R. S. Rivlin (Ed.)

Non-linear Continuum Theories in Mechanics and Physics and their Applications

CIME Summer Schools

50

Bressanone, Italy 1969

Springer

FONDAZIONE CIME ROBERTO CONTI

R. S. Rivlin (Ed.)

Non-linear Continuum Theories in Mechanics and Physics and their Applications

Lectures given at a Summer School of the
Centro Internazionale Matematico Estivo (C.I.M.E.),
held in Bressanone (Bolzano), Italy,
September 3-11, 1969

FONDAZIONE
CIME
ROBERTO CONTI

C.I.M.E. Foundation
c/o Dipartimento di Matematica “U. Dini”
Viale Morgagni n. 67/a
50134 Firenze
Italy
cime@math.unifi.it

ISBN 978-3-642-11089-4 e-ISBN: 978-3-642-11090-0
DOI:10.1007/978-3-642-11090-0
Springer Heidelberg Dordrecht London New York

Reprint of the 1st ed. C.I.M.E., Ed. Cremonese, Roma 1970
With kind permission of C.I.M.E.

Printed on acid-free paper

Springer.com

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

2° Ciclo - Bressanone - dal 3 all'11 Settembre 1969

« NON LINEAR CONTINUUM THEORIES IN MECHANICS AND PHYSICS AND THEIR APPLICATIONS »

Coordinatore: Prof. R. S. RIVLIN

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

P. A. BLYTHE

NON-LINEAR FAR-FIELD THEORIES IN RELAXING GAS FLOWS

tenuto a Bressanone dal 3 all'11 settembre 1969

NON-LINEAR FAR-FIELD THEORIES IN RELAXING GAS FLOWS

by

P.A. Blythe

(Lehigh - University)

Summary

In the introduction the small amplitude non-linear far-field theory for one-dimensional isentropic wave propagation is briefly reviewed. The extension to non-equilibrium situations is then discussed for both high frequency and low frequency disturbances and the limitations of these classical theories are examined. It is shown that a suitable small-energy approach can be used both to remove these limitations and to provide a simplified description over the whole frequency range.

1. Introduction: isentropic far-field theory

The purpose of this lecture is to present a unified non-linear far-field theory* for relaxing or reacting gas flows. Attention will be restricted to small amplitude one-dimensional progressing waves and, for simplicity, only rate processes which involve a single internal mode or reaction will be considered.

The corresponding far-field signalling problem in an inviscid gas which is in thermodynamic equilibrium has been well understood for some time. (Whitham 1950, Lighthill 1955). It is useful first to briefly review this problem before discussing the non-equilibrium situations which are of interest here.

In general the mass-conservation, momentum and energy equations take the form (adiabatic flow)

$$\partial_t \rho + \rho u_x = 0 \tag{1.1}$$

$$\partial_t u + \rho^{-1} p_x = 0 \tag{1.2}$$

$$\partial_t e + p \partial_t (\rho^{-1}) = 0 \tag{1.3}$$

where ρ is the density, p is the pressure, u is the particle speed, e is the internal energy and ∂_t is the convective operator

$$\frac{\partial}{\partial t} + u \frac{\partial}{\partial x} \tag{1.4}$$

---

*i.e., the theory must be capable of providing a valid result for 'large' time.

Here t is the time and x is a spatial co-ordinate measured from some fixed reference point.

In thermodynamic equilibrium $e=e(p,\rho)$ and (1.3) can be re-written

$$\partial_t p - a^2 \partial_t \rho = 0 \tag{1.5}$$

where

$$a^2 = \left(\frac{\partial p}{\partial \rho}\right)_s = (p\rho^{-2} - e_\rho) e_p^{-1} \tag{1.6}$$

and the entropy s is defined, in equilibrium, by

$$\theta ds = de + pd(\rho^{-1}). \tag{1.7}$$

where θ is the translational temperature.

It is sometimes convenient to replace the system (1.1) to (1.3) by (1.5) and the characteristic forms

$$\partial_{\pm} p \pm \rho a \partial_{\pm} u = 0 \tag{1.8}$$

where the operators

$$\partial_{\pm} \equiv \frac{\partial}{\partial t} + (u \pm a)\frac{\partial}{\partial x} \tag{1.9}$$

are associated with the characteristic directions

$$\left(\frac{dx}{dt}\right)_{\pm} = u \pm a. \tag{1.10}$$

It is assumed that the disturbance is set up by the motion of a piston whose path is described by

$$x = x_p f(\omega t), \quad t>0 \tag{1.11}$$

(with the origin chosen such that $f(0)=0$) and for $t \leq 0$

$$u = 0, \quad p = p_0, \quad \rho = \rho_0. \tag{1.12}$$

If, in addition,

$$f'(0) = 0, \tag{1.13}$$

though $f''(0)$ is finite, the disturbance is usually termed an acceleration wave. Unless explicitly stated otherwise these conditions on f will be assumed to hold in the subsequent analysis.

Appropriate non-dimensional variables are

$$\begin{gathered} p' = p/p_0, \quad \rho' = \rho/\rho_0, \quad e' = \frac{e\rho_0}{p_0} \\ u' = u\sqrt{\frac{\rho_0}{p_0}}, \quad a' = a\sqrt{\frac{\rho_0}{p_0}} \\ t' = \omega t \text{ and } x' = x\omega\sqrt{\frac{\rho_0}{p_0}}. \end{gathered} \tag{1.14}$$

The relations (1.1) to (1.10) are invariant under this transformation and it is convenient to omit the primes and to regard (1.1) through (1.10) as dimensionless. Corresponding boundary conditions , again omitting primes, are

$$u = 0, \quad p = \rho = 1, \quad t \leq 0, \tag{1.15}$$

and

$$u = \delta f'(t) \text{ on } x = x_p f(t), \quad t>0. \tag{1.16}$$

The dimensionless amplitude parameter δ, which is a measure of the ratio of the piston speed to the ambient sound speed, is given by

$$\delta = \omega x_p \sqrt{\rho_o}/p_o \qquad (1.17)$$

and the particular aim of the present discussion is to obtain solutions which are valid in the limit $\delta \to 0$.

Substitution of the regular expansion

$$\left.\begin{aligned} u(x,t;\delta) &= \delta u_1(x,t)+.. \\ p(x,t;\delta) &= 1+\delta p_1(x,t)+.. \end{aligned}\right\} \qquad (1.18)$$

etc. into (1.1) to (1.3) shows that the first order perturbation quantities satisfy the linear wave equation

$$\frac{\partial^2 A}{\partial t^2} - a_o^2 \frac{\partial^2 A}{\partial x^2} = 0. \qquad (1.19)$$

and, in particular, that the piston condition on u_1 is (see 1.16)

$$u_1(0,t) = f'(t) \ . \qquad (1.20)$$

Hence the appropriate solution, $\xi = t - x/a_o > 0$, is

$$u_1 = f'(\xi). \qquad (1.21)$$

In addition

$$p_1 = a_o^2 \rho_1 = a_o u_1. \qquad (1.22)$$

However, evaluation of the second order approximation shows that the solution contains secular terms of the form $t g(\xi)$. It is apparent that the expansion (1.18) is not uniformly valid as $t \to \infty$ and that difficulties arise, for $\xi = O(1)$, when $\delta t = O(1)$.

These secular terms are, in fact, due to the displacement of the exact characteristics from their position as predicted by linearized theory (Whitham, 1950). This is easily seen for acceleration waves since the exact solution of the full equations, over a certain time interval, is a simple wave. In the present case it is more useful, with a view to later application, to construct the solution in the small amplitude limit by means of the far field expansions,

$$\left.\begin{aligned} u(x,t;\delta) &= \delta U_1(\xi,\eta)+\ldots, \\ p(x,t;\delta) &= 1+\delta P_1(\xi,\eta)+\ldots, \\ \rho(x,t;\delta) &= 1+\delta R_1(\xi,\eta)+\ldots, \end{aligned}\right\} \qquad (1.23)$$

where

$$\eta = \delta t \ , \qquad (1.24)$$

Substitution in (1.8) and (1.5) shows that, as in linearized theory,

$$P_1 = a_0^2 R_1 = a_0 U_1 \qquad (1.25)$$

but U_1 now satisfies

$$\frac{\partial U_1}{\partial \eta} - \frac{b}{a_0} U_1 \frac{\partial U_1}{\partial \xi} = 0 \ , \qquad (1.26)$$

where

$$b = \left[\frac{1}{a}\left\{\frac{\partial}{\partial \rho} \rho a\right\}_s\right]_0 \qquad (1.27)$$

and the suffix o denotes evaluation at the initial conditions.

The appropriate solution of (1.26), subject to the matching condition

$$U_1(\xi,0) = f'(\xi), \tag{1.28}$$

is

$$U_1 = f(\phi) \tag{1.29}$$

where the characteristic lines ϕ=constant are given by

$$\xi = \phi - \frac{b}{a_o} f'(\phi)\eta \tag{1.30}$$

(choosing $\phi=\xi$ on $\eta=0$). Obviously (1.29) and (1.30) are the small amplitude limit of the exact simple wave solution.

If this solution is unique in x-t space then it does represent a uniformly valid result for all η. However, in general the solution will not be single valued where

$$\xi_\phi = 0$$

or

$$\eta = a_o/bf''(\phi). \tag{1.31}$$

Since, for a gas, $b>0$ equation (1.31) is satisfied for some $\eta>0$ if $f''>0$. It is then necessary to insert a discontinuity or shock in order to make the solution unique. The jump conditions across the shock are defined by the Rankine-Hugoniot relations for the conservation of mass, momentum and energy.

It is convenient to note here the form that these relations take for weak shocks. Correct to first order

in δ it follows that

$$[p] = a_0^2[\rho] = a_0[u] \tag{1.32}$$

and the shock path bisects the characteristics that meet on the shock. This latter condition, in the current notation, becomes

$$U_S = a_0 + \frac{b}{2}\delta[U_1^+ + U_1^-] \tag{1.33}$$

where the superscripts -,+ correspond to conditions ahead of and behind the shock respectively.

