the underlying laws of thermodynamics.

2. THE THERMOMECHANICAL FIELD LAWS

This section outlines the thermomechanical field laws in their local and jump forms. Discussion of thermoelastic–plastic constitutive laws is postponed until after the analysis of Section 3, which legitimizes the approximation of a weak shock by a smooth wave regardless of the constitution of the material.

Adopting a modified version of the "rational thermomechanics" theory of Coleman and Noll (1963), temperature and entropy are assumed to exist for non-equilibrium states (so that rate forms of the governing equations are sensible) and the principal thermodynamical restriction imposed on the constitutive equations is the Clausius–Duhem Inequality. The "rational" approach is justified within the axiomatic foundations of continuum thermomechanics (see, e.g. Gurtin and Williams, 1966, 1967), though there remain some unresolved objections against the physical foundations of the theory (e.g. Jou *et al.*, 1988). For example, temperature and entropy and, therefore, rate forms of the classical thermodynamical laws, are not well-defined during extremely transient processes such as explosions. However, Wallace (1991), who obtains good agreement with experiments using the "rational" approach, argues that rate forms of the thermodynamical laws are sensible for very fast processes such as shocks if (as with metals) the characteristic speed at which a material approaches equilibrium is very fast. Alternatively, Kestin and Rice (1970) and Bataille and Kestin (1975) suggest that entropy and temperature may be defined at a non-equilibrium state if this state can be associated with a fictitious accompanying constrained equilibrium state.

2.1. *Local (smooth) field laws*

Let $\Omega(t)$ denote a Lagrangian control volume at time t. Let Ω_0 denote a reference configuration defined so that points $\mathbf{x}$ in Ω are mapped from points $\mathbf{X}$ in Ω_0 via a piecewise continuous deformation gradient tensor $\mathbf{F} = (\partial\mathbf{x}/\partial\mathbf{X})_t$. Here and throughout this paper, variables held constant when taking partial derivatives are indicated by subscripts.

In regions where all field variables have continuous first derivatives in both space and time, the local forms of conservation of mass, conservation of linear momentum, the first law of thermodynamics, and the Clausius–Duhem Inequality (CDI) are, respectively,

$$\dot{\rho}+\rho\,\vec{\nabla}\cdot\mathbf{v}=0 \quad \text{or} \quad \rho_0=\rho J, \tag{2.1}$$

$$\dot{\mathbf{v}}-\mathbf{b}=\frac{1}{\rho_0}\vec{\nabla}_0\cdot\mathbf{t}=\frac{1}{\rho}\vec{\nabla}\cdot\boldsymbol{\sigma}, \tag{2.2}$$

$$\dot{e}=\mathscr{P}_M+\mathscr{P}_T \quad \text{or} \quad \dot{u}=\mathscr{P}_s+\mathscr{P}_T, \tag{2.3}$$

$$\dot{s}-\left(\frac{r}{T}\right)+\frac{1}{\rho}\vec{\nabla}\cdot\left(\frac{\mathbf{q}}{T}\right)=\dot{s}-\left(\frac{r}{T}\right)+\frac{1}{\rho_0}\vec{\nabla}_0\cdot\left(\frac{\mathbf{Q}}{T}\right)\geqslant 0, \tag{2.4}$$

where

$$\mathscr{P}_M\equiv\mathbf{b}\cdot\mathbf{v}+\frac{1}{\rho_0}\vec{\nabla}_0\cdot(\mathbf{t}\cdot\mathbf{v})=\mathbf{b}\cdot\mathbf{v}+\frac{1}{\rho}\vec{\nabla}\cdot(\boldsymbol{\sigma}\cdot\mathbf{v}), \tag{2.5}$$

$$\mathscr{P}_s\equiv\mathscr{P}_M-\dot{\mathscr{k}}=\frac{1}{\rho_0}\mathbf{t}^{\mathrm{T}}:\dot{\mathbf{F}}=\frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{D}, \tag{2.6}$$

$$\mathscr{P}_T\equiv r-\frac{1}{\rho_0}\vec{\nabla}_0\cdot\mathbf{Q}=r-\frac{1}{\rho}\vec{\nabla}\cdot\mathbf{q}, \tag{2.7}$$

$$\mathscr{k}\equiv\tfrac{1}{2}\mathbf{v}\cdot\mathbf{v}. \tag{2.8}$$

Here, ρ is the mass density; ρ_0 is the initial mass density (when $\mathbf{F}=\mathbf{I}$, the identity tensor); $\mathbf{v}$ is the material velocity; J is the determinant of $\mathbf{F}$; $\mathbf{b}$ is the body force per unit mass; $\boldsymbol{\sigma}$ is the Cauchy stress; $\mathbf{t}$ is the right nominal (or First Piola–Kirchhoff) stress ($\mathbf{t}=J\mathbf{F}^{-1}\cdot\boldsymbol{\sigma}$, where $\mathbf{F}^{-1}$ is the inverse of $\mathbf{F}$); e and u are the specific energy and the specific internal energy, respectively (related by $u=e-\mathscr{k}$, where $\mathscr{k}$ is the kinetic energy per unit mass); s is the specific entropy; T is the absolute temperature; r is the heat source per unit mass; $\mathbf{q}$ is the heat flux and $\mathbf{Q}$ the nominal heat flux ($\mathbf{Q}=J\mathbf{F}^{-1}\cdot\mathbf{q}$); and $\mathbf{D}$ is the symmetric part of $\dot{\mathbf{F}}\cdot\mathbf{F}^{-1}$. The scalars $\mathscr{P}_M$, $\mathscr{P}_s$, $\mathscr{P}_T$ are, respectively, the mechanical power, the stress power, and the thermal heating power, each per unit mass. The operations $\vec{\nabla}\cdot(\,)$ and $\vec{\nabla}_0\cdot(\,)$ denote the spatial and reference divergence, respectively; a superposed dot denotes the material time rate [i.e. for any field variable ψ, $\dot{\psi}\equiv(\partial\psi/\partial t)_{\mathrm{X}}$]; the raised dot $(\cdot)$ represents the vector inner product [e.g. $(\boldsymbol{\sigma}\cdot\mathbf{v})_i = \sigma_{ij}v_j$]; and the double dot (:) denotes the tensor inner product [e.g. $\boldsymbol{\sigma}:\mathbf{D} = \sigma_{ij}D_{ij}$]. (Latin subscripts denote Cartesian components, and the summation convention for repeated subscripts is employed.)

Applying the product rule for the divergence of $\mathbf{q}/T$ and using the first law, the CDI may be written in the alternative form,

$$\zeta + \frac{\mathscr{D}}{T} \geqslant 0, \tag{2.9}$$

where ζ is an innominate measure of the alignment between the heat flux vector and the temperature gradient defined as

$$\zeta = -\frac{1}{\rho T^2}\mathbf{q}\cdot(\vec{\nabla}T) = -\frac{1}{\rho_0 T^2}\mathbf{Q}\cdot(\vec{\nabla}_0 T), \tag{2.10}$$

and the mechanical dissipation $\mathscr{D}$ is given by either of the following equivalent expressions

$$\mathscr{D} = T\dot{s} - \mathscr{P}_T, \tag{2.11a}$$

$$\mathscr{D} = T\dot{s} + \mathscr{P}_s - \dot{u}. \tag{2.11b}$$

If it is assumed or proved that heat always flows from hot to cold (i.e. if $\zeta \geqslant 0$) and if the mechanical dissipation $\mathscr{D}$ is assumed to be independent of the temperature gradient, then the principal restriction imposed by the CDI becomes simply $\mathscr{D} \geqslant 0$.

2.2. *Jump forms of the governing thermomechanical laws*

Suppose, as illustrated in Fig. 1, there is a moving surface of discontinuity Σ, which we term a "shock," across which one or more field variables jump and away from which all field variables have finite continuous first derivatives in both space and time.

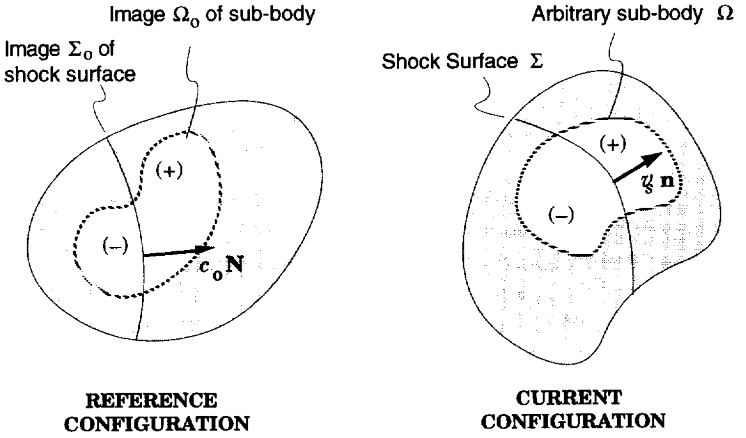
Image Ω_0 of sub-body
Image Σ_0 of shock surface
(+)
(−)
$c_0\mathbf{N}$
REFERENCE CONFIGURATION
Arbitrary sub-body Ω
Shock Surface Σ
(+)
(−)
$\mathscr{c}\,\mathbf{n}$
CURRENT CONFIGURATION

Fig. 1. A surface of discontinuity Σ moving through an arbitrary sub-body Ω; Σ_0 and Ω_0 are the reference images of Σ and Ω, respectively. The vectors indicate the absolute normal propagation velocity of each discontinuity surface as seen by a fixed observer.

The jump of any field variable ψ is denoted by double brackets, defined by

$$\llbracket \psi \rrbracket \equiv \psi^{+} - \psi^{-}, \tag{2.12}$$

where ψ^{+} and ψ^{-} are the limiting values of ψ just ahead of and just behind the shock, respectively. Note that $\llbracket \psi \rrbracket$ is the negative of the temporal jump in ψ. A subscript "A" denotes the average of the $(+)$ and $(-)$ side limiting values:

$$\psi_{A} \equiv \frac{\psi^{+} + \psi^{-}}{2}. \tag{2.13}$$

The jump forms of conservation of mass, conservation of linear momentum, the first law of thermodynamics, and the Clausius–Duhem Inequality are respectively (see, e.g. Chadwick, 1976):

$$\llbracket \rho c \rrbracket = 0 \quad \text{or} \quad \llbracket \rho_0 c_0 \rrbracket = 0, \tag{2.14}$$

$$\llbracket \mathbf{v} \rrbracket = -\frac{\mathbf{N}\cdot\llbracket \mathbf{t} \rrbracket}{\rho_0 c_0} = -\frac{\mathbf{n}\cdot\llbracket \boldsymbol{\sigma} \rrbracket}{\rho c}, \tag{2.15}$$

$$\llbracket e \rrbracket = \llbracket w_M \rrbracket + \llbracket w_T \rrbracket \quad \text{or} \quad \llbracket u \rrbracket = \llbracket w_s \rrbracket + \llbracket w_T \rrbracket, \tag{2.16}$$

$$\llbracket s \rrbracket - \frac{1}{\rho c}\left\llbracket \frac{\mathbf{q}}{T} \right\rrbracket\cdot\mathbf{n} = \llbracket s \rrbracket - \frac{1}{\rho_0 c_0}\left\llbracket \frac{\mathbf{Q}}{T} \right\rrbracket\cdot\mathbf{N} \leqslant 0, \tag{2.17}$$

where

$$c \equiv v_{\mathrm{s}} - \mathbf{v}\cdot\mathbf{n}, \tag{2.18}$$

$$\llbracket w_M \rrbracket \equiv -\frac{\mathbf{N}\cdot\llbracket \mathbf{t}\cdot\mathbf{v} \rrbracket}{\rho_0 c_0} = -\frac{\mathbf{n}\cdot\llbracket \boldsymbol{\sigma}\cdot\mathbf{v} \rrbracket}{\rho c}, \tag{2.19}$$

$$\begin{aligned} \llbracket w_s \rrbracket \equiv \llbracket w_M \rrbracket - \llbracket k \rrbracket &= -\frac{\mathbf{N}\cdot\mathbf{t}_A}{\rho_0 c_0}\cdot\llbracket \mathbf{v} \rrbracket = -\frac{\mathbf{n}\cdot\boldsymbol{\sigma}_A}{\rho c}\cdot\llbracket \mathbf{v} \rrbracket \\ &= \frac{\mathbf{N}\cdot\llbracket \mathbf{t}\cdot\mathbf{t}^{\mathrm{T}} \rrbracket\cdot\mathbf{N}}{2\rho_0^2 c_0^2} = \frac{\mathbf{n}\cdot\llbracket \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} \rrbracket\cdot\mathbf{n}}{2\rho^2 c^2} = \frac{1}{\rho_0}\mathbf{t}_A^{\mathrm{T}} : \llbracket \mathbf{F} \rrbracket, \end{aligned} \tag{2.20}$$

$$\llbracket w_T \rrbracket \equiv \frac{\mathbf{N}\cdot\llbracket \mathbf{Q} \rrbracket}{\rho_0 c_0} = \frac{\mathbf{n}\cdot\llbracket \mathbf{q} \rrbracket}{\rho c}. \tag{2.21}$$

Here, v_{s} is the normal propagation speed of Σ as seen by a fixed observer; $\mathbf{n}$ and $\mathbf{N}$ are the unit normals (in the direction of propagation) to Σ and Σ_0, respectively; c_0 is the normal propagation speed of the reference image of the shock Σ_0; and c is the propagation speed of Σ as seen by an observer moving with a material particle.

