Convexity Conditions and Existence Theorems in Nonlinear Elasticity

JON M. BALL

Communicated by S.S. Antman & C.M. Dafermos

Table of Contents

0. Introduction ............................................. 337
1. Boundary-value problems of non-linear hyperelasticity .......... 342
2. Technical preliminary ..................................... 347
3. Quasiconvexity and the Legendre-Hadamard condition ............ 349
4. Sufficient conditions for quasiconvexity ........................ 356
5. Isotropic convex and polyconvex functions ...................... 363
6. Sequential weak continuity of mappings on Orlicz-Sobolev spaces 367
7. Existence theorems ........................................ 371
8. Applications to specific models of elastic materials ............ 384
9. An example of non-uniqueness: buckling of a rod ................ 398
10. Conclusion remarks ....................................... 399
Acknowledgement ............................................ 399
References ................................................ 337

0. Introduction

The purpose of this article is to present existence theorems for various equilibrium boundary-value problems of nonlinear elasticity in one, two and three dimensions under realistic hypotheses on the material response. Although some of the results may be extended to cover Cauchy elasticity, we shall restrict our discussion to hyperelastic (Green elastic) materials, that is, to elastic materials possessing a stored-energy function. We ignore thermal effects. For such materials a typical boundary-value problem takes the form of finding a vector field $u_0: \Omega \to \mathbb{R}^n$ making the integral

$$ I(u_0, \Omega) = \int_\Omega f(x, u(x), \nabla u(x)) \, dx $$

stationary in a suitable class of functions. Here $\Omega$ is a non-empty, bounded, open subset of $\mathbb{R}^n$, $n = 1, 2, 3$. The integrand $f$ will usually have the form

$$ f(x, u, \nabla u) = y(x, \nabla u) + \phi(x, u). $$

For traction boundary-value problems there will also be a surface integral term.
homogeneous strain, we require that this homogeneous strain be an absolute minimizer for the total energy. Note that if in the above we admitted for consideration inhomogeneous bodies, or if we considered mixed displacement traction boundary-value problems, then the condition would be unacceptable, as we should expect certain buckled states to have lower total energy than the homogeneous strain. As stated, however, the condition has a certain plausibility.

Morrey showed that if \( f(-, u, \cdot) \) is quasiconvex for every \( u \), and if certain continuity and growth hypotheses are satisfied, then for various boundary-value problems there exist minimizers for \( I(u, \Omega) \). Conversely, if \( u \) is a minimizer for \( I(u, \Omega) \) among \( C^1(\Omega) \) functions satisfying given Dirichlet boundary conditions, and if \( u_0 \in \Omega, F_0 = F'(u_0) \), then (0.3) holds. This fact may be used (see Theorem 3.2 to motivate quasiconvexity by showing that it is a necessary condition for the existence of sufficiently regular minimizers for a class of displacement boundary-value problems. The degree of regularity required is, however, fairly severe. Furthermore, if \( \Psi \) is quasiconvex and twice continuously differentiable, then \( \Psi \) satisfies the Legendre-Hadamard or ellipticity condition

\[
\frac{\partial^2 \Psi}{\partial u_1 \partial u_2} \geq \lambda \mu \quad \text{for all } \lambda, \mu \in \mathbb{R}.
\]

(0.4)

(\text{It is not known whether the converse holds.}) Because we have chosen to impose quasiconvexity as a constitutive restriction, we must therefore regard the Legendre-Hadamard condition also as a constitutive restriction.\(^1\)

The statement above that quasiconvexity is sufficient for existence must now be qualified. In fact, Morrey's remarkable existence theorem fails to apply directly to nonlinear elasticity. For compressible materials his growth conditions are too stringent; in particular, they prohibit any singular behaviour of \( \Psi \); such as the natural condition

\[
\Psi(x, F) \to \infty \quad \text{as } \det F \to 0.
\]

(0.5)

Moreover, his work gives no indication of how to treat the unilateral constraint \( \det F = 1 \). Incompressible materials require the constraint \( \det F = 1 \), which also poses problems.

\(^1\) Throughout this article we employ the summation convention for repeated indices.

\(^2\) This contrasts with the views expressed by Truesdell & Noll [1, p. 279], who suggested that the Legendre-Hadamard condition should be regarded not as a constitutive restriction, but as a stability condition. They conjectured that violation of the Legendre-Hadamard condition at a point would lead to wave motion tending to move an elastic body from an unstable to a stable equilibrium configuration and that this process may help explain buckling. While the violation of the Legendre-Hadamard condition at a point may well result in certain kinds of instabilities (cf. Ericksen [3]), it is by no means necessary for buckling. Indeed, in Section 9 we show that buckling can occur when the Legendre-Hadamard condition holds everywhere. Moreover, other kinds of instabilities, such as necking, may well be compatible with the Legendre-Hadamard condition (cf. Antman [5]). Another suggestion of Truesdell & Noll [1, p. 129] concerning internal buckling of a rod would, if true, directly contradict quasiconvexity, but the behaviour described by them is implausible seems typical of buckling.

\(^3\) Antman's material hypotheses for rods and shells are those appropriate under the assumption that (0.4) holds in the three-dimensional theory.

\(^4\) The analogous problem for rods has been studied by Antman [2-5, 7, 8].
To overcome these difficulties we investigate in Section 6 sequential weak continuity properties of functions, defined on Orlicz-Sobolev spaces, having the form

$$\theta: u \rightarrow \phi(\mathbf{F}u(x)),$$

(0.6)

where $\phi$ is a continuous real-valued function defined on the set of all $3 \times 3$ matrices. This map is sequentially continuous from $W^{1,\mathcal{P}}(Q)$ with the weak* topology to $L^p(\Omega)$ with the weak topology if and only if $\theta$ has the form

$$\theta(u) = A + B^*_1(\mathbf{F}u)^p + C^*_1(\mathbf{F}u)^q + D \det \mathbf{F}u,$$

(0.7)

where $A, B^*_1, C^*_1, D$ are constants and $\mathbf{F}u$ is the transpose of the matrix of cofactors of $\mathbf{F}u$. When the domain of $\theta$ is a larger Orlicz-Sobolev space, the problem is more delicate. In this case we give various theorems guaranteeing sequential continuity or closure of $\theta$ relative to various weak topologies.

We combine these results with standard techniques of the calculus of variations to establish the existence of minimizers for $I(u, \Omega)$ in various classes of functions when $\mathcal{W}$ has the form

$$\mathcal{W}(x, F) = g(x, F, \mathbf{F}u, \mathbf{F}^T \mathbf{F}u)$$

(0.8)

with $g(x, \cdot, \cdot, \cdot)$ convex for each $x$. We call such functions $\mathcal{W}$ polyconvex. Note that $F, \mathbf{F}u, \mathbf{F}^T \mathbf{F}u$ govern the deformations of line, surface, and volume elements respectively. If $\mathcal{W}$ is polyconvex, then $\mathcal{W}$ is quasiconvex; in fact polyconvexity is equivalent to a sufficient condition for quasiconvexity given by Morrey. However our existence theorems are valid under weaker growth conditions than Morrey's. Moreover, we can handle the pointwise constraints on $\det \mathbf{F}u$ mentioned above by using our sequential weak continuity results. Since there are few known examples of quasiconvex functions that are not polyconvex, the restriction to polyconvex functions is not serious. It appears that neither the quasiconvexity nor the polyconvexity condition has been considered previously in the context of elasticity.

A wide variety of realistic models of nonlinear elastic materials satisfy the hypotheses of our existence theorems. In particular, these include the Mooney-Rivlin material and certain stored-energy functions similar to, and for incompressible materials identical to, those of Ogden [2, 3]. That these stored-energy functions are polyconvex follows from sufficient conditions for the polyconvexity of isotropic functions given in Section 5, where some related results are also discussed.

Our existence theorems apply to displacement, mixed displacement traction, pure traction, mixed displacement pressure, and pure pressure boundary-value problems. Similar methods work for more general classes of mixed boundary conditions. For the most part we consider only polynomial growth hypotheses, which result in a theory based on Sobolev spaces. For stored-energy functions of slower growth an Orlicz-Sobolev space setting is required; for brevity we treat only the displacement boundary-value problem for such functions. An example is given of a stored-energy function requiring this more elaborate theory.
I. Boundary-Value Problems of Nonlinear Hyperelasticity

This section begins with a brief presentation of the basic equations of nonlinear elasticity. We then go on to consider the various boundary-value problems for which we later prove existence theorems. The calculations here are purely formal, in the sense that we assume that the various quantities appearing have sufficient smoothness to justify any operations required (such as integration by parts). The theory of nonlinear elasticity is discussed at length in the books by Green & Zerna [1], Truesdell & Noll [1] and Truesdell & Noll [1], and the reader is referred to these texts when clarification is necessary.

We consider a material body \( \mathcal{B} \) whose particles are labelled by their positions \( x = (x_1, x_2, x_3) \) with respect to a rectangular Cartesian coordinate system in a reference configuration \( \Omega \) which is a bounded open subset of \( \mathbb{R}^3 \). \( \Omega \) need not be homeomorphic to an open ball. In a given motion the position of the particle \( x \) at time \( t \) is denoted \( u(x, t) \).

The deformation gradient \( F \) is defined by

\[
F = \frac{\partial u}{\partial x}, \quad F_i^j = \frac{\partial u^i}{\partial x^j} = u^i_{,j}.
\]

We suppose that \( u: \Omega \to \mathcal{B}^3 \) is orientation-preserving and locally invertible, so that \( J = \det F > 0 \).

Consideration of the stronger requirement that \( u \) be globally one-to-one is beyond the scope of this article.

The symmetric, positive-definite right and left stretch tensors \( U, V \) and the right and left Cauchy-Green tensors \( C, B \) are defined by

\[
C = U^2 = F^T F, \quad B = V^2 = FFT.
\]

The following relations hold:

\[
F = RU = VR, \quad V = RUR^T, \quad (1.4)
\]

where \( R \) is the orthogonal rotation tensor. The eigenvalues \( r_1, r_2, r_3 \) of \( U \) and \( V \) are positive and are termed the principal stretches of the deformation. The principal invariants of \( B \) and \( C \) are given by

\[
I_0 = I = r_1^2 + r_2^2 + r_3^2, \quad I_2 = \frac{1}{2} (r_1^2 + r_2^2 + r_3^2) + \frac{1}{2} (r_1^2 + r_2^2 + r_3^2), \quad I_3 = r_1^2 + r_2^2 + r_3^2.
\]

The symmetric, positive-definite right and left stretch tensors \( U, V \) and the right and left Cauchy-Green tensors \( C, B \) are defined by

\[
C = U^2 = F^T F, \quad B = V^2 = FFT.
\]

The following relations hold:

\[
F = RU = VR, \quad V = RUR^T, \quad (1.4)
\]

where \( R \) is the orthogonal rotation tensor. The eigenvalues \( r_1, r_2, r_3 \) of \( U \) and \( V \) are positive and are termed the principal stretches of the deformation. The principal invariants of \( B \) and \( C \) are given by

\[
I_0 = I = r_1^2 + r_2^2 + r_3^2, \quad I_2 = \frac{1}{2} (r_1^2 + r_2^2 + r_3^2) + \frac{1}{2} (r_1^2 + r_2^2 + r_3^2), \quad I_3 = r_1^2 + r_2^2 + r_3^2.
\]

We suppose that at each particle \( x \) the material of the body is elastic, so that a constitutive equation of the form

\[
T_R(x, t) = T_R(F(x, t), x)
\]

holds, where \( T_R \) denotes the first Piola-Kirchhoff stress tensor. \( T_R \) is related to the Cauchy stress tensor \( T \) by the constitutive equation

\[
T_R = J T (F^{-1})^T.
\]

We choose the notation common in partial differential equations rather than that of continuum mechanics where our \( x, u \) are customarily denoted \( X, x \) respectively.

The surface traction \( t_R \) measured per unit area in the reference configuration, and the actual stress vector \( t \) measured per unit area of the deformed configuration, are given by

\[
t_R = T_R N; \quad t = T n
\]

respectively, where \( N \) and \( n \) denote the unit outward normals to the boundaries \( \partial Q \) and \( \partial u(Q, t) \) respectively.

The pointwise form of the balance laws of linear and angular momentum are given by

\[
\text{Div} T_R + \rho_R \dot{b} = \rho_R \ddot{u},
\]

\[
T = T^T.
\]

where

\[
(\text{Div} T_R) = \frac{\partial T_R(x)}{\partial x^i},
\]

\( \rho_R(x) \) is the density in the reference configuration, and \( \dot{b} \) is the body force per unit mass.

Throughout this article we assume that the material is hyperelastic; i.e., there exists a real-valued stored-energy function \( \mathcal{W}(x, F) \) such that

\[
T_R = \frac{\partial \mathcal{W}}{\partial F},
\]

\( \mathcal{W} \) is objective if and only if

\[
\mathcal{W}(x, QF) = \mathcal{W}(x, F)
\]

for all proper orthogonal matrices \( Q \). If \( \mathcal{W} \) is objective then

\[
\mathcal{W}(x, F) = \mathcal{W}(x, U),
\]

and it follows from (1.11) that (1.10) is satisfied identically. \( \mathcal{W} \) is isotropic if and only if \( \mathcal{W} \) is objective and

\[
\mathcal{W}(x, QF^T) = \mathcal{W}(x, F)
\]

for all orthogonal matrices \( Q \). In this case

\[
\mathcal{W}(x, F) = \Phi(x, r_1, r_2, r_3),
\]

where \( \Phi \) is symmetric in the \( r_i \).

Deformations of incompressible materials are restricted by the pointwise constraint

\[
J = \det F = 1.
\]

For incompressible materials the above theory has to be modified by replacing \( T \) by the extra stress

\[
T_e = T + p 1,
\]

where \( p \) is an indeterminate hydrostatic pressure. The stored-energy function \( \mathcal{W} \) for an incompressible material need be defined only for \( F \) satisfying (1.16).
We shall be concerned only with equilibrium configurations of $\mathcal{B}$. If $W$ is objective we see from (1.9) that $\mathcal{B}$ is an equilibrium configuration if and only if
\begin{equation}
A_{ij}^{\alpha} u_{i,j}^{\alpha} + q_i + \rho_k b_i = 0, \tag{1.18}
\end{equation}
where
\begin{equation}
A_{ij}^{\alpha}(x, F) = \frac{\partial^2 W(x, F)}{\partial x_i^\alpha \partial x_j^\beta}, \quad q_i = \frac{\partial^2 W(x, F)}{\partial x^i}. \tag{1.19}
\end{equation}

A. The mixed displacement-traction boundary-value problem for a compressible material

In this problem we seek $u$ satisfying (1.18) in $\Omega$ and satisfying the boundary conditions
\begin{equation}
u(x) = u(x) \quad \text{for} \quad x \in \partial \Omega_1, \tag{1.20}
\end{equation}
\begin{equation}
u(x) = \bar{u}(x) \quad \text{for} \quad x \in \partial \Omega_2, \tag{1.21}
\end{equation}
where $\partial \Omega_1 = \partial \Omega_2 \cup \partial \Omega_3 \cup \partial \Omega_4$, $\partial \Omega_2 \cap \partial \Omega_3 = \phi$, and $\bar{u}: \partial \Omega_1 \rightarrow \mathbb{R}^3$, $\bar{u}: \partial \Omega_2 \rightarrow \mathbb{R}^3$ are given functions. The boundary condition (1.21) is a condition of dead loading, i.e., the loads acting on $u(\partial \Omega_2)$ have fixed direction and fixed magnitude per unit area of $\partial \Omega_2$. If $\partial \Omega_2 = \phi$ then we have a pure displacement boundary-value problem, while if $\partial \Omega_2 = \phi$ we have a traction boundary-value problem.

Suppose that the body force $b$ is conservative, so that
\begin{equation}
b = -\text{grad} \Psi, \tag{1.22}
\end{equation}
where $\Psi = \Psi(u)$ is a real-valued potential, and where
\begin{equation}[(\text{grad} \Psi)]^i_{\alpha} = \frac{\partial \Psi}{\partial u^i}. \tag{1.23}
\end{equation}
Define $f_i(x, u, F)$ by
\begin{equation}
f_i(x, u, F) = \Psi(x, F) + \rho_k b_i(\Psi(u)). \tag{1.24}
\end{equation}
Consider the functional
\begin{equation}J_\alpha(u) = \frac{1}{2} \int_\Omega f_i(x, u(x), F(x)) dx - \int_{\partial \Omega_3} u(x) \cdot \bar{n} \cdot \bar{u}(x) dS. \tag{1.25}
\end{equation}
Let $\partial \Omega_3 = \phi$. Then a standard formal calculation shows that $J_\alpha(u_0)$ is stationary with respect to $u$ satisfying (1.20) if and only if the Euler-Lagrange equations (1.18) and the natural boundary conditions (1.18) hold, i.e., if and only if $u_0$ is a solution to the mixed boundary-value problem.

If $\partial \Omega_3 = \phi$ then in general solutions to the boundary value problem will not exist, since a necessary condition for a function $u_0$ to render $J_\alpha$ stationary subject to (1.21) is that
\begin{equation}a = 0 \tag{1.26}
\end{equation}
where
\begin{equation}a = \frac{1}{m(\Omega)} \left( \int_{\partial \Omega} b(u_0) dS + \int_{\partial \Omega} \bar{n} \cdot \bar{u} dS \right). \tag{1.27}
\end{equation}

Condition (1.25) says that the total force on the body due to external loads is zero (cf. TRUESDELL & NOLL [1, p. 127]). To describe the effect of this condition we consider two situations corresponding to different types of existence theorems proved in Section 7.

1. $b(u)$ is not a constant vector. In this case, under suitable hypotheses on $b$ the set of functions $u_0$ satisfying (1.25) will be nonempty, so that under certain conditions it is likely that a function $u_0$ such as to render $J_0$ stationary subject to (1.21) exists. If so, then $u_0$ is a solution to the traction boundary-value problem.

2. $b(u) = b_0$ constant. In this case $a$ is independent of $u_0$, so that (1.25) is a condition on the data of the problem. It proves convenient to consider $J_0(u)$ as a functional defined on functions $u$ satisfying the constraint
\begin{equation}\left[ u - u_0 \right]_\Omega = 0 \tag{1.28}
\end{equation}
where $\epsilon$ is an arbitrary constant vector. The constraint (1.27) removes the indeterminacy resulting from a possible rigid-body translation of $u(\Omega)$. $J_0$ is stationary if $u = u_0$ subject to (1.27) if and only if (1.21) holds and
\begin{equation}\int \text{Div} T_\rho + \rho_k b_0 = a. \tag{1.29}
\end{equation}

To prove this note first that if (1.21) and (1.28) hold, then
\begin{equation}\delta J_0(u_0)(\epsilon) = \frac{d}{d\epsilon} \left. J_0(u_0 + \epsilon v) \right|_{\epsilon = 0} = \frac{d}{d\epsilon} \left. \int_{\partial \Omega} (u_0 + \epsilon v) dS \right|_{\epsilon = 0} = 0, \tag{1.30}
\end{equation}
which is zero if $\int \epsilon v dS = 0$. The converse statement is a direct application of the multiplier rule for isoperimetric problems with $a$ playing the role of the Lagrange multiplier corresponding to the constraint (1.27). A rigorous proof may also be constructed by using a result of SCHWARTZ [1, p. 59].

If $a = 0$ then $u_0$ is an equilibrium solution. If $a \neq 0$ then
\begin{equation}u(x) = u_0(x) + \frac{a}{2} x \tag{1.31}
\end{equation}
is a solution to the dynamic traction boundary-value problem (1.9), (1.21). Note that any equilibrium solution $u$ must also satisfy the zero moment condition
\begin{equation}\int_\Omega a \cdot \rho_k b_0 dS + \int_{\partial \Omega} u \cdot \bar{n} dS = 0. \tag{1.32}
\end{equation}
This condition, unlike (1.25), depends explicitly on the unknown function $u$, and so cannot be imposed a priori.

B. The mixed displacement pressure boundary-value problem for a compressible material

In this problem we seek $u$ satisfying (1.18) in $\Omega$ and satisfying the boundary conditions
\begin{equation}u(x) = \bar{u}(x) \quad \text{for} \quad x \in \partial \Omega_1, \tag{1.33}
\end{equation}
\begin{equation}(\bar{\tau}_\rho)(x) = -\rho_k n \quad \text{for} \quad x \in \partial \Omega_2, \quad (r = 2, ..., M). \tag{1.34}
\end{equation}

\* if $\mathcal{B}$ takes the form of a powerful potential well, for example.
where \( \partial \Omega = \bigcup_{r=1}^{M} \partial \Omega_r \), \( \partial \Omega_r \cap \partial \Omega_s = \emptyset \), \( r \neq s \), \( \bar{u} : \partial \Omega_r \to \mathbb{R}^3 \) is a given function, and 
\( \rho_r (r = 2, ..., M) \) are constant pressures. We assume that for \( r \geq 2 \), \( \overline{\partial \Omega_r} \) is either a closed surface or is bounded by a closed curve lying in \( \overline{\partial \Omega_r} \). Suppose also that there exists a \( C^1(\overline{\Omega}) \) function \( \rho : \Omega \to \mathbb{R} \) taking the value \( \rho_r \) on \( \partial \Omega_r \) for each \( r = 2, ..., M \).