These relations can be used to evaluate the shock path and they become particularly simple when the shock propagates into an undisturbed region for which $U_1^- = 0$. In that case it follows from (1.33) that if $\xi = \xi_S(\phi_S, \eta)$ on the shock, then ξ_S satisfies the differential equation

$$\frac{d\xi_S}{d\eta} = -\frac{1}{2}\frac{b}{a_0}U_1(\phi_S) \tag{1.34}$$

from which, together with (1.30), the solution is easily found. This solution is defined parametrically by

$$\left.\begin{aligned} \eta &= \frac{a_0}{b} f(\phi)/f'^2(\phi) \\ \xi &= \phi - 2f(\phi)/f'(\phi) \end{aligned}\right\} \tag{1.35}$$

The relations (1.29), (1.30) and (1.35) summarize the main results in the small amplitude non-linear far-field limit for equilibrium isentropic flows.

2. Relaxation processes

In general the excitation of any of the internal degrees of freedom, e.g. vibration, molecular dissociation etc., will take a certain finite time (number of collisions) in which the mode adjusts to some new equilibrium state, although the excitation (relaxation) times for the various modes may differ considerably from each other. In fact, it is known that the time scales for the adjustment of the translational and rotational degrees of freedom are usually much less than those for the other internal modes (Herzfeld Litovitz, 1959) and it will be implicitly assumed in the subsequent analysis that the translational and rotational degrees of freedom remain in a local equilibrium state.

It is further assumed that in any situation of interest only one rate dependent process will be of significance. Hence

$$e = e(p,\rho,\sigma) \tag{2.1}$$

where σ is some relaxation variable. For convenience σ can be identified as a measure of the internal energy in the lagging mode. For vibrational excitation in a pure diatomic gas $e(p,\rho,\sigma)$ depends linearly on σ, but in more complex situations this is not necessarily true.

It is supposed that the rate of adjustment of σ is described by an equation of the form

$$\partial_t \sigma = \Lambda F(p,\rho,\sigma) \qquad (2.2)$$

where the rate function F depends only on the local values of p,ρ and σ and perhaps some initial parameters. The dimensionless rate parameter Λ is the ratio of the time scale defined by the piston to some characteristic relaxation time τ, i.e.

$$\Lambda = (\omega\tau)^{-1} \qquad (2.3)$$

[Equation (2.2) is to be regarded as dimensionless with σ and F both normalized by $p_0\rho_0^{-1}$.]

In an equilibrium state, which is identified by the singular limit $\Lambda\to\infty$, $F\equiv 0$. The corresponding equilibrium path is denoted by

$$\sigma = \bar{\sigma}(p,\rho) \ . \qquad (2.4)$$

In this limit the problem reduces to the isentropic case discussed earlier.

A second isentropic limit is defined by $\Lambda=0$. For this case the internal energy σ remains frozen at its initial value. Obviously this limit is also included in the analysis of §1 .

There is,however, an important distinction that must be drawn between the two limits. In the former equilibrium case the appropriate sound speed is defined by

$$\bar{a}^2 = \left(\frac{\partial p}{\partial \rho}\right)_{s,\sigma=\bar{\sigma}} = (p\rho^{-2} - \bar{e}_\rho)\bar{e}_p^{-1} \qquad (2.5)$$

where $\bar{e}=e(p,\rho,\bar{\sigma})$, whereas in the latter frozen case

$$a^2 = \left(\frac{\partial p}{\partial \rho}\right)_{s,\sigma} = (p\rho^{-2} - e_\rho)e_p^{-1} \qquad (2.6)$$

with $e = e(p,\rho,\sigma)$. It can be shown that in general

$$a^2 > \bar{a}^2. \qquad (2.7)$$

For the general non-equilibrium situation the relation (2.1) implies that (1.5) becomes

$$\partial_t p - a^2 \partial_t \rho = -c\Lambda F \qquad (2.8)$$

where a is the frozen sound speed,

$$c = -\left(\frac{\partial p}{\partial \sigma}\right)_{e,\rho} \qquad (2.9)$$

and (2.2) has been used to replace $\partial_t \sigma$. Moreover, the characteristic relations (1.8) become

$$\partial_{\pm} p \pm \rho a \partial_{\pm} u = -c\Lambda F \qquad (2.10)$$

and the characteristic operators are defined by (1.9) with a interpreted as the frozen sound speed. The influence of the rate process on the energy equation and the characteristic relations introduces a source term, $-c\Lambda F$, which depends on the local values of p,ρ and σ.

The linearized signalling problem associated with this system of equations has been considered several times in the literature (Chu, 1957). The regular expansion

$$\left.\begin{aligned} u(x,t;\delta) &= \delta u_1(x,t) + .. \\ p(x,t;\delta) &= 1 + \delta p_1(x,t) + .. \\ \sigma(x,t;\delta) &= \bar{\sigma}_0 + \delta \sigma_1(x,t) + .. \end{aligned}\right\} \qquad (2.11)$$

yields

$$\frac{\partial}{\partial t}\left(\frac{\partial^2 u_1}{\partial t^2} - a_0^2 \frac{\partial^2 u_1}{\partial x^2}\right) + \lambda\left(\frac{\partial^2 u_1}{\partial t^2} - \bar{a}_0^2 \frac{\partial^2 u_1}{\partial x^2}\right) = 0 \qquad (2.12)$$

where

$$\lambda = \Lambda(-F_\sigma c \bar{c}^{-1})_0 \qquad (2.13)$$

is a modified rate parameter. (1.20) again defines the boundary condition on x=0 [Note that it is assumed in (2.11) that the initial conditions correspond to an equilibrium state.]

(2.12) obviously reduces to the standard linearized result in both the frozen (high frequency) limit $\lambda \to 0$ and the equilibrium (low frequency) limit $\lambda \to \infty$. For arbitrary values of λ (2.12) suggests that for $t\lambda \ll 1$ the effective propagation speed is a_0, but for $t\lambda \gg 1$ it is $\bar{a}_0$. This latter statement can be made more precise. The formal solution of (2.12), subject to (1.20) and (1.15), can be obtained by Laplace transforms. An asymptotic evaluation, $t, x \to \infty$ but sufficiently far behind the front, shows that (Clarke, 1965)

$$u_1 \sim \sqrt{\frac{D}{2\pi t}} \int_0^\infty f'(y) \exp\{-D t^{-1}(\bar{\xi} - y)^2\} dy \qquad (2.14)$$

where

$$D = \lambda(\alpha^2 - 1)^{-1}, \qquad (2.15)$$

$$\alpha = a_0/\bar{a}_0 \qquad (2.16)$$

and

$$\bar{\xi} = t - x/\bar{a}_0 \qquad (2.17)$$

is the linearized characteristic associated with the low frequency (equilibrium) signal. The main disturbance is now apparently centered on these latter wavelets.

It is easily verified that according to (2.12) any plane wave is distorted both by dispersion, so that the wave speed depends on the frequency, and by absorption in which the amplitudes of the high frequency components are much more rapidly attenuated than those of the low frequency ones.

However, as in the isentropic case, it can be shown that the regular expansion (2.11) is not necessarily uniformly valid in the far field, and secular terms may again appear in higher order solutions.

The remainder of the lecture will be devoted to a discussion of the modifications that are required in order to obtain a valid far-field result.

3. The high frequency limit

A simple extension of the classical isentropic far-field approach can be used in the high-frequency (near-frozen) limit $\Lambda \to 0$ (see Varley and Rogers, 1967). For ease of discussion it will be assumed that Λ and δ are of a similar magnitude. The corresponding expansion is

$$\left.\begin{aligned} u(x,t;\delta) &= \delta U_1(\xi,\eta)+.. \\ p(x,t;\delta) &= 1+\delta P_1(\xi,\eta)+.. \\ \sigma(x,t;\delta) &= \bar{\sigma}_0+\delta^2 \Sigma_2(\xi,\eta)+.. \end{aligned}\right\} \qquad (3.1)$$

etc. Note that $\sigma-\bar{\sigma}_0$ is a second order quantity (more strictly its magnitude is $0(\Lambda\delta)$). This expansion procedure would appear to be appropriate for large times at distances behind the front which are comparable with the length scale defined by the piston signal but which are much less than the relaxation length. Substitution in (2.10), (2.8) and (2.2) shows that

$$\frac{\partial U_1}{\partial \eta} - \frac{b}{a_0} U_1 \frac{\partial U_1}{\partial \xi} + kU_1 = 0 \qquad (3.2)$$

where

$$k = (1 - \frac{1}{\alpha^2})\frac{\lambda}{2\delta} = 0(1), \qquad (3.3)$$

and b corresponds to (1.27) with the derivative evaluated both at constant S and σ.

The first order perturbation quantities are again related, as in frozen linearized theory, by

$$P_1 = a_0^2 R_1 = a_0 U_1. \qquad (3.4)$$

Equations (3.4) and (3.2) ahould be compared with (1.25) and (1.26) respectively. The attenuation factor kU_1 plays a dominant role in the asymptotic behavior of (3.2) as $\eta \to \infty$.

The inner near-field solution for $\Lambda=0(\delta)$, with $x,t = 0(1)$, is given by the usual frozen linearized result

$$u_1 = f'(\xi) \qquad (3.5)$$

which defines the inner matching condition for U_1.

Subject to (3.5) and the condition at the front, the solution of (3.2) is defined by

$$U_1 = f'(\phi)e^{-k\eta} \tag{3.6}$$

with

$$\xi = \phi - \frac{b}{a_o k} f'(\phi)[1-e^{-k\eta}]. \tag{3.7}$$

Again this solution is not single valued in physical space at points where

$$\xi_\phi = 0$$

or

$$\eta = -\frac{1}{k} \log \left(1 - \frac{ka}{bf''(\phi)}\right) \tag{3.8}$$

However, in contrast to the isentropic solution shocks will not form even for compressive piston motions if

$$\frac{ka_o}{b(f'')_{max}} > 1 \tag{3.9}$$

(Varley and Rogers 1967, Rarity 1967).

If a shock does form its path can be determined, in principle, by the approach outlined in §1. Conditions (1.32) and (1.33) again hold for a weak shock, with a interpreted as the frozen sound speed, together with the additional statement

$$[\sigma] = 0 \tag{3.10}$$

In writing down (3.10) it is implicitly assumed that the shock thickness, across which the translational mode

adjusts to a new equilibrium state, is neglibibly thin in comparison with the relaxation length $a_0\tau$.

For a shock propagating into an undisturbed region it can be shown that its path is described by

$$\eta = -\frac{1}{k}\log\left(1 - \frac{2a_0 k}{b} f(\phi)/f'^2(\phi)\right)$$
$$\xi = \phi - 2f(\phi)/f'(\phi) \tag{3.11}$$

Although (3.11) reduces to (1.35) as $k\to 0$, it follows from (3.6) and (3.11) that for any finite k the amplitude of the shock is exponentially weak as $\eta\to\infty$, even for pistons whose speed is asymptotically constant.