Several of the above alternative versions of the jump laws make use of the physical requirement of continuous displacements (normal displacement continuity is required so material cracking or interpenetration does not occur; tangential displacement continuity so that only finite plastic work is produced by finite motion of a shock), which implies the following restrictions (Hill, 1961; Hadamard, 1903):

$$\llbracket \mathbf{F} \rrbracket = \llbracket \mathbf{F} \rrbracket\cdot\mathbf{N}\mathbf{N}, \tag{2.22}$$

$$[\![\mathbf{v}]\!] = -c_0[\![\mathbf{F}]\!]\cdot\mathbf{N}, \tag{2.23}$$

$$\mathbf{N}\cdot[\![J\mathbf{F}^{-1}]\!] = \mathbf{0}, \tag{2.24}$$

$$\rho c\,\mathrm{dA} = \rho_0 c_0\,\mathrm{dA}_0, \tag{2.25}$$

$$\frac{\mathbf{n}}{\rho c} = \frac{\mathbf{N}}{\rho_0 c_0}\cdot(J\mathbf{F}^{-1}). \tag{2.26}$$

Here, dA is an element of area on the shock surface Σ, dA_0 is the reference image of dA, and **NN** is a dyad. The differential $\rho c\mathrm{dA}$ (or $\rho_0 c_0\,\mathrm{dA}_0$) represents the mass flow rate across the area element.

It is sometimes convenient to express the results of discontinuity analyses in terms of the characteristic segment vector $\boldsymbol{\lambda}$ (Hadamard, 1903), defined as

$$\boldsymbol{\lambda} \equiv [\![\mathbf{F}]\!]\cdot\mathbf{N} = -\frac{[\![\mathbf{v}]\!]}{c_0} = \frac{\mathbf{N}\cdot[\![\mathbf{t}]\!]}{\rho_0 c_0^2} = \frac{\mathbf{n}\cdot[\![\boldsymbol{\sigma}]\!]}{\rho c c_0}. \tag{2.27}$$

2.3. *The Hugoniot*

Using (2.21) and the last expression for $[\![w_s]\!]$ in (2.20), the first law jump equation (2.16) may be written as

$$[\![u]\!] = \frac{1}{\rho_0}\mathbf{t}_\mathrm{A}^\mathrm{T} : [\![\mathbf{F}]\!] + \frac{\mathbf{N}\cdot[\![\mathbf{Q}]\!]}{\rho_0 c_0}. \tag{2.28}$$

The "Hugoniot" is the set of all states $\{\mathbf{F}, \mathbf{t}, u, \mathbf{Q}\}$ on one side of the shock front associated with a given STATE 1, $\{\mathbf{F}_1, \mathbf{t}_1, u_1, \mathbf{Q}_1\}$, on the other side, that are compatible with the internal energy jump equation (2.28) in the form

$$u - u_1 = \frac{1}{2\rho_0}(\mathbf{t}^\mathrm{T} + \mathbf{t}_1^\mathrm{T}) : (\mathbf{F} - \mathbf{F}_1) + \frac{\mathbf{N}\cdot(\mathbf{Q} - \mathbf{Q}_1)}{\rho_0 c_0}. \tag{2.29}$$

The Hugoniot *function H* is thus defined as

$$H(u, \mathbf{t}, \mathbf{F}, \mathbf{Q}) \equiv (u - u_1) - \frac{1}{2\rho_0}(\mathbf{t}^\mathrm{T} + \mathbf{t}_1^\mathrm{T}) : (\mathbf{F} - \mathbf{F}_1) - \frac{\mathbf{N}\cdot(\mathbf{Q} - \mathbf{Q}_1)}{\rho_0 c_0}. \tag{2.30}$$

Any state satisfying $H = 0$ is said to be "on the Hugoniot."

States compatible with the energy jump equation in the form (2.28) will not necessarily be compatible with the energy equation in any of the alternative forms suggested by (2.16) and (2.20) unless the states are additionally compatible with one or more of the other jump equations such as (2.22), (2.23) or (2.15). The form (2.30) for the Hugoniot is used because it employs only thermodynamical quantities (see Courant and Friedrichs, 1948).

The set of solutions to the equation $H = 0$ is generally regarded as the set of all constrained equilibrium states achievable by a material particle once it passes through a shock. (As a simple example, for a shock in an inviscid gas, $H = 0$ defines a simple pressure–volume curve.) It is important to note that these equilibrium states may represent different shock wave strengths, at different shock speeds. For the reference state, the second equation of (2.14) requires $[\![c_0]\!] = 0$, but on the Hugoniot c_0 cannot be treated as a constant because a different set of $\{\mathbf{F}, \mathbf{t}, u, \mathbf{Q}\}$ corresponds to a different c_0 for a given $\{\mathbf{F}_1, \mathbf{t}_1, u_1, \mathbf{Q}_1\}$.

In the next section, it is shown that a *smooth* motion can be constructed so that the state of a material element instantaneously satisfies $H = 0$. Even if a state achieved during a smooth motion instantaneously satisfies $H = 0$, it may not satisfy $H = 0$ at some later time, in which case the smooth motion would only approximate the Hugoniot. If all of the variables in the Hugoniot function are parameterized as functions of time (holding the material particle constant), the Hugoniot H at some time t may be expanded in terms of the Hugoniot function H_1 at time t_1 (when STATE 1 is achieved) as

$$H = H_1 + \dot{H}_1(t-t_1) + \frac{1}{2!}\ddot{H}_1(t-t_1)^2 + \cdots, \tag{2.31}$$

where $\dot{H}_1$ and $\ddot{H}_1$ are the first and second material time rates of H evaluated at STATE 1, respectively. The accuracy of the approximation of a shock by a smooth motion can, therefore, be characterized by the lowest material time rate of H that is non-zero at STATE 1. It is shown in the next section that a smooth wave can be constructed such that the first and second material time rates of the Hugoniot function (2.30) are zero and the third rate is non-zero, thereby showing that the change in thermodynamic state across a general non-steady shock may be approximated until third order in material time rates of fundamental field variables by the change in thermodynamic state for a smooth motion having the same initial state. The error in this approximation may be bounded by applying the Remainder Theorem for the expansion (2.31).

3. APPROXIMATION OF A SHOCK BY A SMOOTH WAVE

In this section, we quantify the degree to which an appropriately constructed smooth wave approximates a general shock wave. The analysis is closely related to a similar, pioneering analysis by Courant and Friedrichs (1948) in which first and higher order increments of the Hugoniot function are compared with increments of the first law. In addition to restricting their attention to fluids, Courant and Friedrichs (1948) implicitly assume that the material is "mechanically conservative," meaning that homogeneous smooth adiabatic ($\mathscr{P}_T = 0$) deformations are also isentropic ($\dot{s} = 0$); this assumption is not reasonable for dissipative materials. An approximation of a shock by an adiabatic smooth wave in a mechanically conservative material predicts $\mathrm{d}^k s = 0$ for all k. Taking increments of the Hugoniot function (using the mechanically conservative version of the first law for increments of u), Courant and Friedrichs (1948) show that $\mathrm{d}s = 0$, $\mathrm{d}^2 s = 0$, and $\mathrm{d}^3 s \neq 0$, thereby showing that entropy generation due to the shock itself occurs at third order. *The derivation of this result depends critically on the implicit assumption that the material is mechanically conservative.* We show below that *for any material*—conservative or not—the entropy generation predicted by a smooth wave approximation of a general shock will agree with the actual entropy generation until third order in material time increments of fundamental field variables, even if entropy generation occurs at first or second order. It is not necessary to assume that the change in entropy across a shock is small in order to justify approximating the shock by a smooth wave—the *order* of accuracy of such an approximation has nothing to do with the entropy production.

3.1. *Error order of a smooth wave approximation to a general shock*

Consider a smooth motion, not necessarily associated with a shock, with which is associated a scalar c_0 and a unit vector $\mathbf{N}$. We seek conditions that must be satisfied by the smooth flow so that there is a transition zone, or sequence of thermomechanical states, that approximates a shock moving with normal velocity $c_0\mathbf{N}$. Just as a pressure–volume diagram describes the states associated with any particular gas process, for a shock in an arbitrary medium, the Hugoniot equation $H = 0$ defines the constrained equilibrium states achievable by a material particle once it passes through the shock (Courant and Friedrichs, 1948). If a smooth wave were to model a shock *precisely*, it would have to satisfy $H \equiv 0$—that is, H and *all* of its material time derivatives would have to vanish. As discussed in Section 2.3, the accuracy of the approximation of a shock by a smooth wave is, therefore, characterized by whether $\dot{H}$ and higher order material time rates are zero at STATE 1. Since the motion is smooth, the material time rate of internal energy is given by the local form of the first law [(2.3) with (2.6) and (2.7)]:

$$\dot{u} = \frac{1}{\rho_0}\mathbf{t}^{\mathrm{T}} : \dot{\mathbf{F}} + r - \frac{1}{\rho_0}\vec{\nabla}_0 \cdot \mathbf{Q}, \tag{3.1}$$

so the material time rate of the Hugoniot function (2.30) becomes, using (3.1):

$$\dot{H} = \frac{1}{2\rho_0}[(\mathbf{t}^{\mathrm{T}} - \mathbf{t}_1^{\mathrm{T}}) : \dot{\mathbf{F}} - \dot{\mathbf{t}}^{\mathrm{T}} : (\mathbf{F} - \mathbf{F}_1)]$$

$$+ r - \frac{1}{\rho_0}\left[\vec{\nabla}_0 \cdot \mathbf{Q} + \frac{\mathbf{N} \cdot \dot{\mathbf{Q}}}{c_0} + \frac{1}{c_0}\left(\dot{\mathbf{N}} - \frac{\dot{c}_0}{c_0}\mathbf{N}\right) \cdot (\mathbf{Q} - \mathbf{Q}_1)\right]. \tag{3.2}$$

Let us first seek the weakest restrictions to place on the flow so that $\dot{H} \approx 0$. A necessary condition for $\dot{H} \approx 0$ is that $\dot{H} = 0$ at the *particular* STATE 1. Since STATE 1 is arbitrary, this requires that

$$r - \frac{1}{\rho_0}\left[\vec{\nabla}_0 \cdot \mathbf{Q} + \frac{\mathbf{N} \cdot \dot{\mathbf{Q}}}{c_0}\right] = 0. \tag{3.3}$$

This restriction is satisfied whenever, for example, $r = 0$ and the heat flux has a zero tangential gradient and appears constant to an observer moving with velocity $c_0\mathbf{N}$. Thus, whenever (3.3) is satisfied, a shock is thermodynamically equivalent to a smooth wave until at least second order in material time rates.

