THE EQUATIONS OF ELASTICITY ARE SPECIAL.
THE EQUATIONS OF ELASTICITY ARE SPECIAL

by

C. M. Dafermos

Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

This research was supported in part by the National Science Foundation under MCS 79-05774, in part by the Office of Naval Research under N0014-76-C-0278-P0002 and in part by the United States Army under ARO-DAAG29-76-G-0294.
1. Introduction

Anyone who is familiar with recent developments in continuum mechanics is aware of the many special features of the equations of hyperelasticity and thermoelasticity. From the viewpoint of analysis, the most aggravating and, at the same time, most challenging characteristic is that internal energy cannot be a globally convex function of deformation gradient without violating the principle of material frame indifference [1].

A consequence of this lack of convexity is that the classical direct methods of the calculus of variations are ineffective for establishing existence, uniqueness and stability of equilibrium solutions. Among the various restrictions on material response that have been laid down for study in the theory of elasticity [1], Strong Ellipticity appears interesting and promising. The work of Ball [2] has produced a breakthrough by establishing the existence of equilibrium solutions in isotropic hyperelasticity under the assumption that the stored-energy function is polyconvex, a condition somewhat stronger than Strong Ellipticity but definitely weaker than convexity and, in particular, not incompatible with frame indifference.

In dynamics, the goal is to relate the Second Law of thermodynamics with stability. The work of Ericksen [3] reveals that the Second Law induces stability of an equilibrium state of a thermoelastic material in the range of convexity of internal energy (e.g. at points where energy attains a minimum.) Similarly, it has been shown [4] that smooth adiabatic processes in
thermoelasticity, residing in the range of convexity of internal energy, depend continuously upon the initial state, even within the class of processes that develop shock waves. Actually, in certain circumstances, the same result can be established under a mere Strong Ellipticity assumption [4].

Here we investigate the implications of the Second Law upon a different manifestation of stability, namely the stability of weak shock waves expressed by the E-condition of Lax [5]. If energy were convex, the results of Lax [6] on general hyperbolic systems of conservation laws endowed with a convex "entropy" would imply that the Second Law is equivalent to the E-condition. A property of this nature was established by Malek-Madani [7] in the context of hyperelasticity under the weaker assumption of Strong Ellipticity. In [7] an "energy criterion" plays the role of the Second Law. The work of Ericksen [8] suggests that certain phase transitions may be interpreted through failure of Strong Ellipticity so it would be advisable to avoid adopting this condition at the outset. Consequently, our project here is to work in the broader framework of adiabatic thermoelasticity, in which the Second Law is expressed in a definitive form by the Clausius-Duhem inequality, and to dispense with the need of imposing Strong Ellipticity in advance. We show that, due to the special structure of the equations of elasticity, the Second Law implies that every weak shock satisfies Lax's E-condition, without regard to convexity or Strong Ellipticity.

It would be interesting to examine the implications of lack of convexity upon other shock admissibility criteria. This program
is carried out, to a certain extent, in [7], within the context of hyperelasticity, and in [9], in the context of one dimensional elasticity.

2. Adiabatic Processes in Thermoelasticity

We consider a body with reference configuration $\mathcal{D} \subset \mathbb{R}^n$, $n = 2$ or $3$, and reference density $\rho(\mathbf{x})$. A thermodynamic process of $\mathcal{D}$ is determined by a motion $\mathbf{x}(\mathbf{x}, t)$ and a (specific) entropy field $\eta(\mathbf{x}, t)$. A motion generates a velocity field $\mathbf{v} = \mathbf{x}$ and a deformation gradient field $F = \nabla x$, $\det F > 0$. An adiabatic process with zero body force is governed by the conservation laws of momentum and energy, viz.,

\begin{align}
\dot{F}_{i\alpha} - v_i,_{\alpha} &= 0 \\
\rho \dot{v}_i - T_{i\alpha},_{\alpha} &= 0 \\
\dot{\rho}E - (T_{i\alpha}v_i),_{\alpha} &= 0 \tag{2.1}
\end{align}