Moreover, it is apparent, both from physical reasoning and by directly computing higher order terms in (3.1), that this high frequency expansion will break down as $\xi\to\infty$, or, more precisely, at distances behind the front which are comparable with the relaxation length. It is easily shown that for $\xi=0(\delta^{-1})$, $\eta=0(1)$ the dominant behavior is described by the linear equation (2.12) (Blythe, 1969) though this result does not necessarily in itself give a uniformly valid description of the limiting asymptotic behavior. Before discussing further this particular difficulty for high frequency disturbances, it is relevant to return to the asymptotic description for $\Lambda=0(1)$.

4. The low frequency far-field limit, $\Delta=0(1)$.

The dominant asymptotic signal according to linear (near-field) theory is defined by (2.14). If $u=0(\Delta(\delta))$ in this region it appears that the only non-trivial stretching of the independent variables is

$$T = \Delta^2(\delta)t, \quad \bar{\chi} = \Delta(\delta)\bar{\xi} \tag{4.1}$$

together with

$$\left.\begin{aligned} u &= \Delta(\delta)V_1(\bar{\chi},T)+.. \\ p &= 1+\Delta(\delta)\pi_1(\bar{\chi},T)+.. \\ \varepsilon = \sigma-\bar{\sigma} &= \Delta^2(\delta)E_2(\bar{\chi},T)+.. \end{aligned}\right\} \tag{4.2}$$

This last relation, which follows directly from the rate equation , implies that the departure from an equilibrium state is small. In this sense the expansion (4.1) and (4.2) defines a low frequency far-field limit. The magnitude of $\Delta(\delta)$ is defined implicitly by (2.14) (see below).

Before substituting these expansions into (2.5), (2.10) and (2.2) it is better to replace σ by ε as a basic dependent variable.

It can be shown that V_1 satisfies

$$\frac{\partial V_1}{\partial T} - \frac{\bar{b}}{\bar{a}_0} V_1 \frac{\partial V_1}{\partial \bar{\chi}} = \mu \frac{\partial^2 V_1}{\partial \bar{\chi}^2} \tag{4.3}$$

where

$$\mu = \frac{1}{2\lambda}(\alpha^2-1) = \frac{1}{2} D^{-1} \tag{4.4}$$

(4.3) is Burger's equation. It has been suggested many times that this provides a satisfactory asymptotic description of the flow field (Lighthill 1956, Jones 1964, Lick 1967). This equation can be transformed into the diffusion equation and it is easily verified that its solution will match with the outer behavior of (2.12) given in (2.14).

In deriving (4.3) it has been assumed that $V_1 = 0(1)$: the magnitude of $\Delta(\delta)$, as noted above, is defined by (2.14). However, it appears that this stretching is not permissible for all piston motions. In fact, if $tf'(t)\to 0$ as $t\to\infty$, $V_1=0(\delta)$ and the non-linear term in (4.3) is negligible in this particular far field region. For piston paths whose decay is slower, e.g.

$$f' \sim t^{-n}, \quad 0 \leq n < 1,$$

$\Delta = 0(\delta^{\frac{1}{1-n}})$.

In the high frequency limit discussed in §3 it is apparent that the solution in the intermediate linearized regime, where $xt=0(\delta^{-1})$, will break down in the same way. Appropriate far field (low-frequency) variables are then

$$T_1 = \delta\Delta^2 t, \quad \bar{\chi}_1 = \delta\Delta\bar{\xi} \tag{4.5}$$

However, this asymptotic solution is always shock free. (Even if any shock forms at the front its strength will become exponentially weak for all bounded piston speeds.) In particular, when the piston speed attains a constant limiting value the associated steady state profile is fully-

dispersed: all convective steepening can be balanced solely by the dissipative nature of the rate process. Yet, it is well known that stable steady partly dispersed wave forms, in which the relaxation region is preceeded by a Rankine-Hugoniot shock, do exist and it is informative to discuss this limitation in these asymptotic solutions.

Throughout the analysis so far it has been assumed that the energy σ is of a similar magnitude to the total internal energy e, or equivalently that

$$a/\bar{a}_0 - 1 = O(1). \tag{4.6}$$

This latter restriction, for steady state waves, always implies that $U_w - \bar{a}_0 = o(1)$, where U_w is the wave speed, but for partly dispersed waves to exist

$$U_w > a_0 . \tag{4.7}$$

This latter condition cannot hold for small amplitude waves ($\delta \to 0$) if (4.6) is satisfied.

5. The small energy limit

Situations in which both $\alpha - 1$ and u are "small" are obviously of some interest. In this limit it is possible to obtain a simplified description of the far field in which both fully-dispersed and partly-dispersed wave-profiles can be discussed in a unified manner.

For ease of discussion, the magnitude parameter δ will also be used as a characteristic measure of σ.

This statement should not be taken to imply any relation between the internal energy and the piston speed. If necessary a second parameter δ_1, with $\sigma=0(\delta_1)$, can be introduced and the subsequent analysis will hold provided terms $0(\delta,\delta_1)$ etc. are retained.

The appropriate far-field expansion is again of the type outlined in §2, with a slight modification in the energy term. ξ and η are used as independent variables and

$$\left.\begin{aligned} u &= \delta U_1(\xi,\eta)+.. \\ p &= 1+\delta P_1(\xi,\eta)+.. \\ \sigma &= \delta(e_0+\delta e_1(\xi,\eta)+.. \end{aligned}\right\} \qquad (5.1)$$

Note that

$$e = \sigma\delta^{-1} = 0(1) \ , \qquad (5.2)$$

Substitution in (2.2), (2.8) and (2.10) gives

$$P_1 = a_0^2 R_1 = a_0 U_1 \qquad (5.3)$$

which are the usual linearized relations but, U_1 and e_1 now satisfy

$$\frac{\partial U_1}{\partial \eta} - \frac{b}{a_0} U_1 \frac{\partial U_1}{\partial \xi} = -\frac{c_0}{2a_0} \frac{\partial e_1}{\partial \xi} \qquad (5.4)$$

$$\frac{\partial e_1}{\partial \xi} = 2 \frac{a_0}{c_0} kU_1 - \lambda e_1 \qquad (5.5)$$

Here

$$k = (1 - \frac{1}{\alpha^2})\frac{\lambda}{2\delta}$$

is to be regarded as $0(1)$.

In this first order approximation the rate equation (5.5) is linear, though it now contains both 'forward' and 'backward' terms. The only non-linear convective term occurs in (5.4).

In the near-frozen limit $\Lambda \to 0$ $(\lambda, k \to 0)$ equation (5.4) reduces to the expected result (1.26), and iteration using (5.5) gives the Varley-Rogers limit (3.2). In the low frequency or near-equilibrium limit, $\Lambda \to \infty (\lambda, k \to \infty)$, equations (5.4) and (5.5) give

$$\frac{\partial U_1}{\partial \eta} - \left(\frac{b}{a_o} U_1 - \frac{k}{\lambda}\right)\frac{\partial U_1}{\partial \xi} = 0.$$

Since

$$\frac{k}{\lambda} = \frac{\alpha - 1}{\delta} + 0(\delta), \quad b = \bar{b} + 0(\delta)$$

this last result reduces to

$$\frac{\partial U_1}{\partial \eta} - \frac{\bar{b}}{\bar{a}_o} U_1 \frac{\partial U_1}{\partial \bar{\xi}} = 0 \tag{5.6}$$

neglecting terms $0(\delta)$. (5.6) is the classical equilibrium result also defined by (1.26). By including terms $0(\Lambda^{-1})$ it can be shown that U_1 satisfies Burger's equation (4.3) when only the dominant terms with respect to δ are retained.

Under the transformation

$$\left.\begin{aligned} U_1 &= \frac{k}{\lambda}\frac{a_o}{b} w, \quad e_1 = \frac{k^2}{\lambda^2}\frac{2a_o^2}{c_o b} E \\ \xi &= \psi/\lambda, \quad \eta = Y/k \end{aligned}\right\} \tag{5.7}$$

(5.4) and (5.5) reduce to

$$\frac{\partial w}{\partial Y} - w\frac{\partial w}{\partial \psi} = -\frac{\partial E}{\partial \psi} \qquad (5.8)$$

$$\frac{\partial E}{\partial \psi} = w-E \qquad (5.9)$$

and are free of parameters. The piston condition on Y=0 becomes

$$w = \frac{\lambda b}{ka_o} f'\left(\frac{\psi}{\lambda}\right) . \qquad (5.10)$$

Although for geometrically similar paths the solutions will in general be similar only for fixed values of the parameters λ and ka_o/b, a considerable simplification occurs in one particular case. For a centered expansion wave the condition at the origin is

$$u \sim -\frac{a_o}{b}\left(1 - \frac{x}{a_o t}\right)$$

which re-expressed in far field variables gives

$$w \sim -\psi/Y. \qquad (5.11)$$

The differential equations (5.8) and (5.9), the front condition and the initial condition (5.11) are now independent of all parameters. This similarity form has been discussed in Blythe (1969) where a numerical solution, using a characteristics method, was presented

It is sometimes convenient to eliminate E from (5.8) and (5.9) . The resulting second order equation is

$$\frac{\partial}{\partial \psi}\left(\frac{\partial w}{\partial Y} - w\frac{\partial w}{\partial \psi}\right) + \frac{\partial w}{\partial Y} - (w-1)\frac{\partial w}{\partial \psi} = 0. \qquad (5.12)$$

The structure of this equation should be compared with that of the classical linearized result (2.12). Here the linear

operators of (2.12), associated with the high and low frequency sound speeds respectively, are replaced by corresponding non-linear convective operators. The linearized form of (5.12),

$$w_{\psi Y}+w_Y+w_\psi = 0, \tag{5.13}$$

is the telegraph equation. Moore and Gibson (1960) deduced (5.13) from the usual linearized form (2.12) in the limit $\alpha-1<<1$. In Moore and Gibson's derivation $t=0((\alpha-1)^{-1})$ but it is apparent that in order for this equation to be applicable in this domain

$$\delta^{-1}(\alpha-1)>>1.$$

The simplest solutions of (5.8) and (5.9) are those of steady state form

$$\begin{aligned} w &= w(\psi+CY), \\ E &= E(\psi+CY), \end{aligned} \tag{5.14}$$

where the wave speed associated with C, in (x,t) space, is

$$U = a_o[1+\delta\frac{Ck}{\lambda}]\approx a_o[1+(\alpha-1)C] \tag{5.15}$$

Solutions of this form correspond to the asymptotic state due to a compressive piston moving at constant speed.