Consider the functional
\[
J_1(u) = \int f_2(x, u(x), F(x)) \, dx,
\]
where
\[
f_2(x, u, F) = f_1(x, u, F) + \rho J + \frac{1}{2} \rho \partial_{ij} \varepsilon^{ij} \varepsilon^{kl} u_i u_j u_k u_l N_i N_j dS.
\]
and where \( J \) is defined in (1.16). By the divergence theorem
\[
J_1(u) = \int f_2 \, dx + \int \frac{1}{2} \rho \partial_{ij} \varepsilon^{ij} \varepsilon^{kl} u_i u_j u_k u_l N_i N_j dS.
\]
Suppose now that \( \partial \Omega_r \neq \emptyset \). Then for \( n \) satisfying \( n = 0 \) on \( \partial \Omega_r \), we obtain
\[
\frac{d}{dn} J_1 (u + \varepsilon n) \bigg|_{n=0} = - \int \frac{1}{2} (\partial_\nu \varepsilon^{ij} \varepsilon^{kl} u_i u_j u_k u_l N_i N_j dS.
\]
The fourth integral in (1.37) is zero by Kelvin's theorem applied to each \( \partial \Omega_r \), and the last integral is zero since \( \rho \mathbf{n} \cdot \mathbf{n} = 0 \) on \( \partial \Omega_r \). Thus
\[
\frac{d}{dn} J_1 (u + \varepsilon n) \bigg|_{n=0} = \int (\partial_\nu \varepsilon^{ij} \varepsilon^{kl} u_i u_j u_k u_l N_i N_j dS.
\]
Thus \( J_1(u_0) \) is stationary if and only if \( u_0 \) is a solution to the mixed displacement pressure boundary-value problem. The calculation above is a slightly simplified version of that of Sewell [1]; see also Beatty [1].

If \( \partial \Omega_r = \emptyset \) then we proceed in a fashion similar to A.

The functional that we study in later sections include functionals of the form \( J_0 \) and \( J_1 \). For the purposes of the existence theorems it is not necessary to assume that \( \pi \) is objective. If, however, this assumption is not made, the resulting minimizers, if smooth, will not necessarily satisfy (1.10).

For incompressible materials the admissible functions are restricted by the additional constraint (1.16).

Finally in this section we discuss briefly the analogous problems of one and two-dimensional plane strain we consider deformations having the form
\[
u = (u_1(x_1, x_2), u_2(x_1, x_2), \lambda x_3), \quad \lambda > 0 \text{ constant.}
\]
The closure of $\mathcal{P}(\Omega)$ in $W^{1,p}(\Omega)$ is denoted $W^{1,p}_0(\Omega)$. $W^{1,p}_{loc}(\Omega)$ denotes the space of functions $u$ which together with their weak derivatives $\frac{\partial u}{\partial x^i}$ $(1 \leq i \leq m)$ are locally integrable. When dealing with the spaces in this paragraph we assume that $p < \infty$ unless otherwise stated.

Weak and strong $\ast$ convergence of sequences are denoted by $\rightharpoonup_{w}$ and $\rightharpoonup_{s}$, respectively. In the case of the Banach space $W^{1,p}(\Omega)$ we define the weak $\ast$ topology to be that induced by the natural embedding of $W^{1,p}(\Omega)$ in the product space $(L^p(\Omega))^{\ast \ast}$, where each factor has the weak $\ast$ topology. Thus a sequence $u_n \rightharpoonup_{s} u$ in $W^{1,p}(\Omega)$ if and only if $u_n \rightharpoonup_{w} u$ in $L^p(\Omega)$ and $\frac{\partial u}{\partial x^i} \rightharpoonup_{w} \frac{\partial u}{\partial x^i}$ in $L^p(\Omega)$ $(1 \leq i \leq m)$.

If $A$ is a real-valued, continuous, even, convex function of $t \in \mathbb{R}$ satisfying

$$A(t) \geq 0 \quad \text{for } t > 0, \quad A(t) \to 0 \text{ as } t \to 0 \quad \text{as } t \to \infty \text{ as } t \to \infty \text{ then we call } A \text{ an } N- \text{function.}$$

If $A$ is an $N$-function, its conjugate function $A' (t)$ is defined by $A(t) = \sup \{ts - A(s) \colon s \in \mathbb{R}\}$. $A'$ is also an $N$-function and satisfies $A(t) = A(t)$.

Furthermore Young's inequality,

$$t s \leq A(t) + A'(s) \quad \text{for all } s, t \in \mathbb{R}.$$ 

holds for all $s, t \in \mathbb{R}$. If $A, B$ are $N$-functions then we write $A \leq B$ if and only if there exist positive numbers $t_0$ and $k$ such that

$$A(t) \leq B(k t)$$

for all $t \geq t_0$. We write $A \sim B$ if and only if $A < B$ and $B < A$, and $A \sim B$ if and only if

$$\lim_{t \to \infty} \frac{A(t)}{B(t)} = 0 \quad \text{for every } t > 0.$$ 

for every $t > 0$. If $A$ is an $N$-function the Orlicz class $L^A(\Omega)$ consists of all (equivalence classes of) real-valued measurable functions $u$ on $\Omega$ such that

$$\int \frac{A(u(x))}{A'}(x) dx < \infty.$$ 

The Orlicz space $L^A(\Omega)$ is the linear hull of $L^A(\Omega)$. $L^A(\Omega)$ is a Banach space with respect to the Luxemburg norm

$$\|u\|_{L^A(\Omega)} = \inf \{k > 0 \colon \int \frac{A(u(x))}{A'}(x) dx \leq k \}.$$ 

If $A \sim B$ then $L^A(\Omega) \subseteq L^B(\Omega)$, while if $A \leq B$ then $L^B(\Omega) \subseteq L^A(\Omega)$. The space $E^A(\Omega)$ is defined as the closure of the bounded functions in the $L^A(\Omega)$ norm. We have that $E^A(\Omega) \subseteq L^1(\Omega)$ and $L^1(\Omega) \subseteq E^A(\Omega)$. The dual of $E^A(\Omega)$ can be identified by means of the scalar product $\int u(x) dx$ with $L^1(\Omega)$. The norm on $L^1(\Omega)$ dual to $\| \cdot \|_{(A)}$ on $E^A(\Omega)$ is defined as

$$\|u\|_{L^1(\Omega)} = A(\|u\|_{L^\infty(\Omega)}).$$ 

$E^A(\Omega)$ is a Banach space under the norm $\| \cdot \|_{E^A(\Omega)} = \| \cdot \|_{(A)}$. The Orlicz-Sobolev space $W^{1,A}(\Omega)$ is defined as the set of functions $u \in E^A(\Omega)$ such that the weak derivatives $\frac{\partial u}{\partial x^i} \in L^A(\Omega)$ $(1 \leq i \leq m)$.

The dual of $E^A(\Omega)$ is $E^A(\Omega)$.

Finally we define some regularity conditions for $\Omega$.

(i) $\Omega$ has the segment property if there exists a locally finite open covering $\{\Omega_j\}$ of $\partial \Omega$ and corresponding vectors $\{y_j\}$ such that $x + t y_j \in \Omega$ for all $x \in \Omega \cap \partial \Omega$, and for all $t \in (0,1)$.

(ii) $\Omega$ satisfies the cone condition if there exists a fixed cone $k \Omega \subseteq \mathbb{R}^m$ such that each point $x \in \partial \Omega$ is the vertex of a cone $k \Omega$ that lies in $\Omega$ and is congruent to $k \Omega$.

(iii) $\Omega$ satisfies a strong Lipschitz condition if each $x \in \partial \Omega$ has a neighbourhood $\Omega_x \subseteq \mathbb{R}^n$ such that in some co-ordinate system, with origin at $x$, $\Omega_x \cap \Omega_y$ is represented in $\Omega_x$ by $\Omega_y = (\xi_1, \ldots, \xi_{n-1})$ with $F$ a Lipschitz continuous function.

3. Quasiconvexity and the Legendre-Hadamard Condition

We consider integrals of the form

$$I(u, \Omega) = \int f(x, u(x), \nabla u(x)) dx,$$

where $(x, u(x)) \in \Omega \times \mathbb{R}^m$, $\|\nabla u(x)\| = \|\frac{\partial u}{\partial x}(x)\| = \|u(x)\|(1 \leq i \leq n, 1 \leq \alpha \leq m)$, and where the real-valued function $f$ is defined and continuous on a given relatively open subset $S$ of $\Omega \times \mathbb{R}^m \times M^{m \times n}$. Here $M^{m \times n}$ denotes the space of real $n \times m$ matrices, with the induced norm of $M^{m \times n}$. We assume that for each $x \in \Omega$ there exist $u \in \mathbb{R}^m$, $F \in M^{m \times n}$ with $(x, u, F) \in S$. 

noted $\| \cdot \|_A$ and is equivalent to $\| \cdot \|_{(A)}$. Hölder's inequality is valid in the form

$$\int \frac{\partial u(x)}{\partial x^i} dx \leq F(x, u(x), \nabla u(x)),$$

for all $u \in E^A(\Omega)$, $w \in L^A(\Omega)$.

The Orlicz-Sobolev space $W^{1,A}(\Omega)$ is defined as the set of functions

$$u \in E^A(\Omega) \quad \text{such that the weak derivatives } \frac{\partial u}{\partial x^i} \in L^A(\Omega), \quad (1 \leq i \leq m).$$ 

$W^{1,A}(\Omega)$ is a Banach space under the norm

$$\|u\|_{W^{1,A}(\Omega)} = \|u\|_{L^A(\Omega)} + \int \frac{\partial u}{\partial x^i} dx,$$

and similarly for $W^{2,A}(\Omega)$. The closure of $\mathcal{P}(\Omega)$ in $W^{1,A}(\Omega)$ is written $W^{1,A}(\Omega)$. In the special case when $A(t) \sim |t|^p (p > 1)$ we have the equalities

$$L^p(\Omega) = E^p(\Omega) = L^p(\Omega), \quad W^1, K(\Omega) = W^1, P(\Omega) = W^1, A(\Omega),$$

and $A(t) \sim |t|^p$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

Throughout this article we shall be dealing with vector and matrix functions $\Psi$. When we write $\Psi \in X$, where $X$ is any one of the spaces introduced above, we mean that each component of $\Psi$ belongs to $X$, and we define $\|\Psi\|_X$ to be $\sum \|\Psi_i\|_X$ where the sum is over all components of $\Psi$ of $\Psi$.

Finally we define some regularity conditions for $\Omega$.
Definition 3.1 (cf. Morrey [1]). Let $U$ be an open subset of $M^{n \times m}$. Let $g : U \to \mathbb{R}$ be continuous. Then $g$ is said to be quasiconvex if at $F_0 \in U$ if and only if

$$
\int g(F_0 + P'(y))\,dy \geq g(F_0)\,m(D)
$$

(3.1)

for every bounded open subset $D \subseteq \mathbb{R}^m$ and for every $\xi \in \mathscr{D}(D)$ which satisfies $F_0 + P'(y) \in U$ for all $y \in D$. $g$ is quasiconvex on $U$ if it is quasiconvex at each $F_0 \in U$.

Note that if $g$ is quasiconvex at $F_0 \in U$, and if $\xi \in W^{1,\infty}(D)$ satisfies $F_0 + P'(y) \in K$ for almost all $y \in D$ and for some compact subset $K$ of $U$, then (3.1) holds for $\xi$. In fact by the definition of $W^{1,\infty}(D)$ there exists a sequence of functions $\xi_n \in \mathscr{D}(D)$ that converges to $\xi$ in $W^{1,1}(D)$. Since $K$ is compact there exists an integer $N$ and a compact set $K_n$ with $K \subset K_n \subset U$ such that $F_0 + P'(y) \in K_n$ for almost all $y \in D$ and all $n \geq N$. If $n \geq N$ then (3.1) holds for $\xi_n$. By the compactness of $K_n$, the continuity of $g$, and the bounded convergence theorem we obtain (3.1) for $\xi$.

Let $A = \{w \in W^{1,1}(\Omega) : (x, w(x), Pw(x)) \in S$ for almost all $x \in \Omega$, and $I(w, \Omega)$ exists and is finite $\}$. Then $g$ is quasiconvex if and only if

$$
I(w, \Omega) \leq I(w, \Omega)
$$

for all $w \in A$ with $w - u \in L^p(\Omega)$ and $\|w - u\|_{\text{CM}}$ sufficiently small. Let $w_0 \in A$ and suppose that $u$ and $Pw$ have representatives, again denoted by $u$ and $Pw$, that are continuous at $x_0$ with $(x_0, u(x_0), Pw(x_0)) \in S$. Let $U = \{F : x_0 \in U, F \in S\}$. Then $f(x, u(x_0), F) = \int f(x, u(x_0), \xi) \,d\xi$ is quasiconvex at $F \in U$.

Proof. Let $u$ satisfy the hypotheses of the theorem, let $D$ be a bounded open subset of $\mathbb{R}^m$ and let $\xi \in L^p(\Omega)$, then (3.1) holds for $\xi$. For $c > 0$ define $u_c : \Omega \to \mathbb{R}^m$ by

$$
u_c(x) = u(x) + c\left(\frac{x - x_0}{c}\right) \text{ if } x \in D$$

and $u_c(x)$ otherwise. For $c$ small enough the set $x_0 + cD$, on which $u$ and $u_c$ differ, is contained in $\Omega$, and thus $u - u_c \in L^p(\Omega)$. Also for $x \in x_0 + cD$ we have

$$
f(x, u(x), Pw(x)) = f\left(x, u(x) + c\left(\frac{x - x_0}{c}\right), Pw(x) + Pw\left(\frac{x - x_0}{c}\right)\right),
$$

which by our continuity assumptions and our assumptions is bounded uniformly above on the set $x_0 + cD$ for $c$ small enough. Thus $u_c \in A$ and $I(u, \Omega) \leq I(u_c, \Omega)$. Making the change of variables $y = x - x_0$ we obtain

$$
\int g(F_0 + P'(y))\,dy \geq g(F_0)\,m(D)
$$

where $F_0 = u(x_0) + Pw(x_0)$.

Note that if $g$ is quasiconvex at $F_0 \in U$, and if $\xi \in W^{1,\infty}(D)$ satisfies $F_0 + P'(y) \in K$ for almost all $y \in D$ and for every $\xi \in \mathscr{D}(D)$ which satisfies $F_0 + P'(y) \in U$ for all $y \in D$, Then (3.1) holds for all such $D, \xi$, i.e., $g$ is quasiconvex at $F_0$.

Proof. Apply the theorem to the integrand $g(F)$ with $\Omega = D$ and $u(x) = (F_0 \mid x^2)$. Then (3.1) holds for all such $D, \xi$, i.e., $g$ is quasiconvex at $F_0$.

Theorem 3.1 is essentially the same as a result stated by Silverman [1], following earlier work of Busemann & Shephard [1, p. 31].

We next show that for integrands that are independent of $\phi$ and $\mu$, the existence of a sufficiently regular minimizer to certain Dirichlet problems implies that the quasiconvexity condition holds.

Theorem 3.2. Let $U \subseteq M^{n \times m}$ be open and let $g : U \to \mathbb{R}$ be continuous. Suppose that either (i) $n = 1, \Omega$ is arbitrary, or (ii) $\Omega$ is a hypercube, $\Omega = \{x \in \mathbb{R}^n : 0 < x^* < 1, 1 \leq x \leq m\},$ say. Let $u_0 : \Omega \to \mathbb{R}^m$ be defined by

$$
u_0(x) = F_0 \mid x^2 + z^2,$$

where $F \in U$ and $z \in \mathbb{R}^m$ are constants. Let

$$
I(u) = \int g(Pw(x))\,dx
$$

and let $A_1 = \{u \in C(\bar{\Omega}) : Pw(x) \in U$ for all $x \in \Omega, J(u)$ exists and is finite, and $u = u_0$ on $\partial \Omega\}$. Suppose there exists $u \in A_1$ such that

$$
J(u) \leq J(\nu_0)
$$

for all $u \in A_1$.

Then $g$ is quasiconvex at $F \in U$.

Proof. Let $w = \nu_0 - u_0$. Then we are in $C(\bar{\Omega})$ and $w = 0$ on $\partial \Omega$. In case (i) there exists $x_0 \in \Omega$ with $Pw(x_0) = 0$. Therefore $Pw(x_0) = F$ and the result follows from Theorem 3.1. In case (ii) we have that $Pw(x_0) \to 0$ as $r \to \infty$ for any sequence $\{x_n\} \subseteq \Omega$ with $x_n \to 0$ as $r \to \infty$. By Theorem 3.1, $g$ is quasiconvex at $F + Pw(x_0)$ in $U$. Taking the limit $r \to \infty$ in the quasiconvexity condition gives the result.

Remarks. For $n = 1$ or $m = 1$ quasiconvexity is equivalent to convexity (see Morrey [1, 2]). The analogue of Theorem 3.2 for $n = 1, \Omega$ arbitrary, is false. As an example, let $m = n = 2, \Omega = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$. Define $g : M^{2 \times 2} \to \mathbb{R}$ by $g(F) = \rho(F)$, where $\rho = \text{tr}(FF^T)$ and where $\rho : \mathbb{R} \to \mathbb{R}$ is zero for $r \geq 1$ and positive for $0 \leq r < 1$. We show that for any $\omega_0 \in C(\bar{\Omega})$ there exists an absolute minimizer for $J(u) = \int g(Pw(x))\,dx$ among $C(\bar{\Omega})$ functions $u$ satisfying $u = u_0$ on $\partial \Omega$. Let $\xi \in C(\bar{\Omega})$ satisfy $\xi \equiv 0$ on $\partial \Omega$ and let $J(\xi) = \infty$ for all $x \in \Omega$. e.g., we may take $\xi = \text{h}_1(x_1, h_2(x_2), h_3(x_3))$, where $h_1$ and $h_2$ are $C(\bar{\Omega})$ functions satisfying $h_1 = h_2 = 0$ on $\partial \Omega$. Suppose $\xi$ is a minimizer for $J(u)$ with $u = h_1(x_1, h_2(x_2), h_3(x_3))$, where $h_1$, $h_2$, and $h_3$ are $C(\bar{\Omega})$ functions satisfying $h_1 = h_2 = 0$ on $\partial \Omega$. Then for all $x \in \Omega$ we have

$$
|P(u_0 + \xi)(x)| \geq c\cdot - |P(u_0)|_{\text{CM}} > 1,
$$
so that \( J(u_0 + k \xi) = 0 \) and \( u_0 + k \xi \) is an absolute minimizer. But \( g \) is not quasi-convex, as may be seen by putting \( u_0 = F \) in the above argument, where \( F \in M^{2 \times 2} \) is constant with \( |F| < 1 \). With a little more work one can show that the absolute minimizer \( u_0 \) corresponding to \( u_0 \) may be chosen so that \( \det F u(x) \geq c > 0 \) for all \( x \in \Omega \).

**Definition 3.2.** Let \( U \) be an open subset of \( M^{m \times n} \). A function \( g: U \rightarrow \mathbb{R} \) is rank \( 1 \) convex on \( U \) if it is convex on all closed line segments in \( U \) with end points differing by a matrix of rank \( 1 \), i.e.,

\[
g(F + (1 - \lambda) a \otimes b) \leq (1 - \lambda) g(F) + \lambda g(F + a \otimes b)
\]

for all \( F \in U, \lambda \in [0, 1], a \in B^m, b \in B^n, \) with \( F + \mu a \otimes b \in U \) for all \( \mu \in [0, 1] \). Here \( a \otimes b \) denotes the outer product of \( a \) and \( b \).

**Theorem 3.3.** Let \( U \) be an open subset of \( M^{m \times n} \) and let \( g: U \rightarrow \mathbb{R} \). The following conditions (i)-(iv) are equivalent:

(i) \( g \) is rank \( 1 \) convex on \( U \);

(ii) for each fixed \( F \in M^{m \times n} \), \( b \in B^n \) the function \( a \mapsto g(F + a \otimes b) \) is convex on all closed line segments in the set \( \{ a: F + a \otimes b \in U \} \);

(iii) for each fixed \( F \in M^{m \times n} \), \( a \in B^m \) the function \( b \mapsto g(F + a \otimes b) \) is convex on all closed line segments in the set \( \{ b: F + a \otimes b \in U \} \);

(iv) the inequality

\[
g(H) \leq (1 - \lambda) g(H + c \otimes d) + \lambda g(H - \frac{\lambda}{1 - \lambda} c \otimes d)
\]

holds for all \( \lambda \in [0, 1] \) and for all \( H \in M^{m \times n}, c \in B^m, d \in B^n \) satisfying \( H + \mu c \otimes d \in U \) for all \( \mu \in \left[ \frac{1}{1 - \lambda}, 1 \right] \).

If \( g \in C(U) \), then (i)-(iv) are equivalent to

(v) for each \( F \in U \) there exists \( A(F) \in M^{m \times n} \) such that

\[
g(F + a \otimes b) \geq g(F) + A(F) a \otimes b
\]

whenever \( F + \lambda a \otimes b \in U \) for all \( \lambda \in [0, 1] \).

If \( g \in C^1(U) \), then \( A(F) = \frac{\partial g(F)}{\partial F} \).