and the Clausius-Duhem inequality

\[ \rho \dot{\eta} \geq 0, \tag{2.2} \]

where $T$ is the Piola-Kirchhoff stress and $E$ is the (specific) energy, sum of kinetic energy and internal energy $\epsilon$,

\[ E = \frac{1}{2} v_i v_i + \epsilon. \tag{2.3} \]
In thermoelasticity theory $\varepsilon, T$ and temperature $\theta$ are determined by $F$ and $n$ via constitutive relations

$$\varepsilon = \hat{\varepsilon}(F, n; X),$$

$$T_{i\alpha} = \hat{T}_{i\alpha}(F, n; X) = \rho \frac{\partial \hat{\varepsilon}}{\partial F_{i\alpha}},$$

$$\theta = \hat{\theta}(F, n; X) = \frac{\partial \hat{\varepsilon}}{\partial n} > 0.$$ (2.6)

For every proper orthogonal matrix $R$, the function $\hat{\varepsilon}$ should satisfy on its domain the condition

$$\hat{\varepsilon}(RF, n; X) = \hat{\varepsilon}(F, n; X),$$ (2.7)

which is dictated by the principle of material frame indifference, and possibly additional restrictions expressing material symmetry.

One of the implications of (2.7) is [1, §52] that $\hat{\varepsilon}$ cannot be a globally convex function of $F$, for fixed $n$ and $X$. Weaker than convexity and not necessarily incompatible with (2.7) is rank-one convexity (Hadamard's condition). Uniform rank-one convexity or Strong Ellipticity requires that at every point in the domain of $\hat{\varepsilon}$ and for any unit vector $N$ in $\mathbb{R}^n$ the matrix $P(N)$,

$$P_{ij}(N) = \frac{\partial^2 \hat{\varepsilon}}{\partial F_{i\alpha} \partial F_{j\beta}} N_\alpha N_\beta,$$ (2.8)
is positive definite. This condition implies that the system (2.1) is hyperbolic. The characteristic speeds, in the direction \( N \), are 0, with multiplicity \( n(n-1) + 1 \), and \( \pm \sqrt{\lambda_i} \), \( i = 1, \ldots, n \), where \( \lambda_1, \ldots, \lambda_n \) are the eigenvalues of \( P(N) \).

3. Lax's E-condition and the Second Law of Thermodynamics

In view of (2.5) and (2.6) every smooth process that satisfies the conservation laws (2.1) satisfies automatically the Clausius-Duhem inequality (2.2), as an equality. However, smooth solutions of (2.1) generally break down and shock waves develop. It is in the class of discontinuous solutions that the role of (2.2) becomes important.

Functions of bounded variation in the sense of Tonelli-Cesari [10] constitute the natural class in which solutions of (2.1) should be sought. Functions in this class are endowed with geometric structure that closely resembles that of piecewise smooth functions. In particular, within the class of solutions of bounded variation one may distinguish shock waves across which the classical Rankine-Hugoniot jump conditions are satisfied, namely,

\[
\begin{align*}
    s[F_{i\alpha}] + [v_i]N_\alpha &= 0 \\
    \rho s[v_i] + [T_{i\alpha}v_i]N_\alpha &= 0 \\
    \rho s[E] + [T_{i\alpha}v_i]N_\alpha &= 0
\end{align*}
\]
where \( \mathbf{N} \) is a unit vector pointing in the direction of propagation and \( s \) is the speed of propagation of the shock. In (3.1) and throughout, a bracket \([\ ]\) denotes the jump (left minus right) of the enclosed quantity across the shock. The jump condition associated with (2.2) reads

\[
\rho s[\eta] \leq 0.
\] (3.2)

A shock that satisfies (3.2) will be called admissible.