The differential equations satisfied by w and E are

$$(C-w)w' = -E' = E-w . \tag{5.16}$$

whose non-trivial solution is defined by

$$w' = \frac{E-w}{C-w} = \frac{\frac{1}{2}w^2-(C+1)w+K}{C-w} \tag{5.17}$$

Since w'=w=0 at upstream infinity apparently

$$K=0 \tag{5.18}$$

However, solutions of (5.17) are unique only if

$$C<0 \tag{5.19}$$

with the piston speed given by

$$w_p = 2(C+1)>0. \tag{5.20}$$

((5.17) cannot be used to study expansion waves with $w_p<0$. It is easily shown that the overall entropy change would be negative for this case).

Note from (5.15), that the restrictions (5.19) and (5.20) imply

$$a_o>U_w>\bar{a}_o \tag{5.21}$$

which is the usual condition for a fully-dispersed wave (Lighthill, 1956).

If C>0, (5.17), with K=0, does not represent a single valued solution. For compression waves a Rankine-Hugoniot shock must be inserted at the front. From the weak shock relations it follows that

$$w = 2C \tag{5.22}$$

immediately behind the shock. Hence from (5.17), with E=0,

it again follows that K=0.

Apart from an arbitrary constant, (5.17) integrates to give

$$\psi+CY = \frac{C}{1+C}\log\left\{\frac{2C}{w}(C+1 - \frac{w}{2})^{-(1+\frac{2}{C})}\right\}. \tag{5.23}$$

For centered expansion waves it is expected that the asymptotic disturbance will be the equilibrium solution

$$w = -\bar{\psi}/Y = 1-\psi/Y+0(\delta). \tag{5.24}$$

It is easily verified, neglecting terms $0(\delta)$, that (5.24) is an exact solution of the full equation (5.12).

Although other exact analytical solutions of (5.4) and (5.5) are not readily found, it is apparent that these equations do provide a uniform small amplitude far-field limit with respect to the rate parameter Λ. In addition, they will describe the structure of both partly-dispersed and fully-dispersed wave forms. Some further discussion of the properties of these equations can be found in Blythe (1969) (see also Spence & Ockendon 1969).

P.A. Blythe

References

Blythe, P.A. 1969 Non-linear wave propagation in a relaxing gas. J. Fluid Mechs. 37, 31.

Chu, B.T. 1957 Wave propagation and the method of characteristics in reacting gas mixtures with applications to hypersonic flow. WADC TN 57-213.

Clarke, J.F. 1965 On a first order wave theory for a relaxing gas flow. College of Aeronautics, Rep. Aero. 182.

Herzfeld, K.F. & Litovitz, T.A. 1959 Absoprtion and dispersion of ultrasonic waves. Academic Press.

Jones, J.G. 1964 On the near-equilibrium and near-frozen regions in an expansion wave in a relaxing gas. J. Fluid Mechs. 19, 81.

Lick, W.J. 1967 Wave propagation in real gases. Advances in Applied Mech. 10, Fas. 1. Academic Press.

Lighthill, M.J. 1955 Section E of General Theory of High Speed Aerodynamics (Vol. VI of High Speed Aerodynamics and Jet Propulsion). Oxford University Press.

Lighthill, M.J. 1956 Viscosity in waves of finite amplitude. Article in Surveys in Mechanics. Cambridge University Press.

Moore, F.K. & Gibson, W.E. 1960 Propagation of weak disturbances in a gas subject to relaxation effects. J. Aero. Sci. 27, 117.

Rarity, B.S. 1967 On the breakdown of characteristic solutions in flows with vibrational relaxation. J. Fluid Mechs. 27, 49.

Spence, D.A. & Ockendon, H. 1969 To be published in J. Fluid Mechs.

Varley, E. & Rogers, T.G. 1967 The propagation of high frequency finite acceleration pulses and shocks in visco-elastic materials. Proc. Roy. Soc. A 296, 498.

Whitham, G.B. 1950 The behaviour of supersonic flow past a body of revolution, far from the axis. Proc. Roy. Soc. A 201, 89.

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C. I. M. E.)

J. MEIXNER

THERMODYNAMICS OF DEFORMABLE MATERIALS

Corso tenuto a Bressanone dal 3 all'11 Settembre 1969

THERMODYNAMICS OF DEFORMABLE MATERIALS

by

J. Meixner [1)]

Center for the Application of Mathematics, Lehigh University, Bethlehem

1. General remarks on the state of thermodynamics

It is now a little over a hundred years that the conceptual basis of thermostatics has been laid. Clausius introduced the entropy concept and obtained the laws of thermostatics for fluids which are in internal equilibrium. Not much later, i.e. 94 years ago, Gibbs completed the formal mathematical structure of thermostatics and developed the minimum and maximum principles and the stability conditions for fluids. There came many applications of thermostatics, of which chemical reactions, blackbody radiation and the thermostatics of surface tension may be mentioned.

Thermostatics of solid deformable materials has been restricted to small strains and stresses for a long time. It is only through the last few decades that thermostatics of large deformations has found attention. And there are still some important problems which as yet have not found a satisfactory solution. The thermostatics of plastic deformation is, in my opinion, still in a very primitive state. But even if we confine our attention, as we shall do here, to deformations where plasticity does not yet occur, or more precisely, to simple deformable materials, no full thermostatic stability condition is known. It goes without saying that such stability conditions must exist. But none of the many inequalities in mechanics of deformable materials which have been postu-

1) On Leave of absence from Rheinisch-Westfälische Technische Hochschule, Aachen.

lated or motivated can be considered as such (or as the mechanical part thereof) because , when applied to fluids, they do not go over into the well-established stability conditions for fluids.

Thermodynamics proper, which deals with irreversible processes (to avoid confusion with the usual usage of the word thermodynamics we shall speak of thermodynamics of processes) has found attention during the centuries, and its first law has been given 300 years ago; it is Newton's viscosity law. Fourier's law of heat conduction was given in 1835. Some decades later we had Fick's law of diffusion. The beginning of a systematic theory of thermodynamics of processes should be credited to Jaumann and Natanson (from 1910) .

But it was not before 1940 that thermodynamics of processes came to life again, and there was an enormous growth of literature, but also a conflict of approaches and of ideas. What started in 1941 and was developed mainly during the following decade I shall in the following call classical thermodynamics of irreversible processes (classical TIP for short) . Main contributers were C. Eckart (1940); myself (from 1941), Prigogine (since 1947), de Groot (since 1945) and Mazur and Wergeland. But I should not forget to mention Onsager's work of 1931 which gave an important contribution to the classical TIP*

The value of classical TIP has been completely denied by another group of scientists of which Truesdell, Coleman and Noll are perhaps the

---

*) Presentations of classical TIP including bibliographies of the relevant Literature are given in [1) 2) 3)] .

most important representatives. Coleman [4) 5)] has developed a new thermodynamics which is partly founded on ideas which have grown and been used in the non-equilibrium mechanics of large deformations during the past 15 years *) . It is now in such a state that it has joined, as Truesdell remarks, "the older science of mechanics and the younger science of electromagnetism as a fully formulated mathematical discipline" and it is also, as Truesdell says, so simple that it can be taught to beginners [7)] .

A new thermodynamics of processes has been developed during the last five years**) . It started from the analysis of very special but on the whole also very well understood thermodynamic systems, namely electrical networks. There was one point which , for good reasons, has not been paid much attention to by the network engineers nor the network mathematicians. This was the question how the entropy of a network in a given state should be defined if the network is to be contained in a black box which means that we derive information about the network only be operating and measuring at the terminals, in other words that we describe the network by its impedance matrix with respect to the accessible terminals. It turned out, that the entropy of a network, if it can be defined at all, has not a unique value, but that there exists an infi-

---

*) A very comprehensive presentation of this development up to 1964 has been given in [6)] . Most of the more recent literature on this subject is found in the "Archive of Rational Mechanics and Analysis" and in the "International Journal of Engineering Science".

**) Some earlier papers [8)11)] prepared the ground. The main ideas have been given and worked out in [12)-17)]

nite number of network realizations for a given black box behavior or for a given impedance matrix and that they have , except for an impedance matrix of the reactance type, with the same applied voltage histories, many numerically different entropy histories. This is, of course, only true if there are actual processes, i.e. voltage histories which are not constant in time. An escape from this disturbing finding would be possible if one could exclude electrical networks from the set of thermodynamic systems. Since this would contradict general physical ideas, the conclusion is unavoidable : If electrical networks cannot be assigned a unique value of the entropy in a non-equilibrium state, then a unique value of the entropy in other thermodynamic systems should not exist either in a non-equilibrium state.

In such a situation the best thing to do is to go back to the original sources. What we find there is, indeed, quite puzzling. Clausius[18] states the second part of the second law in the form

$$S(B) - S(A) > \int_A^B \frac{\delta Q}{T} \tag{1.1}$$

which is to hold for any process without exchange of matter and which leads in course of time from an equilibrium state A to an equilibrium state B. By δQ we understand the supplied heat during an interval of the process while S(A) and S(B) are the thermostatic entropies in the equilibrium states A and B. Clausius himself emphasizes quite strongly that A and B must be equilibrium states in order that (1) hold.

Some ten years later, Clausius [19] gives without motivation much less a proof the inequality

$$(1.2) \qquad dS \geq \frac{\delta Q}{T}$$

meaning that S ist the entropy difference of two adjacent non-equilibrium states. Thereby he assumes that an entropy can be defined also for non-equilibrium states and that (1.2) can be considered in a way as a differential formulation of the inequality (1.1). This is in complete contrast to his earlier statements about the entropy concept and the Second Law.

So we see that an entropy in non-equilibrium states has no foundation in Clausius' work.

Up to the present time the inequality (1.2) has been taken over from Clausius without much criticism and, as far as we know, without an attempt to remove the deficiency in the definition of a non-equilibrium entropy. Classical TIP as well as Coleman's thermodynamics are based on the inequality (1.2) which, for fields, is usually transformed into the Clausius-Duhem inequality

$$\rho \frac{ds}{dt} + \nabla \cdot \underline{q} > 0$$

where ρ, s, $\underline{q}$ are the density, the specific entropy and the heat flow, respectively.