If \( g \in C^2(U) \), then (i)-(v) are equivalent to

(vi) (Legendre-Hadamard condition)

\[
\frac{\partial^2 g}{\partial F^2} a \otimes b \geq 0 \quad \text{for all} \quad a \in B^m, b \in B^n, F \in U.
\]

Proof. The equivalence of (i), (ii) and (iii) is clear (cf. SILVERMAN [1, Thm. 4]). The equivalence of (i) and (iv) is proved by making the change of variables \( c = (1 - \lambda) a, d = b, H = F + (1 - \lambda) a \otimes b \). Let \( g \in C(U) \). That (v) implies (i) follows by a well known condition for convexity. To show that (ii) implies (v) one can use the arguments of MORREY [1, p. 47] to establish (3.3) for \( a \otimes b \) belonging to some neighbourhood of zero in \( M^{m \times n} \), and then deduce (v) from the convexity of the function \( \alpha \mapsto g(F + a \otimes b) \). The remaining assertions of the theorem are obvious.
important open question (cf. the comments of Morrey [2, p. 122]). The conditions are known to be equivalent only in certain special cases, for example in the quadratic case $f(F) = a_F F^T F$ with $a_F$ constant and $n$, $n$ arbitrary (Morrey [1, 2], Van Hove [1]), and for certain parametric integrands when $n = m + 1$ (Morrey [1, 2]). In particular nothing interesting is known about the case $m = n > 1$, which occurs in nonlinear elasticity.

To discuss this problem, consider a continuous integrand $f(F, u)$, defined on all of $M^{m \times m}$ and independent of $x$ and $u$. Suppose that $f$ is rank 1 convex on $M^{m \times m}$. It is well known that if $D$ is a bounded open set in $\mathbb{R}^n$ then any function $\xi \in W^{1,n}(D)$ can be approximated in $W^{1,\infty}(D)$ by piecewise affine functions (Ekeland & Temam [1, p. 286]). Thus a natural method of attack is to follow the lead of the proof of Theorem 3.4 and to seek domains $D$ with a partition into a finite number of disjoint open sets $D_k$ and a set of measure zero, such that the quasi-convexity condition

$$\int_D f(F_0 + F\xi(x)) \, dx \geq f(F_0) \, m(D)$$

(3.4)

holds for any $F_0 \in M^{m \times m}$ and for any $\xi \in W^{1,n}(D)$ that is affine on each $D_k$ (cf. Silverman [1, Thm. 2]).

For ease of illustration we consider the case $m = 2$, $n$ arbitrary; similar comments apply for $m = 3$. First let $D$ be the interior of a triangle in $\mathbb{R}^2$ with vertices $a_1, a_2, a_3$, and let $e$ be an interior point of $D$. Let $D_1, D_2, D_3$ be the interiors of the triangles $a_2, e, a_1$, $a_1, e, a_2$, respectively. Let $n_1$ be the unit outward normal to $\partial D$ on the side $a_2 a_1$, let $l_1 = |a_2 - a_1|$, and let $n_2, n_3, l_2, l_3$ be defined analogously. Let $\xi \in W^{1,n}(D)$ be affine on each $D_k$ with $\xi(e) = e$. Then

$$\lambda_1 l_1 = \frac{l_1}{2m(D_k)} \, e \otimes n_1 \quad \text{on} \quad D_k,$$

and (3.4) becomes

$$\sum_{k=1}^{3} \lambda_k f \left( \frac{l_k}{2m(D_k)} \, e \otimes n_k \right) \geq f(F_0),$$

where $\lambda_k = m(D_k)/m(D)$. But this inequality follows from rank 1 convexity of $f$ because

$$\sum_{k=1}^{3} \lambda_k l_k n_k = 0.$$

A similar argument shows that (3.4) holds for piecewise affine functions if $D$ is the interior of a convex polygon and the $D_k$'s are triangles formed by joining a single interior point of $D$ to adjacent vertices of the polygon.

A different situation arises if we introduce more interior nodes into the partition of $D$. For example let $D$ be an equilateral triangle $A_1 A_2 A_3$ of side 1 partitioned into 16 congruent equilateral subtriangles of side $\frac{1}{2}$ (see Fig. 1). Let $n_1, n_2, n_3$ be the unit outward normals shown, and let $e_1, e_2, e_3$ be the position vectors of the three interior nodes $B_1, B_2$ and $B_3$. Let $c_1, c_2, c_3$ be given and let $\xi \in W^{1,n}(D)$ be affine on each subtriangle with $\xi(e_j) = -\frac{8}{\sqrt{3}} \, e_j$. The values of $V\xi$ in each subtriangle are shown in Fig. 1. The corresponding quasi-convexity inequality is easy to write down, but does not seem to follow from rank 1 convexity of $f$. This suggests that the Legendre-Hadamard condition does not imply quasi-convexity. Unfortunately the search for a counter-example is hampered by the fact that, as we have mentioned, any such $f$ cannot be quadratic in $F u$. It is conceivable that if $m = n$ then the Legendre-Hadamard condition implies quasi-convexity for objective functions.

Finally we mention in passing an implication of the configuration in Figure 1 for finite element methods. Let $D$ be as in Figure 1 and suppose that one wishes to solve a boundary-value problem for $u$ subject to the pointwise constraint $\det V u(x) = 1$ for all $x \in \partial D$ (the two-dimensional analogue of the incompressibility constraint (1.16)) and, for example, the boundary conditions $u(x) = x$ for all $x \in \partial D$. Then the only function $\xi \in W^{1,n}(D)$ that is affine on each subtriangle with $\det V(x + \xi(x)) = 1$ for all $x \in \partial D$ is $\xi = 0$. Indeed, any such map $x \to x + \xi(x)$ deforms each subtriangle into another triangle with equal area. It follows that, for example, $B_1$ can be displaced only along the line through $B_1$ parallel to $A_1 A_2$, and also along the line through $B_1$ parallel to $A_1 A_2$. Hence $B_1$ is fixed by the map, and
similarly for $B_2, B_3$. A similar argument applies when the number of congruent subtriangles in the partition of $D$ is increased. Thus either nonlinear interpolation functions or nonconforming elements must be used. A related difficulty for incompressible fluids is discussed by Témam [1].

4. Sufficient Conditions for Quasiconvexity

The quasiconvexity condition is not a pointwise condition on the function $f$, and is therefore difficult to verify in particular cases. In this section we shall be concerned with more accessible conditions that are sufficient for quasiconvexity. These conditions apply to functions $f$ for which the Legendre-Hadamard condition is not known to be equivalent to quasiconvexity.

Throughout the rest of the article we assume, unless the contrary is stated, that $m = n = 1, 2$ or 3.

We first study those functions $\phi(F)$ that belong to the null-space of the Euler-Lagrange operator; i.e., those functions for which the corresponding Euler-Lagrange equations are identically satisfied. For smooth $\phi$ the following result is a special case of ERICKSEN [1], EDELEN [1, 2] and RUND [1, 2].

Theorem 4.1. Let $\phi: M^{**(k)} \to \mathcal{R}$ be continuous and such that both $\phi$ and $-\phi$ are rank 1 convex on $M^{**(k)}$, so that

$$\phi(F + (1 - \lambda) a \otimes b) = \lambda \phi(F) + (1 - \lambda) \phi(F + a \otimes b)$$

for all $F \in M^{**(k)}, a, b \in \mathbb{R}^n, \lambda \in [0, 1]$. Then $\phi$ has the form

$$\phi(F) = a + bF$$ for $\lambda = 1$,

$$\phi(F) = \alpha + \beta_1 F_1' + \beta_2 F_2' + \gamma \det F$$ for $\lambda = 2$, and

$$\phi(F) = A + \beta_1 F_1' + \beta_2 F_2' + D \det F$$ for $\lambda = 3,$

where $a, b, \alpha, \beta_1, \beta_2, \gamma, A, B_1, B_2, C_1, D$ are arbitrary constants, and where $\det F$ denotes the adjugate matrix of $F$ (i.e., the transpose of the matrix of cofactors).

Proof. We just treat the case $\lambda = 3$; the cases $\lambda = 1, 2$ are easier. Suppose first that $\phi$ is $C^2$. Then by Theorem 3.3 (vi), (4.1) is equivalent to

$$A^{ij}_{ij}(F) = \frac{\partial^2 \phi(F)}{\partial F^i \partial F^j} = 0$$

for all $a, b \in \mathbb{R}^3$ and for all $F \in M^{**(k)}$, where

$$A^{ij}_{ij}(F) = A^{ij}_{ij}(\phi(F)) = \frac{\partial^2 \phi(F)}{\partial F^i \partial F^j}$$

It follows that $A^{ij}_{ij}(F)$ is alternating; i.e., $A^{ij}_{ij} = -A^{ij}_{ji}$ and $A^{ij}_{ij} = 0$ if $i = j$ or $i \neq j$. Since $A^{ij}_{ij} = A^{ji}_{ij} = A^{ij}_{ji}$, it follows that $\phi$ is affine in each of $F_1', F_2'$ and $F_3'$, so that

$$\phi(F) = \phi(F_1') F_1' + \phi(F_2') F_2' + \phi(F_3') F_3' + \phi(F)$$

where $\phi$ denotes the matrix of off-diagonal elements of $F$, and where $\phi_i, \theta_i, \chi$ are $C^2$. Since $A^{ij}_{ij} = 0$, etc., we obtain the equations

$$\frac{\partial \phi_0}{\partial F_1' = \frac{\partial \phi_2}{\partial F_1' = \frac{\partial \phi_1}{\partial F_1' = \frac{\partial \theta_1}{\partial F_1'} = 0.$$
Theorem 4.2 (BUSEMANN; EWALD & SHEPHARD[1]).

(i) \( \mathcal{F} \) is convex on \( M \) if and only if it has a convex lower bound and the inequality

\[
\mathcal{F}(z_i) \leq \sum_{i=1}^{k} \lambda_i \mathcal{F}(z_i)
\]

holds for all \( z_1, \ldots, z_k \) and \( z_0 = \sum_{i=1}^{k} \lambda_i z_i \), lying in \( M \). A suitable convex extension to \( \mathcal{F} \) is given by

\[
g_{\lambda}(z) = \inf_{z_i \in M} \sum_{i=1}^{k} \lambda_i \mathcal{F}(z_i), \quad z_0 \in \mathcal{F}.
\]

(ii) Let \( \mathcal{F} \) be open. Then either of the following conditions is necessary and sufficient for \( \mathcal{F} \) to be convex on \( M \):

(a) \( \mathcal{F} \) has a convex lower bound and the inequality

\[
\mathcal{F}(z_i) \leq \sum_{i=1}^{k} \lambda_i \mathcal{F}(z_i)
\]

holds for all \( z_1, \ldots, z_k \) and \( z_0 = \sum_{i=1}^{k} \lambda_i z_i \), lying in \( M \).

(b) for each point \( z_0 \in M \) there exist numbers \( a_i(z_0) \) (1 = 1, ..., \( s \)) such that

\[
\mathcal{F}(z) \geq \mathcal{F}(z_0) + \sum_{i=1}^{s} a_i(z_0)(z_i - z_0),
\]

for all \( z \in M \).

We now define finite-dimensional Euclidean spaces \( E \) and \( E_1 \) by

\[
E = E_1 \times \mathcal{R},
\]

where

\[
E_1 = \begin{cases} \text{empty} & \text{if } n = 1, \\
M^{2 \times 2} & \text{if } n = 2, \\
M^{2 \times 3} \times M^{3 \times 3} & \text{if } n = 3.
\end{cases}
\]

Thus \( E \) may be identified with \( \mathcal{R}^{n+s} \), where \( s(1) = 1, s(2) = 5 \) and \( s(3) = 19 \).

Define the map \( T: M^{n+s} \to E \) by

\[
T(F) = F \text{ if } n = 1, \\
T(F) = (F, \det F) \text{ if } n = 2, \\
T(F) = (F, \adj F, \det F) \text{ if } n = 3.
\]

Let \( U \subseteq M^{n+s} \). By the theorem on the invariance of domain the set \( T(U) \subseteq E \) is open if and only if \( U \) is open. However, \( T(U) \) is not in general convex even if \( U \) is convex (except in the case \( n = 1 \)). The following result shows, in particular, that in certain important cases when \( U \) is open, so is \( \Co T(U) \).

Theorem 4.3. Let \( K \subseteq \mathcal{R} \) be nonempty and convex, and let \( U = \{ F \in M^{n+s}: \det F \in K \} \). Then \( \Co T(U) = E_1 \times K \).
If \( n = 3 \):
there exists a convex function \( C(F, A, b) \) on \( C(O) \) with
\[
g(F) \geq C(F, \text{adj } F, \det F)
\]
for all \( F \in O \),
and the inequality
\[
g \left( \sum_{i=1}^{10} \lambda_i F_i^{(i)} \right) \leq \sum_{i=1}^{10} \lambda_i g(F_i^{(i)})
\]
holds for all \( \lambda_i \geq 0 \) with \( \sum \lambda_i = 1 \), and for all \( F_i^{(i)} \in O \) satisfying
\[
\sum_{i=1}^{10} \lambda_i \text{adj } F_i^{(i)} = \text{det } \left( \sum_{i=1}^{10} \lambda_i F_i^{(i)} \right)
\]
and
\[
\sum_{i=1}^{10} \lambda_i \det F_i^{(i)} = \det \left( \sum_{i=1}^{10} \lambda_i F_i^{(i)} \right)
\]
(4.5)

(ii) If \( n = 2 \):
for each \( F \in O \) there exist numbers \( a^i(F), a(F) \) such that
\[
g(F) \geq a^i(F) (F_i^2 - F_i^1) + a(F) (\det \bar{F} - \det F)
\]
for all \( F \in O \).

If \( n = 3 \):
for each \( F \in O \) there exist numbers \( a^i(F), b^i(F), c(F) \) such that
\[
g(F) \geq a^i(F) (F_i^2 - F_i^1) + b^i(F) (\text{adj } F_i^1 - \text{adj } F_i^2) + c(F) (\det \bar{F} - \det F)
\]
for all \( F \in O \).

(iii) If \( n = 2 \):
for each \( F \in O \) there exist numbers \( A^i(F), a(F) \) such that
\[
g(F + \pi) \geq g(F) + A^i(F) \pi_i + a(F) \det \pi
\]
for all \( F + \pi \in O \).

If \( n = 3 \):
for each \( F \in O \) there exist numbers \( A^i(F), B^i(F), c(F) \) such that
\[
g(F + \pi) \geq g(F) + A^i(F) \pi_i + B^i(F) (\text{adj } F_i^1) + c(F) \det \pi
\]
for all \( F + \pi \in O \).

Proof. That polyconvexity of \( g \) is equivalent to (i) or (ii) follows immediately from Theorem 4.2(ii) and Theorem 4.3. That (iii) and (ii) are equivalent follows by setting \( \bar{F} = F + \pi \) and rewriting the right-hand sides of (4.7) and (4.9). (\( \epsilon(F) \) has the same value in both conditions.)

Of course, if \( G \) is \( C^1 \), then the coefficients on the right-hand sides of (4.6) and (4.7) are given by the derivatives of \( G \) with respect to its arguments. Condition (iii) is the form given by Morrey [2, p. 123], who proved the following theorem:
Define

\[ J_s(\sigma) = \int_B G(\sigma(x)) \, dx. \]

Suppose that \( J_s \) exists and is finite or \( +\infty \) for all \( \sigma \in \mathcal{X}_s \). Let \( G \) be such that \( J_s \) is a convex function on the convex set \( \mathcal{X}_s \).

**Definition 4.3.** If \( g : U \to \mathbb{R} \) then \( g \) is said to satisfy condition \( (P_{\mu,n}) \) at \( u \) if and only if there exists \( G = G_0 + \text{Co} T(U) \to \mathbb{R} \) with the above properties and such that

\[ g(F) = G(F, \text{adj} F, \det F) \]

for all \( F \in U \).

**Theorem 4.6.** Let \( U \) be such that \( \text{Co} T(U) \) is open. Then \( g \) satisfies \( (P_{\mu,n}) \) at \( u \) if and only if there exists \( G = G_0 + \text{Co} T(U) \to \mathbb{R} \) with the above properties and such that

\[ g(F) = \frac{1}{\det F} \int_{\mathbb{R}^n} G(F + t \varphi, \text{adj} F, \det F) \, dt \]

for all \( F \in U \).

Proof. (i) Since \( g \) is polyconvex there exists a convex function \( G : \text{Co} T(U) \to \mathbb{R} \) satisfying (4.13). Since \( \text{Co} T(U) \) is open \( G \) is continuous, and hence \( G(\sigma(t)) \) is measurable for each \( \sigma \in \mathcal{X}_s \). By Theorem 4.2 (i) we may suppose that \( G \) is bounded below on \( \text{Co} T(U) \). Thus \( J_s \) exists and is finite or \( +\infty \) for all \( \sigma \in \mathcal{X}_s \), and \( J_s \) is clearly convex.

(ii) Suppose first that \( G = C^1 \) on \( \text{Co} T(U) \). Let \( \zeta \in \mathcal{S}(\Omega) \) satisfy \( F_0 + \varphi(x)e \in U \) for all \( x \in \Omega \). Let \( \sigma = \Sigma(u) \), \( \delta = \Sigma(u) + \zeta \). Then \( e, \delta \in \mathcal{X}_s \). A standard argument shows that if

\[ \Theta(t) = \int_B \left[ G(\sigma(0) - t(\delta - \sigma)) \right] \, dx, \]

then \( \Theta \in C^1([0, 1]) \) with \( \Theta(0) = \Theta(1) = \Theta(0) \). Since \( g \) is polyconvex, we have a convex function \( \Theta(\sigma) \) has the obvious derivative. By the convexity of \( J_s \) we have that \( \Theta(0) \leq \Theta(1) \). Hence

\[ g(F_0) \leq \int_B \left[ G(F_0 + \varphi(x)e) - \frac{\partial G}{\partial F_0}(\Sigma(u)) \zeta^t + \frac{\partial G}{\partial (\text{adj} F_0)}(\Sigma(u))(\text{adj} F_0)^t \right] \, dx, \]

and so \( g \) is quasiconvex at \( F_0 \in U \).

The result follows for general continuous \( G \) by a mollifier argument.

Condition \( P_{\mu,n} \) does not in general imply polyconvexity. Indeed let \( g(F) \) be given by (4.10) with \( a^0_i \) satisfying (4.12) but not (4.11) for any \( B^i_0 \), so that \( g \) is not polyconvex. Let \( \gamma = 2 \) with \( \mu \) arbitrary and let \( U = M^{3 \times 3} \). Define \( G : E \to \mathbb{R} \) by \( G(F, A, \delta) = g(F) \) for all \( (F, A, \delta) \in E \). Let \( u \) be as above. For any \( \zeta \in W^{1,2}_0(\Omega) \) we have (VAN HOVE [1], MORREY [2])

\[ \int_B a^0_i \zeta^t A_i^j \zeta^j \, dx \geq 0. \]

It follows that \( J_s \) is convex on \( \mathcal{X}_s \), and hence \( g \) satisfies \( P_{\mu,n} \) at \( u \).

There are some grounds connected with the results of Section 6, for believing that condition \( P_{\mu,n} \) at affine \( u \) and the quasiconvexity condition are equivalent, but I have been unable to prove or disprove this.

To complete this section I remark that many of the results may be extended without undue difficulty to arbitrary \( m \) and \( n \); the polyconvexity condition is then a requirement of convexity with respect to the basis elements of the null space of the Euler-Lagrange operator. (See BALL [2].)

### 5. Isotropic Convex and Polyconvex Functions

The purpose of this section is to give a method for producing a wide variety of nontrivial isotropic polyconvex functions. These functions will prove valuable in Section 8 when we apply our existence theorems to certain models which have been proposed for rubbers. We begin by discussing isotropic convex functions of \( n \times n \) matrices for arbitrary \( n \geq 1 \). We recall that the singular values of an \( n \times n \) matrix \( F \) are by definition the eigenvalues of the positive semidefinite symmetric matrix \( \sqrt{FF^t} \). When \( F \) is the deformation gradient these eigenvalues are the principal stretches of the deformation. When examining the results below the reader should bear in mind equation (1.15).

**Notation.** Vectors in \( \mathcal{A} \) are denoted by \( x = (x_1, \ldots, x_n) \) and the inner product of two vectors \( x, y \in \mathcal{A} \) is written \((x, y)\). \( \mathcal{A}^n \) denotes the positive orthant \( \{x : x_i \geq 0 \text{ for } 1 \leq i \leq n\} \) of \( \mathcal{A} \). \( \mathcal{P} \) denotes the permutation group on \( n \) symbols (an element \( P \) of \( \mathcal{P} \) acts on an \( n \) vector by permuting its entries).

We shall prove the following theorem:

**Theorem 5.1.** Let \( n \geq 1 \). Let \( \Phi(v_1, \ldots, v_n) \) be a symmetric real-valued function defined on \( \mathcal{A}^n \). For \( F \in M^{n \times n} \) define

\[ W(F) = \Phi(v_1, \ldots, v_n). \]

where \( v_1, \ldots, v_n \) are the singular values of \( F \). Then

(i) \( W \) is convex on \( K = \{V \in M^{n \times n} : V \text{ is positive-semidefinite and symmetric}\} \) if and only if \( \Phi \) is convex.

(ii) \( W \) is convex on \( M^{n \times n} \) if and only if \( \Phi \) is convex and nondecreasing in each variable \( v_i \).

**Remarks:** Part (i), which is probably known to matrix theorists, was stated by HILL [3] for \( n = 3 \). HILL's proof relies on a property of the trace of the product of two symmetric matrices, a recent proof of which has been given by THEOBALD [1]. HILL assumed that \( W \) is differentiable everywhere, an assumption which rules out several simple and useful functions and which can be surprisingly tedious to verify. Proofs using differentiability obscure the geometric nature of the result.