Courant and Friedrichs (1948) assume at the outset that the heat source and the heat flux are zero; we shall now make the slightly weaker assumptions that the heat source is zero and the heat flux is constant, since otherwise, except for the special case of steady-state shock propagation ($\dot{c}_0 = 0$) with a non-rotating reference configuration image ($\dot{\mathbf{N}} = \mathbf{0}$), the term involving $(\mathbf{Q}-\mathbf{Q}_1)$ in (3.2) will contribute to the higher order rates of H even at STATE 1.

With these assumptions of zero heat source and constant heat flux, the Hugoniot function and its first, second, and third derivatives become

$$H = (u-u_1) - \frac{1}{2\rho_0}(\mathbf{t}^{\mathrm{T}} + \mathbf{t}_1^{\mathrm{T}}) : (\mathbf{F}-\mathbf{F}_1), \tag{3.4}$$

$$\dot{H} = \frac{1}{2\rho_0}[(\mathbf{t}^{\mathrm{T}} - \mathbf{t}_1^{\mathrm{T}}) : \dot{\mathbf{F}} - \dot{\mathbf{t}}^{\mathrm{T}} : (\mathbf{F}-\mathbf{F}_1)], \tag{3.5}$$

$$\ddot{H} = \frac{1}{2\rho_0}[(\mathbf{t}^{\mathrm{T}} - \mathbf{t}_1^{\mathrm{T}}) : \ddot{\mathbf{F}} - \ddot{\mathbf{t}}^{\mathrm{T}} : (\mathbf{F}-\mathbf{F}_1)], \tag{3.6}$$

$$\dddot{H} = \frac{1}{2\rho_0}[\dot{\mathbf{t}}^{\mathrm{T}} : \ddot{\mathbf{F}} - \ddot{\mathbf{t}}^{\mathrm{T}} : \dot{\mathbf{F}} + (\mathbf{t}^{\mathrm{T}} - \mathbf{t}_1^{\mathrm{T}}) : \dddot{\mathbf{F}} - \dddot{\mathbf{t}}^{\mathrm{T}} : (\mathbf{F}-\mathbf{F}_1)]. \tag{3.7}$$

Here, the superposed double and triple dots denote the second and third material time rates, respectively.

At STATE 1, (3.4) through (3.7) show directly that H and its rates become

$$H_1 = 0, \tag{3.8}$$

$$\dot{H}_1 = 0, \tag{3.9}$$

$$\ddot{H}_1 = 0, \tag{3.10}$$

$$\dddot{H}_1 = \frac{1}{2\rho_0}(\dot{\mathbf{t}}_1^{\mathrm{T}} : \ddot{\mathbf{F}}_1 - \ddot{\mathbf{t}}_1^{\mathrm{T}} : \dot{\mathbf{F}}_1). \tag{3.11}$$

Since STATE 1 is arbitrary, these results show that a general shock is thermodynamically equivalent to a smooth wave until third order in material time rates at which point contributions due to the shock itself first appear. *This result is independent of the material constitution*, as the preceding analysis shows.

The similar, but one-dimensional, analysis by Courant and Friedrichs (1948) indirectly employs the material constitutive equation, which is not necessary. Courant and Friedrichs (1948) write the first law (in the current notation) as $\dot{u} = \mathscr{P}_s + T\dot{s}$, and they therefore implicitly assume the material is mechanically conservative during smooth deformations. That is, referring to (2.11b), they assume the dissipation $\mathscr{D}$ is zero. They also restrict attention to shear-free motion. These assumptions are neither necessary nor—for solids or viscous fluids—reasonable. Of course, one might suspect that perhaps Courant and Friedrichs (1948) are analyzing a *fictitious* reversible process with the same thermodynamic end states for the purpose of calculating entropy changes (see Resnick and Halliday, 1977). However, if this were the case, the fictitious stresses/strains would not necessarily correspond to the *actual* total stresses/strains (the fictitious strains would be the *elastic* strains for an elastic–plastic material or, alternatively, the stresses would be the so-called "thermodynamic" stresses), and therefore would not necessarily be restricted to satisfy the Hugoniot function. If the Hugoniot function is enforced, then the first law must correspond to the constrained dissipative process, not a fictitious reversible process. As a consequence of Courant and Friedrichs' (1948) implicit assumption of a mechanically conservative material, they find that the entropy increments $\mathrm{d}s$ and d^2s are zero while $\mathrm{d}^3s \neq 0$; this result is apparently the source of the common but erroneous statement that a weak shock may be approximated to at least second order by a smooth wave only if the entropy generation of the real shock is small. In general, however, $\mathrm{d}s$, or at least d^2s, will be non-zero for a locally adiabatic ($\mathscr{P}_T = 0$) motion on account of dissipation, but that motion will nevertheless approximate a shock.

Courant and Friedrichs (1948) also state that the result is valid only for weak shocks (meaning ones for which the jump in traction or velocity is small), but, as shown above, the error of the smooth wave approximation is precisely third order regardless of shock strength. In short, our result shows that *a shock of arbitrary strength may be approximated until third order by a smooth wave even if entropy generation is significant at first or second order*. Of course, if the shock strength is large enough, even a third order approximation might be poor.

Neglecting heat flux, Morland (1959) and Germain and Lee (1973) reach a similar conclusion to the one italicized above in the *one-dimensional* context for *specific* elastic–plastic constitutive models. The latter authors interpret this result as an indeterminacy in the Hugoniot attributable to neglect of rate effects, which may be a valid hypothesis because, for a real shock, rate effects must be enforced artificially in a smooth wave approximation of the Hugoniot which, recalling Kestin and Rice's (1970) concept of constrained equilibrium, represents a fictitious evolution of equilibrium states. For shocks hypothesized to exist in the solution of a boundary value problem for a rate-independent material, the results of this section may be interpreted as consistency relations.

The conclusions of this section are unrelated to several similar-sounding statements by Wallace (1991). When Wallace states that entropy is second order in strains, he is simply referring to the fact that entropy generation does not commence until yield is achieved, and yield occurs at a non-zero strain. When Wallace states that the normal stress is lower order in entropy than the shear stress, he is merely enforcing the physically plausible assumption that the Grüneisen (thermal deformation coefficient) tensor is isotropic at small elastic strains. Wallace correctly points out that entropy generation for elastic–plastic solids commences at second order in strains rather than at the third order liquid Hugoniot prediction. For the one-dimensional shocks studied by Wallace, $T\dot{s} = \dot{u} - \mathscr{P}_s + \mathscr{P}_p$, where $\mathscr{P}_p$ represents the plastic power and is proportional to the shear stress. Hence, when the shear stress is zero, $T\dot{s} = \dot{u} - \mathscr{P}_s$, which makes the material equivalent to the implicitly mechanically conservative material studied by Courant and Friedrichs (1948), and, therefore, the entropy generation occurs at third order when shear stress is neglected. Wallace blames the discrepancy (in the order of entropy generation) on the liquid Hugoniot assumption that shear stress is zero; instead, it is neglect of *dissipation*, and not neglect of shear stress *per se*, that leads to differences in the order of entropy. Any material (such as a porous metal) that is capable of exhibiting dissipation in isotropic deformations may generate entropy at first or second order even if the shear stress is zero. Nevertheless, the results of this section prove that the smooth wave approximation is valid until third order in material rates even if entropy generation occurs at second (or first) order.

A common justification for the liquid Hugoniot assumption is that, for high pressure shocks, the effect of shear stress is negligible. However, for elastic–plastic shocks, the analysis of Section 5 suggests that the existence of such shocks is determined, at least in part, by thermal stresses due to plastic dissipation, which (for plastically incompressible materials) in turn is typically generated only in the presence of shear stresses. Thermal stresses are often small compared to the total stress (especially for weak shocks, as indicated by Wallace, 1991) primarily because the thermal expansion coefficient is small, not necessarily because the change in entropy is small, though entropy generation *is* smaller in weaker shocks. Neglecting low order changes in entropy would be tantamount to neglecting plastic work—certainly not a reasonable assumption. It is shown in Section 5 that, even if thermal stresses and strains (which result from thermomechanical coupling) are small, they can play a key role in the existence of certain elastic–plastic shocks and therefore must not be neglected.

3.2. *The smooth wave approximation improves for steady-state and non-rotating shock propagation*

In the previous section we showed, by analyzing the Hugoniot function for a smooth wave, that the first law of thermodynamics across a smooth wave agrees with the jump form of this law that is satisfied across a general shock until third order in material time rates of fundamental field variables. In performing this analysis, we did not enforce the conditions of momentum conservation and compatibility. These certainly must also be satisfied, but it appears that they do not permit any stronger results than those of Section 3.1 for the general shock case. However, for a steady-state shock that propagates with a non-rotating reference configuration image, we show here that enforcement of momentum conservation and compatibility reveal that a smooth wave thermodynamically approximates a shock through *all* orders of material time rates of fundamental field variables.

Assuming no body forces, the local forms of momentum conservation and compatibility in terms of reference configuration quantities, which must be satisfied by a smooth wave, are, respectively

$$\vec{\boldsymbol{\nabla}}_0 \cdot \mathbf{t} = \rho_0 \dot{\mathbf{v}}, \tag{3.12}$$

$$\dot{\mathbf{F}} = \dot{\mathbf{x}} \overleftarrow{\boldsymbol{\nabla}}_0 = \mathbf{v} \overleftarrow{\boldsymbol{\nabla}}_0 . \tag{3.13}$$

For a smooth wave that propagates without rotating under steady-state conditions with constant speed c_0 in the direction of its normal $\mathbf{N}$, the material time derivative of any field quantity ψ within the wave is simply

$$\dot{\psi} = -c_0 \psi \overleftarrow{\boldsymbol{\nabla}}_0 \cdot \mathbf{N}, \tag{3.14}$$

and we have required

$$\dot{\mathbf{N}} = \mathbf{0}, \quad \dot{c}_0 = 0. \tag{3.15}$$

A final requirement is that field quantities may vary through the smooth wave (i.e. in the direction of $\mathbf{N}$), but *not* along it (it is intuitively obvious that for a smooth wave to emulate a shock, the smooth wave should have this feature). Under this condition, the reference configuration del operator simplifies to

$$\nabla_0 = \frac{\mathrm{d}}{\mathrm{d}v}\mathbf{N}, \tag{3.16}$$

where v measures distance in the $\mathbf{N}$ direction. Combining (3.14) and (3.16) gives

$$\dot{\psi} = -c_0\left(\frac{\mathrm{d}\psi}{\mathrm{d}v}\mathbf{N}\right)\cdot\mathbf{N} = -c_0\frac{\mathrm{d}\psi}{\mathrm{d}v}. \tag{3.17}$$

Now, application of (3.17) and then (3.16) shows that

$$-\frac{1}{c_0}\dot{\mathbf{t}}^{\mathrm{T}}\cdot\mathbf{N} = \frac{\mathrm{d}\mathbf{t}^{\mathrm{T}}}{\mathrm{d}v}\cdot\mathbf{N} = \nabla_0\cdot\mathbf{t}. \tag{3.18}$$

Also, application of (3.16) and then (3.17) to (3.13) gives

$$\dot{\mathbf{F}}\cdot\mathbf{N} = \mathbf{v}\overleftarrow{\nabla}_0\cdot\mathbf{N} = \left(\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}v}\mathbf{N}\right)\cdot\mathbf{N} = -\frac{1}{c_0}\dot{\mathbf{v}}. \tag{3.19}$$

Using (3.18) and (3.19), (3.12) becomes:

$$\boxed{\mathbf{N}\cdot\dot{\mathbf{t}} = \rho_0 c_0^2\dot{\mathbf{F}}\cdot\mathbf{N}}\,; \tag{3.20}$$

taking the material time rate of this using (3.15) gives

$$\mathbf{N}\cdot\ddot{\mathbf{t}} = \rho_0 c_0^2\ddot{\mathbf{F}}\cdot\mathbf{N}. \tag{3.21}$$