It can be shown [11] that strict rank-one convexity is equivalent to nonexistence of isentropic shocks of speed 0 (if one allows \([\eta] \neq 0\), this is no longer true). On the contrary, as the following two propositions indicate, positive eigenvalues of \( P(\mathbf{N}) \) and weak shocks of nonzero speed go hand in hand.

**Proposition 3.1.** Assume that every neighborhood of a state \((F^-,v^-,n^-)\) contains some point \((F^+,v^+,n^+)\) which can be connected to \((F^-,v^-,n^-)\) with a (not necessarily admissible) shock propagating with speed \( s \), \(|s| > s_0 > 0\). Then there is a unit vector \( \mathbf{N} \) in \( \mathbb{R}^n \) such that the matrix \( P(\mathbf{N}) \), defined by (2.8), at \((F^-,v^-,n^-)\) has at least one eigenvalue \( \lambda \geq s_0^2 \).

**Proof.** We fix a point \((F^+,v^+,n^+)\) which can be connected to \((F^-,v^-,n^-)\) with a weak shock propagating in the direction, say, \( \hat{\mathbf{N}} \) with speed \( s \). From (3.1), \( s^2 [F_{i\alpha}][F_{i\alpha}] = [v_i][v_i] \) so that \([F]\) and \([v]\) are of the same order of magnitude. We now
show that, in contrast, \([n]\) is of higher order than \(\|v\|\).

Using (3.1), (2.3), (2.5) and (2.6),

\[
0 = \rho s[E] + [T_{ia}v_i]N_\alpha = \frac{1}{2} \rho s[v_i^+v_i^-] + \rho s[\epsilon] + [T_{ia}v_i]N_\alpha
\]

\[
= \frac{1}{2} \rho s(v_i^+v_i^-)[v_i] + sT_{ia}[F_{ia}] + \frac{1}{2} \rho s \frac{\partial^2 \xi}{\partial F_{ia} \partial F_{j\beta}} [F_{ia}][F_{j\beta}]
\]

\[
+ \rho s\bar{\sigma}[n] + [T_{ia}v_i]N_\alpha
\]

where

\[
\bar{\sigma} = \sigma(E^+, n), \quad \frac{\partial^2 \xi}{\partial F_{ia} \partial F_{j\beta}} = \frac{\partial^2 \xi}{\partial F_{ia} \partial F_{j\beta}} (F, n^+)
\]

with \(n^+\), \(n^-\) and \(F\) on the straight line segment in \(R^n\) which joins \(E^-\) with \(E^+\). Hence, on account of (3.1) and (2.8),

\[
-\rho s\bar{\sigma}[n] = \frac{1}{2} \rho s(v_i^+v_i^-)[v_i] + T_{ia}[v_i]N_\alpha + [T_{ia}v_i]N_\alpha
\]

\[
+ \frac{1}{2} \frac{\rho}{s} F_{ij}(\bar{\sigma})[v_i][v_j]
\]

\[
= \frac{1}{2} \rho s(v_i^+v_i^-)[v_i] + v_i^+[T_{ia}]N_\alpha + \frac{1}{2} \frac{\rho}{s} F_{ij}(\bar{\sigma})[v_i][v_j]
\]

\[
= \frac{1}{2} \rho s(v_i^+v_i^-)[v_i] - \rho s[v_i^+] + \frac{1}{2} \frac{\rho}{s} F_{ij}(\bar{\sigma})[v_i][v_j]
\]

\[
= -\frac{1}{2} \rho s[v_i][v_i] + \frac{1}{2} \frac{\rho}{s} F_{ij}(\bar{\sigma})[v_i][v_j],
\]

(3.4)
which shows that $|\eta|$ is at most of order $|\eta|^2$.