Although kinetic theories seem to prove the existence of an entropy also during a process, this proof is open to criticism. In Boltzmann's kinetic theory of gases the H-theorem is proved for the one particle distribution functions. But one can derive other H-theorems for 2-particle distribution functions etc. and each such H-theorem defines

an entropy. That is to say, that quite a multitude of entropy values can be defined for a material which has been brought, within an irreversible process, to a well-defined non-equilibrium state. In other words, no unique entropy value exists in an non-equilibrium state.

It is therefore necessary to develop the laws of thermodynamics of processes without taking recourse to the doubtful concept of a non-equilibrium entropy. The main ideas of such a theory will be developed in the following sections.

2. Thermostatics of deformable materials

We introduce at first a few kinematic concepts. Consider a homogeneous deformable body at rest in an equilibrium state of vanishing stresses. Assume that it undergoes after time t_0 a deformation without diffusion which will always be excluded in the following. We define the velocity $\underline{v}(\underline{x}, t)$ through the average linear momentum of the atomic constituents which are at time t within a sufficiently small element of volume with the point $\underline{x}$ as its center of gravity. When we solve the differential equation

$$\frac{d}{dt}\,\underline{x} = \underline{v}(\underline{x}, t) \tag{2.1}$$

with the initial condition $x(t_0) = X$, we obtain

$$\underline{x} = \underline{x}(X, t)\ . \tag{2.2}$$

This is called the trajectory of a material point X which is defined by its coordinates at time t_o and which is not to be confused with atoms. The equation (2.2) is a mapping of the material points $\underline{X}$ in the unstressed equilibrium state into their positions $\underline{x}$ at time t.

In this section we are interested in equilibrium states only and therefore assume that t in (2.2) refers to such a state . We usually omit, therefore, t in the argument. The deformation gradient of this mapping is defined as

$$F_{i\alpha}(X) = \frac{\partial x_i(X_1 t)}{\partial X_\alpha} \tag{2.3}$$

It assumes the value

$$F_{i\alpha} = \delta_{i\alpha} \tag{2.4}$$

if $\underline{x} \equiv \underline{X}$. The density ρ in the state $\underline{x}$ is expressed by

$$\rho^+ = \rho \det |F_{i\alpha}| \tag{2.5}$$

where ρ^+ is the density in the state $\underline{X}$.

There are unique factorizations`

$$\underline{\underline{F}} = \underline{\underline{R}}\ \underline{\underline{U}} = \underline{\underline{V}}\ \underline{\underline{R}} \tag{2.6}$$

with an orthogonal matrix $\underline{\underline{R}}$, $\det \underline{\underline{R}} = 1$ and with symmetric matrices $\underline{\underline{U}}$, $\underline{\underline{V}}$ which have the same positive eigenvalues.
We also introduce the right and left Cauchy-Green stretch tensors

$$\underline{\underline{C}} = \underline{\underline{\tilde{F}}} \cdot \underline{\underline{F}} = \underline{\underline{U}}^2, \quad \underline{\underline{B}} = \underline{\underline{F}} \cdot \underline{\underline{\tilde{F}}} = \underline{\underline{V}}^2 \tag{2.7}$$

where the tilde means the adjoint tensor.

We assume now that the material under consideration has a unique static specific entropy s_{st} which depends on the specific energy u, on the deformation gradient $\underline{\underline{F}}$ and on some scalar internal variables $\xi_1, \ldots, \xi_n$ which need no specification. Then

$$s_{st} = s_{st}(u, \underline{\underline{F}}, \xi_1, \ldots, \xi_n) . \tag{2.8}$$

Therefrom we obtain

$$ds_{st} = \frac{1}{T_{st}} du + \frac{\partial s_{st}}{\partial F_{i\alpha}} dF_{i\alpha} + H_i d\xi_i . \tag{2.9}$$

Here and in the following the summation convention is adopted. T_{st} is the temperature of the static state. The H_i are the affinities in the entropy representation. In order to identify the coefficient of $dF_{i\alpha}$, we have to evaluate the work done in an infinitesimal reversible transformation without heat supply ($ds_{st} = 0$) and with frozen internal variables. Thus one obtains

$$\frac{\partial s_{st}(u, \underline{\underline{F}}, \xi_i)}{\partial F_{i\alpha}} = -\frac{1}{\rho T_{st}} (\underline{\underline{F}}^{-1} \underline{\underline{\sigma}}_{st})_{i\alpha} . \tag{2.10}$$

A slight simplification of (2.8) is possible if we note that s_{st} is a scalar invariant and that it should not change its value if we apply a further deformation $\underline{\underline{Q}}$ which is a pure rotation. Then we have

$$s_{st}(u, \underline{\underline{F}}, \xi_i) = s_{st}(u, \underline{\underline{Q}}\,\underline{\underline{F}}, \xi_i) \tag{2.11}$$

and choosing $\underline{\underline{Q}} = \underline{\underline{R}}^{-1}$ (see (2.6)), we see that s is a function of U, or equivalently of C. So

$$s_{st} = s_{st}(u, \underline{\underline{C}}, \xi_i) \tag{2.12}$$

and

$$ds_{st} = \frac{1}{T_{st}} du - \frac{1}{2\rho T_{st}} \operatorname{Trace} \underline{\underline{F}}^{-1} \cdot \underline{\underline{\sigma}}_{st} \cdot \underline{\underline{\tilde{F}}}^{-1} d\underline{\underline{C}} + H_i d\xi_i \,. \tag{2.13}$$

We have added the subscript st to the entropy s, to T and $\underline{\underline{\sigma}}$ in order to emphasize that these quantities refer to an equilibrium state. It is clear that T_{st} and $\underline{\underline{\sigma}}_{st}$ are functions of u, $\underline{\underline{F}}$ (or $\underline{\underline{C}}$) and ξ_i as well as s_{st} is. These functions will also play a part in the non-equilibrium theory where they are understood just as the same functions of u, $\underline{\underline{F}}$, ξ_i but not implying that they can be identified with the non-equilibrium temperature or the non-equilibrium stress tensor.

3. The balance equations and the fundamental inequality

The conservation laws for mass, linear and angular momentum and energy are expressed by balance equations.

The balance equation for mass, also called continuity equation, can be written in various forms. They are

$$\rho^{+} = \det \underline{\underline{F}} \,, \tag{3.1}$$

$$(3.2)\qquad \frac{d\rho}{dt} + \rho\,\nabla\cdot\underline{v} = 0\ , \qquad \frac{d\rho}{dt} + \rho\ \mathrm{Trace}\,(\underline{\underline{F}}^{-1}\,\dot{\underline{\underline{F}}}) = 0\ .$$

The d/dt and the dot over $\underline{\underline{F}}$ denote the material derivative

$$(3.3)\qquad \frac{d}{dt} = \frac{\partial}{\partial t} + \underline{v}\cdot\nabla$$

The balance equation for linear momentum is

$$(3.4)\qquad \rho\,\frac{dv_i}{dt} = \frac{\partial \sigma_{ik}}{\partial x_k}$$

if external forces are ignored.

We assume for simplicity that there is no exchange between angular momentum of motion and intrinsic angular momentum due to internal rotation of molecules or of spins. Then the stress tensor is symmetric

$$(3.5)\qquad \underline{\underline{\sigma}} = \underline{\underline{\tilde{\sigma}}}\ .$$

The energy balance equation can be changed into a balance equation for the internal energy alone if proper use of the balance equation (3.4) is made. It reads

$$(3.6)\qquad \rho\,\frac{du}{dt} = \mathrm{Tr}\ \sigma\ \dot{F}\ F^{-1} - \nabla\cdot\underline{q}$$

By $\underline{q}$, the heat flow vector is understood.

These balance equations indicate a first hand choice of parameters of state. They contain u but not T, $\underline{q}$ but not ∇T. We shall, therefore, prefer to use $u(t)$ and $\underline{q}(t)$ as independent variables when

setting up constitutive equations.

From the balance equations (3.2) and (3.6) and from the entropy differential (2.9) one infers the identity

$$\rho \frac{ds_{st}}{dt} + \nabla\cdot\left(\frac{1}{T}\,\underline{q}\right) = \left(\frac{1}{T_{st}} - \frac{1}{T}\right)\rho\dot{u} + \nabla\frac{1}{T}\,\underline{q} + Tr\left(\frac{\underline{\underline{\sigma}}}{T} - \frac{\underline{\underline{\sigma}}_{st}}{T_{st}}\right)\dot{\underline{\underline{F}}}\,\underline{\underline{F}}^{-1} \tag{3.7}$$

if internal variables are from now on ignored. In this identity T is quite an arbitrary function of $\underline{x}$, t.

If we assume $T = T_{st}$ and understand by $s_{st}(u, \underline{\underline{F}})$ the specific entropy during the process, which may be reasonable for small departures from an equilibrium state, then (3.7) is the entropy balance equation of classical TIP as applied to deformable materials with internal equilibrium. In this case the Clausius-Duhem inequality (1.3) would require that the right member of (3.7) is non-negative and can therefore be interpreted as entropy production.

Without these assumptions nothing can be said about the sign of the right member of (3.7). However, we obtain an interesting statement if we integrate the equation (3.7) for a fixed small mass element along its motion for a process which starts at $t_o = -\infty$ from an equilibrium state and leads at $t = +\infty$ again to an equilibrium state. Since $\rho\, dV$, with dV being the volume element, is a constant during the motion, we obtain by integrating the left member of (3.7) from $-\infty$ to ∞

$$(3.8) \qquad m\left[s_{st}(\infty) - s_{st}(-\infty)\right] + \int_{-\infty}^{\infty} dt \oint \frac{1}{T}\, \underline{q} \cdot d\underline{a}$$

with m being the mass of the mass element, and if we understand by T the temperature during the process, or the temperature in the Clausius inequality (1.1) we arrive at the conclusion that the expression (3.8) is non-negative. Therefore the same is true for

$$\int_{-\infty}^{\tau} \left[\left(\frac{1}{T_{st}} - \frac{1}{T}\right) \dot{u} + \nabla \frac{1}{T} \cdot \frac{1}{\rho}\, \underline{q} + \frac{1}{\rho}\, T\, r\left(\frac{\underline{\underline{\sigma}}}{T} - \frac{\underline{\underline{\sigma}}_{st}}{T_{st}}\right) \dot{\underline{\underline{F}}}\, \underline{\underline{F}}^{-1} \right] dt \geqslant 0.$$

provided $\tau = \infty$. If it is permitted to go to the limit of infinitesimal mass elements this inequality remains true for any finite τ because one can continue the process after any time τ in such a fashion that the integral from $t = \tau$ to $t = \infty$ goes to zero as the mass element tends to zero. It is only to be assumed that for fixed values of u and $\underline{\underline{F}}$ and with $\underline{q} = 0$ after $t = \tau$ the material element approaches a unique thermostatic equilibrium state.