Returning to compatibility (3.13), we apply (3.16) to obtain

$$\dot{\mathbf{F}} = \mathbf{v}\overleftarrow{\nabla}_0 = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}v}\mathbf{N}, \tag{3.22}$$

which shows immediately that

$$\boxed{\dot{\mathbf{F}} = (\dot{\mathbf{F}}\cdot\mathbf{N})\mathbf{N}}\,. \tag{3.23}$$

Taking the material time rate of (3.23) gives, again using (3.15):

$$\ddot{\mathbf{F}} = (\ddot{\mathbf{F}}\cdot\mathbf{N})\mathbf{N}. \tag{3.24}$$

Finally we observe that (3.23) and (3.24) show, recalling from Section 2 the definition of tensor inner product (:)

$$\ddot{\mathbf{t}}^{\mathrm{T}}:\dot{\mathbf{F}} = \mathbf{N}\cdot\ddot{\mathbf{t}}\cdot\dot{\mathbf{F}}\cdot\mathbf{N}, \tag{3.25a}$$

$$\dot{\mathbf{t}}^{\mathrm{T}}:\ddot{\mathbf{F}} = \mathbf{N}\cdot\dot{\mathbf{t}}\cdot\ddot{\mathbf{F}}\cdot\mathbf{N}. \tag{3.25b}$$

Now recall from (3.8)–(3.11) of Section 3.1 that at an arbitrary STATE 1, we showed that $H_1 = \dot{H}_1 = \ddot{H}_1 = 0$, but for a smooth wave modeling a general shock we were only able to conclude that

$$\dddot{H}_1 = \frac{1}{2\rho_0}(\dot{\mathbf{t}}_1^{\mathrm{T}}:\ddot{\mathbf{F}}_1 - \ddot{\mathbf{t}}_1^{\mathrm{T}}:\dot{\mathbf{F}}_1). \tag{3.26}$$

Now, however, for a smooth wave modeling a steady-state non-turning shock, application of (3.20), (3.21) and (3.25) to (3.26) reveals

$$\dddot{H}_1 = \frac{1}{2\rho_0}[(\rho_0 c_0^2 \dot{\mathbf{F}}_1 \cdot \mathbf{N}) \cdot (\ddot{\mathbf{F}}_1 \cdot \mathbf{N}) - (\rho_0 c_0^2 \ddot{\mathbf{F}}_1 \cdot \mathbf{N}) \cdot (\dot{\mathbf{F}}_1 \cdot \mathbf{N})] = 0. \tag{3.27}$$

Results (3.8)–(3.10) and (3.27) lead to some fascinating conclusions with respect to the comparison of shocks and smooth waves. First, these equations show that information on whether a shock is propagating under steady-state conditions, and whether its reference configuration image rotates during propagation, does not enter the material time rates of the Hugoniot function at STATE 1 until third order. This means that an arbitrarily propagating non-steady shock can be accurately (i.e. to third order) represented thermodynamically by a general smooth wave.

Second, (3.27) shows the surprising new result that if a shock wave propagates under steady-state conditions and with a reference configuration image that does not rotate during propagation, then $\dddot{H}_1 \equiv 0$, meaning that such types of shock can be represented *at least to fourth order accuracy* by a steady-state, non-rotating smooth wave. In fact, it is easy to show by extending the approach used above that for smooth waves propagating under steady-state, non-rotating conditions, *all* orders of material time derivative of H vanish at STATE 1! While not guaranteeing that such a smooth wave can exactly represent a shock wave, this does show that such a representation will be either exact or extremely good in such cases.

3.3. *Comparison of momentum conservation and compatibility for a smooth wave versus a shock under steady-state and non-rotating propagation*

The previous two sub-sections have focused on showing how the first law of thermodynamics compares across a shock versus across a smooth wave. Here we show, for the case of steady-state and non-rotating shock propagation, that the jump forms of momentum conservation and compatibility, which must be satisfied across a shock, are identical to their forms that govern a smooth wave. We thus conclude that under these steady-state, non-rotating conditions, a smooth wave provides a precise representation of a shock with respect to these two fundamental conditions also.

For a shock, the jump conditions resulting from momentum conservation and compatibility (i.e. the assumption of continuous displacements), (2.15) and (2.23), can be combined to give

$$\mathbf{N} \cdot [\![\mathbf{t}]\!] = \rho_0 c_0^2 [\![\mathbf{F}]\!] \cdot \mathbf{N}. \tag{3.28}$$

Since this equation is valid for any shock wave, we may rewrite it for a sequence of shock waves each having the same $\{\mathbf{F}_1, \mathbf{t}_1, u_1, \mathbf{Q}_1\}$ and ρ_0, $\mathbf{N}$, but different end states $\{\mathbf{F}, \mathbf{t}, u, \mathbf{Q}\}$:

$$\mathbf{N} \cdot (\mathbf{t} - \mathbf{t}_1) = \rho_0 c_0^2 (\mathbf{F} - \mathbf{F}_1) \cdot \mathbf{N}, \tag{3.29}$$

where c_0 and $\mathbf{N}$ are functions, in general, of $\mathbf{t}$ and $\mathbf{F}$, and hence of time t. Taking first and second material time rates of (3.29) thus gives:

$$\mathbf{N} \cdot \dot{\mathbf{t}} + \dot{\mathbf{N}} \cdot (\mathbf{t} - \mathbf{t}_1) = \rho_0 c_0^2 \dot{\mathbf{F}} \cdot \mathbf{N} + (\mathbf{F} - \mathbf{F}_1) \cdot (2\rho_0 c_0 \dot{c}_0 \mathbf{N} + \rho_0 c_0^2 \dot{\mathbf{N}}), \tag{3.30}$$

$$\mathbf{N}\cdot\ddot{\mathbf{t}}+2\dot{\mathbf{N}}\cdot\dot{\mathbf{t}}+\ddot{\mathbf{N}}\cdot(\mathbf{t}-\mathbf{t}_1) = \rho_0 c_0^2\ddot{\mathbf{F}}\cdot\mathbf{N}+4\rho_0 c_0\dot{c}_0\dot{\mathbf{F}}\cdot\mathbf{N}+2\rho_0 c_0^2\dot{\mathbf{F}}\cdot\dot{\mathbf{N}}$$
$$+(\mathbf{F}-\mathbf{F}_1)\cdot[2\rho_0(\dot{c}_0^2+c_0\ddot{c}_0)\mathbf{N}+4\rho_0 c_0\dot{c}_0\dot{\mathbf{N}}+\rho_0 c_0^2\ddot{\mathbf{N}}]. \tag{3.31}$$

Similarly to the above analysis, the jump condition (2.22), also a consequence of displacement continuity, is valid for any shock wave and thus requires for a sequence of shock waves each having the same $\{\mathbf{F}_1, \mathbf{t}_1, u_1, \mathbf{Q}_1\}$ and ρ_0, $\mathbf{N}$, but different end states $\{\mathbf{F}, \mathbf{t}, u, \mathbf{Q}\}$:

$$(\mathbf{F}-\mathbf{F}_1) = (\mathbf{F}-\mathbf{F}_1)\cdot\mathbf{NN}. \tag{3.32}$$

Taking first and second material time rates of this gives

$$\dot{\mathbf{F}} = \dot{\mathbf{F}}\cdot\mathbf{NN}+(\mathbf{F}-\mathbf{F}_1)\cdot(\dot{\mathbf{N}}\mathbf{N}+\mathbf{N}\dot{\mathbf{N}}), \tag{3.33}$$

$$\ddot{\mathbf{F}} = \ddot{\mathbf{F}}\cdot\mathbf{NN}+2\dot{\mathbf{F}}\cdot(\dot{\mathbf{N}}\mathbf{N}+\mathbf{N}\dot{\mathbf{N}})+(\mathbf{F}-\mathbf{F}_1)\cdot(\ddot{\mathbf{N}}\mathbf{N}+2\dot{\mathbf{N}}\dot{\mathbf{N}}+\mathbf{N}\ddot{\mathbf{N}}). \tag{3.34}$$

Now, for the case of steady-state propagation of a shock whose reference configuration image does not rotate, conditions (3.15) apply. In this case, (3.30), (3.31), (3.33) and (3.34) are seen to reduce directly to the smooth wave restrictions derived in Section 3.2, namely (3.20), (3.21), (3.23) and (3.24), respectively. This shows that these two sets of conditions agree exactly on *both* sides of the shock and smooth wave, and it is easily shown that this is true for *all* orders of material time rates of these conditions. Thus, for steady-state and non-rotating propagation, a smooth wave across which (3.20), (3.21), (3.23) and (3.24) are satisfied precisely represents how momentum conservation and compatibility restrict stress and deformation paths across a shock.

3.4. *Discussion*

The analysis of Section 3 *legitimizes the approximation of a shock by a smooth wave, regardless of the material constitution.* Because this result is valid up to third order for general non-steady shocks, we must still limit the shock strength, or the error—even at third order—might be too large. We note in passing that a smooth approximation of a shock will fail to capture a portion of the temperature increase that is generated by the passage of a shock, even in "mechanically conservative" (e.g. perfectly elastic) materials, unless the smooth wave constitutive description includes rate effects that are appropriately scaled to accommodate the smooth wave expansion of the time scale. Such effects are implicitly neglected in classical (constrained equilibrium) thermodynamics. However, the magnitude of these effects on predictions of stress, temperature, etc. (which can be quantified only by inclusion of non-classical thermodynamic rate terms) are here assumed negligible in comparison with the larger, smooth-wave classical contributions, as evidenced, for example, by the successful application of classical thermodynamics to shocks in metals by Wallace (1991).

The remainder of this work focuses on specific conclusions that can be drawn from smooth wave approximations (that incorporate thermodynamics) of steady, non-rotating elastic–plastic shocks, for which we have just shown the smooth wave approximations to be possibly exact or at least extremely accurate. Large deformation elastic–plastic kinematics and constitutive equations are reviewed in Section 4. Then, in Section 5, large deformation thermoelastic–plastic shocks are shown to be possible in certain situations only if thermomechanical coupling is not neglected.

4. LARGE DEFORMATION ELASTIC–PLASTIC KINEMATICS AND CONSTITUTIVE LAWS

Discussion of elastic–plastic kinematics and constitutive laws has been postponed until now to emphasize that the results of Section 3 are independent of the constitution of the material. The present section outlines the constitutive equations that will be used in Section 5 for the analysis of large deformation thermoelastic–plastic shocks.

4.1. *Deformation gradient decomposition*

As pointed out by Lee and Liu (1967) and extended to general three-dimensional deformations by Lee (1969), the deformation gradient $\mathbf{F}$ may be multiplicatively decomposed:

$$\mathbf{F} = \mathbf{F}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{p}}, \tag{4.1}$$

where our interpretation is that $\mathbf{F}^{\mathrm{e}}$ represents the macroscopically thermoelastic part of the deformation and $\mathbf{F}^{\mathrm{p}}$ represents the part of the deformation due to both dissipative rearrangement of the atoms (plastic deformation) and microscopic changes in the defect structure (recoverable by annealing). The intermediate (generally non-Euclidean) configuration, which corresponds to $\mathbf{F}^{\mathrm{e}} = \mathbf{I}$, is called the elastic reference configuration and, by definition, is at an elastic reference temperature and is stress-free. The stress and other state variables are assumed to be derivable from thermodynamical constitutive laws identical to those for classical thermoelasticity except that the reference configuration for the thermoelastic part of the response is the intermediate configuration and parameters in the constitutive equations may permissibly change with plastic deformation. The rotational parts of $\mathbf{F}^{\mathrm{e}}$ and $\mathbf{F}^{\mathrm{p}}$ are assigned so that the rotation of the elastic lattice is zero in the elastic reference configuration.