On the strength of the above results, combining (3.1)$_1$ with (3.1)$_2$ we obtain

$$P_{ij}(\hat{\eta})[v_j] - s^2[v_i] = 0(|\eta|^2) \quad (3.5)$$

with $P(\hat{\eta})$ evaluated at $(\xi^-,\eta^-)$. Dividing through by $|\eta|$ in (3.5) and passing to the limit as $|\eta| \to 0$, using the compactness of the unit sphere in $\mathbb{R}^n$ we arrive at the assertion. The proof is complete.

**Proposition 3.2.** Assume that $P(\hat{\eta})$, at a state $(\xi^0,\eta^0)$ and a vector $\eta$, has a simple positive eigenvalue $\lambda_0$. Set

$$\mu_0 = \sqrt{\lambda_0} \quad \text{or} \quad \mu_0 = -\sqrt{\lambda_0} \quad \text{and fix any } \eta^0 \text{ in } \mathbb{R}^n.$$  

Then there are two smooth maps from an interval $(-\delta,\delta)$ to $\mathbb{R}$ and $\mathbb{R}^{n^2+n+1}$ which carry $\tau \in (-\delta,\delta)$ into $s(\tau)$ and $(\xi,\eta,\hat{\eta})(\tau)$, respectively, with the property $s(0) = \mu_0, (\xi,\eta,\hat{\eta})(0) = (\xi^0,\eta^0,\hat{\eta}^0)$ and such that $(\xi,\eta,\hat{\eta})(\tau)$ can be connected to $(\xi^0,\eta^0,\hat{\eta}^0)$ by a shock propagating in the direction $\eta$ with speed $s(\tau)$, i.e.,

$$s(F_{ia}-F_{ia}) + (v_i-v_i^0)N_\alpha = 0$$

$$\rho s(v_i-v_i^0) + (T_{ia}-T_{ia}^0)N_\alpha = 0 \quad (3.6)$$

$$\rho s(E-E^0) + (T_{ia}v_i-T_{ia}^0v_i^0)N_\alpha = 0.$$ 

Furthermore, $(\xi,\eta,\hat{\eta})(\tau)$, $\tau \in (-\delta,\delta)$, are the only states in some neighborhood of $(\xi^0,\eta^0,\hat{\eta}^0)$ that can be connected to $(\xi^0,\eta^0,\hat{\eta}^0)$
by a shock of speed close to $\mu_0$.

**Proof.** First we observe that, since $p > 0$ and $\partial \rho E / \partial \eta = \rho \phi > 0$, the Jacobian matrix $\partial (\rho, \rho \phi, \rho E) / \partial (F, y, \eta)$ is nonsingular. Thus, we may visualize (3.6) as an equation $sZ - G(Z) = 0$ where $Z = (F - F^0, \rho \phi - \rho \phi^0, \rho E - \rho E^0)$. We observe that $G(0) = 0$. Furthermore, since $\lambda_0$ is a simple eigenvalue of $P(N)$, $\mu_0$ is a simple characteristic speed of (2.1) so that $\mu_0$ is a simple eigenvalue of the matrix $G_z(0)$. The assertion of the proposition then follows from a standard theorem in bifurcation theory [12].

We now assume that the conditions of Proposition 3.2 hold and we discuss the admissibility of the shocks associated with the shock curve $(F, y, \eta)(\tau)$.

For $\tau$ near 0, $P(N)$, evaluated at $(F, \eta)(\tau)$, will have a simple eigenvalue $\lambda(\tau)$ near $\lambda_0$ and, therefore, (2.1) will have a characteristic speed $\mu(\tau)$ near $\mu_0$. The shock which connects $(F, y, \eta)(\tau)$ with $(F^0, y^0, \eta^0)$ will satisfy Lax's E-condition [5] if

$$\mu(\tau) < s(\tau) < \mu_0. \quad (3.7)$$

This condition may be motivated in a variety of ways. For example, it can be interpreted as a stability statement for the systems derived by linearizing (2.1) on both sides of the shock.