This inequality will be called the fundamental inequality.

4. Constitutive equations

Besides the balance equations which hold for all materials at all times we have now to introduce laws which distinguish one material from another. In thermostatics we have the equations of state. In thermodynamics we have to introduce an appropriate definition of the state **at**

time t before setting up something like equations of state. The problem is to give as much information on the state at time t as is necessary to make predictions on the behavior of the thermodynamic system at later times. It is, for that purpose, irrelevant what kind of information we invoke if it is only complete. There are two extremes: knowledge of a sufficient set of state variables (u, $\underline{q}$, $\underline{\underline{F}}$ and suitable internal variables) at time t, or knowledge of u, $\underline{q}$, $\underline{\underline{F}}$ only at time t and at all previous times, i.e. the history of these variables with respect to time t. If we adopt the second point of view, then we characterize the state at time t by explaining the treatment in $-\infty < s \leqq t$ by which it has been produced. Of course, we permit here a loss of generality because one might admit that the state in a mass element depend also on the hystory of its neighborhood. Thus we single out a special class of materials which we call simple materials. For such materials all quantities depend on the histories u(s), $\underline{q}(s)$, $\underline{\underline{F}}(s)$ in $-\infty < s \leqq \infty$ and consequently there exist constitutive equations

$$\frac{1}{T_{st}(t)} - \frac{1}{T(t)} = \mathcal{F}_0 \{\dots\} , \tag{4.1}$$

$$\nabla \frac{1}{T(t)} = \mathcal{F}_1 \{\dots\} , \tag{4.2}$$

$$\frac{\underline{\underline{\sigma}}(t)}{T(t)} - \frac{\underline{\underline{\sigma}}_{st}(t)}{T_{st}(t)} = \mathcal{F}_2 \{\dots\} \tag{4.3}$$

with

$$\{\dots\} = \{u(s) , \underline{q}(s) , \underline{\underline{F}}(s) ; \quad -\infty < s < t\} \tag{4.4}$$

with functionals $\mathcal{F}_i$ which must satisfy the fundamental inequality for all admissible thermodynamic processes.

These functionals can be somewhat simplified by applying the principle of material frame indifference. Further simplifications are possible for isotropic materials . But on the whole a description of the material properties in processes by such functionals is not very satisfactory. For one thing one would need a whole library to put down these properties for a single material considering the big manifold of possible histories. And it also seems that the contents of such libraries for two pieces of a material which are produced under the same technical condition would not quite agree. So if one must use such a library in a technical application, one would also need data on the fluctuations of the entries in the library which exist for materials produced under the same technical conditions.

Although the practical value of such constitutive equations seems to be relatively small, the theory as such is pretty good. And it permits specialization to special classes of materials which are simpler than the simple materials and also to special conditions. For instance the functionals reduce to linear functionals if only small departures from a reference state are to be considered. Then the powerful superposition principle is available and the material is characterized by three after effect functions associated with each equilibrium reference state in the relevant domain.

Simplifications of a different kind occur if the processes are slow on some time scale τ . If we assume some kind of fading memory, which roughly means that the previous values $u(s)$, $\underline{q}(s)$, $\underline{\underline{F}}(s)$ enter

the functionals with decreasing weight when S decreases , then it may suffice for slow processes to know u(t), $\underline{F}(t)$ only, or in a better approximation $\dot{u}(t)$, $\dot{\underline{F}}(t)$, $\underline{q}(t)$ in addition in order to know the values of all other variables. Then the functionals reduce to functions of u(t), F(t) or u(t), F(t) ; $\dot{u}(t)$, $\underline{q}(t)$, $\dot{\underline{F}}(t)$. If the functionals $\mathfrak{F}_i$ are exactly of this form, then we speak of a material of the differential type and of complexity 0 or 1, respectively. Otherwise, the material is approximated by a material of the differential type and of complexity 0 or 1 for sufficiently slow processes.

The materials of differential type and of complexity one are of particular interest. They satisfy the fundamental inequality trivialy because not only the integral in (3.9) is non-negative but already the integrand itself *) . As a consequence the left member in (3.7) is non-negative and the static entropy function $S_{st}(u, \underline{F})$ satisfies the Clausius-Duhem-inequality (1.3) . In this case the constitutive equations are just the same as those given in classical TIP . This result explains clearly, also in more general cases, the position of classical TIP within the framework of the new thermodynamics of processes.

5. Discussion

We try to clarify somewhat the notion of memory . Of course, the material itself has no memory and we cannot consult it at time f for

---

*) The proof has been given by Jürgen Keller (Aachen) and is contained in the Appendix of [16]

its past treatment at times $t - s$ $(0 < s < \infty)$. Although the state at time t is determined by the histories $u(t-s)$, $\underline{q}(t-s)$, $\underline{\underline{F}}(t-s)$ in $0 \leqq s < \infty$ the converse need not be true. There may be families of histories with an infinite number of specimens which lead to the same state at time t and the most one can infer from the present state about the past treatment is that it is one specimen of such a family. We illustrate this by a simple example which is the well-known dynamical equation of state

$$\sigma + \tau_1 \frac{d\sigma}{dt} = E_o \left(\varepsilon + \tau_2 \frac{d\varepsilon}{dt} \right). \tag{5.1}$$

Here σ is the one-dimensional stress and ε is the one-dimensional strain. The coefficient $E_o > 0$ is the elastic modulus in equilibrium and τ_1, τ_2 are the relaxation times at constant strain or stress, respectively, with $\tau_2 > \tau_1 > 0$. The stress can be written as a linear functional of the strain by integrating (5.1) with appropriate initial conditions, $\varepsilon \to 0$ and σ bounded as $t \to -\infty$. The result is

$$\sigma(t) = E_o \frac{\tau_2}{\tau_1} \varepsilon(t) - E_o \frac{\tau_2 - \tau_1}{\tau_1^2} \int_{-\infty}^{t} \exp\left(\frac{s-t}{\tau_1}\right) \varepsilon(s)\, ds. \tag{5.2}$$

It is easily seen that the solution of (5.1) with prescribed $\varepsilon(t)$ in $t \geqq t_o$ and given $\sigma(t_o)$ is unique. The past history $\varepsilon(s)$ in $-\infty < s \leqq t_o$ has then to satisfy (5.2) with $t = t_o$. If $\varepsilon(s)$ and $\varepsilon^*(s)$ are two solutions of (5.2), then one obtains

$$\int_o^{\infty} \exp\left(-\frac{s}{\tau_1}\right) \left[\varepsilon^*(t_o - s) - \varepsilon(t_o - s)\right] ds = 0. \tag{5.3}$$

If $\varepsilon(s)$ in $-\infty$ in $-\infty < s \leqslant t_o$ is one possible history leading to $\sigma(t_o)$, $\varepsilon(t_o)$, then it is obvious from (5.3) that there is an infinite number of other possible histories $\varepsilon^*(s)$ leading to $\sigma(t_o)$, $\varepsilon(t_o)$.

We can say that the histories are parametrized by the "state parameter"

$$\xi(t) = \int_{-\infty}^{t} \exp\left(\frac{s-t}{\tau_1}\right) \varepsilon(s)\, ds \quad . \tag{5.4}$$

Then

$$\sigma(t) = E_o \frac{\tau_2}{\tau_1} \varepsilon(t) - E_o \frac{\tau - \tau_1}{\tau_1^2} \xi(t) \quad , \tag{5.5}$$

$$\xi(t) + \tau_1 \dot{\xi}(t) = \tau_1 \varepsilon(t) \, . \tag{5.6}$$

The memory can therefore not be considered as an objective property of the material; it is rather the knowledge of the past treatment that the experimenter needs if he wishes to make predictions on the future behavior of the material under given external action, and it has therefore subjective character. This is the more so because the experimenter can at will use additional information about the state of the material at time t, for instance the value of an internal variable $\xi(t)$. Apparently, this together with the histories $u(t-s)$, $\underline{q}(t-s)$, $\underline{\underline{F}}(t-s)$ would give redundant information on the state at time t. This fact enables one to ignore part of the "memory". Practically this works out so that the original memory function may be reduced to a shorter range memory function.

We speak of the lowest level of description if the information about the state at time t is contained in the knowledge of the histories $u(t-s)$, $\underline{q}(t-s)$, $\underline{\underline{F}}(t-s)$ in $0 \leqslant s < \infty$. Assume that practical-

ly the memory reaches back to time $t - \tau^{(o)}$ so that the histories before this time have a negligible influence in the functionals (4.1) to (4.3). A higher level of description would consist in the introduction of internal variables ξ_k and their histories $\xi_k(t-s)$ with $k = 1, 2, \ldots, n$. Then there would be additional functional equations for the internal variables of which (5.6) is a very special case. But the memory in all the functionals would reach back practically only to time $t - \tau^{(n)}$ with $\tau^{(o)} > \tau^{(1)} > \ldots \tau^{(n)}$. If we now consider processes which are slow on the time scale $\tau^{(n)}$ - they can be much faster than processes which are slow on the time scale $\tau^{(o)}$ - it is possible to replace the memory by the values of $u(t)$, $\underline{\underline{F}}(t)$, $\xi_k(t)$ and of $\dot{u}(t)$, $\underline{q}(t)$, $\dot{\underline{\underline{F}}}(t)$ $\dot{\xi}_k(t)$; that is to approximate the material on this higher level of description by a material of differential type and of complexity one. Then one can again prove by Keller's theorem [16] that the integrand of a modified fundamental inequality, which has additional terms pertaining to the internal variables, is itself non-negative. Therefrom it is concluded that (2.8) is a possible entropy function during the process and the laws of classical TIP are recovered.