The harder implication in (ii) is due to THOMPSON & FREED [3]. Our method of proof is broadly similar in approach, but rather different in detail. The result extends work of VON NEUMANN [1] on gauge-functions and matrix norms (see especially his Remark 5). For further information and references on singular value inequalities see AMIR-MOEZ [1], AMIR-MOEZ & HORN [1], THOMPSON [1], THOMPSON & FREED [1-3].
Lemma 5.1. (Von Neumann [1]; see also Mirsky [1, 2].) Let \( A, B \in M^{**} \) have singular values \( \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n \geq 0 \) and \( \beta_1 \geq \beta_2 \geq \cdots \geq \beta_n \geq 0 \). Then
\[
\left| \text{tr}(AB) \right| \leq (\alpha, \beta).
\] (5.2)

Lemma 5.2. (Von Neumann [1].) Under the hypotheses of Lemma 5.1
\[
\max_{Q, R} \text{tr}(AQB) = (\alpha, \beta),
\] (5.3)
where the maximum is taken over all pairs \( Q, R \) of orthogonal matrices.

Proof. There exist orthogonal matrices \( Q_1, Q_2, R_1, R_2 \) such that \( A = Q_1 \text{diag}(\alpha) R_1 \), \( B = Q_2 \text{diag}(\beta) R_2 \). Choose \( Q = R_1^* Q_1^T \), \( R = R_1^* Q_2^T \). Then
\[
\text{tr}(AQB) = (\alpha, \beta).
\]

But for any orthogonal \( Q, R \) the matrices \( AQ \) and \( BR \) have singular values \( \alpha \) and \( \beta \) respectively. Hence \( \text{tr}(AQB) \leq (\alpha, \beta) \) by Lemma 5.1. \( \square \)

Lemma 5.3. Let \( r_1 \geq r_2 \geq \cdots \geq r_n \geq 0 \). Then \( (r, v) \) is a convex function of \( F \), where \( r_1 \geq r_2 \geq \cdots \geq r_n \geq 0 \) are the singular values of \( F \).

Proof. In Lemma 5.2 put \( A = F, B = \text{diag } r \). Then
\[
(r, v) = \max_{Q, R} \text{tr}(FQBR),
\] (5.4)
and each \( \text{tr}(FQBR) \) is a convex function of \( F \). \( \square \)

Remark. By letting \( r = (1, \ldots, 1, 0, \ldots, 0) \) in Lemma 5.3 it follows that for each place
\[
1 \leq k \leq n, \frac{1}{k} v_k \text{ is a convex function of } F.
\]

Lemma 5.4. Let \( c_1 \geq c_2 \geq \cdots \geq c_n \geq 0 \). Define the sets
\[
L = \{ y \in \mathbb{R}^n : (r, y) \leq (r, c) \text{ for all } r_1 \geq r_2 \geq \cdots \geq r_n \geq 0 \text{ and all } P \in \mathbb{R}_+ \},
\]
\[
M = \{ y \in L : \sum_{i=1}^n y_i = \sum_{i=1}^n c_i \},
\]
\[
L_0 = \text{Co } \{ P(c_1, c_2, \ldots, c_1, 0, \ldots, 0) : P \in \mathbb{R}_+, 1 \leq l \leq n \},
\]
\[
M_0 = \text{Co } \{ P(c_1, c_2, \ldots, c_n) : P \in \mathbb{R}_+ \}.
\]

Then \( L = L_0 \) and \( M = M_0 \).

Proof. \( L \) is a convex set containing \( 0 \) and the points \( P(c_1, c_2, \ldots, c_n, 0, \ldots, 0) \). Thus \( L_0 \subseteq L \). To show that \( L_0 \subseteq L \) we prove that any closed half-space containing \( L_0 \) also contains the closed convex set \( L \). Let such a half-space be \( \pi = \{ y \in \mathbb{R}^n : (x, y) \geq \mu \} \), where \( x \in \mathbb{R}^n, \mu \in \mathbb{R} \) are fixed. Let the coordinates of \( x \) in some order be
\[
\bar{x}_1 \geq \bar{x}_2 \geq \cdots \geq \bar{x}_n > \bar{x}_{n+1} \geq \cdots \geq x_n,
\]
where the symbols to the right of 0 are omitted if \( k = n \). Let \( y \in L \) and let
\[
\bar{y}_1 \geq \bar{y}_2 \geq \cdots \geq \bar{y}_n \geq 0,
\]
be the coordinates of \( y \) in some order. Then
\[
(y, x) \leq \bar{y}_1 \bar{x}_1 + \cdots + \bar{y}_n \bar{x}_n \leq c_1 \bar{x}_1 + \cdots + c_n \bar{x}_n \leq \mu,
\]
so that \( y \in \pi \). Hence \( L_0 \subseteq L \).

\( M \) is a convex set containing the points \( P \). Thus \( M_0 \subseteq M \). Let \( y \in M \) and let \( \bar{y}_1 \geq \bar{y}_2 \geq \cdots \geq \bar{y}_n \geq 0 \) be the coordinates of \( y \) in some order. Then
\[
(y, x) \leq \sum_{i=1}^n \bar{y}_i \bar{x}_i = \sum_{i=1}^n \left( \sum_{j=1}^{n-1} c_j \right) \bar{x}_j < \cdots < \sum_{i=1}^n \bar{y}_i (\bar{x}_i - \bar{x}_{i+1}) + \bar{x}_n \sum_{i=1}^n c_i,
\]

Choosing \( r_i = x_i - x_{i+1} (1 \leq i < n), r_n = 0 \) in the definition of \( L \) we obtain
\[
(y, x) \leq \sum_{i=1}^n c_i (\bar{x}_i - \bar{x}_{i+1}) + \bar{x}_n \sum_{i=1}^n c_i \leq \mu.
\]

Thus \( y \in \pi \) and \( M_0 \subseteq M \). \( \square \)

Proof of Theorem 5.1.

(i) That the convexity of \( W \) implies the convexity of \( \Phi \) is immediate. Thus let \( \Phi \) be convex and symmetric. Let \( U, V \in K \) have singular values \( (a_1, a_2, \ldots, a_n, 0) \) and \( (b_1, b_2, \ldots, b_n, 0) \) respectively. Note that for some orthogonal \( Q, R \) we have
\[
\sum_{i=1}^n a_i = \sum_{i=1}^n (b_i Q^T u_i + (1 - \lambda) R^T v_i),
\] (5.5)

By (5.5) and Lemma 5.3 we have \( a \in M \). Hence by Lemma 5.4 \( a \in M \). Therefore
\[
W(A) = \Phi(a) = \Phi(\sum_{i=1}^n \lambda_i P_i e),
\] (5.6)

where \( \lambda_i P_i e, \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1 \). Thus
\[
W(A) \leq \sum_{i=1}^n \lambda_i W(P_i) \leq \sum_{i=1}^n \lambda_i \Phi(P_i e) = \sum_{i=1}^n \lambda_i \Phi(e) = \Phi(e) \leq \lambda W(U) + (1 - \lambda) W(V),
\] (5.7)

where we have used the symmetry of \( \Phi \). Hence,

\[
W(A) \leq \lambda W(U) + (1 - \lambda) W(V)
\]
as required.
The geometrical basis for Theorem 5.1 is easily seen by considering the special examples. Let $\mathcal{A} = \lambda F + (1 - \lambda) G$ have singular values

\[ v_1 \geq v_2 \geq \cdots \geq v_n \geq 0, \quad a_1 \geq a_2 \geq \cdots \geq a_n \geq 0 \]

respectively. Let $\epsilon = \lambda u + (1 - \lambda) v$. By Lemma 5.3 $\epsilon \in \mathcal{L}$. Thus $\epsilon \in \mathcal{L}$, and since $\Phi$ is nondecreasing in each variable we have

\[ W(A) = \Phi(a) = \Phi(\epsilon) \leq \lambda W(F) + (1 - \lambda) W(G). \quad \square \]

Theorem 5.2. If $n = 2$

\[ W(F) = \psi(v_1, v_2, a_1, a_2), \quad (5.8) \]

where $v_1, v_2$ are the singular values of $F$, and where $\psi: \mathcal{R}^2 \times K \to \mathcal{R}$ is convex and satisfies

(a) $\psi(x_1, x_2, \delta) = \psi(x_2, x_1, \delta)$ for all $x_1, x_2 \in \mathcal{R}^n, \delta \in K$,

(b) $\psi(x_1, x_2, \delta)$ is nondecreasing in $x_1, x_2$.

If $n = 3$

\[ W(F) = \psi(v_1, v_2, v_3, a_1, a_2, a_3), \quad (5.9) \]

where $v_1, v_2, v_3$ are the singular values of $F$, and where $\psi: \mathcal{R}^3 \times K \to \mathcal{R}$ is convex and satisfies

(a) $\psi(P, x, y, \delta) = \psi(x, y, \delta)$ for all $P, \delta \in \mathcal{R}^3$, and all $x, y \in \mathcal{R}^3$,

(b) $\psi(x_1, x_2, x_3, y_1, y_2, y_3, \delta)$ is nondecreasing in each $x_i, y_j$.

Then $W$ is polyconvex on $U$.

Proof. We give the proof for $n = 3$. Define $G: \mathcal{E} \times K \to \mathcal{R}$ by

\[ G(A, \delta) = \psi(v_1, v_2, v_3, a_1, a_2, a_3), \quad (5.10) \]

where $v_1, a_1$ are the singular values of $A$. $G$ is well-defined by (a). Clearly

\[ W(F) = G(F, \text{adj } F, \text{det } F) \quad (5.11) \]

for all $F \in U$. It remains to show that $G$ is convex. Let $F, H, A, B \in \mathcal{M}^n \times \mathcal{M}^n$, $\lambda \in [0, 1]$. Let $F, H, A, B, \lambda F + (1 - \lambda) H, A + (1 - \lambda) B$ have singular values

\[ v_1 \geq v_2 \geq \cdots \geq v_n \geq 0, \quad a_1 \geq a_2 \geq \cdots \geq a_n \geq 0 \]

respectively. Let $w = \lambda u + (1 - \lambda) v$, $\epsilon = \lambda \delta + (1 - \lambda) b$. Using Lemma 5.3, Lemma 5.4, (a) and (b) we see that

\[ G(\lambda F + (1 - \lambda) H, \lambda A + (1 - \lambda) B, \lambda \delta + (1 - \lambda) b) = \psi(w, \lambda \delta + (1 - \lambda) b) \]

\[ \leq \psi(w, \lambda \delta + (1 - \lambda) b) \]

\[ \leq \psi(w, \epsilon) \leq \lambda G(F, A, \delta) + (1 - \lambda) G(H, B, b). \]

A special case of a function $\psi$ satisfying the hypotheses of Theorem 5.2 for $n = 3$ is

\[ \psi(x, a, \delta) = \psi_1(x) + \psi_2(a) + \psi_3(\delta), \quad (5.12) \]

where the $\psi_i$ are convex, and where $\psi_1, \psi_2, \psi_3$ are symmetric and nondecreasing in each variable. We use this example in Section 8.

6. Sequential Weak Continuity of Mappings on Orlicz-Sobolev Spaces

Definition 6.1. Let $X$ and $Y$ be real Banach spaces. A map $f: X \to Y$ is sequentially weakly continuous if and only if $x_n \to x$ in $X$ implies $f(x_n) \to f(x)$ in $Y$.

In general a nonlinear sequentially weakly continuous map $f: X \to Y$ is not continuous with respect to the weak topologies on $X$ and $Y$ (see Ball [13]).
Sequentially weak \( \ast \) continuous maps are defined analogously. In this section we study the case when \( X \) is an Orlicz-Sobolev space and \( \mathcal{F} = L^1(\Omega) \). The reader unfamiliar with the theory of Orlicz and Orlicz-Sobolev spaces can replace them everywhere by the corresponding Lebesgue and Sobolev spaces, the results for which are indicated in parentheses in some of the theorems which follow. There will be no great loss of generality in doing so, at least as far as most of the examples in Section 8 are concerned. Some of the results proved here in the framework of Orlicz-Sobolev spaces are proved for ordinary Sobolev spaces in Ball [2].

We first obtain necessary conditions.

Theorem 6.1. (Morrey [1]) Let \( m \) and \( n \) be arbitrary. Let \( \psi: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R} \) be continuous. Then

\[
I(u, \Omega) = \int \psi(x, u(x), \nabla u(x)) \, dx
\]

is sequentially weak \( \ast \) semicontinuous on \( W^{1, \psi}(\Omega) \) (i.e., \( u, \cdot \rightarrow u \) in \( W^{1, \psi}(\Omega) \) implies \( I(u, \Omega) \leq \lim \inf I(u_n, \Omega) \) if and only if \( \psi(x, u, \cdot) \) is quasiconvex on \( M^{m \times n} \) for each \( x \in \Omega, u \in \mathbb{R}^n \).

Corollary 6.1.1. Let \( n = 1, 2 \) or \( 3 \) and let \( \phi: M^{m \times n} \rightarrow \mathbb{R} \) be continuous. Then the map \( u \rightarrow I(u, \Omega) = \int \phi(\nabla u(x)) \, dx \) is sequentially weak \( \ast \) continuous map from \( W^{1, \psi}(\Omega) \) to \( \mathbb{R} \) if and only if \( \phi \) has the form (4.2).

Proof. If the given map is sequentially weak \( \ast \) continuous then by the theorem both \( \phi \) and \( - \phi \) are quasiconvex, so by Corollary 4.1.1 \( \phi \) has the form (4.2). Conversely, let \( \phi \) have the form (4.2) and let \( u, \cdot \rightarrow u \) in \( W^{1, \psi}(\Omega) \). The sequence \( \phi(\nabla u_n(x)) \) is bounded in \( L^n(\Omega) \), so there exists a subsequence \( u_n \), \( u \), such that \( \phi(\nabla u_n(x)) \rightarrow \phi(\nabla u(x)) \) in \( L^n(\Omega) \). Let \( \alpha: \mathbb{R}^n \rightarrow \mathbb{R} \) be continuous and define \( \phi_j(x, F) = \phi(\alpha(F)) \) so that \( \phi_j \) is quasiconvex. By the theorem

\[
\int \phi(\nabla u_j(x)) \, dx = \lim \int \phi(\nabla u(x)) \, dx.
\]

The arbitrariness of \( \alpha \) implies that \( \theta = \phi \) (\( \phi \) is quasiconvex), and hence \( \phi(\nabla u(x)) \rightarrow \phi(\nabla u(x)) \) in \( L^n(\Omega) \). The results follow.

Corollary 6.1.2. Let \( n = 1, 2 \) or \( 3 \) and let \( A \) be an \( N \)-function (cf. Section 2). Let \( \phi: M^{m \times n} \rightarrow \mathbb{R} \) be continuous and such that \( u \rightarrow \phi(\nabla u(x)) \) is a sequentially continuous map of \( W^{1, \psi}(\Omega) \) with the weak \( \ast \) topology into \( L^n(\Omega) \) with the weak topology. Then \( \phi \) has the form (4.2).

Proof. The hypotheses imply that \( u \rightarrow I(u, \Omega) \) is a sequentially weak \( \ast \) continuous map from \( W^{1, \psi}(\Omega) \) to \( \mathbb{R} \).

Remark: Morrey's proof of Theorem 6.1 may be easily adapted to show that if \( \Omega \) is a bounded open subset of \( \mathbb{R}^n \) (in arbitrary) and if \( \phi: \mathbb{R} \rightarrow \mathbb{R} \), then the map \( \theta \rightarrow \phi(\theta(\cdot)) \) is sequentially weakly continuous from \( L^n(\Omega) \rightarrow L^n(\Omega) \) if and only if \( \phi \) is affine. For details see Ball [2].

We turn now to sufficient conditions. Our results are based on the following elementary identities for \( C^1 \) functions \( u \):

\[
\begin{align*}
\text{(6.1)} & \quad n = 2: \quad \det \nabla u = \left( u^1 u^2, -u^1 u^2, 1 \right) \\
\text{(6.2)} & \quad n = 3: \quad (\text{adj} \, \nabla u)^T = \left( u^2 u^3 - u^3 u^2, u^3 u^1 - u^1 u^3, u^1 u^2 - u^2 u^1 \right).
\end{align*}
\]

(In (6.2) there is no implied summation, and the indices are to be taken modulo 3.)

Lemma 6.1.

(i) \( n = 2 \): If \( u \in W^{1, 2}(\Omega) \) then \( \det \nabla u \in L^1(\Omega) \) and formula (6.1) holds in \( \mathcal{D}'(\Omega) \).

(ii) \( n = 3 \): (a) If \( u \in W^{1, 2}(\Omega) \) then \( \text{adj} \, \nabla u \in L^1(\Omega) \) and formula (6.2) holds in \( \mathcal{D}'(\Omega) \).

Proof. (i) That \( \det \nabla u \in L^1(\Omega) \) is obvious. Formula (6.1) holds in \( \mathcal{D}'(\Omega) \) if and only if

\[
\int \det \nabla u \, \phi(x) \, dx = \int \left( u^1 u^2, -u^1 u^2, 1 \right) \phi(x) \, dx
\]

for all \( \phi \in \mathcal{D}(\Omega) \).

But (6.4) holds trivially if \( u \in C^\infty(\Omega) \), and \( C^\infty(\Omega) \) is dense in \( W^{1, 2}(\Omega) \) in its norm topology. Since both sides of (6.4) are continuous functions of \( u \in W^{1, 2}(\Omega) \), (6.4) holds for \( u \in W^{1, 2}(\Omega) \).

(ii) The proof of (a) is identical to that of (i). Let \( w \) be defined by \( w = \text{adj} \, \nabla u \).

To prove (b) we first note that \( u_j \in E_{1,2}(\Omega), w_j \in E_1(\Omega) \) so that \( \det \nabla u_j \in L^1(\Omega) \). We now show that \( \nabla u_j \rightarrow 0 \) in a weak sense, i.e.,

\[
\int \nabla u_j \phi(x) \, dx = 0
\]

for all \( \phi \in \mathcal{D}(\Omega) \).

If \( u \in C^\infty(\Omega) \), then \( \nabla u \rightarrow 0 \) and (6.5) holds. Since \( C^\infty(\Omega) \) is dense in \( W^{1, 2}(\Omega) \), (6.5) holds for any \( u \in W^{1, 2}(\Omega) \).

To show that (6.3) holds in \( \mathcal{D}'(\Omega) \) it is thus sufficient to prove that

\[
\int w_j \phi(x) \, dx = - \int w \phi_j(x) \, dx
\]

for all \( \phi \in \mathcal{D}(\Omega) \).

whenever \( w \in E_1(\Omega) \) and satisfies (6.5).

By the results of Donaldson & Trudinger [1, Thm. 22] there exists a sequence \( u_n \in C^\infty(\Omega) \) with \( u_n \rightarrow u \) in \( W^{1, 2}(\Omega) \). Let \( \rho \in \mathcal{D}(\Omega) \), \( \rho \geq 0 \), \( \int \rho(x) \, dx = 1 \), and define \( \rho_n \in \mathcal{D}(\Omega) \) by \( \rho_n(x) = k \rho(kx) \). Extend \( w \) by zero outside \( \Omega \), so that \( w \in E_{1,2}(\Omega) \). Donaldson & Trudinger [1, Lemma 2.1] show that the convolutions \( \rho_n \) are in \( E_{1,2}(\Omega) \) and \( \rho_n \ast w \rightarrow w \) in \( E_{1,2}(\Omega) \) as \( k \rightarrow \infty \). Fix \( \phi \in \mathcal{D}(\Omega) \). Then if \( k \) is large enough, (6.5) implies that

\[
\text{div}(\rho_n \ast w)(x) = \int \rho_n(x-y) \phi(y) \, dy = 0
\]
Lemma 6.2.

Note that for $A(t) = \mathbb{I}W$, the distributions $\text{Det} \ V_u$ define $\frac{d}{dt}(\mathbb{I}W)$ for all $X \in \text{supp} \phi$. Since $\mathbb{A} \prec \mathbb{B}$ we obtain (6.6) by letting $k \to \infty$. 

Remark. Note that $A(t) = \mathbb{I}t^2$ if and only if $A = \mathbb{A}$. This may be proved directly from the definition of $\mathbb{A}$.

The functions $\text{det} \ V_u (n = 2)$ and $\text{adj} \ V_u$ and $\text{det} \ V_u (n = 3)$ can be given a meaning as distributions under weaker conditions than those of Lemma 6.1. We thus define the distributions $\mathcal{D} \ V_u (n = 2)$, $\mathcal{A} \ V_u$ and $\text{det} \ V_u (n = 3)$ by

$$
\mathcal{D} \ V_u = (u^i u^j)_{1,2} (u^l u^r)_{1,2},
$$

$$
\mathcal{A} \ V_u = (u^i u^j u^k)_{1,2,3} + (u^l u^r u^s)_{1,2,3},
$$

$$
\text{det} \ V_u = [u^i (\text{adj} \ V_u)]_{1,2},
$$

when these distributions are meaningful. Obviously if $u$ satisfies the hypotheses of Lemma 6.1 then these distributions may be identified with the $L^1(\Omega)$ functions $\text{det} \ V_u (n = 2)$, $\text{adj} \ V_u$ and $\text{det} \ V_u (n = 3)$ respectively.