The deformation decomposition $\mathbf{F} = \mathbf{F}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{p}}$ implies that the velocity gradient ($\mathbf{L} = \dot{\mathbf{F}} \cdot \mathbf{F}^{-1}$) may be decomposed additively into elastic and plastic parts:

$$\mathbf{L} = \mathbf{L}^{\mathrm{e}} + \mathbf{L}^{\mathrm{p}}, \tag{4.2}$$

where

$$\mathbf{L}^{\mathrm{e}} = \dot{\mathbf{F}}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{e}-1} \quad \text{and} \quad \mathbf{L}^{\mathrm{p}} = \mathbf{F}^{\mathrm{e}} \cdot \dot{\mathbf{F}}^{\mathrm{p}} \cdot \mathbf{F}^{\mathrm{p}-1} \cdot \mathbf{F}^{\mathrm{e}-1}. \tag{4.3}$$

The plastic part of the deformation is assumed incompressible; that is, the determinant of $\mathbf{F}^{\mathrm{p}}$ is unity, implying that the trace of $\mathbf{L}^{\mathrm{p}}$ is zero and the density in the elastic reference configuration is equal to the initial density ρ_0 (i.e. the density when $\mathbf{F} = \mathbf{I}$).

Enforcing symmetry of the Cauchy stress $\boldsymbol{\sigma}$, (4.2) shows that the stress power (2.6) may be decomposed into elastic and plastic parts:

$$\mathscr{P}_s = \mathscr{P}_e + \mathscr{P}_p, \tag{4.4}$$

where

$$\mathscr{P}_s = \frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{D} = \frac{1}{\rho_0}\mathbf{P}:\dot{\mathbf{E}}, \tag{4.5}$$

$$\mathscr{P}_e = \frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{D}^{\mathrm{e}} = \frac{1}{\rho_0}\mathbf{P}^{\mathrm{e}}:\dot{\mathbf{E}}^{\mathrm{e}}, \tag{4.6}$$

$$\mathscr{P}_p = \frac{1}{\rho}\boldsymbol{\sigma}:\mathbf{D}^{\mathrm{p}}. \tag{4.7}$$

Here, $\mathbf{D}^{\mathrm{e}}$ and $\mathbf{D}^{\mathrm{p}}$ are the symmetric parts of $\mathbf{L}^{\mathrm{e}}$ and $\mathbf{L}^{\mathrm{p}}$, respectively; $\mathbf{P}$ and $\mathbf{P}^{\mathrm{e}}$ are Second Piola–Kirchhoff stresses defined as $\mathbf{P} = J\mathbf{F}^{-1}\cdot\boldsymbol{\sigma}\cdot\mathbf{F}^{-\mathrm{T}}$ and $\mathbf{P}^{\mathrm{e}} = J\mathbf{F}^{\mathrm{e}-1}\cdot\boldsymbol{\sigma}\cdot\mathbf{F}^{\mathrm{e}-\mathrm{T}}$; and $\mathbf{E}$ and $\mathbf{E}^{\mathrm{e}}$ are the total and elastic Lagrange strains defined, respectively, as

$$\mathbf{E} = \tfrac{1}{2}(\mathbf{F}^{\mathrm{T}}\cdot\mathbf{F}-\mathbf{I}) = \tfrac{1}{2}(\mathbf{U}\cdot\mathbf{U}-\mathbf{I}), \tag{4.8}$$

$$\mathbf{E}^{\mathrm{e}} = \tfrac{1}{2}(\mathbf{F}^{\mathrm{eT}}\cdot\mathbf{F}^{\mathrm{e}}-\mathbf{I}) = \tfrac{1}{2}(\mathbf{U}^{\mathrm{e}}\cdot\mathbf{U}^{\mathrm{e}}-\mathbf{I}), \tag{4.9}$$

where $\mathbf{U}$ and $\mathbf{U}^{\mathrm{e}}$ are the positive definite stretch tensors from the polar decompositions $\mathbf{F} = \mathbf{R}\cdot\mathbf{U}$ and $\mathbf{F}^{\mathrm{e}} = \mathbf{R}^{\mathrm{e}}\cdot\mathbf{U}^{\mathrm{e}}$, respectively.

4.2. *Thermoelastic–plastic constitutive equations*

The internal energy is taken to be a constitutive function of the elastic deformation gradient $\mathbf{F}^{\mathrm{e}}$, the specific entropy s, and an array κ of macroscopic laboratory variables that characterize the microscopic elastic fields associated with the defect structure. This choice of independent variables, which is also used by Perzyna (1971), Mandel (1972), Wallace (1985), and Mróz and Raniecki (1976), does not exclude the type of material response described by internal variable theories (e.g. Rice, 1971) because κ (or any other parameters such as the elastic moduli in the constitutive equation) may permissibly depend on plastic deformation history via an independent kinetic relation. In accordance with Truesdell and Toupin's (1960) principle of equipresence, the temperature T and the Cauchy stress $\boldsymbol{\sigma}$ (or, equivalently, the Second Piola–Kirchhoff stress $\mathbf{P}^{\mathrm{e}}$) are also taken to be constitutive functions of $\mathbf{F}^{\mathrm{e}}$, s, and κ. It is often assumed (e.g. Germain and Lee, 1973) or argued (Lubliner, 1972) that the internal energy may be decomposed into a thermoelastic free energy $\tilde{u}$ plus a hardening energy ξ:

$$u = \tilde{u}+\xi, \tag{4.10}$$

where

$$\tilde{u} = \tilde{u}(\mathbf{F}^{\mathrm{e}}, s), \tag{4.11}$$

$$\xi = \xi(\kappa). \tag{4.12}$$

Using (4.4), (4.10), (4.11) and (4.12) in (2.11b) and (2.9), noting that the rates $\dot{s}$ and $\dot{\mathbf{F}}^{\mathrm{e}}$ may be controlled independently of s, $\mathbf{F}^{\mathrm{e}}$ and κ, and imposing the principle of material frame indifference for elastic lattice rotations, it is straightforward to show that

$$T = \left(\frac{\partial\tilde{u}}{\partial s}\right)_{\mathbf{E}^{\mathrm{e}}}, \tag{4.13a}$$

$$\frac{1}{\rho_0}\mathbf{P}^{\mathrm{e}} = \left(\frac{\partial \tilde{u}}{\partial \mathbf{E}^{\mathrm{e}}}\right)_s, \tag{4.13b}$$

$$\mathscr{D} = \mathscr{P}_p - \dot{\xi} \geqslant 0, \tag{4.14}$$

where it has been assumed that the dissipation $\mathscr{D}$ is independent of the temperature gradient. Because ξ represents the *elastic* strain energy (recoverable by annealing) associated with defect structures such as dislocation pile-ups, the dissipation $\mathscr{D}$ is the portion of the plastic power that is not stored elastically in the defect structure. For moderate temperatures, the experiments of Farren and Taylor (1925) and Dillon (1967) show that $\mathscr{D} \approx 0.9\mathscr{P}_p$ at the onset of plastic deformation from a fully annealed state and $\mathscr{D} \to \mathscr{P}_p$ as the total plastic work increases.

By the chain rule, (4.13) and (4.6) show that

$$\dot{\tilde{u}} = \mathscr{P}_{\mathrm{e}} + T\dot{s}, \tag{4.15}$$

which will later be used in the incremental form

$$\mathrm{d}\tilde{u} = \frac{1}{\rho}\boldsymbol{\sigma} : (\mathrm{d}\mathbf{F}^{\mathrm{e}} \cdot \mathbf{F}^{\mathrm{e}-1}) + T\,\mathrm{d}s. \tag{4.16}$$

The specific heat at constant strain ς_v, the thermal stress coefficient tensor $\mathbf{B}$ and the isothermal elastic stiffness tensor $\mathbb{E}$ are regarded as known material functions defined, respectively, as

$$\varsigma_v = \left(\frac{\partial \tilde{u}}{\partial T}\right)_{\mathbf{E}^{\mathrm{e}}}, \tag{4.17}$$

$$\mathbf{B} = -\left(\frac{\partial \mathbf{P}^{\mathrm{e}}}{\partial T}\right)_{\mathbf{E}^{\mathrm{e}}}, \tag{4.18}$$

$$\mathbb{E} = \left(\frac{\partial \mathbf{P}^{\mathrm{e}}}{\partial \mathbf{E}^{\mathrm{e}}}\right)_T. \tag{4.19}$$

The Grüneisen tensor γ, which appears frequently in shock wave analysis, is related to $\mathbf{B}$ and ς_v by $\gamma \equiv \mathbf{B}/\rho_0\varsigma_v$.

If ς_v, $\mathbf{B}$ and $\mathbb{E}$ are approximately constant over the range of temperatures and elastic stretches of interest, direct integration of (4.17), (4.18) and (4.19) leads eventually to the following classical constitutive equations for the entropy, stress, and thermoelastic internal energy:

$$s = s_1 + \frac{1}{\rho_0}\mathbf{B} : (\mathbf{E}^{\mathrm{e}} - \mathbf{E}^{\mathrm{e}}_1) + \varsigma_v \ln\left(\frac{T}{T_1}\right), \tag{4.20a}$$

$$\mathbf{P}^{\mathrm{e}} = \mathbf{P}^{\mathrm{e}}_1 - (T - T_1)\mathbf{B} + \mathbb{E} : (\mathbf{E}^{\mathrm{e}} - \mathbf{E}^{\mathrm{e}}_1). \tag{4.20b}$$

$$\tilde{u} = \tilde{u}_1 + \frac{1}{\rho_0}\mathbf{P}^e_1 : (\mathbf{E}^e - \mathbf{E}^e_1) + \frac{T_1}{\rho_0}\mathbf{B} : (\mathbf{E}^e - \mathbf{E}^e_1) + \varsigma_v (T - T_1) + \frac{1}{2\rho_0}(\mathbf{E}^e - \mathbf{E}^e_1) : \mathbb{E} : (\mathbf{E}^e - \mathbf{E}^e_1), \quad (4.21)$$

where the subscript "1" denotes a known STATE 1 lying in the temperature and strain range of interest.

In the special case of small elastic stretches and elastic isotropy (while the plastic part of the material response remains arbitrarily anisotropic), the material tensors $\mathbf{B}$ and $\mathbb{E}$ are isotropic, and the above constitutive equations specialize to

$$s = s_1 + \frac{1}{\rho_0}\mathbf{B} : (\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1) + \varsigma_v \ln\left(\frac{T}{T_1}\right), \quad (4.22a)$$

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_1 - (T - T_1)\mathbf{B} + \mathbb{E} : (\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1). \quad (4.22b)$$

$$\tilde{u} = \tilde{u}_1 + \frac{1}{\rho_0}\boldsymbol{\sigma}_1 : (\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1) + \frac{T_1}{\rho_0}\mathbf{B} : (\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1) + \varsigma_v (T - T_1) + \frac{1}{2\rho_0}(\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1) : \mathbb{E} : (\boldsymbol{\varepsilon}^e - \boldsymbol{\varepsilon}^e_1). \quad (4.23)$$

Here $\boldsymbol{\varepsilon}^e$ is the rotated elastic strain (also known as the elastic Signorini or Finger strain), defined as $\boldsymbol{\varepsilon}^e \equiv \mathbf{R}^e \cdot \mathbf{E}^e \cdot \mathbf{R}^{e^T} = \frac{1}{2}(\mathbf{V}^e \cdot \mathbf{V}^e - \mathbf{I})$, where $\mathbf{V}^e$ is the elastic stretch tensor from the polar decomposition $\mathbf{F}^e = \mathbf{V}^e \cdot \mathbf{R}^e$.

4.3. *Plastic potentials*

For the special case of small deformations, Mróz and Raniecki (1976) show that, for rate-independent plasticity, there must exist a scalar yield function ϕ relating the thermodynamic forces $\boldsymbol{\sigma}$ and $\mathscr{K}$, where $\mathscr{K} \equiv \rho\ \mathrm{d}\xi/\mathrm{d}\kappa$ such that the yield surface is described by

$$\phi(\boldsymbol{\sigma}, \mathscr{K}, T) = 0. \quad (4.24)$$

The yield surface is convex in $\{\boldsymbol{\sigma}, -\mathscr{K}\}$ space.