If (2.1) were genuinely nonlinear [5] and $\tilde{E}$ were a uniformly convex function of $(F, \eta)$ (in which case $-\eta$ would be a uniformly convex function of $(F, \rho \phi, \rho E)$ [13]), (3.2) and (3.7) would be equivalent by a theorem of Lax [6]. We wish to investigate the
relationship between (3.7) and the Second Law of thermodynamics without imposing any convexity restrictions upon \( \hat{c} \).

In order to avoid the degenerate case where \( s'(\tau)^* \) changes sign infinitely many times in every neighborhood of 0, we make the assumption

\[
s^{(\ell)}(0) \neq 0, \text{ for some } \ell \geq 1, \tag{3.8}
\]

which is generically satisfied. For \( \ell = 1 \), (3.8) expresses the genuine nonlinearity of the \( \mu_0 \) characteristic field at \((F^0, \eta^0, n^0)\) and it will be satisfied if and only if

\[
\frac{\partial^3 \hat{c}(F^0, \eta^0)}{\partial F_\alpha \partial F_\beta \partial F_\gamma} N_\alpha N_\beta N_\gamma r_1 r_j r_k \neq 0, \tag{3.9}
\]

where \( \tau \) is the eigenvector of \( P(N) \) associated with the eigenvalue \( \lambda_0 \).

We are now prepared to state the main result:

**Proposition 3.3.** Under assumption (3.8) and for \( \tau \) sufficiently close to 0, Lax's E-condition (3.7) is equivalent to the strict Clausius-Duhem inequality

\[
s(\tau)(\eta(\tau) - \eta^0) < 0. \tag{3.10}
\]

\((*)\) Here and throughout a prime (') denotes derivative with respect to \( \tau \).
Proof. Differentiating (3.6) with respect to \( \tau \),

\[ \frac{dF_i}{d\tau} + v_i'N_a = -s'(F_i - F_i^0) \quad (3.11) \]

\[ \rho s v_i' + T_i' N_a = -\rho s'(v_i - v_i^0) \quad (3.12) \]

\[ \rho s E' + T_i' v_i' N_a + T_i' v_i'^2 N_a = -\rho s'(E - E^0). \quad (3.13) \]

Using (2.3), (2.5) and (2.6), (3.13) yields

\[ \rho s v_i v_i' + T_i' F_i + (\rho s \theta) n' + T_i' v_i' N_a + T_i' v_i'^2 N_a = -\rho s'(E - E^0) \]

whence, by virtue of (3.11) and (3.12),

\[ \rho s \theta n' = -s'(\rho(E - E^0) - T_i(F_i - F_i^0) - \rho v_i (v_i - v_i^0)). \quad (3.14) \]

On account of (3.6), (3.14) gives

\[ \rho s^2 \theta n' = -s'(-T_i v_i - T_i^0 v_i^0) N_a + T_i (v_i - v_i^0) N_a - \rho s v_i (v_i - v_i^0) \]

\[ = -s'(-v_i^0 (T_i - T_i^0) N_a - \rho s v_i (v_i - v_i^0)) \]

\[ = -s'(\rho s v_i^0 (v_i - v_i^0) - \rho s v_i (v_i - v_i^0)) \]

\[ = \rho s s' (v_i - v_i^0)(v_i - v_i^0). \]

Hence,
\( s' \approx n' = s'(v_i - v_i^o)(v_i - v_i^o). \) \hspace{1cm} (3.15)

In view of (3.8), for \( \bar{\tau} \) near 0, \( s' \) does not change sign between 0 and \( \bar{\tau} \). It then follows from (3.15) that \( n' \) also does not change sign between 0 and \( \bar{\tau} \) and that \( s_n' \) and \( s' \) have the same sign on this interval. Thus, (3.10) is equivalent to \( s(\tau) < s(0) = \nu_0, \ \tau \in (0, \tau). \) By reversing the roles of left and right state one shows that (3.10) is also equivalent to \( s(\tau) > \nu(\tau). \) The proof is complete.
REFERENCES