Bibliography

[1] J. Meixner and H. G. Reik : Thermodynamik der irreversiblen Prozesse. Handbuch der Physik, Vol. III/2, 413-523, Ed. S. Flügge, Berlin-Göttingen-Heidelberg : Springer 1959.

[2] S. R. de Groot and P. Mazur : Non-equilibrium thermodynamics. Amsterdam : North-Holland Publishing Co. 1962.

[3] R. Haase : Thermodynamik der irreversiblen Prozesse. Darmstadt : Dr. Dietrich Steinkopff 1963.

[4] B. D. Coleman : Archive Rational Mech. Anal. 17 , 1, 1964.

[5] B. D. Coleman : Archive Rational Mech. Anal. 17 , 1, 230, 1964.

[6] C. Truesdell and W. Noll : The non-linear field theories of mechanics. Handbuch der Physik Vol. III/2, 1-602. Ed. S. Flügge, Berlin-Heidelberg-New York : Springer 1965.

[7] C. Truesdell : Thermodynamics for Beginners, p. 373-389, in : Irreversible aspects of continuum mechanics and transfer of physical characteristics in moving fluids. IUTAM Symposia Vienna, June 1966. Ed. H. Parkus and L. I. Sedov. Wien and New York : Springer 1968.

[8] H. König and J. Meixner : Math. Nachr. 19 , 265-323 (1959)

[9] J. Meixner ; Zeitschrift für Physik 156 , 200-210, 1959.

[10] J. Meixner : Thermodynamic Theory of Relaxation Phenomena, p. 73-89, in : Non-equilibrium thermodynamics, variational techniques and stability. Ed. R. Donnely, R. Herman and I. Prigogine, Chicago University Press, 1965.

[11] J. Meixner : Thermodynamics of electrical networks, p. 13-25, in : Proceedings of the Symposium on Generalized Net-

works, Microwave Symposia Series XVI. Polytechnic Press of the Polytechnic Institute of Brooklyn. Brooklyn, N. Y. 1966.

[12] J. Meixner : J. Applied Mechanics 33, Series E, 481-488, 1966.

[13] J. Meixner : Zeitschrift für Physik 193, 366, 1966

[14] J. Meixner : Rheologica Acta 7, 8, 1968.

[15] J. Meixner : Beziehungen zwischen Netzwerktheorie und Thermodynamik. Sitzungsber. der Arbeitsgemeinschaft für Forschung des Landes NRW. Köln and Opladen : Westdeutscher Verlag 1968.

[16] J. Meixner : Zeitschr. f. Physik 219, 79, 1969.

[17] J. Meixner : Archive Rational Mech. Anal. 33, 33, 1969.

[18] R. Clausius : Die mechanische Wärmetheorie, I. Band, 2. Aufl., p. 206 and p. 224. Braunschweig : Vieweg & Sohn, 1976.

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C. I. M E)

A. C PIPKIN

NON-LINEAR PHENOMENA IN CONTINUA

Corso tenuto a Bressanone dal 3 all'11 settembre 1970

NON-LINEAR PHENOMENA IN CONTINUA

by
A. C. PIPKIN
Brown University

0. Introduction

These lectures will consist of a sequence of problems, mainly from the theories of large elastic deformation and viscoelastic flow. In these theories, properties of materials are described by equations that are not linear. They are usually so badly non-linear that it is just as well to regard the functions that appear in them as almost arbitrary.

The prospect of a non-linear problem is frightening because the standard methods based on superposition are inapplicable. The fright is psychological; no one solves linear problems either, by pencil and paper, unless the problem has a great deal of symmetry. The use of symmetry is usually not acknowledged. It is regarded as cheating.

All of the problems to be considered here are highly symmetrical. The symmetries most easily recognized are translational or rotational. Scale invariance is still important in non-linear problems, but less easy to perceive. Invariance that removes a variable before you even thought of it is of course not recognized, but it doesn't need to be.

Symmetry is used in solving problems by invoking the fact that the complete solution must have all of the symmetry that the statement of the problem has. In the ideal case, only one candidate has all of the necessary symmetries, so it is the solution. Often, a stronger hypothesis is used: we claim that the problem has only

one solution, so that particular solution by itself has all the symmetries of the data. The phrase, "By symmetry", means "By symmetry, assuming uniqueness".

Even this stronger hypothesis is ordinarily not enough to single out just one possibility. Constraints are useful. Anything that limits the class of candidates makes it easier to pick out the solution. In continuum mechanics, the constraint of incompressibility is widely applicable.

Linearization is, of course, the main way of solving less symmetrical problems (by pencil and paper). I will not say much about perturbation methods in spite of their obvious importance, because the equations are usually too long to remember or to work out in the midst of a short lecture.

I will also say little about numerical methods. When material properties are given in terms of empirically determined functions, it goes without saying that solving problems requires numerical work. The task of the analyst is to reduce this to a minimum, not to avoid it entirely.

With apologies to all whose work I have stolen, I give no references except to the most recent literature. References to most of the original sources can be found in several books [1,2,3, 4]*. Whatever is original will be easy to recognize: the mistakes are all mine.

---

* Numbers in square brackets indicate references listed at the end of the paper.

1. Elementary Problems of Finite Elastic Deformation

The theory of finite elastic deformations is a relatively simple non-linear theory because all relevant material properties are bound up in a single scalar function W, the elastic stored energy density. This gives the theory a great deal of coherence compared to, say, the theory of large viscoelastic deformations. It also makes possible the use of energy methods to simplify the analysis and to give a neater picture of what is involved in a problem.

Treloar's [5] book is the best reference on physical aspects of the subject. There are several fine books [1,2,3,4] on the mathematical theory. These books rightly emphasize the general theory and the techniques used in solving hard problems. To try to dispel the idea that finite elasticity is a very difficult subject, I will confine attention to dead easy problems.

1.1. Strings

Many of the basic notions of finite elasticity theory can be illustrated in terms of its simplest special case, the theory of elastic strings. A string resists extension but has no bending stiffness. Its mechanical behavior is described by a function $f(\lambda)$, the tensile force f required to stretch the string to λ times its initial, unstressed length. The elastic stored energy $w(\lambda)$, measured per unit of initial length, is related to $f(\lambda)$ through the assumption that the stored energy is equal to the work done in stretching the string; $dw = fd\lambda$, or $f(\lambda) = w'(\lambda)$.

If a string is stretched inhomogeneously, so that a particle initially at a distance X from one end moves to a distance x(X) from that end (measured along the string), then the stretch at the particle X is $\lambda = dx/dX$. We assume that the tensile force there is $f[\lambda(X)]$, just as if the string were stretched uniformly.

Let us determine the equilibrium configuration of a string that is stretched out between two points A and B on the surface of a smooth, rigid, convex body. Let L_o be the unstretched length. The total energy of the stretched string is

$$E = \int_0^{L_o} w[x'(X)]dX \ . \qquad (1.1.1)$$

(Notice the convenience of specifying w as energy per unit initial length.)

To find the configuration that minimizes E, first choose any admissible path from A to B, and vary the function x(X), which describes how the string is stretched along that path. The Euler equation characterizing the minimizing function is

$$\frac{d}{dX}\frac{dw}{dx'} = 0 \ . \qquad (1.1.2)$$

Thus, since dw/dx' is the tensile force f, the tension is constant.

Under loose assumptions, λ must be constant if f is constant. Then if L is the length of the path considered, λ is equal to L/L_o at every particle, and the energy is $E = w(L/L_o)L_o$.

Now consider variations of the path. Under obvious assumptions, E (or w) is minimized by choosing the shortest path,

which minimizes the stretch.

Thus, in equilibrium, the string lies along a geodesic from A to B, and it is stretched uniformly. We can say that the problem is solved, although neither the solution nor even the problem itself has been described in all detail.

Notice that nothing could have been gained by specifying the form of $w(\lambda)$, and that we conceivably might even have failed to understand the solution if confused by a complicated specific function, or a specific surface shape.

1.2. Membranes

The two-dimensional analog of the string is the membrane. Consider a patch of membrane, initially a unit square. Suppose that it is stretched into a rectangle of dimensions λ_1 and λ_2. If the membrane is homogeneous and isotropic, this can be done by applying uniformly distributed normal forces to its edges. Let f_1 and f_2 be their resultants; f_1 acts on the edge of length λ_2. If $w(\lambda_1,\lambda_2)$ is the stored energy, and it is equal to the work done in deforming the membrane, then

$$dw = f_1 d\lambda_1 + f_2 d\lambda_2 \ . \qquad (1.2.1)$$

Hence, $f_\alpha = \partial w/\partial\lambda_\alpha$. The principal tensions, or forces per unit current length, are $t_1 = f_1/\lambda_2 = \lambda_1 f_1/A$ and $t_2 = \lambda_2 f_2/A$, where A is the area per unit initial area.

Consider the inflation of a spherical balloon from an initial radius R to a final radius r We wish to find the relation

between r and the excess internal pressure Δp.

Each element of membrane is stretched equally in orthogonal directions, the extension ratios being $\lambda_1 = \lambda_2 = r/R = \lambda$ (say). The energy of the inflated balloon is $w(\lambda,\lambda)$ per unit of initial area, so its total energy is $4\pi R^2 w(\lambda,\lambda)$. By setting the change of energy equal to the work $\Delta p\, dV$ done in increasing the enclosed volume V, we obtain

$$\Delta p\ 4\pi r^2 dr = 4\pi R^2 dw \ . \tag{1.2.2}$$

With $r = R\lambda$ and $dr = Rd\lambda$, this gives

$$\Delta p = \frac{1}{R\lambda^2}\frac{d}{d\lambda} w(\lambda,\lambda) \ . \tag{1.2.3}$$

For a soap bubble, w is proportional to the area λ^2 (per unit initial area), so the pressure drops like $1/\lambda$ as the radius grows. Rubber behaves this way at moderate extensions. As λ increases from unity, the pressure at first increases, but then at larger λ the energy w is roughly proportional to λ^2, so the pressure begins to decrease. It is a matter of common experience that it takes more pressure to inflate a balloon a little than to enlarge it further after it is started. At very large stretches, the pressure increases again as the long-chain rubber molecules approach their maximum extensibility.

Two balloons connected by a pipe, and thus under the same pressure excess Δp, need not be equally inflated. Experience indicates that under more than slight inflation, unequal inflation is more stable than equal inflation. This is a minor example of the non-uniqueness that can be expected in non-linear problems.