Let $A$ be an $N$-function. Following DONALDSON & TRUDINGER [1], we let

$$
g_a(t) = \frac{A^{-1}(t)}{t^{1-n}}, \quad t \geq 0,
$$

where $A^{-1}$ denotes the inverse function to $A$ on $[0, \infty)$. If $A$ satisfies

$$
\int_0^t g_a(s) ds < \infty, \quad \int_0^t g_a(s) ds = \infty,
$$

then we define the $N$-function $A^*$ by

$$
(A^*)^{-1}(t) = \int_0^t g_a(s) ds.
$$

Note that for $A(t) = |t|^p$, $1 < p < n$, we have

$$
g_a(t) = t^{\frac{1}{p} - \frac{1}{n} - 1}, \quad A^*(t) = |t|^p.
$$

Lemma 6.2. Let $\Omega$ satisfy the cone condition.

(i) $n = 2$: Let $A$ be an $N$-function satisfying either $\int_0^t g_a(s) ds < \infty$, or both (6.12) and $\mathbb{A} \prec \mathbb{A}$. If $u \in L^1 L^1(\Omega)$ (e.g., $u \in W^1 L^1(\Omega)$), then $u^i u^j$ and $u^l u^r$ belong to $L^1(\Omega)$, so that $\mathbb{D} \ V_u$ exists as an element of $\mathcal{D}^*(\Omega)$. 

(ii) $n = 3$: (a) Let $A$ be as in (i). Then $u \in L^1 L^1(\Omega)$ (e.g., $u \in W^1 L^1(\Omega)$), then $u^i u^j u^k$ and $u^l u^r u^s$ belong to $L^1(\Omega)$, so that $(\text{adj} \ V_u)^d$ exists as an element of $\mathcal{D}^*(\Omega)$.

Proof. (i) If $\int_0^t g_a(s) ds < \infty$ then the imbedding theorem of DONALDSON & TRUDINGER [1, Thm. 3.2] implies that $u^i u^j \in L^1(\Omega)$, while if $A$ satisfies (6.12) the same theorem implies that $u^i u^j \in L^1(\Omega)$. The result follows by Young's inequality.

(ii) This is proved similarly. 

Remark. If $\Omega$ is an arbitrary bounded open set then Lemma 6.2 holds with $L^1(\Omega)$ replaced by $L^1(\Omega)$.

The main result of this section is the following:

Theorem 6.2.

(i) $n = 2$: Let $A$ be an $N$-function satisfying either $\int_0^t g_a(s) ds < \infty$, or both (6.12) and $\mathbb{A} \prec \mathbb{A}$. If $u \in L^1 L^1(\Omega)$ (e.g., $u \in W^1 L^1(\Omega)$, $p > \frac{1}{q}$), then $\mathbb{D} \ V_u \to \mathbb{D} \ V_u \in \mathcal{D}^*(\Omega)$.

(ii) $n = 3$: (a) Let $A$ be as in (i). Then $u \in L^1 L^1(\Omega)$ (e.g., $u \in W^1 L^1(\Omega)$, $p > \frac{1}{q}$) then $(\text{adj} \ V_u)^d \to (\text{adj} \ V_u)^d \in \mathcal{D}^*(\Omega)$.

Proof. (i) Fix $\phi \in \mathcal{D}(\Omega)$ and let $\Omega'$ be an open set with $\Omega' \supset \text{supp} \phi$ and such that the imbedding theorems of DONALDSON & TRUDINGER hold for $\Omega'$. Since $|u||u||L^1(\Omega)|$, is bounded and $u \to u$ in $L^1(\Omega')$, it follows (see KRASNOSELSKII & RUTICKII [1, p. 132]) that $u \to u$ in $L^1(\Omega')$. Therefore the Hölder inequality and the boundedness of $|u||u||L^1(\Omega)|$ imply that

$$
\int_0^t u^i u^j u^k u^l u^r u^s dx = \int_0^t (u^i u^j u^k u^l u^r u^s) dx + \int_0^t (u^i u^j u^k) dx
$$

tends to zero as $t \to \infty$. Hence $u^i u^j u^k \to u^i u^j u^k$ in $\mathcal{D}(\Omega)$. Similarly $u^l u^r u^s \to u^l u^r u^s$ in $\mathcal{D}(\Omega)$. The result follows. The proof of (ii) is similar. 

Corollary 6.2.1.

(i) $n = 2$: Let $A$ be as in Theorem 6.2(i) and let $u \in W^1 L^1(\Omega)$. Then

$$
\text{det} \ V_u = (u^i u^j)_{1,2} (u^l u^r)_{1,2} \in \mathcal{D}(\Omega).
$$

(ii) $n = 3$: (a) Let $A$ be as in (i). Then $\text{det} \ V_u = (u^i u^j u^k)_{1,2,3} \in \mathcal{D}(\Omega)$.

We may take for $\Omega$ a finite subcover by open balls of the closure of $\text{supp} \phi$. 


Let $A$ be as in Theorem 6.2(iiia) and let $u \in W^1L_4(\Omega)$. Then
\[
(\text{Adj } Vu)^n = (u^{n+1}u^{n+1})_{x+1} - (u^n u^{n+1})_{x+1} \quad \text{in } \mathcal{D}'(\Omega).
\] (6.16)

(b) Let $A, B$ be as in Theorem 6.2(iiib) and let $u \in W^1L_4(\Omega), \text{Adj } Vu \in L_\phi(\Omega)$. Then
\[
\text{Det } Vu = [u^3(\text{Adj } Vu)^3]_{ij} = [u^3(\text{Adj } Vu)^3]_{ij} \quad \text{in } \mathcal{D}'(\Omega).
\] (6.17)

Proof. (i) Let $\phi \in \mathcal{D}(\Omega)$ and let $\Omega'$ be an open set with $\Omega = \Omega' \cap \text{supp } \phi$ and satisfying the segment property. Then by the results of Gossez [1, Thm. 1.3] there exists a sequence $u_n \in C^\infty(\Omega')$ with $u_n \rightarrow u$ in $W^1L_4(\Omega')$. Clearly $\text{Det } Vu_n(\phi) = ([u^n u^n]_{x+1} - (u^n u^n)_{x+1} | \phi)$. Letting $r \rightarrow 0$ we obtain from the theorem that $\text{Det } Vu(\phi) = ([u^n u^n]_{x+1} - (u^n u^n)_{x+1} | \phi)(\phi)$, and the result follows.

(ii) The proof of (a) is identical to that of (i). The proof of (b) is similar to that of Lemma 6.2(iiib), the principal change being the use of Lemma 1.6 of Gossez [1] to show that if $u \in L_\phi(\Omega)$, then $\rho_{x} w \rightarrow u$ in $L_\phi(\Omega)$. We omit the details. □

Corollary 6.2.2.*

(i) $n = 2$: The map $u \rightarrow \text{det } Vu: W^{1,1}(\Omega) \rightarrow L^2(\Omega)$ is sequentially weakly continuous if $p > 2$.

(ii) $n = 3$: (a) The map $u \rightarrow \text{adj } Vu: W^{1,1}(\Omega) \rightarrow L^2(\Omega)$ is sequentially weakly continuous if $p > 2$.

(b) The map $u \rightarrow \text{det } Vu: W^{1,1}(\Omega) \rightarrow L^2(\Omega)$ is sequentially weakly continuous if $p > 3$.

Proof. We just prove (iiib). Let $p > 3$. It is clear from the Hölder inequality that if $u \in W^{1,1}(\Omega)$, then $\text{det } Vu \in L^2(\Omega)$. Let $u_n \rightarrow u$ in $W^{1,1}(\Omega)$. Then $\text{adj } Vu_n$ is bounded in the reflexive space $L^2(\Omega)$, and hence by the theorem (part (ii)) $\text{adj } Vu_n \rightarrow \text{adj } Vu$ in $L^2(\Omega)$. But $\text{det } Vu_n$ is bounded in the reflexive space $L^2(\Omega)$ and thus by part (ii) of the theorem $\text{det } Vu_n \rightarrow \text{det } Vu$ in $L^2(\Omega)$. □

Warning. The distributions $\text{det } Vu(\text{adj } Vu)$ and $\text{det } Vu(\text{Adj } Vu)$ need not be the same even if the former is a continuous function, as the following example shows.

Example 6.1, $n = 2$ or 3.

Let $r = |x|$ and let $R(r)$ be a smooth real-valued function on $(0, 1]$.

Let $\Omega = \{x < 1\}$ and define $u: \Omega \rightarrow \mathbb{R}^n$ by
\[
u(x) = \frac{R(r)}{r} x, \quad r > 0, u(0) \text{ arbitrary.}
\] (6.18)

Then for $r > 0$ we have
\[
u_r = \frac{R(r)}{r^2} x + \frac{R(r)}{r} x, \quad \text{det } Vu = \frac{R(r)}{r^3} (n-1).
\] (6.19)

7. Existence Theorems

We have already stated the result of Morrey, Theorem 6.1, which shows that for a continuous integrand quasiconvexity is necessary and sufficient for sequential weak lower semicontinuity in $W^{1,1}(\Omega)$. Morrey [1, 2] has also given sufficient conditions for sequential weak lower semicontinuity in $W^{1,1}(\Omega)$, $s \geq 1$. For purposes of comparison, and for future use, we give an extension of his results due to Meyers [1].

Theorem 7.1. Let $f: \Omega \times \mathbb{R}^n \times M^{n,n} \rightarrow \mathbb{R}$ be continuous, and let $f(x, u, \cdot)$ be quasiconvex for all $x \in \Omega$, $u \in \mathbb{R}^n$. Suppose there exist real constants $C_i > 0$ ($i = 1, 2$), $s \geq 1$, $0 < p < 1$ and a function $f \in L^1(\Omega)$ such that
\[
\int_B f(x, u, \cdot) dx = \int_B f(x, u, \cdot) dx - \int_B f(x, u, \cdot) dx - \int_B f(x, u, \cdot) dx.
\] (6.21)

But
\[
\int_B f(x, u, \cdot) dx = \int_B f(x, u, \cdot) dx - \int_B f(x, u, \cdot) dx.
\] (6.21)

Provided $r \rightarrow 0$ as $r \rightarrow 0$ and $u, \text{Vu} \in L^1(\Omega)$ we therefore obtain from (6.21) that
\[
\int_B f(x, u, \cdot) dx = \int_B f(x, u, \cdot) dx - \int_B f(x, u, \cdot) dx.
\] (6.21)

Hence formula (6.1) does not hold in this case, so that $\text{det } Vu \neq \text{Det } Vu$. Note also that if $p < 2$ there is no sequence of $C^\infty(\Omega)$ functions $u_i$ such that $u_i \rightarrow u$ in $W^{1,1}(\Omega)$ and $\text{det } Vu_i \rightarrow \text{det } Vu$ in $L^1(\Omega)$, since such a sequence would satisfy
\[
\int \text{det } Vu \, dx = \int \text{det } Vu \, dx = 4\pi.
\]

whereas
\[
\int \text{det } Vu \, dx = 3\pi.
\]

When $n = 3$, a similar calculation shows that $u \in W^{1,1}(\Omega)$ for $p < 3$, adj $Vu \in L^1(\Omega)$ for $p < 1$, but that (6.17) does not hold.

In the above example $\text{Det } Vu$ has an atom at $x = 0$. I do not know whether $\text{det } Vu = \text{Det } Vu$ if $\text{Det } Vu$ is a function.

* Note added in proof. This corollary follows from results of V. G. Reshetnyak [1, Thm. 4, [2, Thm. 2]. Theorem 4 in Ball [2] is also essentially a consequence of Reshetnyak's work, to which I would have referred had I seen it in time.
We shall use the following lower semicontinuity theorem, which is a special case of a result given by EKELAND & TÉMAM [1, Thm. 2.1, p. 226]. For related results see CESARI [1, 2, 3].

**Theorem 7.2.** Let \( G_1: \Omega \times (\mathbb{R}^n \times \mathbb{R}^*) \to \mathcal{A} \) be of Carathéodory type (Definition 7.1) with \( n = m + e \) and satisfy

\[
G_1(x, u, a) \geq \Phi(|a|)
\]

for some \( \Phi \)-function \( \Phi \). Suppose that \( G_1(x, u, \cdot) \) is convex on \( \mathbb{R}^* \) for all \( x \in \Omega, u \in \mathbb{R}^n \). Let \( a_n \to a \) in \( L^1(\Omega) \) and let \( \{u_n\} \) be a sequence of measurable functions with \( u_n \to u \) almost everywhere in \( \Omega \). Then

\[
\liminf_{n \to \infty} \int \Omega G_1(x, u_n(x), a_n(x)) \, dx \leq \int \Omega G_1(x, u(x), a(x)) \, dx.
\]

We now describe the main ingredients of our existence theory, starting first with those relevant for compressible materials.

For each \( x \in \Omega \) let \( W(x) \) be a nonempty convex open subset of \( \mathbb{R}^{m+n} \), such that for each \( a \in E \) the set \( W^{-1}(a) = \{x \in \Omega: a \in W(x)\} \) is measurable. We impose, as a local constraint on our variational problem, that in a sense to be made precise later, \( T[p(x)] \) maps \( W(x) \) almost everywhere in \( \Omega \), where we are using the notation of Section 4. In applications to nonlinear elasticity \( W(x) \) will often have the form

\[
W(x) = \{a \in E: c_1(x, a) < k_1(x), k_2(x) < c_2(x, a)\}
\]

where \( c_1, c_2: \Omega \to \mathbb{R}^* \) are of Carathéodory type and convex with respect to \( a \in E \), and where \( k_1, k_2: \Omega \to \mathbb{R}^* \) are measurable. Examples of relevant choices of \( c_1, k_1, c_2, k_2 \) are

1. \( c_1, c_2, k_1 = \infty, k_2 = -\infty \) (no constraint).
2. \( (n = 3) \) \( c_1(x, F, A, \delta) = \delta, k_1 = 0 \) (corresponding to the continuity condition \( \det F \geq 0 \) and an additional unilateral constraint, absent if \( k_2 = \infty \), on the measure of strain \( c_1(x, F, A, \det F)^* \)).

We now make continuity, growth and polyconvexity hypotheses on the integrand. Because of the nature of the growth conditions, and because we wish to consider situations in which the distributions \( \operatorname{adj} p(x) \) and \( \operatorname{adj} F \) have different aspects, we make these hypotheses on the associated function \( G(x, u, a) \) (cf. (4.3)).

Let \( \mathcal{Q} = \{x \in \Omega: a \in E\} \). We set \( G: \mathcal{Q} \to \mathcal{A} \) to be such that

\[
G(x, u, a) = G_1(x, u(x), a(x))
\]

is sequentially weakly lower semicontinuous on \( W^{-1}(a) \).

Remarks.

1. Let \( x_0 \in \Omega \). Then conditions (ii) and (iii) imply in particular that

\[
|f(x, u, F) - f(x, 0, 0)| \leq K_1 \eta(x) + |F|, \tag{7.1}
\]

and

\[
|f(x, 0, 0) - f(x, 0, 0)| \leq K_2 \eta(x-x_0), \tag{7.2}
\]

Combining (7.1) and (7.2) we see that

\[
|f(x, u, F)| \leq K_0 \eta(x) + |F| \tag{7.3}
\]

for all \( x, u, F \), where \( K_0 > 0 \) is a constant. Conditions (i), (ii) and (iii) are Meyer's continuity and growth conditions for the function \( f \), while (i) implies a further hypothesis of his theorem.

2. If \( n \leq m \) then the growth conditions with respect to \( u \) may be weakened by use of the Sobolev imbedding theorems; the reader is referred to MEYER [1] for details. An extension to an Orlicz-Sobolev space setting could also probably be made.

In order to prove existence theorems by use of Theorem 7.1, it is necessary to make, in addition to conditions (i)-(iii), a coercivity assumption on \( f \). Typically we might assume that \( a > 1 \) and

\[
f(x, u, F) \geq K_3 |F|^{a} + f(x), \quad \text{where} \quad K_3 > 0.
\]

The conditions (i)-(iv) are extremely restrictive with regard to applications to nonlinear elasticity. Firstly, (i) precludes any singular behaviour of \( f \) (for example \( f(x, u, F) = 0 \)). Secondly (ii) and (iv) together rule out integrands typified by the example

\[
f(F) = |F|^{a} + |\det F|
\]

with \( n > m \); we shall see that many such integrands belong to a physically interesting class included in our existence theorems.

**Definition 7.1.** Let \( D \subseteq \mathbb{R}^d \) be open. A map \( G_1: D \times \mathbb{R}^* \to \mathcal{A} \) is said to be of Carathéodory type if

(a) for almost all \( x \in D \), \( G_1(x, \cdot) \) is continuous on \( \mathbb{R}^* \); and

(b) for all \( a \in \mathbb{R}^* \), \( G_1(x, a) \) is measurable on \( D \).
n = 1: there exists an N-function $\Phi$ and a function $b \in L^1(\Omega)$ such that
\[ G(x, u, F) \geq b(x) + A(F) \quad \text{for all } (x, u, F) \in \mathcal{S}; \]  
(7.7)
n = 2: there exist N-functions $A, B$ with $A$ satisfying either $\int g_A(t) dt < \infty$ or both (6.12) and $A \ll A^*$, and a function $b \in L^1(\Omega)$, such that
\[ G(x, u, F, \delta) \geq b(x) + A(|F|) + B(\delta) \quad \text{for all } (x, u, F, \delta) \in \mathcal{S}; \]  
(7.8)
n = 3: there exist N-functions $A, B, C$ satisfying either $\int g_A(t) dt < \infty$ or the conditions (6.12), $A \ll A^*$ and $B \ll A^*$, and a function $b \in L^1(\Omega)$, such that
\[ G(x, u, F, \delta) \geq b(x) + A(|F|) + B(|H|) + C(\delta) \quad \text{for all } (x, u, F, H, \delta) \in \mathcal{S}; \]  
(7.9)

If we define $G_1 : \Omega \times (\mathbb{R}^n \times E) \to \mathbb{R}$ by
\[ G_1(x, u, a) = G(x, u, a) - b(x) \quad \text{if } (x, u, a) \in \mathcal{S}, \]
\[ = + \infty \quad \text{otherwise,} \]
then clearly $G_1$ is of Carathéodory type and satisfies (7.4) for some N-function $\Phi$.

We define the admissibility set $\mathcal{S}$ by
\[ n = 1 : \mathcal{S} = \{ u \in W^1 L_4(\Omega) : F u(x) \in W(x) \text{ almost everywhere in } \Omega \}, \]
\[ n = 2 : \mathcal{S} = \{ u \in W^1 L_4(\Omega) : \text{Det } F u(x) \in L_4(\Omega), \text{Det } F u(x) \in W(x) \text{ almost everywhere in } \Omega \}, \]
\[ n = 3 : \mathcal{S} = \{ u \in W^1 L_4(\Omega) : \text{Adj } F u(x) \in L_4(\Omega), \text{Adj } F u(x) \in W(x) \text{ almost everywhere in } \Omega \}. \]
The equivalence classes in $\mathcal{S}$ under the equivalence relation
\[ u \sim v \quad \text{if and only if} \quad u - v \in W^1 L_4(\Omega), \]
are termed the Dirichlet classes in $\mathcal{S}$.

If $u \in \mathcal{S}$ then it follows from results of Ekeland & Temam [1, Prop. 1.1, p. 218] that $J(u)$ exists and is finite or $+ \infty$, where
\[ J(u) = \int G(x, u(x), F u(x)) dx \quad \text{if } n = 1, \]
\[ = \int G(x, u(x), F u(x), \text{Det } F u(x)) dx \quad \text{if } n = 2, \]
\[ = \int G(x, u(x), F u(x), \text{Adj } F u(x), \text{Det } F u(x)) dx \quad \text{if } n = 3. \]

We are now in a position to present our first existence theorem, which includes as a special case the displacement boundary-value problem of nonlinear hyperelasticity. For $n = 1$, of course, the result is well known.

**Theorem 7.3.** Let $G$ satisfy $(H_1)$-$(H_7)$ above. Let $\mathcal{C}$ be a Dirichlet class in $\mathcal{S}$, and suppose that there exists $u_0 \in \mathcal{C}$ with $J(u_0) < + \infty$. Then there exists $u_0 \in \mathcal{C}$ that minimizes $J(u)$ in $\mathcal{C}$.

**Proof.** We just give the proof for $n = 3$, the other cases being easier. It is sufficient to establish the existence of a minimizer in $\mathcal{C}$ for
\[ J(u) = \int G(x, u(x), F u(x), \text{Adj } F u(x), \text{Det } F u(x)) dx. \]  
(7.10)

$J(u)$ is bounded below on $\mathcal{C}$. Let $u_0$ be a minimizing sequence from $\mathcal{C}$. By (7.9) the quantities
\[ \int A(|F u_j(x)|) dx, \int B(|\text{Adj } F u_j(x)|) dx, \int C(\text{Det } F u_j(x)) dx \]
are bounded independently of $r$. The Poincaré inequality for $W^1 L_4(\Omega)$ (Gossez [1, p. 202]) implies (cf. Krasnosel’ski & Rutickii [1, p. 131], Ekeland & Temam [1, p. 223]) that for a subsequence $(u_j)$ we have
\[ u_j \rightarrow u_0 \quad \text{in } W^1 L_4(\Omega), \quad \text{Det } F u_j \rightarrow \delta \quad \text{in } L^\infty(\Omega), \]
and
\[ \text{Adj } F u_j \rightarrow H \quad \text{in } L^\infty(\Omega), \quad \text{Det } F u_j \rightarrow \delta \quad \text{in } L^\infty(\Omega). \]

By Theorem 6.2(ii) $H = \text{Adj } F u_0$ and $\delta = \text{Det } F u_0$. By Theorem 7.2,
\[ J(u_0) \leq \lim J(u_j). \]

Since $G_1(x, u, a) = + \infty$ if $a \notin W(x)$ it follows that $(F u_j(x), \text{Adj } F u_j(x), \text{Det } F u_j(x)) \in W(x)$ almost everywhere. Thus $u_0 \in \mathcal{C}$ and the result follows.