Elastic deformation occurs when $\phi < 0$ or $\phi = 0$ with $\dot{\phi} < 0$. Plastic loading occurs when $\phi = 0$ and $\dot{\phi} = 0$. Then for rate-independent plasticity, the yield surface serves as a potential for the plastic strain rate and the defect structure rate:

$$\dot{\boldsymbol{\varepsilon}}^p = \dot{\lambda}\frac{\partial \phi}{\partial \boldsymbol{\sigma}} \quad \text{and} \quad \dot{\kappa} = \dot{\lambda}\frac{\partial \phi}{\partial \mathscr{K}}, \quad (4.25)$$

where the scalar λ is required (by convexity and the second law) to be non-negative. Normality and convexity of the yield surface may be expressed by

$$(\boldsymbol{\sigma} - \boldsymbol{\sigma}^0) : \boldsymbol{\varepsilon}^p \geqslant (\mathscr{K} - \mathscr{K}^0)\dot{\kappa}, \quad (4.26)$$

where $\{\boldsymbol{\sigma}^0, \mathscr{K}^0\}$ is any state inside or on the yield surface at the current temperature. In other words, the actual thermodynamic forces are the ones that maximize the dissipation $\mathscr{D}$ with respect to thermodynamic forces not violating the yield condition. Suppose $\{\boldsymbol{\sigma}^0, \mathscr{K}^0\}$ corresponds to some previously achieved state; then, because the temperature has been assumed low enough that $\mathscr{K}$ is independent of temperature, it is expected that the right hand side of (4.26) will be non-negative. Alternatively, if a sufficient amount of plastic deformation has occurred, the right hand side of (4.26) will be nearly zero because (Dillon, 1967) $\dot{\kappa} \to 0$ as $W_p \to \infty$. Therefore, it will be assumed that

$$(\boldsymbol{\sigma}-\boldsymbol{\sigma}^0):\dot{\boldsymbol{\varepsilon}}^{\mathrm{p}} \geqslant 0. \tag{4.27}$$

That is, the maximum plastic work inequality (MPWI) is expected to hold for thermoelastic–plastic deformations if the temperature remains well below the annealing temperature.

To include materials that experience arbitrarily anisotropic yield and large deformations, the yield condition is expressed in terms of the rotated Cauchy stress and, following Drugan and Shen (1990), the MPWI is stated as

$$\frac{1}{\rho}[(\mathbf{R}^{\mathrm{T}}\cdot\boldsymbol{\sigma}\cdot\mathbf{R})-(\mathbf{R}^{\mathrm{T}}\cdot\boldsymbol{\sigma}\cdot\mathbf{R})^0]:(\mathbf{R}^{\mathrm{T}}\cdot\mathbf{D}^{\mathrm{p}}\cdot\mathbf{R}) \geqslant 0. \tag{4.28}$$

This equation expresses normality and convexity of the rotated plastic rate of deformation $\mathbf{R}^{\mathrm{T}}\cdot\mathbf{D}^{\mathrm{p}}\cdot\mathbf{R}$ to a yield surface phrased in terms of the rotated Cauchy stress, $\mathbf{R}^{\mathrm{T}}\cdot\boldsymbol{\sigma}\cdot\mathbf{R}$. Here $(\mathbf{R}^{\mathrm{T}}\cdot\boldsymbol{\sigma}\cdot\mathbf{R})^0$ is any rotated stress on or within the yield surface.

5. RESTRICTIONS ON LARGE DEFORMATION STEADY THERMOELASTIC-PLASTIC SHOCKS

5.1. *Conditions for general materials*

As justified rigorously by the results of Section 3, we require the stress and deformation paths experienced by a material particle during passage of a shock that propagates under steady-state conditions with a non-rotating reference configuration image to satisfy the requirements of compatibility and mass and momentum conservation in the incremental forms these conditions take through a steady-state non-rotating smooth wave:

$$\mathrm{d}\mathbf{F} = \mathrm{d}\mathbf{F}\cdot\mathbf{N}\mathbf{N}, \tag{5.1}$$

$$\mathbf{N}\cdot\mathrm{d}\left(\frac{\mathbf{F}^{-1}}{\rho}\right) = 0, \tag{5.2}$$

$$\mathrm{d}\mathbf{F}\cdot\mathbf{N} = \frac{\mathbf{N}\cdot\mathrm{d}\mathbf{t}}{\rho_0 c_0^2}. \tag{5.3}$$

Specifically, (5.1) and (5.3) follow from (3.23) and (3.20), respectively, exactly in the case of steadily propagating non-rotating shocks, and approximately [as shown by (3.33) and (3.30), respectively] for slightly non-steady and/or rotating shocks. Equation (5.2) is obtained simply by applying (5.1) and mass conservation (2.1) to the standard formula for the inverse of $\mathbf{F}$, $F_{ij}^{-1} = C_{ji}/J$, where C_{ij} is the cofactor (signed minor sub-determinant) of the element F_{ij}.

An alternative, physical argument supporting the enforcement of (5.1)–(5.3) on the stress–deformation path experienced by a material particle during shock passage is as follows. In a material whose constitutive model is identical to the rate-independent one to be considered below except that it possesses a small amount of rate dependence (all real materials exhibit some rate dependence), a shock in the present model would become a narrow transition layer through which field quantities vary rapidly but continuously in the rate-dependent model. Since conditions (5.1)–(5.3) represent only the requirements of mass and momentum conservation and compatibility, they are valid for any constitutive model and in particular would have to be satisfied through this rapid transition layer were it propagating under steady-state, non-rotating conditions or nearly so.

The deformation during shock passage is taken to satisfy the same constitutive equations as satisfied outside the shock. Again, justification for this was rigorously provided in Section 3, which showed that for a general shock, entropy production by the shock does not appear until third order in material time rates in the first law of thermodynamics, so that for shocks of low to moderate strength the use of purely mechanical constitutive equations to describe material response within the shock will be extremely accurate. Section 3.2 showed that this is potentially exact for a steady-state shock with a non-rotating reference configuration image.

The constitutive class is restricted such that (Drugan and Shen, 1990) if a Cauchy stress $\boldsymbol{\sigma}^*$ is chosen constant across the shock, then $\mathbf{R}^{\mathrm{T}}\cdot\boldsymbol{\sigma}^*\cdot\mathbf{R}$ is on or within the yield surface for all states in the transition layer. Any particular choice $\boldsymbol{\sigma}^*$ satisfying this restriction will be called "admissible," and the MPWI (4.28) may thus be written as

$$\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):\mathbf{D}^{\mathrm{p}} \geqslant 0. \tag{5.4}$$

Employing the symmetry of $\boldsymbol{\sigma}$ and the decomposition (4.2), $\mathbf{D}^{\mathrm{p}}$ may be replaced by $(\mathbf{L}-\mathbf{L}^{\mathrm{e}})$, and (5.4) becomes

$$\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}\cdot\mathbf{F}^{-1})-\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}^{\mathrm{e}}\cdot\mathbf{F}^{\mathrm{e}-1}) \geqslant 0. \tag{5.5}$$

Generalizing the key idea of Drugan and Rice (1984) and Drugan and Shen (1990), the MPWI is integrated at a fixed material particle from the time it enters the shock on the (+) side to the time it exits on the (−) side:

$$\int_{\mathbf{F}^+}^{\mathbf{F}^-}\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}\cdot\mathbf{F}^{-1})-\int_{\mathbf{F}^{\mathrm{e}+}}^{\mathbf{F}^{\mathrm{e}-}}\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}^{\mathrm{e}}\cdot\mathbf{F}^{\mathrm{e}-1}) \geqslant 0. \tag{5.6}$$

By using jump restrictions and the constitutive law, the first integral may be integrated explicitly and the second integral may be integrated almost explicitly, with the indefinite term being the integral of $T\,\mathrm{d}s$. These calculations are quite similar to those of Drugan and Shen (1990) and will, therefore, be only briefly outlined below, though new terms arising from thermodynamics will be discussed in detail.

Using (5.1), the integrand of the first term may be rewritten as

$$\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}\cdot\mathbf{F}^{-1}) = \frac{1}{\rho}\mathbf{N}\cdot\mathbf{F}^{-1}\cdot(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*)\cdot\mathrm{d}\mathbf{F}\cdot\mathbf{N}$$

$$= \frac{1}{\rho_0}\mathbf{N}\cdot(\mathbf{t}-\mathbf{t}^*)\cdot\mathrm{d}\mathbf{F}\cdot\mathbf{N}, \tag{5.7}$$

where

$$\mathbf{t}^* \equiv \frac{\rho_0}{\rho}\mathbf{F}^{-1}\cdot\boldsymbol{\sigma}^*. \tag{5.8}$$

If $\boldsymbol{\sigma}^*$ is chosen constant across the shock, then $\mathbf{N}\cdot\mathbf{t}^*$ is also constant by virtue of (5.2). Thus, using (5.3),

$$\frac{1}{\rho}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^*):(\mathrm{d}\mathbf{F}\cdot\mathbf{F}^{-1}) = \frac{1}{\rho_0^2c_0^2}\mathbf{N}\cdot(\mathbf{t}-\mathbf{t}^*)\cdot\mathrm{d}\mathbf{t}^{\mathrm{T}}\cdot\mathbf{N}$$

$$= \frac{1}{2\rho_0^2c_0^2}\mathbf{N}\cdot\mathrm{d}[(\mathbf{t}-\mathbf{t}^*)\cdot(\mathbf{t}-\mathbf{t}^*)^{\mathrm{T}}]\cdot\mathbf{N}. \tag{5.9}$$

For the second integral in (5.6), (4.16) shows that

$$-\frac{1}{\rho}\boldsymbol{\sigma}:(\mathrm{d}\mathbf{F}^{\mathrm{e}}\cdot\mathbf{F}^{\mathrm{e}-1}) = -\mathrm{d}\tilde{u}+T\mathrm{d}s. \tag{5.10}$$

Using the polar decomposition $\mathbf{F}^{\mathrm{e}} = \mathbf{V}^{\mathrm{e}}\cdot\mathbf{R}^{\mathrm{e}}$ and assuming $\mathbf{V}^{\mathrm{e}} \approx \mathbf{I}$ (while $\mathrm{d}\mathbf{V}^{\mathrm{e}}$ is *not* neglected), Drugan and Shen (1990) show that

$$\frac{1}{\rho}\boldsymbol{\sigma}^*:(\mathrm{d}\mathbf{F}^{\mathrm{e}}\cdot\mathbf{F}^{\mathrm{e}-1}) \approx \frac{1}{\rho_0}\boldsymbol{\sigma}^*:\mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}}, \tag{5.11}$$

where (recall) $\boldsymbol{\varepsilon}^{\mathrm{e}}$ is the elastic Signorini strain.

Using (5.9)–(5.11) in (5.6), the MPWI integrates across the shock almost explicitly to give

$$-\frac{1}{2\rho_0^2c_0^2}\mathbf{N}\cdot[\![(\mathbf{t}-\mathbf{t}^*)\cdot(\mathbf{t}-\mathbf{t}^*)^{\mathrm{T}}]\!]\cdot\mathbf{N}+[\![\tilde{u}]\!]-\frac{1}{\rho_0}\boldsymbol{\sigma}^*:[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!]+\int_{s^+}^{s} T\mathrm{d}s \geqslant 0. \tag{5.12}$$

With the specific choices $\boldsymbol{\sigma}^* = \mathbf{0}$, $\boldsymbol{\sigma}^-$, and $\boldsymbol{\sigma}^+$, this expression becomes, respectively:

$$\boldsymbol{\sigma}^* = \mathbf{0} \Rightarrow [\![\tilde{u}]\!]-[\![w_s]\!]+\int_{s^+}^{s^-} T\mathrm{d}s \geqslant 0, \tag{5.13}$$

$$\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^- \Rightarrow [\![\tilde{u}]\!]-\frac{1}{\rho_0}\boldsymbol{\sigma}^-:[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \geqslant +\tfrac{1}{2}c_0^2\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}-\int_{s^+}^{s^-} T\mathrm{d}s, \tag{5.14}$$

$$\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^+ \Rightarrow [\![\tilde{u}]\!]-\frac{1}{\rho_0}\boldsymbol{\sigma}^+:[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \geqslant -\tfrac{1}{2}c_0^2\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}-\int_{s^+}^{s^-} T\mathrm{d}s, \tag{5.15}$$

where we have used (2.20) and (2.27). Any particular choice for $\boldsymbol{\sigma}^*$ is "admissible" if

$\boldsymbol{\sigma}^*$ is within or on the yield surface for all states in the transition zone. The choice $\boldsymbol{\sigma}^* = \mathbf{0}$ corresponds to enforcement of non-negative plastic work. The choice $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^-$ is admissible if the material is non-hardening or if it has commenced isotropic softening by the time the shock arrives. The choice $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^+$ is admissible if the material hardens in such a way that the current yield locus incorporates all prior yield loci, which includes non-hardening, isotropic hardening, and many types of non-isotropic hardening, including the formation of yield surface vertices.