We now give a modified version of Theorem 7.3 for the case in which $n = 3$ and $G$ is independent of $\delta$. The proof is similar and is omitted.

**Theorem 7.4.** Let $n = 3$, $\mathcal{C}$ and $\mathcal{S}$ replace $E$ by $E_1 = M^{n \times 3} \times M^{n \times 3}$. Let $G : \mathcal{S} \to \mathbb{R}$ satisfy $(H_1)$-$(H_6)$ and the following hypothesis:
\[ \text{(H}_6) \quad \text{There exist N-functions } A, B \text{ with } A \text{ satisfying either } \int g_A(t) dt < \infty \text{ or both (6.12) and } A \ll A^*, \text{ and a function } b \in L^1(\Omega) \text{ such that} \]
\[ G(x, u, F, H) \geq b(x) + A(|F|) + B(|H|) \quad \text{for all } (x, u, F, H, \delta) \in \mathcal{S}; \]

Define
\[ \mathcal{S} = \{ u \in W^1 L_4(\Omega) : \text{Adj } F u(x) \in L_4(\Omega), \text{Adj } F u(x) \in W(x) \text{ almost everywhere in } \Omega \}, \]

Let
\[ J(u) = \int G(x, u(x), F u(x), \text{Adj } F u(x)) dx. \]  
(7.11)

Let $\mathcal{C}$ be a Dirichlet class in $\mathcal{S}$, and suppose that there exists $u_0 \in \mathcal{C}$ with $J(u_0) < + \infty$. Then there exists $u_0 \in \mathcal{C}$ that minimizes $J(u)$ in $\mathcal{C}$.

**Remark.** Similar modifications to Theorem 7.3 can be made when $n = 2$, and when $n = 3$ and $G$ is independent of both $H$ and $\delta$; the details are left to the reader.
Next we give an existence theorem for the displacement boundary-value problem in three dimensions for an incompressible hyperelastic body.

Theorem 7.5. Theorem 7.4 remains valid if \( \mathcal{A} \) is redefined by \( \mathcal{A} = \{ u \in W^{1, L}(\Omega): \text{Adj} \, Vu \in L_b(\Omega), (Vu(x), \text{Adj} \, Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, \text{Det} \, Vu(x) = 1 \text{ almost everywhere in } \Omega \} \), provided that if \( \int f(t) \, dt = \infty \), we make the extra assumption \( B \leq A^* \).

Proof. Let \( \{u_j\} \) be a minimizing sequence from \( \mathcal{A} \). Then \( \{\text{Det} V u_j\} \) is bounded in \( L^\ast(\Omega) \) independently of \( r \). As in the proof of Theorem 7.3 we may extract a subsequence \( \{u_j\} \subset \mathcal{A} \) with, among other properties,

\[
\text{Det} \, Vu_j \to \delta \quad \text{in } L^\ast(\Omega).
\]

Clearly \( \delta(x) = 1 \) almost everywhere. By Theorem 6.2(ii), \( \delta = \text{Det} \, Vu \), and the result follows. \( \square \)

Remark. A variety of 'weakly closed' constraints can be treated in this way. Analogous results hold for \( n = 2 \).

For the remainder of this section we restrict our attention to the cases \( n = 2, 3 \), which enable us to work in Sobolev spaces, rather than in Orlicz-Sobolev spaces. It would be possible to extend most of our results to an Orlicz-Sobolev space setting.

For ease of reference we now restate hypotheses \((H_5)-(H_6)\) 'in modified form.

For \( n = 2 \):

there exists an \( N \)-function \( B \), real constants \( K_1 > 0, K_2 \geq 0, \gamma > 1, \mu \geq 1 \), and a function \( b \in \mathcal{L}(\Omega) \) such that

\[
G(x, u, F, \delta) \geq b(x) + K_1 |F|^\gamma + K_2 |u|^\mu + B(\delta)
\]

for all \( (x, u, F, \delta) \in \mathcal{A} \).

\( n = 3 \):

there exists an \( N \)-function \( C \), real constants \( K_1 > 0, K_2 \geq 0, \gamma > 1, \mu > 1, \sigma \geq 1 \), and a function \( b \in \mathcal{L}(\Omega) \) such that

\[
G(x, u, F, H, \delta) \geq b(x) + K_1 |F|^\gamma + |H|^\mu + K_2 |u|^\sigma + C(\delta)
\]

for all \( (x, u, F, H, \delta) \in \mathcal{A} \).

Mixed displacement traction boundary-value problems

be considered as well.

Theorem 7.6 (cf. Section 1A). Let \( \Omega \) satisfy a strong Lipschitz condition. Let \( \partial \Omega = \partial \Omega_1 \cup \partial \Omega_2 \), let \( \partial \Omega_1 \cap \partial \Omega_2 = \emptyset \), and let \( \partial \Omega_1 \) and \( \partial \Omega_2 \) be measurable as subsets of

\( \partial \Omega \) with \( \partial \Omega_1 \) having positive measure. Let \( \bar{u}: \partial \Omega_1 \to \mathbb{R}^3 \) be measurable and let \( f_R \in \mathcal{L}(\partial \Omega_1) \) with \( \sigma \geq 1 \). Let \( G: \mathcal{A} \to \mathbb{R} \) satisfy hypotheses \((H_1)-(H_4)\) and \((H_8)\). If

\( n = 2 \), let \( \gamma > \frac{3}{2}, \sigma = \frac{3}{2} \) if \( \gamma < 2, \sigma > 1 \) if \( \gamma = 2, \sigma = 1 \) if \( \gamma > 2, K_2 = 0 \), and

\[
\mathcal{A} = \{ u \in W^{1, \gamma}(\Omega): \text{Adj} \, Vu \in L_b(\Omega), (Vu(x), \text{Adj} \, Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, u = \bar{u} \text{ almost everywhere in } \partial \Omega_1 \}.
\]

If \( n = 3 \), let \( \gamma > \frac{3}{2}, \sigma = \frac{3}{2} \) if \( \gamma < 3, \sigma > 1 \) if \( \gamma = 3, \sigma = 1 \) if \( \gamma > 3, K_2 = 0 \), and

\[
\mathcal{A} = \{ u \in W^{1, \gamma}(\Omega): \text{Adj} \, Vu \in L_b(\Omega), \text{Det} \, Vu \in L_c(\Omega), (Vu(x), \text{Adj} \, Vu(x), \text{Det} \, Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, u = \bar{u} \text{ almost everywhere in } \partial \Omega_1 \}.
\]

Let

\[
J_0(u) = J(u) - \int (Vu(x) - \bar{u}(x))^2 \, dS,
\]

(7.15)

where the integral is defined in the sense of trace.

Suppose that there exists \( u_0 \in \mathcal{A} \) with \( J_0(u_0) < \infty \). Then there exists \( u_0 \in \mathcal{A} \) that minimizes \( J_0(u) \) in \( \mathcal{A} \).

Proof. We give the proof just for \( n = 3 \).

By the trace theorems (cf. Morrey [2], Necas [1]) \( \bar{u} \in L^\ast(\partial \Omega_1) \) and there exists \( k > 0 \) such that

\[
|u|_{W^{1, \gamma}(\Omega)} \leq k \|u\|_{L^\ast(\partial \Omega_1)} \quad \text{for all } u \in W^{1, \gamma}(\Omega).
\]

Since \( \partial \Omega_1 \), has positive measure, a result of Morrey [2, p. 82] implies that there exists \( k > 0 \) such that

\[
\int |u|^\gamma \, dx \leq k \left[ \int |Vu(x)|^\gamma \, dx + \left( \int \|u\|^\mu \, dx \right)^\frac{\gamma}{\mu} \right] \quad \text{for all } u \in W^{1, \gamma}(\Omega).
\]

(7.17)

By \((H_8)\), (7.16) and (7.17) we have for arbitrary \( u \in \mathcal{A} \)

\[
J_0(u) \geq \int b(x) |d_x| + K_1 \int |Vu(x)|^\gamma \, dx + K_2 \int |\text{Adj} \, Vu(x)|^\mu \, dx + \int C(\text{Det} \, Vu(x)) \, dx
\]

\[
- \int (u(x) \cdot \bar{u}(x))^\gamma \, dS
\]

\[
\geq \left( \frac{1}{2} - c \right) \int |Vu(x)|^\gamma \, dx + K_1 \int |\text{Adj} \, Vu(x)|^\mu \, dx
\]

\[
+ \left( \frac{K_1}{2k_1 - c} \right) \int |u(x)|^\gamma \, dx - K_1 \int \|u\|_{L^\ast(\partial \Omega_1)}^\gamma \, dx + c k \|u\|_{L^\ast(\partial \Omega_1)}^\gamma \, dx
\]

\[
- \frac{1}{2} \int \|u\|_{L^\ast(\partial \Omega_1)}^\gamma \, dx + \int C(\text{Det} \, Vu(x)) \, dx,
\]

for \( e > 0 \) and \( d > 0 \). Choosing \( e \) and \( d \) small enough with \( d \leq k(\gamma/\mu)^\gamma \), we obtain

\[
J_0(u) \geq c + r_0 \int |Vu(x)|^\gamma \, dx + K_1 \int |\text{Adj} \, Vu(x)|^\mu \, dx + \int C(\text{Det} \, Vu(x)) \, dx,
\]

(7.18)

where \( c \) and \( r_0 > 0 \) are constants.

\* In the sense of trace.
Let \( \{u_i\} \) be a minimizing sequence for \( J_0 \). It follows from (7.18) that a subsequence \( \{u_j\} \) satisfies
\[
u_j - u_0 \in W^{1,1}(\Omega), \quad u_j - u_0 \in L^2(\partial \Omega),
\]
and \( u_j \to u_0 \) almost everywhere in \( \Omega \) and \( \partial \Omega \),
\[
\text{Adj} \, \nu_j \to \text{Adj} \, \nu_0 \quad \text{in} \quad L^2(\Omega), \quad \text{Det} \, \nu_j \to \text{Det} \, \nu_0 \quad \text{in} \quad L_c(\Omega).
\]
For \( \tilde{J} \) given by (7.10) it follows that
\[
J(u_0) \leq \liminf_{j \to \infty} J(u_j),
\]
while
\[
l_\delta \{ u_j(x) \cdot \nu_j(x) \} = \lim_{j \to \infty} l_\delta \{ u_j(x) \cdot \nu_j(x) \}.
\]
Noting that \( u_0 = \tilde{u} \) almost everywhere in \( \partial \Omega \), we see, as in the proof of Theorem 7.3, that \( u_0 \in \mathcal{A} \). The result follows.

**Remark.** In Theorem 7.6, and in the results below, the hypotheses on \( \tilde{u} \) are concealed in the assumption that \( \mathcal{A} \) is nonempty.

**Theorem 7.7.** If \( n = 2 \) let \( \gamma \geq 2, \sigma > 1 \) if \( \gamma = 2, \sigma = 1 \) if \( \gamma > 2, K_2 = 0 \). If \( n = 3 \) let \( \gamma \geq 2, \sigma = 2 < \gamma/2 \) if \( 2 \leq \gamma < 3, \sigma = 1 \) if \( \gamma = 3, \sigma = 1 \) if \( \gamma > 3, K_2 = 0 \).

Let the other hypotheses of Theorem 7.6 remain unchanged. Then Theorem 7.6 remains valid with \( \text{Adj} \, \nu_0 \), \( \text{Det} \, \nu_0 \) replaced everywhere by \( \text{Adj} \, \nu_j \), \( \text{Det} \, \nu_j \) respectively.

**Proof.** This is immediate from Lemma 6.1.

**Remark.** In Theorem 7.7, and in the results below that concern the distributions \( \text{Det} \, \nu_0 \), \( \text{Adj} \, \nu_0 \) it it only necessary for \( G \) to be defined on the set \( \{ (x,u,a) : x \in \Omega, u \in \mathbb{R}^n, a \in C^0(\mathbb{R}^{M^d}) \cap W \} \) (see Section 4 and the remark after Example 6.1).

Next we give the analogous of Theorems 7.6 and 7.7 for incompressible materials. The proof is similar to that of Theorem 7.5 and is omitted. An analogue of Theorem 7.4 may also be simply proved.

**Theorem 7.8.** Let \( n = 3 \). Let \( \Omega, \partial \Omega_1, \partial \Omega_2, \tilde{u}, \tilde{f} \) be as in Theorem 7.6. In the definitions of \( W(x) \) and \( W \) replace \( E \) by \( E_1 \). Let \( G : \mathcal{S} \to \mathcal{R} \) satisfy hypotheses \((H_1)-(H_4)\) and \((H_5)\).

Either let \( \gamma, \mu, \sigma, K_2 \) be as in Theorem 7.6 and let
\[
\mathcal{A} = \{ u \in W^{1,1}(\Omega) : \text{Adj} \, \nu \in L^2(\Omega), (\nu(x), \mathcal{A} \text{Adj} \, \nu(x)) \in W(x) \text{ almost everywhere in } \Omega, u = \tilde{u} \text{ almost everywhere in } \partial \Omega_1, \text{Det} \, \nu(x) = 1 \text{ almost everywhere in } \Omega \}
\]
and \( \sigma = \{ u \in W^{1,1}(\Omega) : \text{Adj} \, \nu \in L^2(\Omega), (\nu(x), \mathcal{A} \text{Adj} \, \nu(x)) \in W(x) \text{ almost everywhere in } \Omega, u = \tilde{u} \text{ almost everywhere in } \partial \Omega_1, \text{Det} \, \nu(x) = 1 \text{ almost everywhere in } \Omega \}\).

Suppose that there exists \( u_0 \in \mathcal{A} \) such that \( J_0(u_0) < \infty \). Then there exists \( \tilde{u} \in \mathcal{A} \) that minimizes \( J_0(\tilde{u}) \) in \( \mathcal{A} \).

Let
\[
J_0(u) = J(u) - \int_{\partial \Omega} \tilde{f}(x) \cdot \nu(x) \, dS,
\]
where \( \tilde{f} \) is given by (7.11).

**Pure traction boundary-value problems**

**Theorem 7.9 (cf. Section 1, A1).** Let \( \Omega \) satisfy a strong Lipschitz condition. Let \( \tilde{f} \in L^2(\partial \Omega) \). \( \sigma \geq 1 \). Let \( G : \mathcal{S} \to \mathcal{R} \) satisfy hypotheses \((H_1)-(H_4)\) and \((H_5)\) with \( K_2 > 0 \). Let \( K = \min(\gamma, \sigma) \).

\[
\begin{align*}
\text{sup} & = 2: \quad \gamma > \frac{3}{2}, \quad \sigma = K_2^2 \quad \text{if } 1 \leq K < 2, \quad \sigma > 1 \quad \text{if } K = 2, \quad \sigma = 1 \quad \text{if } K > 2. \\
\text{If } s = 1 \text{ let } R_{K} \quad & = \text{min} (\gamma, \sigma, K_2) \\
& = \min (s, \gamma, \sigma, K_2) \quad \text{if } K = 0 > 0 \text{ is a certain constant. Let } \\
\mathcal{A} & = \{ u \in W^{1,1}(\Omega) : \text{Adj} \, \nu \in L^2(\Omega), (\nu(x), \text{Det} \, \nu(x)) \in W(x) \text{ almost everywhere in } \Omega \}.
\end{align*}
\]

**Proof.** Let \( n = 3 \). By using the hypothesis \( K_2 > 0 \) instead of (7.17) we obtain the a priori bound
\[
J_0(u) \leq c + c_0 \| \mathcal{A} \|_{L^2(\partial \Omega)} + K_2 \int_{\partial \Omega} \| \text{Adj} \, \nu \|^2 \, d\Omega + c_1 \int_{\partial \Omega} \| \nu \|^2 \, d\Omega,
\]
where \( c_1 > 0 \). If \( s = 1 \) then \( K_0 \) is such that \( \| u \|_{W^{1,1}(\Omega)} \geq K_0 > 0 \) for all \( u \in W^{1,1}(\Omega) \). By using the Poincaré inequality (MORREY [2, p. 82]) we can complete the proof in the same way as for Theorems 6.6, 6.7.
Theorem 7.10 (cf. Section 1, A2). Let $\Omega$ satisfy a strong Lipschitz condition. Let $I_{x} \in L^{r}(\partial \Omega)$. Let $G: \mathbb{R} \to \mathbb{R}$ satisfy hypotheses $(H_{1})-(H_{2})$ with $K_{2}=0$.

If $n=2$, let $\gamma \geq \frac{4}{3}$, $\sigma = \frac{\gamma}{2}$ if $\gamma < 2$, $\sigma = 1$ if $\gamma = 2$, $\sigma = 1$ if $\gamma > 2$, $e$ be a constant vector and

$$\mathfrak{S} = \{u \in W^{1,1}(\Omega) : \text{Det} Vu \in L^{r}(\Omega), (Vu(x), \text{Det} Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, \int_{\Omega} |u(x)| \, dx = \epsilon \}.$$  

If $n=3$, let $\gamma \geq \frac{4}{3}$, $\sigma = \frac{\gamma}{2}$ if $\gamma < 3$, $\sigma = 1$ if $\gamma = 3$, $\sigma = 1$ if $\gamma > 3$, $e$ be a constant vector and

$$\mathfrak{S} = \{u \in W^{1,1}(\Omega) : \text{Adj} Vu \in L^{r}(\Omega), \text{Det} Vu \in L^{r}(\Omega), (Vu(x), \text{Adj} Vu(x), \text{Det} Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, \int_{\Omega} |u(x)| \, dx = \epsilon \}.$$  

Let $\mu_{r} \in L^{1}(\Omega)$ and let $J_{0}$ be given by

$$J_{0}(u) = J(u) - \int_{\Omega} \rho_{r}(x) \mu_{r} \cdot u(x) \, dx - \int_{\Omega} u \cdot \mathbf{t}_{r} \, dS. \quad (7.24)$$

Suppose there exists $u_{1} \in \mathfrak{S}$ with $J_{0}(u_{1}) < \infty$. Then there exists $u_{2} \in \mathfrak{S}$ that minimizes $J_{0}(u)$ in $\mathfrak{S}$.

If, in addition, we assume (7.23), then the result holds with $\text{Adj} Vu$, $\text{Det} Vu$ replaced everywhere by $\text{Adj} Vu$, $\text{Det} Vu$, respectively.

Proof. If $u \in \mathfrak{S}$, then

$$\int_{\Omega} \left[ |u(x)| - \frac{1}{m(\Omega)} \epsilon \right] \, dx = 0.$$  

Thus by a version of the Poincaré inequality (MORREY [2, p. 83]) there are constants $k_{1}, k_{2} > 0$, such that

$$\int_{\Omega} |u|^{\gamma} \, dx \leq k_{1} k_{2} \int_{\Omega} |Vu|^{r} \, dx \quad (7.25)$$

for all $u \in \mathfrak{S}$. Applying (7.25) and the simple estimate

$$\int_{\Omega} \rho_{r} b_{0} \cdot u \, dx \geq - \left[ \frac{b_{0}}{\gamma} \norm{Vu}_{L^{r}(\Omega)}^{r} + \frac{1}{\gamma} \norm{\rho_{r} b_{0}}_{L^{r}((\Omega)} \right], \quad (7.26)$$

we again obtain the bound given by $n=3$ by (7.18). The rest of the proof follows the usual pattern. \[\square\]

Remark. We re-emphasise (cf. Section 1, A2 and Theorem 7.13) that for non-linear elasticity, when $G$ is independent of $u$, it is necessary to impose the extra condition (1.31) in order to show that $u_{0}$ is in some sense a solution of the equilibrium equations.

Mixed displacement pressure boundary-value problems

For these problems we restrict our attention to the case $n=3$. The case $n=2$ can be treated similarly.

Theorem 7.11 (cf. Section 1B). Let $n=3$. Let $\Omega$ satisfy a strong Lipschitz condition. Let $\partial \Omega = \partial \Omega_{1} \cup \Sigma$, $\partial \Omega_{1} \cap \Sigma = \emptyset$, with $\partial \Omega_{1}$ having positive measure as a subset of $\partial \Omega$. Let $G: \mathbb{R} \to \mathbb{R}$ be measurable, and let $p \in W^{1,1}(\Omega)$. Let $G: \mathbb{R} \to \mathbb{R}$ satisfy hypotheses $(H_{1})-(H_{2})$ and $(H_{3})$. Let $K_{2}=0$. If $p(x) \equiv \text{constant almost everywhere}$, let $\gamma \geq \frac{4}{3}$, $\sigma = \frac{\gamma}{2}$ if $\gamma < 3$, $\sigma = 1$ if $\gamma = 3$, $\sigma = 1$ if $\gamma > 3$, $e$ be a constant vector and

$$\mathfrak{S} = \{u \in W^{1,1}(\Omega) : \text{Adj} Vu \in L^{r}(\Omega), \text{Det} Vu \in L^{r}(\Omega), (Vu(x), \text{Adj} Vu(x), \text{Det} Vu(x)) \in W(x) \text{ almost everywhere in } \Omega, \int_{\Omega} |u(x)| \, dx = \epsilon \}.$$  

Let $J_{1}(u) = J(u) + P(u)$. \quad (7.27)

where (cf. 1.35))

$$P(u) \triangleq \int_{\Omega} \left[ p \text{ Det} Vu + \frac{1}{2} p \left( \text{Adj} Vu \right)^{\mu} \right] \, dx. \quad (7.28)$$

Suppose that there exists $u_{1} \in \mathfrak{S}$ with $J_{1}(u_{1}) < \infty$. Then there exists $u_{2} \in \mathfrak{S}$ that minimizes $J_{1}(u)$ in $\mathfrak{S}$.

If, in addition

$$\gamma \geq 2, \quad \frac{1}{\gamma} + \frac{1}{\mu} \leq 1, \quad (7.29)$$

then the result holds with $\text{Adj} Vu$, $\text{Det} Vu$ replaced everywhere by $\text{Adj} Vu$, $\text{Det} Vu$, respectively.

Proof. It suffices to establish the bound

$$J_{1}(u) \leq c_{1} + \int_{\Omega} |\text{Det} Vu|^{\gamma} \, dx + \frac{K_{1}}{\gamma} \int_{\Omega} C(\text{Det} Vu) \, dx \quad (7.30)$$

for all $u \in \mathfrak{S}$, where $c_{1} > 0$ are constants.

We use the Sobolev inequality

$$\norm{u}_{W^{1,1}(\Omega)} \geq k_{2} \norm{u}_{L^{r}(\Omega)} \quad \text{for all } u \in W^{1,1}(\Omega), \quad (7.31)$$

where $k_{2} > 0$ and $r = 3 \gamma / 3 - \gamma$ if $\gamma < 3$, $1 < r < \infty$ if $\gamma = 3$, $\infty$ if $\gamma > 3$.

* In applications we shall have $\Sigma = \bigcup_{i=1}^{r} \partial \Omega_{i}$ with the $\partial \Omega_{i}$ as in Section 1B.
For $u \in \mathcal{A}$, we can use (7.17), (7.31) to obtain

\[ J_\varepsilon(u) \geq \frac{K_1}{2} \int \left( \frac{1}{\varepsilon} + K_1 \frac{|u|^2}{2} \right) dx + K_1 \int |\text{Adj} Fu|^2 dx + C(\text{Det} Fu) dx \]

\[ + \left( \frac{K_1}{2} - \varepsilon \right) \int |u|^2 dx - K_1 \left( \int b(x) dx + eK_2 \|u\|_{L^2(\Omega)} \right) \]

\[ - J \varepsilon \|p\|_{L^2(\Omega)} \int |\text{Det} Fu| dx - \frac{3}{K_1} \int \|p\|_{L^2(\Omega)} \left[ \frac{d^u}{\mu} (\text{Adj} Fu) dx + \frac{1}{\varepsilon} \|u\|_{L^2(\Omega)} \right] \]

where $\varepsilon > 0$, $d > 0$.

If $\rho$ is constant, then clearly (7.30) follows. If $\rho$ is not constant and $\gamma > \mu$ so that we obtain (7.30) by choosing $\varepsilon$ and $d$ small enough. If $\rho$ is not constant and $\gamma = \mu$ and (7.30) follows similarly. □

The above proof is valid if $\varepsilon$ is taken to be zero, but the value of $k_3$ so obtained is then smaller. There is no difficulty in giving the analogous results to Theorem 7.11 for the pure pressure boundary-value problem and for incompressible materials.

Solutions to the equilibrium equations

We now turn to the question of whether the minimizers whose existence we have established satisfy the corresponding Euler-Lagrange equations. There is at present no available regularity theory for our problems under acceptable hypotheses, and we therefore confine our discussion to whether the minimizers are weak solutions: Unfortunately there are technical problems associated with the two most important cases, namely (i) when the material is compressible and $W(x)$ is given by

\[ W(x) = \{ a \in E : \rho > 0 \} \]

and (ii) when the material is incompressible.

We therefore consider the simpler situation when $W(x) = E$ for all $x$, so that there is no local constraint:

We replace (H3) by the following stronger hypotheses on $G$:

(H3) $n = 2$: $G$ is continuously differentiable in $u, F, H, \rho$ for all $x \in \Omega$, and there exist real constants $K_1 > 0$, $K_2 \geq 0$, $B_1 > 0$, $C_1 > 0$, $\rho_i \geq 0$, $1 < \gamma \leq 7$, $\tau \geq 0$, $s \geq 1$, and a function $b \in L^2(\Omega)$ such that

\[ G(x, u, F, H, \rho) \geq b(x) + K_1 |F|^2 + K_2 |u|^2 + B_1 |\rho|^s, \]

\[ \frac{\partial G}{\partial u} \geq \rho_1 + C_1 [u]^F + |H|^\tau + |\rho|^s, \]

\[ \frac{\partial G}{\partial F} \geq \rho_2 + C_2 [u]^F + |F|^\tau + |\rho|^s. \]

for all values of the arguments. If $\gamma < 2$, we assume further that $\tau = \frac{2 \gamma}{2 - \gamma}$.

Let $f$ be given by

\[ n = 2: f(x, u, F) = G(x, u, F, \text{det} F) \]

\[ n = 3: f(x, u, F) = G(x, u, F, \text{adj} F, \text{det} F) \]

Two typical results are the following:

**Theorem 6.12.** In the hypotheses of Theorems 7.7, 7.9 replace (H3) by (H3). Suppose also that (cf. (7.23))

If $n = 2$: $\gamma \geq 2$,

If $n = 3$: $\gamma \geq 2$, $\frac{1}{\gamma} + \frac{1}{\mu} \leq 1$.

Then the minimizing functions $u = u_0$ satisfy the Euler-Lagrange equation

\[ \int \left[ \frac{\partial f}{\partial u} + \frac{\partial f}{\partial u_0} \right] dx = \int v \cdot T_k dS \]

for all $v \in C_0^\infty(\mathbb{R}^3)$ with $v |_{\partial \Omega} = 0$.

**Proof.** We give the proof for the case in which $n = 3$. Fix $v \in C_0^\infty(\mathbb{R}^3)$ with $v |_{\partial \Omega} = 0$. Since $1 < \gamma \leq 7$ it follows that $u_0 + \varepsilon v \in \mathcal{A}$ for any $\varepsilon$. Also we clearly have

\[ \frac{d}{d \varepsilon} \left\{ \int (u_0 + \varepsilon v) \cdot T_k dS \right\} = \int v \cdot T_k dS. \]
It thus suffices to show that
\[ \frac{d}{d\varepsilon} f(x, u(x) + \varepsilon v(x), F_u(x) + \varepsilon F_v(x)) \bigg|_{\varepsilon=0} \leq \Theta(x), \]  
for fixed \( u, v \in \mathcal{A} \) with \( I(u) < \infty \), and for all \( \varepsilon \in (0,1) \). Here and below \( c \) denotes a constant. The estimate (7.46) now follows from (7.37)-(7.40) and Hölder's inequality. For brevity we display the calculation only for the last term of (7.47).

\[ I \left( \int \rho \left[ \frac{\partial \psi}{\partial u} \cdot \frac{\partial \psi}{\partial u} + |\nabla \psi|^2 + |\nabla F_u|^2 + |\nabla F_v|^2 \right] \right) \leq c \left[ \rho_k + |\partial \psi|^2 + |\nabla F_u|^2 + |\nabla F_v|^2 \right] \leq \Theta(x), \]  
for some \( \Theta \in L^2(\Omega) \), where we have used (7.36) and the facts that if \( \gamma \geq 3 \) then \( \varepsilon \in L^2(\Omega) \) while if \( 2 \leq \gamma < 3 \) then \( \varepsilon \in L^2(\Omega) \).

Theorem 7.13 (cf. Section 1, A2). In the hypotheses of Theorem 7.10 replace (H_a) by (H_b) and (7.42) hold, and let \( f = \psi \cdot (F, H) \) be independent of \( u \) so that
\[ \int_\Omega \psi'(x, F_u(x)) \, dx - \int_\Omega \psi'(x, F_u(x)) \, dx = \int_{\partial \Omega} \psi \, dS. \]  
Suppose in addition that
\[ \int_{\partial \Omega} \psi \, dS = 0, \]  
Then the minimizing function \( u = u_0 \) satisfies the Euler-Lagrange equations
\[ \int_\Omega \left[ \rho_k \, b_0 \cdot \nabla \psi + \frac{\partial \psi}{\partial \nabla u} \cdot \psi \right] \, dx + \int_{\partial \Omega} \psi \, dS = 0 \]  
for all \( \psi \in C^0(\partial\Omega) \).

Existence under other hypotheses

We next give a sample theorem under the hypothesis (7.45), for the displacement-boundary-value problem for a homogeneous material. This result almost certainly has generalizations to integrands with \( x, u \) dependence.

Theorem 7.14. Let \( n = 3 \). Let \( U \subseteq M^{3 \times 3} \) be such that \( C_0T(U) \) is open. Let \( u_1 \in W^{1,1}(\Omega), \) adj \( F_u_1 \in L^1(\Omega), \) det \( F_u_1 \in L^1(\Omega), \) for all \( (F, H, \delta) \in C_0T(U), \) where \( \gamma \geq 2, \mu > 1, \frac{1}{\gamma} + \frac{1}{\mu} < 1, \nu > 1. \)

In the notation of Definition 4.3 let \( g: U \to \mathbb{R} \) satisfy
\[ G(F, H, \delta) \leq h + K_1 |F|^\gamma + |H|^\nu + |\delta|^\gamma \]  
for all \( (F, H, \delta) \in C_0T(U) \) where \( h, K_1 > 0 \) are constants, and
\[ G(F, H, \delta) \to \infty \]  
a.s. \( (F, H, \delta) \to \partial(C_0T(U)) \)

Let \( \mathcal{A} = \{ u \in W^{1,1}(\Omega): \nabla - u_1 \in W^{1,1}(\Omega), \) adj \( F_u \in L^1(\Omega), \) det \( F_u \in L^1(\Omega), \Sigma(u) \in C_0T(U) \} \) almost everywhere in \( \Omega \).

Define for \( u \in \mathcal{A} \)
\[ I(u) = \int_\Omega g(F_u) \, dx. \]  
Then if \( I(u) < \infty \), there exists \( u_0 \in \mathcal{A} \) that minimizes \( I(u) \) in \( \mathcal{A} \).

Proof. Let \( (u_n) \) be a minimizing sequence. By (7.53), \( \{ \Sigma(u_n) \} \) is bounded in the reflexive Banach space \( \mathcal{B} = L^2(\Omega)^\gamma \times L^1(\Omega)^\nu \times L^1(\Omega) \). By the Poincaré inequality
and our standard arguments there exists a subsequence \( \{u_j\} \) such that
\[
\begin{align*}
u_j - u_0 \quad \text{in} \quad W^{1,1}(\Omega), \\
\Sigma(u_j) - \Sigma(u_0) \quad \text{in} \quad \mathcal{X}^n, \subseteq \mathcal{B}.
\end{align*}
\]
with \( u_0 \in \mathcal{X}^n \). Define \( G_i : \mathcal{E} \to \mathcal{B} \) by
\[
G_i(F, H, \delta) = G(F, H, \delta) \quad \text{if} \quad (F, H, \delta) \in \text{Co } T(U)
\]
\[= \infty \quad \text{otherwise}.
\]
Let
\[
J(\sigma) = \int G_i(\sigma(x)) \, dx.
\]
(7.56)
Then \( J : \mathcal{X}^n \to \mathcal{B} \). Since \( g \) satisfies \( P_{n,n} \) at \( u_0 \) it follows that \( J \) is convex. Applying Fatou's lemma to the integrand
\[
G_i(F, H, \delta) - b - K_i(\|F\| + \|H\| + \|\delta\|)
\]
we see that \( J \) is (strongly) lower semicontinuous. Hence (cf. EKELAND & TEMAM [1, p. 33]) \( J \) is sequentially weakly lower semicontinuous. Thus
\[
J(\Sigma(u_0)) \leq \liminf_{j \to \infty} J(\Sigma(u_j)),
\]
and the proof is complete. \( \square \)

The most general integrands for which our methods establish the existence of minimizers are given by the sum of polyconvex functions satisfying a suitable subset of hypotheses \( (H_1)-(H_9) \) quasiconvex functions satisfying the hypotheses of Theorem 6.1, and, where appropriate, functions satisfying condition \( P_{n,n} \) as in the above theorem. By suitably combining the growth conditions of each of the terms in this sum various existence theorems may be given. At present both the scarcity of examples of quasiconvex functions that are not polyconvex and the abundance of physically useful polyconvex functions make these theorems of little interest. We therefore leave the routine formulation of the results to the reader.

8. Applications to Specific Models of Elastic Materials

Many forms of the stored-energy function have been proposed for nonlinear elastic materials, particularly for various rubbers. An excellent review of the literature can be found in the papers of ODGREN [2, 3]. We now examine the extent to which these models satisfy the hypotheses of our existence theorems. By concentrating on a certain class of models below we do not mean to imply that other models are inferior for empirical reasons. Neither should the omission of a model from the discussion be construed as suggesting that it fails to satisfy our existence hypotheses. Our purpose is simply to indicate the flexibility of the hypotheses to serve for a variety of stored-energy functions and also to discuss the position occupied with respect to these hypotheses by certain well known models.

We assume that \( n=3 \) unless otherwise stated, and consider for simplicity only isotropic materials, for which the stored-energy function has the form (see (1.15))
\[
W^*(x, F) = \Phi(x, v_1, v_2, v_3).
\]
(8.1)

For compressible materials, in the case when the local constraint set \( W(x) \) has the form
\[
W(x) = E_1 \times K(x),
\]
(8.2)
with \( K(x) \subseteq \mathcal{B} \) nonempty, open and convex, it follows from Theorem 4.3 that \( \text{Co } U(x) = W(x) \), where \( U(x) = \{F \in M^{3 \times 3} : \text{det } F \in K(x) \} \). Necessary and sufficient conditions for \( \Psi(x, \cdot) \) to be polyconvex on \( U(x) \) are given by Theorem 4.4. For simplicity we shall assume in the compressible case that \( K(x) = \{\text{det } F > 0\} \) (continuity condition), while in the incompressible case \( W(x) = E_1 \) (no local constraint).

We consider a modification of a class of stored-energy functions introduced by ODGREN [2, 3]. For \( a \geq 1 \) let
\[
\psi(x) = \frac{1}{a} + 3 \geq 4, \quad \chi(a) = (v_1 v_2 v_3) + (v_1 v_3 v_2) + (v_1 v_2 v_3) - 3.
\]
(8.3)

Consider first incompressible materials, and let
\[
\Psi^*(x, F) = B(x) + \sum_{l=1}^m a_l(x) \psi(x) + \sum_{l=1}^n c_l(x) \chi(\beta_l),
\]
(8.4)
where \( a_1 \geq \cdots \geq a_m \geq 1, \beta_1 \geq \cdots \geq \beta_n \geq 1, \) and where \( B, a_l, c_l \) are functions in \( L^1(\Omega) \) satisfying
\[
a_l(x) \geq k > 0, \quad c_l(x) \geq k > 0, \quad \text{for almost all } x \in \Omega
\]
(8.5)
for some constant \( k \).

By Theorem 5.2 \( \Psi^*(x, F) \) is polyconvex on \( U(x) \). Since \( v_1^2 + v_2^2 + v_3^2 \) is a continuous function of \( F^* \) it follows that
\[
\psi^* \geq \psi^{**} \geq d(\alpha) |F^*|^{1/2}
\]
(8.6)
for some constant \( d(\alpha) > 0 \). Similarly
\[
(v_1 v_2 v_3) + (v_1 v_3 v_2) + (v_1 v_2 v_3) \leq e(\alpha) |\text{adj } F^*|^{1/2}, \quad e(\alpha) > 0.
\]
(8.7)
It follows from (8.4)-(8.7) that \( \Psi^* \) considered as a function of \( x, F, \text{adj } F \) satisfies hypotheses \( (H_1)-(H_9) \), and \( (H_9) \) with
\[
\gamma = a_1, \quad \mu = \beta_1, \quad K_2 = 0.
\]
(8.8)
Thus if
\[
a_1 \geq \frac{1}{2}, \quad \frac{1}{a_1} + \frac{1}{\beta_1} \leq 3,
\]
(8.9)
we obtain from Theorems 7.8 and the analogues of Theorems 7.9-7.11 for incompressible materials the existence of minimizers for the various boundary-value problems in terms of the distributions \( F_u, \text{adj } F_u \). These minimizers satisfy the incompressibility condition \( \text{det } F_u = 1 \) almost everywhere. Note that in order to obtain existence for variable pressures \( p \) in the mixed displacement-pressure problem we require (Theorem 7.11) either
\[
\frac{1}{a_1} + \frac{1}{\beta_1} < 1
\]
(8.10)
* Because, for example, it is a finite-valued convex function.
Then the modified stored-energy function satisfies \((H_1)-(H_4)\) and \((H_5)\) with \(\gamma = x_1\), \(\mu = \beta_1, K_2 = 0\), and thus satisfies the hypotheses of our existence theorems under the conditions \((8.9)-(8.12)\).

It is clear that a wide variety of stored-energy functions having the form \((5.11)\) (with \(x\)-dependence if necessary) can be treated by our theory in an way analogous to that for the models discussed above. We end this section by exhibiting such a stored-energy function, which satisfies the hypotheses of Theorem 7.3 but not those of Theorem 7.6, etc. (for \(\partial \Omega = \emptyset\), and thus requires the Orlicz-Sobolev space apparatus. Our example is of a stored-energy function with slow growth. For functions of very fast growth the Orlicz-Sobolev space setting would also be necessary for any proof that the Euler-Lagrange equations are satisfied. We need two lemmas:

**Lemma 8.1.** Let \(C, D\) be \(N\)-functions. Then

\[
\frac{C^{-1}(t)}{D^{-1}(t)} \to 0 \quad \text{as} \quad t \to \infty
\]  

if and only if \(D \leq C\).

**Proof.** We just prove the 'only if' part, the 'if' part being easier. Set \(s = C(t)\) for \(\lambda = 0\). Then by \((8.18)\) \(\lambda t/D^{-1}(C(t)) \to 0\) as \(t \to \infty\). Hence \(k \leq D^{-1}(C(t))\) for \(t\) large enough. By the convexity of \(D\) we have for \(t\) large enough

\[
\frac{C(t)}{D^{-1}(C(t))} \to 0 \quad \text{as} \quad t \to \infty.
\]

**Lemma 8.2.** Let \(g, k\) be non-negative continuous functions on \(\mathbb{R}_+\) such that \(k(s) \to 0\) as \(s \to \infty\) and such that

\[
\int_0^\infty g(t) dt = \infty.
\]

Let

\[
\int_0^\infty \int_0^\infty 0(s) = \int_0^\infty g(t) dt
\]

Then \(\theta(s) \to 0\) as \(s \to \infty\).

**Proof.** This follows immediately from L'Hôpital's rule. \(\square\)

**Existence Theorems in Non-Linear Elasticity**

For compressible materials Ogden [3] considered the effect of adding a term \(\Gamma(\det F)\) to \((8.4)\) \([\dagger]\). Suppose that

\[
\Gamma(t) \geq C(t) \quad \text{for all} \quad t > 0, \quad \text{where} \quad C \quad \text{is an \(N\)-function},
\]

\[
\Gamma \quad \text{is convex on} \quad (0, \infty), \quad \Gamma(\infty) \to \infty \quad \text{as} \quad t \to 0+.
\]

\[\dagger\]

For related experimental work see Treloar [1].

\[\dagger\] Outer values of the constants \(B, \alpha, \beta, \mu, \nu, \gamma, \) but for incompressible materials replaced the term \(c_1 x_2(2)\) by \(c_1 x_1^2 x_2^2 + 2 x_1 x_2 + 3 \delta,\) these terms are identical if \(x_1 x_2 = 1\), and since \(x_1 x_2 \gamma\) is in practice very close to 0 the alteration is insignificant for experimental correlations.
where $r$ is as in (8.17) and $a, c$ are positive constants. Since $A < A^*$, let $B(t) = t^4$. Then $B < A$ so that $\tilde{A} < B^* = B^*$. It is therefore sufficient to prove that $B^* < A^*$. We have that
\begin{equation}
(8.21)
g_A(t) = \frac{A \cdot \gamma(t)}{t^8} = \frac{k(t)}{t^8},
\end{equation}
where
\begin{equation}
k(t) = \frac{1}{(\log A^{-1}(t))} \to 0 \text{ as } t \to \infty.
\end{equation}
Also $g_B(t) = \frac{1}{t^8}$, and $\int_0 g_A(t) \, dt = \int_0 g_B(t) \, dt = \infty$. Therefore
\begin{equation}
(8.22)
A^* < \frac{1}{B^*} = \frac{1}{A^*}.
\end{equation}
By Lemmas 8.1 and 8.2 we deduce $B^* < A^*$ as required.

9. An Example of Nonuniqueness; Buckling of a Rod

In this section we establish nonuniqueness for the mixed displacement zero traction boundary-value problem corresponding to buckling of a rod of a homogeneous, incompressible, Mooney-Rivlin material having uniform cross-section. We do this by exhibiting an admissible displacement field with total energy lower than that of the trivial solution, and then applying Theorem 7.8 to ensure the existence of a nontrivial minimizer for the total energy. Under suitable conditions a similar but more complex analysis can be carried out for incompressible rods consisting of material not of Mooney-Rivlin type. The extension to compressible materials, however, is not so easy because an explicit trivial solution is not available.