5.2. *Specialization of the shock restrictions for elastic isotropy*

Equations (5.13)–(5.15) hold for arbitrary elastic and plastic material anisotropy and large displacement gradients provided the elastic stretches are small. In the case of elastic isotropy, the tensors $\mathbb{E}$ and $\mathbf{B}$ are isotropic, and the constitutive equations (4.22) and (4.23) may be used to simplify the left-hand sides of (5.14) and (5.15) as well as the integral of $T\,\mathrm{d}s$. Alternatively, in the special case of small displacement gradients, $\mathbf{R}^{\mathrm{e}} \approx \mathbf{I}$, and therefore the constitutive equations (4.22) and (4.23) are valid with the tensors $\mathbb{E}$ and $\mathbf{B}$ arbitrarily anisotropic.

Taking the two states (the current state and the thermodynamic reference state) in (4.23) to be states (+) and (−) and vice versa gives:

$$[\![\tilde{u}]\!] - \frac{1}{\rho_0}\boldsymbol{\sigma}^- : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] = \frac{T^-}{\rho_0}\mathbf{B} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] + \varsigma_v [\![T]\!] + \frac{1}{2\rho_0}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!], \tag{5.16}$$

$$[\![\tilde{u}]\!] - \frac{1}{\rho_0}\boldsymbol{\sigma}^+ : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] = \frac{T^+}{\rho_0}\mathbf{B} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] + \varsigma_v [\![T]\!] - \frac{1}{2\rho_0}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!]. \tag{5.17}$$

Hence, (5.14) and (5.15) respectively, become

$$\tfrac{1}{2}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \geqslant \tfrac{1}{2}\rho_0 c_0^2 \boldsymbol{\lambda}\cdot\boldsymbol{\lambda} - T^-\mathbf{B} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] - \rho_0 \varsigma_v [\![T]\!] - \int_{s^+}^{s^-} \rho_0 T\,\mathrm{d}s, \tag{5.18}$$

$$\tfrac{1}{2}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \leqslant \tfrac{1}{2}\rho_0 c_0^2 \boldsymbol{\lambda}\cdot\boldsymbol{\lambda} + T^+\mathbf{B} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] + \rho_0 \varsigma_v [\![T]\!] + \int_{s^+}^{s^-} \rho_0 T\,\mathrm{d}s, \tag{5.19}$$

where (recall) the first inequality holds if the choice $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^-$ is admissible whereas the second inequality holds if $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^+$ is admissible.

Using (4.20) together with symmetry of $\mathbf{B}$ shows that

$$\int_{s^+}^{s^-} \rho_0 T\,\mathrm{d}s = \int_{(+)}^{(-)} T\mathbf{B} : \mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}} - \rho_0 c_v [\![T]\!]. \tag{5.20}$$

Thus, the inequalities (5.18) and (5.19) become

$$\tfrac{1}{2}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \geqslant \tfrac{1}{2}\rho_0 c_0^2 \boldsymbol{\lambda}\cdot\boldsymbol{\lambda} + \int_{(+)}^{(-)} (T^- - T)\mathbf{B} : \mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}}, \tag{5.21}$$

$$\tfrac{1}{2}[\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] : \mathbb{E} : [\![\boldsymbol{\varepsilon}^{\mathrm{e}}]\!] \leqslant \tfrac{1}{2}\rho_0 c_0^2 \boldsymbol{\lambda}\cdot\boldsymbol{\lambda} + \int_{(+)}^{(-)} (T - T^+)\mathbf{B} : \mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}}. \tag{5.22}$$

These versions show that the shock restrictions depend on *changes* in temperature, not the absolute temperature as the versions (5.14) and (5.15) might misleadingly suggest. The restrictions (5.21) and (5.22) show most clearly that it is mechanical coupling (thermal stresses and strains) that prevents the maximum plastic work inequality from being integrated explicitly. Suppose, however, that the integral of $T\mathbf{B}:\mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}}$ is approximated by the trapezoidal rule:

$$\int_{(+)}^{(-)} T\mathbf{B}:\mathrm{d}\boldsymbol{\varepsilon}^{\mathrm{e}} = -\frac{T^{+}+T^{-}}{2}\mathbf{B}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket+\eta. \tag{5.23}$$

In most situations, the error of the approximation, η, is expected to be quite small because (recalling that, for elastic isotropy, $\mathbf{B}$ is isotropic) η is of the order of $(\llbracket \mathrm{trace}\,(K\beta\boldsymbol{\varepsilon}^{\mathrm{e}})\rrbracket)^3$, where K is the elastic bulk modulus and β is the volumetric coefficient of thermal expansion. The error η is identically zero if T varies proportionally with $\mathbf{B}:\boldsymbol{\varepsilon}^{\mathrm{e}}$ across the shock transition layer. With the trapezoidal approximation (5.23), (5.21) and (5.22) may be written as

$$\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket:\mathbb{E}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket \geqslant \rho_0 c_0^2\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}+\llbracket T\rrbracket\mathbf{B}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket-\eta, \tag{5.24}$$

$$\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket:\mathbb{E}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket \leqslant \rho_0 c_0^2\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}+\llbracket T\rrbracket\mathbf{B}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket+\eta. \tag{5.25}$$

These restrictions may be phrased in terms of the Cauchy stress by using (4.22) to write

$$\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket:\mathbb{E}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket-\llbracket T\rrbracket\mathbf{B}:\llbracket\boldsymbol{\varepsilon}^{\mathrm{e}}\rrbracket = \llbracket\boldsymbol{\sigma}\rrbracket:\mathbb{M}:\llbracket\boldsymbol{\sigma}\rrbracket+\llbracket T\rrbracket\boldsymbol{\beta}:\llbracket\boldsymbol{\sigma}\rrbracket, \tag{5.26}$$

where $\mathbb{M}\equiv\mathbb{E}^{-1}$ and $\boldsymbol{\beta}\equiv\mathbb{M}:\mathbf{B}$. Using (5.26) and (2.27), (5.24) and (5.25) become

$$\mathbf{n}\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\mathbf{n}-x(\llbracket\boldsymbol{\sigma}\rrbracket:\mathbb{M}:\llbracket\boldsymbol{\sigma}\rrbracket+\llbracket T\rrbracket\boldsymbol{\beta}:\llbracket\boldsymbol{\sigma}\rrbracket) \leqslant x\eta, \tag{5.27}$$

$$\mathbf{n}\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\mathbf{n}-x(\llbracket\boldsymbol{\sigma}\rrbracket:\mathbb{M}:\llbracket\boldsymbol{\sigma}\rrbracket+\llbracket T\rrbracket\boldsymbol{\beta}:\llbracket\boldsymbol{\sigma}\rrbracket) \geqslant -x\eta, \tag{5.28}$$

where (recall) the first inequality holds if the choice $\boldsymbol{\sigma}^*=\boldsymbol{\sigma}^+$ is admissible and the second inequality holds if the choice $\boldsymbol{\sigma}^*=\boldsymbol{\sigma}^-$ is admissible. Here, x is the wave speed eigenvalue, defined $x\equiv\rho^2c^2/\rho_0=\rho c^2/J=\rho_0c_0^2/\Gamma^2$, where $\Gamma\equiv J\mathbf{N}\cdot\mathbf{F}^{-1}\cdot\mathbf{n}$. Assuming $\llbracket\rho_0\rrbracket=0$, (2.14) shows that $\llbracket x\rrbracket=0$.

5.3. *Specialization for non-hardening materials*

Suppose the choices $\boldsymbol{\sigma}^*=\boldsymbol{\sigma}^+$ and $\boldsymbol{\sigma}^*=\boldsymbol{\sigma}^-$ are *both* admissible, as is the case, for example, for a non-hardening material with an isotropic yield surface. Then (5.27) and (5.28) together imply $\eta\geqslant 0$ and, more importantly, if η is negligible, (5.27) and (5.28) imply the equality

$$\boxed{\mathbf{n}\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\llbracket\boldsymbol{\sigma}\rrbracket\cdot\mathbf{n}-x(\llbracket\boldsymbol{\sigma}\rrbracket:\mathbb{M}:\llbracket\boldsymbol{\sigma}\rrbracket+\llbracket T\rrbracket\boldsymbol{\beta}:\llbracket\boldsymbol{\sigma}\rrbracket)=0}\,. \tag{5.29}$$

The equation (5.29) is the large deformation thermodynamical generalization of Drugan and Shen's (1987) and (1990) purely mechanical analyses. As discussed by Drugan and Shen (1987), the restriction (5.29) holds not only across the entire shock, but also across any sub-section of the stress and temperature path through the shock, in which case, the double brackets, $[\![\ ]\!]$, are replaced by angled brackets, $\langle\ \rangle$, defined as the jump across the subsection.

Recall that for small displacement gradients, $\mathbb{M}$ and $\boldsymbol{\beta}$ may permissibly be anisotropic, but for large displacement gradients, the derivation of (5.29) required an assumption of isotropic thermoelastic response, meaning that the tensors $\mathbb{M}$ and $\boldsymbol{\beta}$ are isotropic, in which case, (5.29) may be written as

$$\mathbf{n}\cdot[\![\mathbf{S}]\!]\cdot[\![\mathbf{S}]\!]\cdot\mathbf{n}-2[\![p]\!]\mathbf{n}\cdot[\![\mathbf{S}]\!]\cdot\mathbf{n}+[\![p]\!]^2 - x\left(\frac{1}{2G}[\![\mathbf{S}]\!]:[\![\mathbf{S}]\!]+\frac{1}{K}[\![p]\!]^2-\beta[\![T]\!][\![p]\!]\right)=0, \tag{5.30}$$

where $\mathbf{S}$ is the deviatoric stress, p is the spherical pressure [i.e. $p = -\frac{1}{3}\text{trace}(\boldsymbol{\sigma})$], G is the elastic shear modulus, K is the elastic bulk modulus, and β is the coefficient of volumetric expansion. Several specific applications of the preceding results will now be discussed.

5.4. *Special case: stress path throughout shock is below yield*

If the stress state is below yield for all states in the transition layer and if the thermal power $\mathscr{P}_T$ can be neglected, then the transition is isentropic and, therefore, the wave speed eigenvalue x is a principal value of the "isentropic acoustic tensor", $\mathbf{A}_s = \mathbf{n}\cdot\mathbb{E}_s\cdot\mathbf{n}$, where $\mathbb{E}_s$ is the isentropic elastic stiffness tensor, defined as

$$\mathbb{E}_s = \mathbb{E}+\frac{T_1}{\rho_0\varsigma_v}\mathbf{BB}. \tag{5.31}$$

The temperature T_1 in this definition corresponds to the reference STATE 1 in the expansions (4.20) and (4.21). Noting that $\mathbf{A}_s$ is the sum of a symmetric second order tensor plus a dyad, the spectral analysis of Brannon and Drugan (1993) may be used to find the eigenvalues. For elastic isotropy, $\mathbf{B} = K\beta\mathbf{I}$, and therefore

$$\mathbf{A}_s = \mathbf{A}_T+\frac{T_1K^2\beta^2}{\rho_0\varsigma_v}\mathbf{nn} = \mathbf{A}_T+(\varsigma_p-\varsigma_v)K\mathbf{nn}, \tag{5.32}$$

where ς_p is the specific heat at constant stress and $\mathbf{A}_T$ is the isothermal acoustic tensor, $\mathbf{n}\cdot\mathbb{E}\cdot\mathbf{n}$. The eigenvalues of $\mathbf{A}_s$ are G and $2G+\lambda^*$, where $\lambda^* = \lambda+(\varsigma_p-\varsigma_v)K$. Because $\varsigma_p \approx \varsigma_v$ for most dense solids (see, e.g. Ashby, 1989), the isothermal eigenvalues differ very little from the isentropic eigenvalues.