In the stress-free reference configuration the rod occupies the region
\begin{equation}
Q = D \times (0, 1),
\end{equation}
where the cross-section $D$ is a nonempty bounded open set in $\mathbb{R}^2$ satisfying a strong Lipschitz condition. We suppose that the density in the reference configuration is a constant $\rho > 0$, and that the plane $x_2 = 0$ contains the line of centroids of the rod in the reference configuration, so that
\begin{equation}
\int_0 \xi \, dS = 0, \quad dS = dx_1 \, dx_2.
\end{equation}
Let $\tilde{\Omega}_1 = \tilde{D} \times (0, 1), \tilde{\Omega}_2 = \tilde{D} \times (0, 1)$. Let $\Omega = \int_0 \tilde{dS}$ be the area of $D$. For $\lambda > 0$ we consider the equilibrium mixed boundary-value problem with boundary conditions
\begin{equation}
u = \bar{u} \quad \text{on } \tilde{\Omega}_1,
\end{equation}
\begin{equation}
t = 0 \quad \text{on } \tilde{\Omega}_2.
\end{equation}

The stored-energy function has the form (cf. 8.11)
\begin{equation}
\mathcal{W}(F) = a (I_{22} - 3) + c (I_{11} - 3),
\end{equation}
where $a, c > 0$ are constants. In the notation of Theorem 7.8 we set
\begin{equation}
\gamma = \mu = 2, \quad K = 0, \quad \mathcal{W}(x) = E_1 = M^{1x3} \times M^{1x3},
\end{equation}
\begin{equation}
\mathcal{A} = \{ u \in W^{1,2}(Q) : \text{adj } F u \in L^2(Q), u = \bar{u} \text{ almost everywhere in } \partial \tilde{\Omega}_1, \det F u = 1 \text{ almost everywhere in } \Omega \},
\end{equation}
\begin{equation}
\mathcal{J}(u) = \int \mathcal{W}(F u(x)) \, dx.
\end{equation}
From (9.4) we obtain
\begin{equation}
\mathcal{F} \bar{u} = \text{diag}(\lambda - 1, - \lambda, - 1), \quad \det \mathcal{F} \bar{u} = 1,
\end{equation}
and
\begin{equation}
\mathcal{J}(u) < \mathcal{J}(\bar{u}) \quad \text{for a suitable hydrostatic pressure.}
\end{equation}
It is easily shown that $\bar{u}$ satisfies the equilibrium equations and boundary conditions for a suitable hydrostatic pressure. By Theorem 7.8 (with $u = \bar{u}$) there exists $u_0 \in \mathcal{A}$ that minimizes $\mathcal{J}(u)$ in $\mathcal{A}$.

We perform this construction in a manner reminiscent of derivations of rod theories in engineering. Let $y_0, \theta_0$ be real valued functions with $y_0 \in C^2([0, 1]), \theta_0 \in C^2([0, 1])$ and
\begin{equation}
y_0(0) = y_0(1) = \theta_0(0) = \theta_0(1) = \theta_0(1) = 0.
\end{equation}
Let $\varepsilon > 0$, $\varepsilon \to 0$, and define
\begin{equation}
u_{\varepsilon}(x) = \frac{y_0(x_1, x_2, x_3)}{\varepsilon} + \lambda - 1 \varepsilon_1 \cos \theta(x_3) + \epsilon \varepsilon_1 - \lambda - 1 \varepsilon_1 \sin \theta(x_3),
\end{equation}
where $g$ is a function to be chosen. $u_\varepsilon$ represents a deformation in which points $(x_0, x_1, x_2)$ are mapped to $(0, y(x_1, x_2), x_3)$, and in which cross-sections normal to the $x_3$-axis in the reference configuration stay plane and are so inclined that their normals remain parallel to the $x_2x_3$ plane and make an angle $\theta(x_3)$ with the $x_3$-axis (Fig. 3).

From (9.12) we obtain
\begin{equation}
\mathcal{F} u_{\varepsilon} = \begin{pmatrix}
\varepsilon \delta_{11} & \varepsilon_1 & \varepsilon \delta_{13} \\
0 & \lambda - 1 \cos \theta & \varepsilon_1 - \lambda - 1 \varepsilon_1 \sin \theta(x_3) \\
0 & - \lambda - 1 \sin \theta(x_3) & \lambda - 1 \varepsilon_1 \cos \theta + \epsilon \varepsilon_1 \sin \theta(x_3)
\end{pmatrix}
\end{equation}

* Note that we do not have to satisfy the zero traction condition on $\partial \Omega_1$ because our existence theorem incorporates this as a natural boundary condition.
Therefore \[
\det P u_* = g [x^1 \cos \theta + x^2 \sin \theta - x^1 x^2 \theta^0].
\] (9.14)

Since \(\lambda > 0\) the expression in brackets in (9.14) is positive for all \(x \in \Omega\) if \(\epsilon\) is small enough. Thus we may choose
\[
g = [x^1 \cos \theta + x^2 \sin \theta - x^1 x^2 \theta^0]^{-1},
\]
whence
\[
\det P u_* = 1 \quad \text{for all} \quad x \in \Omega.
\] (9.16)

It follows from (9.11)-(9.16) that \(u_* \in \mathcal{A}\) for \(\epsilon\) small enough. Let
\[
B = P u_* V u_*^T.
\] (9.17)

A routine but tedious calculation shows that
\[
B_1 = \lambda^{-1} + 2 \lambda^{-1} x_3 \theta_0^0 + \epsilon^2 [x^1 \lambda^{-2} \theta_0^0 - 2 \lambda^{-2} \theta_0 y_0 + (3 x_3^2 + x_1^2) \lambda^{-4} \theta_0^2 + \lambda^{-4} x_2^2 x_1^2 \theta_0^2] + o(\epsilon^2),
\]
\[
B_2 = \lambda^{-1} + \epsilon^2 [x_2^2 \lambda^{-1} \theta_0^0 - 2 \lambda^{-2} \theta_0 y_0 + (3 x_3^2 + x_1^2) \lambda^{-4} \theta_0^2 + \lambda^{-4} x_2^2 x_1^2 \theta_0^2] + o(\epsilon^2),
\]
and hence that
\[
I_0 = 2 \lambda^{-1} + 2 \epsilon x_2 \theta_0^0 (x_1 - x_2) + \epsilon^2 [x^1 \lambda^{-2} \theta_0^0 + (3 x_3^2 + x_1^2) \lambda^{-4} + \lambda^{-4} x_2^2 x_1^2] \theta_0^2 + x_2^2 x_1^2 \lambda^{-4} \theta_0^0 - 2 \lambda^{-2} \theta_0 y_0 + o(\epsilon^2),
\]
\[
I_1 = \lambda^{-2} + 2 \epsilon x_2 \theta_0^0 (x_1 - x_2) + \epsilon^2 [x^1 \lambda^{-2} \theta_0^0 + (3 x_3^2 + x_1^2) \lambda^{-4} + \lambda^{-4} x_2^2 x_1^2] \theta_0^2 + x_2^2 x_1^2 \lambda^{-4} \theta_0^0 - 2 \lambda^{-2} \theta_0 y_0 + o(\epsilon^2),
\]
(9.19)

where \(o(\epsilon)\) denotes a function of \(\epsilon\) such that \(o(\epsilon)/\epsilon\to 0\) uniformly in \(\Omega\) as \(\epsilon\to 0\).

From (9.2), (9.3), (9.7) we obtain (with the standard meaning for \(o(\epsilon^2)\))
\[
\bar{J}(u_*^0) = \bar{J}(u) + \epsilon^2 \left[ d_0 \theta_0^0 + d_1 \theta_0^2 + d_2 \theta_0^2 - d_3 \theta_0 y_0 + d_4 y_0 y_2 \right] d x_3 + o(\epsilon^2),
\] (9.20)

where
\[
d_0 = A \lambda^{-2} (a \lambda + c),
\]
\[
d_1 = a \left[ (3 k_3 + k_4) \lambda^{-4} + k_2 \lambda^{-1} \right] + c \left[ (k_1 + k_2) \lambda^{-2} + 3 k_3 \lambda^{-3} \right],
\]
\[
d_2 = k_3 \lambda^{-3} (a \lambda + c),
\]
\[
d_3 = 2 A \lambda^{-3} (a \lambda + c),
\]
\[
d_4 = A \lambda^{-1} (a \lambda + c),
\]
and
\[
k_1 = \int \frac{x_3^2}{d S}, \quad k_2 = \int \frac{x_3^2}{d S}, \quad k_3 = \int \frac{x_3^2}{d S}.
\] (9.22)

The Euler-Lagrange equations corresponding to the quadratic part of (9.20) are
\[
2 d_4 \theta_0^0 - d_4 \theta_0^2 = \gamma,
\] (9.23)
\[
d_2 \theta_0^0 - d_2 \theta_0^2 + d_0 \theta_0 - \frac{d_3}{k_3} \theta_0 = 0,
\] (9.24)
where \(\gamma\) is a constant. Setting \(\gamma = 0\) and combining (9.23), (9.24) we obtain
\[
\theta_0'' - \alpha_1 \theta_0 + \alpha_2 \theta_0 = 0,
\] (9.25)
where
\[
\alpha_1 = \frac{d_2}{d_4}, \quad \alpha_2 = \frac{d_0 - 2 \lambda^3/4 d_4}{d_2} A \lambda^{3-1}/k_3.
\] (9.26)

With the above as motivation we seek a solution, antisymmetric about \(x_3 = \frac{1}{2}\), to the equation
\[
\theta_0'' = \theta_0 - \alpha_2 \theta_0 = 0
\] (9.27)
subject to boundary conditions
\[
\theta_0 = \theta_0'' = 0 \quad \text{at} \quad x_3 = 0, 1,
\] (9.28)
for some \(\alpha_2 > \alpha_1\). If \(\theta_0\) is such a solution and
\[
y_0(x_3) = \frac{d_3}{2 d_4} \int_0^{x_3} \theta_0(s) d s,
\] (9.29)
then \(y_0\) satisfies (9.11) and we have from (9.20), (9.26)-(9.29) that
\[
\bar{J}(u_*) = \bar{J}(u) + \epsilon^2 d_4 \left[ \int \frac{\theta_0}{d} d x_3 + o(\epsilon^2) \right].
\] (9.30)

It follows that for \(\epsilon\) small enough \(\bar{J}(u_*) < \bar{J}(u)\) as required.

To solve (9.27), (9.28) we first note that for all \(\lambda \in [0, 1]\) we have
\[
\alpha_1 \geq \frac{3 \epsilon k_1}{k_3 (a \lambda + c)} > 0, \quad \alpha_2 \leq 0.
\] (9.31)

Set
\[
\bar{\alpha}_2 = \alpha_2 + \tau, \quad 0 < \tau < \frac{3 \epsilon k_1}{k_3 (a \lambda + c)}.
\] (9.32)
Existence Theorems in Non-Linear Elasticity

It is not hard to show that \( \kappa \) and \( \mu \) are continuous functions of \( \lambda \) in \([0, 1]\), and that \( \kappa \) is not constant in \([0, 1]\). Therefore there exists an interval in the range of \( \kappa : [0, 1] \to \mathbb{R}_+ \) with length \( \delta > 0 \). If \( \kappa_0 + \kappa_1 \) then we may take

\[
\delta = |\kappa_0 - \kappa_1| + r(t),
\]

where \( r(t) \to 0 \) as \( t \to 0 \).

From Figure 4 it is thus clear that if

\[
\frac{\delta l}{2} > 2\pi,
\]

then there exists \( 0 < \lambda < 1 \) such that (9.36) is satisfied. The corresponding \( \theta_0 \) is the function required. To satisfy (9.39) one need only choose \( l \) large enough. Thus sufficiently long rods of arbitrary cross-section will exhibit nonuniqueness for some \( \lambda \in (0, 1) \). (An obvious refinement of this argument shows that if \( 0 < \lambda_0 < 1 \) then for \( l \) large enough there will be nonuniqueness for some \( \lambda \) with \( \lambda_0 < \lambda < 1 \).)

Usually \( \kappa_0 \neq \kappa_1 \) so that by choosing \( t > 0 \) small enough we get nonuniqueness for some \( \lambda \in (0, 1) \) whenever

\[
|\kappa_0 - \kappa_1| > \frac{4\pi}{l}.
\]

This is a condition expressed entirely in terms of \( l \) and the cross-sectional parameters \( k_1, k_2, k_3 \) and \( A \).

Example. Let \( D \) be the disc \( x_1^2 + x_2^2 < a^2 \). Then

\[
k_1 = k_2 = \frac{\pi a^4}{4}, \quad k_3 = \frac{\pi a^6}{24}, \quad A = \pi a^2
\]

so that

\[
\kappa_0 = \frac{1}{a} \sqrt{9 + \sqrt{105}}, \quad \kappa_1 = \frac{1}{a} \sqrt{30}.
\]

Condition (9.40) therefore becomes

\[
\frac{a}{l} < 0.9, \quad v < 0.09.
\]

The condition (9.40) is somewhat crude; indeed it is possible that no such condition is necessary. Improved estimates, and lower bounds on the supremum of \( 0 < \lambda < 1 \) for which nonuniqueness occurs, may be obtained by more detailed calculations based on Figure 4. I have not included these results since they are messy and since my method is severely limited in scope due to the type of trial deformation considered in (9.12). Even in situations where we envisage nonuniqueness occurring by Euler buckling, the deformation of cross-sections implied by (9.12) is unrealistic.

There are numerous formal stability calculations in the literature for rods in tension or compression, and for other problems in three dimensional nonlinear elasticity. For the most part these calculations are based on the theory of small deformations superposed upon large; the status of this theory with respect to

The indicial equation for (9.27) is

\[
m^4 - 2m^2 + 3 = 0
\]

with roots \( m = \pm \kappa, \pm i\mu \), where

\[
\kappa^2 = \frac{a_1^2 + \sqrt{a_1^4 - 4a_2^2}}{2}, \quad \mu^2 = \frac{\sqrt{a_1^4 - 4a_2^2}}{2},
\]

and \( \kappa > \mu > 0 \). \( \kappa, \mu \) are real by (9.31), (9.32). Hence

\[
\theta_0(x_3) = A_4 \sinh \kappa \left( x_3 - \frac{l}{2} \right) + A_4 \sin \mu \left( x_3 - \frac{l}{2} \right)
\]

will satisfy (9.27), (9.28) provided that

\[
\frac{2}{\mu} \tan \frac{\mu l}{2} = \frac{2}{\kappa l} \tanh \frac{\kappa l}{2}.
\]

Also \( \theta_0 \) given by (9.35) is antisymmetric about \( x_3 = \frac{l}{2} \). The graphs of the functions \( \zeta = \frac{1}{\eta} \tan \eta, \zeta = \frac{1}{\eta} \tanh \eta \) are sketched in Figure 4. It is easy to prove the indicated monotonicity properties.

Let \( \kappa_0, \kappa_1 \) be the values of \( \frac{a_1^2 + \sqrt{a_1^4 - 4a_2^2}}{2} \) at \( \lambda = 0, 1 \), respectively, so that

\[
\kappa_0 = \frac{3k_2}{2k_3} + \sqrt{\frac{9k_2^2 + 4k_3^2}{4k_3^2}}, \quad \kappa_1 = \left( \frac{4k_2 + k_1}{k_3} \right)^{1/4}.
\]
nonlinear elasticity has yet to be established. The reader is referred for details and references to FOSDICK & SHIELD [1], HOLDEN [1], KNOPS & WILKES [1], SENSING [1], WESOLOWSKI [1], WILKES [1]. Nonuniqueness for the pure traction boundary-value problem of a rectangular block of Neo-Hookean material loaded uniformly on each face has been established explicitly by RIVLIN [1, 2, 3]. For various one and two-dimensional rod and shell theories rigorous proofs of nonuniqueness have been given by ANTMAN [2, 3, 6].

10. Concluding Remarks

The main implication of this work for constitutive inequalities is that the quasiconvexity condition (and in particular the Legendre-Hadamard condition) is consistent with realistic models of hyperelastic solids. In the one-dimensional case, when convexity and quasiconvexity of \( W(x, \cdot) \) are the same, Theorem 3.2 shows that the existence of \( C^1(\Omega) \) minimizers for various homogeneous displacement boundary-value problems implies that \( W \) is quasiconvex, while the same result holds in three dimensions if \( \Omega \) is a cube. If \( W \) is not quasiconvex then minimizers may exist that are not \( C^1 \). Some examples in one dimension are discussed by ERIKSEN [3]. It should be noted that we have not proved that \( C^1(\Omega) \) minimizers exist in general for displacement boundary-value problems when \( \partial \Omega \) is suitably regular under any reasonable hypotheses on \( W \). The existence theorems proved in this article take the form that existence is established for a given material for all suitable boundary data. In general such unqualified existence is not to be expected for real materials, since rupture will occur under extreme conditions of deformation. We may also not be interested in solutions having at some points deformation gradients that lie outside the range in which the material behaves elastically. One way of partially circumventing these difficulties is to choose the local constraint set \( W(x) \) so as to prohibit such behaviour, and then to check \( \text{a posteriori} \) whether the minimizer \( u_0 \) is such that \( W_{u_0}(x), \partial W(x) \) for any \( x \). One would then like \( \text{a priori} \) conditions on the size of the boundary data to prevent this happening. The derivation of any such conditions would require delicate estimates. The reader is referred to the papers by ERIKSEN [4] and KNOWLES & STERNBERG [1] for further discussion of some of these points.

In general weak lower semicontinuity will not hold if the quasiconvexity or polyconvexity hypotheses are replaced by a hypothesis of convexity of the function \( W' \) restricted to positive definite symmetric tensors \( U \). It is nevertheless instructive to see how an attempt to establish lower semicontinuity in this case breaks down. The difficulty is that if \( u \to u_0 \) in the Sobolev space \( W^{1,2}(\Omega) \), then the weak limit in \( L^2(\Omega) \) of the sequence \( U_i = \sqrt{\nabla u_i} \sqrt{\nabla u_i} \), will not necessarily be \( \sqrt{\nabla u_0} \sqrt{\nabla u_0} \), and indeed may not arise from any displacement. This is because \( \sqrt{\nabla u_i} \sqrt{\nabla u_i} \) is not of the form (4.2) and hence not sequentially weakly continuous. The difficulty is also connected with the nonlinearity of the Riemann-Christoffel tensor based on \( C \). Similar mathematical problems arise from attempts to establish existence under the Coleman & Noll condition [1], or Hill's inequalities [2, 3]. These conditions do not imply the Legendre-Hadamard condition. The Coleman & Noll condition cannot apply to all hyperelastic materials because it is violated for nearly incompressible materials such as rubber (see HILL [2], OGDEN [1, 3], RIVLIN [2], SIDOROFF [1]).

For bodies that are not homeomorphic to an open ball the various minimizers whose existence we have established may represent deformations topologically isolated from those desired on physical grounds. One might require, for example, all admissible deformations to be accessible from a given deformation by a homotopy of globally invertible configurations. In this article we have not studied such global constraints (although they may be of a weakly closed type), but have concentrated on local constraints such as the local invertibility condition \( \text{det} F > 0 \). Local invertibility is a relatively weak requirement; indeed a hollow sphere may be everted without violation of the condition in any intermediate deformation (SMALE [1]).

Finally I remark on the implications of the results of Section 6 for theories of elasticity incorporating pointwise constraints on the deformation gradient \( F \). These results suggest strongly that the only nontrivial homogeneous constraints giving rise to a well posed theory have the form (see (4.2))

\[
\phi(F) = A + B^T F^2 + C^T \text{adj} F + D \text{det } F = 0, \tag{10.1}
\]

where \( A, B^T, C^T, D \) are constants. It is not hard to show that the only \textit{objective} constraints of this form (i.e., \( \phi \) satisfying \( \phi(QF) = \phi(F) \) for all orthogonal \( Q \)) are those with \( B^T = C^T = 0 \), so that \( \text{det} F \) is specified. In particular, as we have seen, the incompressibility condition \( \text{det} F = 1 \) gives rise to a well posed theory. Note, however, that the constraint of inextensibility (TauerDELL & NOLL [1, p. 72]) is not included. It seems possible, therefore, that solutions do not in general exist for boundary-value problems of inextensible elasticity, and that a higher order theory is required to make such constraints well behaved.

Acknowledgement

The research reported here was begun in the summer of 1974 when I held part of a Science Research Council research fellowship at the Lefschetz Center for Dynamical Systems, Brown University. I am greatly indebted to CONSTANTINE DAPERMAS, without whose enthusiastic interest and consistently excellent advice this article would not have been written. I would also like to thank STUART ANTMAN for several useful suggestions, and for his careful criticism of the manuscript. Finally, I am glad to acknowledge helpful discussions at various stages of the project with ZVI ARTSTEIN, KEN BROWN, DAVID EDMUNDS, J. L. ERIKSEN, ROBIN KNOPS and A.C. PIPKIN.

References


For bodies that are homomorphic to an open ball the same may occur for mixed problems; the situation for displacement boundary-value problems is unclear.

\( ^* \) A visualization of the eversion due to STAPF can be found in PHILLIPS [1].


On the stability of con-formal mappings in multidimensional spaces,


2. On the theory and numerical analysis of the Navier-Stokes equations, Lecture notes in Mathematics, No. 9, University of Maryland.