5.5. *Special case: axisymmetric stress—existence and stability*

Suppose that, throughout the shock transition layer, the stress tensor is of the uniaxial form $\boldsymbol{\sigma} = \sigma_n\mathbf{nn}+\sigma_t\,(\mathbf{I}-\mathbf{nn})$, where σ_n and σ_t are scalars. Then the deviatoric stress is of the form $\mathbf{S} = \tau\,(3\mathbf{nn}-\mathbf{I})$, where $\tau \equiv \frac{1}{3}(\sigma_n-\sigma_t)$. Note that the only way to change $\mathbf{S}$ is to change its Euclidean magnitude—its direction is constant. If one hypothesizes a sub-section of the shock transition layer throughout which a pressure-insensitive non-hardening yield condition is satisfied, then a change in $\mathbf{S}$ is necessarily accompanied by a change in the direction of $\mathbf{S}$ which, as just noted, is impossible for generalized uniaxial stress states. Therefore, $\langle \mathbf{S} \rangle = \mathbf{0}$ across any sub-section of the shock transition zone that is at yield throughout. In other words, the jump in stress across such a sub-section must be isotropic, meaning that the jump in stress must be of the form $\langle \boldsymbol{\sigma} \rangle = -\langle p \rangle \, \mathbf{I}$. The only way the jump in stress can be non-isotropic is if every state throughout the sub-section is *below yield*; such a stress jump must therefore propagate at an elastic wave speed.

Suppose a shock transition zone is *hypothesized* to consist of a "plastic zone" throughout which the yield condition is satisfied and an "elastic zone" throughout which the stress state is below yield. Using $\langle \mathbf{S} \rangle = \mathbf{0}$ in (5.30) shows that, within the plastic zone, either $\langle p \rangle = 0$ or the plastic zone must have the propagation eigenvalue

$$x = \frac{K\langle p \rangle}{\langle p \rangle - K\beta\langle T \rangle}. \tag{5.33}$$

This plastic eigenvalue depends on both the jump in pressure and the jump in temperature. If thermomechanical coupling is neglected (i.e. if $\beta = 0$ or $\langle T \rangle = 0$), then the eigenvalue reduces to the known (Morland, 1959; Milne *et al.*, 1988) purely mechanical prediction: $x = K$. If the shock profile is to be stable, the hypothesized accompanying elastic zone must *also* propagate with the eigenvalue $x = K$, but since K is not an elastic eigenvalue as defined by (5.32), an accompanying elastic zone must be ruled out. Therefore, the shock must be *entirely* elastic (i.e. entirely below yield); *entirely* elastic–plastic (i.e. entirely at yield); or a mix of elastic and elastic–plastic, with the jump in stress being zero across the elastic–plastic sub-section(s). By (2.27), the latter scenario is possible only if the shock is quasi-static.

Suppose that a shock in a thin bar is to be modeled by simply taking the normal stress as non-zero with all other stress components as zero. Then a jump in stress across any sub-section of the shock transition layer is necessarily accompanied by a jump in deviatoric stress. Therefore the entire sub-section must be below yield, and the shock must propagate at an elastic wave speed. For elastic isotropy, this elastic wave speed is $\sqrt{E/\rho}$, where E is Young's modulus. This elastic wave speed is not one of the elastic wave speeds derived in Section 5.4, which were obtained by employing the jump condition (2.22) which is not sensible for a shock in a thin bar. For a shock in a thin bar, lateral strain may jump. Only continuity of the normal component of displacement may be sensibly enforced. Taking only the normal component of stress as non-zero and enforcing continuity of normal displacement, the entire analysis leading up to (5.29) may be modified to show that (5.29) is valid for the thin bar stress state even though (5.1) is not valid.

6. SUMMARY AND CONCLUSIONS

An approximation of a general shock by a smooth wave is legitimized in Section 3 by showing that changes across a shock of internal energy and, therefore, of variables such as stress and entropy (defined by constitutive derivatives of energy) agree until third order in material time rates of fundamental field variables with the changes predicted from a smooth wave analysis. This result, which is independent of the constitution of the material, corrects a common misconception that entropy changes across the shock must be small in order to approximate the shock by a smooth wave; even if entropy changes are significant at first order (as with elastic–plastic waves) the change in entropy predicted by a smooth wave approximation will agree until third order with the actual change across the shock. For the special class of steady shock waves whose reference configuration images propagate without rotating, the Taylor expansion of the Hugoniot function for a smooth wave of the same type, evaluated at the state just ahead of the wave, is shown to have coefficients of zero for *every* term, which concurs with the shock value of zero for the Hugoniot. This shows that a smooth wave provides a *very* accurate, possibly exact, representation of the sequence of states experienced by a material particle during passage of a shock wave of this class. For this special shock wave class, we also showed that the rate forms of the conservation of momentum and compatibility are identical for shocks and smooth waves, both at the states just ahead of and just behind the waves, and for all orders of material time rate.

Having legitimized the approximation of a shock by a smooth wave, several new large and small deformation steady thermoelastic–plastic shock wave existence conditions are deduced by integrating the maximum plastic work inequality across the shock. Inclusion of thermal effects in elastic–plastic shock analyses leads to shock existence conditions that are not quite as restrictive as those obtained using a purely mechanical analysis, except when the material is non-hardening. If thermomechanical coupling is neglected, the existence conditions reduce to those derived by Drugan and Rice (1984) and Drugan and Shen (1987, 1990). Consequently, neglect of thermomechanical coupling may rule out the existence of shocks, or, as shown by Drugan and Shen (1987), may require that the shocks propagate only at an elastic wave speed. Thus we have rigorously shown, possibly exactly for steady-state non-rotating shocks and to an accurate approximation for general shocks of low to moderate strength, that propagating surfaces of discontinuity in solutions of purely mechanical elastic–plastic boundary-value problems are required by the underlying laws of thermodynamics to satisfy the restrictions derived by Drugan and Rice (1984) and Drugan and Shen (1987, 1990) when thermal effects are neglected, and we have derived the generalizations of these restrictions when thermal effects are included in the formulation.

ACKNOWLEDGEMENTS

Support by the Mechanics and Materials Program of the National Science Foundation under Grants MSS-8552486 and MSS-9215688 is gratefully acknowledged. This work was completed while one of the authors (WJD) held a SERC Visiting Fellowship (Grant GR/J 98967) in the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, U.K. It is a pleasure to thank Professor J. R. Willis for his interest and for helpful discussions.

REFERENCES

Ashby, M. F. (1989) On the engineering properties of materials. *Acta Metall.* **37,** 1273–1293.
Bataille, J. and Kestin, J. (1975) *J. Méc.* **14,** 365–384.

Brannon, R. M. and Drugan, W. J. (1993) Influence of non-classical elastic–plastic constitutive features on shock wave existence and spectral solutions. *J. Mech. Phys. Solids* **41,** 297–330.

Chadwick, P. (1976) *Continuum Mechanics: Concise Theory and Problems.* Wiley, New York.

Coleman, B. D. and Noll, W. (1963) The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Rat. Mech. Anal.* **13,** 167–178.

Courant, R. and Friedrichs, K. O. (1948) *Supersonic Flow and Shock Waves* (Third Printing: 1985). Springer, Berlin.

Dillon, O. W., Jr. (1967) Plastic deformation waves and heat generated near the yield point of annealed aluminum. In *Mechanical Behavior of Materials under Dynamic Loads* (ed. U. S. Lindholm), pp. 21–60, ARO SRI Symposium, San Antonio. Springer-Verlag, New York.

Drugan, W. J. and Rice, J. R. (1984) Restrictions on quasi-statically moving surfaces of strong discontinuity in elastic–plastic solids. In *Mechanics of Material Behavior* (ed. G. J. Dvorak and R. T. Shield), pp. 59–73. Elsevier Science, Amsterdam.

Drugan, W. J. and Shen, Y. (1987) Restrictions on dynamically propagating surfaces of strong discontinuity in elastic–plastic solids. *J. Mech. Phys. Solids* **35,** 771–787.

Drugan, W. J. and Shen, Y. (1990) Finite deformation analysis of restrictions on moving strong discontinuity surfaces in elastic–plastic materials: quasi-static and dynamic deformations. *J. Mech. Phys. Solids* **38,** 553–574.

Farren, W. S. and Taylor, G. I. (1925) The heat developed during plastic extension of metals. *Proc. R. Soc. London A* **107,** 422–451.

Germain, P. and Lee, E. H. (1973) On shock waves in elastic–plastic solids. *J. Mech. Phys. Solids* **21,** 359–382.

Gurtin, M. E. and Williams, W. O. (1966) On the Clausius–Duhem inequality. *ZAMP* **17,** 626–633.

Gurtin, M. E. and Williams, W. O. (1967) An axiomatic foundation for continuum thermodynamics. Carnegie Institute of Technology Report.

Hadamard, J. (1903) *Lecons sur la Propagation des Ondes et les Equations de l'Hydrodynamique.* Hermann, Paris.

Hill, R. (1961) Discontinuity relations in mechanics of solids. In *Progress in Solid Mechanics,* **2** (ed. R. Hill and I. N. Sneddon), pp. 244–276. North Holland, Amsterdam.

Jou, D., Casas-Vázquez, J. and Lebon, G. (1988) Extended irreversible thermodynamics. *Rep. Prog. Phys.* **51,** 1105–1179.

Kestin, J. and Rice, J. R. (1970) *A Critical Review of Thermodynamics* (ed. E. B. Stuart *et al.*), pp. 275–298. Mono Book Corp., Baltimore.

Lax, P. D. (1973) Hyperbolic systems of conservation laws and the mathematical theory of shock waves. *Regional Conference Series in Applied Mathematics, No.* 11. SIAM, Philadelphia.

Lee, E. H. (1969) Elastic–plastic deformation at finite strains. *J. Appl. Mech.* **36,** 1–6.

Lee, E. H. and Liu, D. T. (1967). Finite strain elastic–plastic theory particularly for plane wave analysis. *J. Appl. Phys* **38,** 19–27.

Lubliner, J. (1972) *Int. J. Non-Linear Mech.* **7,** 237–254.

Mandel, J. (1972) *Plasticité Classique et Viscoplasticité* (CISM, Udine, 1971). Springer-Verlag, Vienna.

Milne, P. C., Morland, L. W. and Yeung, W. (1988) Spherical elastic–plastic wave solutions. *J. Mech. Phys. Solids* **36,** 15–28.

Morland, L. W. (1959) The propagation of plane irrotational waves through an elastoplastic medium. *Phil. Trans. Roy. Soc. A* **251,** 341–383.

Mróz, Z. and Raniecki, B. (1976) On the uniqueness problem in coupled thermoplasticity. *Int. J. Engng Sci.* **14,** 211–221.

Perzyna, P. (1971) Thermodynamics of rheological materials with internal changes. *J. Méc.* **10,** 391–408.

Resnick, R. and Halliday, D. (1977) *Physics—Part* 1, 3rd Edn. Wiley, New York.

Rice, J. R. (1971) Inelastic constitutive relations for solids: an internal-variable theory and its application to metal plasticity. *J. Mech. Phys. Solids* **19,** 433–455.

Truesdell, C. and Toupin, R. A. (1960). The classical field theories. *Handbuch der Physik, III*/1 (ed. S. Flügge). Springer-Verlag, Berlin.

Wallace, D. C. (1985) *Thermoelastic–Plastic Flow in Solids.* Los Alamos National Laboratory Monograph LA-10119 UC-34, Los Alamos.

Wallace, D. C. (1991) *Structure of Shocks in Solids and Liquids: Six Reprints with an Introduction.* (Includes supplementary discussions and errata for papers originally appearing in *Phys. Rev.* **B22** (1980), **B24** (1981) and **A25** (1982)). Los Alamos National Laboratory Monograph LA-12020 UC-700 and UC-706, Los Alamos.