1.3. Blocks

Let $W(\lambda_1\ \lambda_2,\lambda_3)$ be the elastic stored energy of a body, initially a unit cube, that is stretched into a block of dimensions $\lambda_1,\lambda_2,\lambda_3$. If the block is homogeneous and isotropic, this deformation will require normal forces only. Let f_α be their resultants. The work-energy relation is

$$dW = f_1 d\lambda_1 + f_2 d\lambda_2 + f_3 d\lambda_3 \ . \tag{1.3.1}$$

Thus, the nominal stresses, or engineering stresses, are given in terms of W by $f_\alpha = \partial W/\partial\lambda_\alpha$. The true stresses σ_α, the forces per unit current area, are

$$\sigma_1 = f_1/\lambda_2\lambda_3 = \lambda_1 f_1/J \ , \tag{1.3.2}$$

and so on. J is equal to $\lambda_1\lambda_2\lambda_3$, the volume per unit initial volume.

Since it is much easier to change the shape of a piece of rubber than to change its volume, in many problems it is permissible to use the idealization that the material is incompressible. For the block considered here, this means that $\lambda_1\lambda_2\lambda_3$ equals unity in all admissible deformations. With the increments $d\lambda_\alpha$ then subject to the constraint $\Sigma d\lambda_\alpha/\lambda_\alpha = 0$, it no longer follows from (1.3.1) that f_α is equal to $\partial W/\partial\lambda_\alpha$. Instead, by introducing the constraint with a Lagrange multiplier p, we obtain

$$f_\alpha = \frac{\partial W}{\partial\lambda_\alpha} - \frac{p}{\lambda_\alpha} \tag{1.3.3}$$

and, since J=1,

$$\sigma_\alpha = \lambda_\alpha \frac{\partial W}{\partial \lambda_\alpha} - p \ . \tag{1.3.4}$$

The multiplier p represents the reaction pressure arising from the constraint.

Analytically simple forms of W that are often used for qualitative illustrations are the neo-Hookean form

$$W = \tfrac{1}{2} G(I_1-3) \tag{1.3.5}$$

and the Mooney form

$$W = C_1(I_1-3)+C_2(I_2-3), \tag{1.3.6}$$

where

$$I_1 = \Sigma\lambda_\alpha^2 \quad \text{and} \quad I_2 = \Sigma\lambda_\alpha^{-2} \ . \tag{1.3.7}$$

Neither makes any sense unless the material is incompressible. If it is, both yield a minimum energy W=0 at λ_α =1, provided that the constants are positive.

In isotropic materials, W is a symmetric function of the stretches, and thus a function of, say, I_1,I_2, and J, or just I_1 and I_2 if the material is incompressible. With the notation $\partial W/\partial I_\alpha = W_\alpha$, (1.3.4) becomes

$$\sigma_\alpha = 2W_1\lambda_\alpha^2 - 2W_2\lambda_\alpha^{-2} - p \ . \tag{1.3.8}$$

1.4. Strings, Membranes, and Blocks

The energy functions $w(\lambda)$ and $w(\lambda_1,\lambda_2)$, for strings and membranes, are, of course, related to the energy density for a three-dimensional body of the same material.

The membrane energy $w(\lambda_1,\lambda_2)$ is related to $W(\lambda_1,\lambda_2,\lambda_3)$ by

$$w(\lambda_1,\lambda_2) = hW(\lambda_1,\lambda_2,1/\lambda_1\lambda_2) , \qquad (1.4.1)$$

where h is the initial thickness of the membrane; provided, that the material is incompressible, so that the stretch λ_3 is equal to $1/\lambda_1\lambda_2$.

Similarly, for a string of incompressible material, initially of cross-sectional area A, the energy per unit initial length is

$$w(\lambda) = AW(\lambda,\lambda^{-1/2},\lambda^{-1/2}) . \qquad (1.4.2)$$

With the neo-Hookean form of W, $w(\lambda)$ is

$$w(\lambda) = (AG/2)(\lambda^2+2\lambda^{-1}-3) , \qquad (1.4.3)$$

and the force-stretch relation is

$$f(\lambda) = w'(\lambda) = AG(\lambda-\lambda^{-2}). \qquad (1.4.4)$$

The modulus G can be adjusted to agree with data near $\lambda=1$. If this is done, agreement with data for rubber is still reasonably good up to $\lambda=2$ (100% stretch). The linear increase of f for large λ is not observed; the force actually increases sharply when the rubber approaches its maximum extensibility at λ of the order of 5 or 10.

1.5. Stability of a Block

If a block of incompressible material is stable in the undeformed state $\lambda_\alpha = 1$ when no forces are applied to it, it remains

stable under arbitrarily large positive or negative pressure loading. For, the total energy of the block and the loading system is W-pV; but the second term is irrelevant because V cannot vary if the block is incompressible.

In contrast, consider a block under equal normal dead loads; the force resultants f_α are maintained at the value f, however the block may deform. We admit no deformations except homogeneous deformations, which are described by the stretches λ_α. Then for sufficiently large loads f, the undeformed state $\lambda_\alpha=1$ is unstable, if the material is incompressible.

To see why, consider the energy, including that of the loads:

$$E = W-f\Sigma\lambda_\alpha. \tag{1.5.1}$$

Let values at an equilibrium state λ_α^o be marked with a zero superscript, and let $\Delta\lambda_\alpha = \lambda_\alpha-\lambda_\alpha^o$. Then near equilibrium,

$$E = E^o + \Sigma E_\alpha^o\Delta\lambda_\alpha + \frac{1}{2}\Sigma\Sigma E_{\alpha\beta}^o\Delta\lambda_\alpha\Delta\lambda_\beta + \dots , \tag{1.5.2}$$

where $E_\alpha = \partial E/\partial\lambda_\alpha$. The constraint condition $\lambda_1\lambda_2\lambda_3 = 1$ gives

$$0 = \operatorname{Ln}\Pi(\lambda_\alpha^o+\Delta\lambda_\alpha) = \Sigma(\Delta\lambda_\alpha/\lambda_\alpha^o) - \frac{1}{2}\Sigma(\Delta\lambda_\alpha/\lambda_\alpha^o)^2 + \dots \tag{1.5.3}$$

Hence, with a Lagrange multiplier μ ,

$$E-E^o = \Sigma(E_\alpha^o+\mu/\lambda_\alpha^o)\Delta\lambda_\alpha + \frac{1}{2}\Sigma\Sigma(E_{\alpha\beta}^o-\mu\delta_{\alpha\beta}/\lambda_\alpha^o\lambda_\beta^o)\Delta\lambda_\alpha\Delta\lambda_\beta+ \dots \tag{1.5.4}$$

Since E must be stationary at equilibrium, the sum linear in $\Delta\lambda_\alpha$ must vanish:

$$E^{o}_{\alpha} + \mu/\lambda^{o}_{\alpha} = 0 \ . \qquad (1.5.5)$$

Then, with this value of μ,

$$E-E^{o} = \frac{1}{2}\,\Sigma\Sigma(E^{o}_{\alpha\beta}+E^{o}_{\alpha}\delta_{\alpha\beta}/\lambda^{o}_{\beta})\Delta\lambda_{\alpha}\Delta\lambda_{\beta}+ \ \ldots \ . \qquad (1.5.6)$$

The equilibrium is stable (toward homogeneous alternatives) if the quadratic form is strictly positive, and unstable if it can be negative. If E has the form (1.5.1), and $\lambda^{o}_{\alpha} = 1$, the matrix of coefficients is

$$E^{o}_{\alpha\beta}+E^{o}_{\alpha}\delta_{\alpha\beta}/\lambda^{o}_{\beta} = W^{o}_{\alpha\beta}+ (W^{o}_{\alpha}-f)\delta_{\alpha\beta}. \qquad (1.5.7)$$

Whatever values W^{o}_{α} and $W^{o}_{\alpha\beta}$ may have, this is negative definite for sufficiently large f, and the undeformed state is unstable.

Now consider the alternative, stable, equilibrium state that must exist when the undeformed state is unstable. The equilibrium equations, obtained by eliminating μ from (1.5.5), are

$$\lambda_1(W_1-f) = \lambda_2(W_2-f) = \lambda_3(W_3-f) \ . \qquad (1.5.8)$$

If we take $\lambda_2 = \lambda_3$, the latter of these two equations is satisfied since W is a symmetric function of the stretches. In that case, $\lambda_1 = \lambda$, say, and $\lambda_2 = \lambda_3 = \lambda^{-1/2}$. Then

$$\frac{d}{d\lambda}\, W(\lambda,\lambda^{-1/2},\lambda^{-1/2}) = W_1-\lambda^{-3/2}W_2 \ , \qquad (1.5.9)$$

and the first equation in (1.5.8) can be written as

$$dW/d\lambda = (1-\lambda^{-3/2})f \ . \qquad (1.5.10)$$

The left-hand member is the force $f(\lambda)$ in a "string" of unit initial cross-sectional area. If it grows without bound as λ

increases, (1.5.10) surely has a solution λ , provided that f is so large that the right-hand member initially (at $\lambda=1$) grows more rapidly than the left-hand member does.

Without doubt, this artificial example is a little silly, but I do think that it shows again that little or nothing can be gained by stating the form of W at the outset. It seems that we must know what function W is in order to solve (1.5.10), but I doubt that I could have reached that stage of the analysis if I had known W in the first place.

It should also be pointed out that it is trivial to solve (1.5.10) graphically, given experimental values of $dW/d\lambda$. Analytical solution of (1.5.10) is neither necessary nor even desirable.

1.6. Expansion of a Spherical Shell

For an example of an inhomogeneous deformation of a three-dimensional body, let us consider the expansion of a spherical shell, initially of internal radius R_0 and external radius R_1, under internal and external pressures p_0 and p_1, respectively. We take the material to be incompressible. Notice how the symmetry and the constraint, taken together, make the problem trivial.

Since the material is incompressible, the final radius r(R) of a membrane initially of radius R is determined by the condition that the volume of the shell between r(R) and $r_0 = r(R_0)$ is the same as that between R and R_0:

$$r^3 = r_o^3 + R^3 - R_o^3 . \qquad (1.6.1)$$

We are mainly interested in determining the relation between r_o and $p_o - p_1$.

We treat the shell as a set of nested membranes, each of initial thickness dR. The energy per unit initial area of each such membrane is $w(\lambda,\lambda) = W(\lambda,\lambda,\lambda^{-2})dR$, where $\lambda = r/R$. We know that the pressure drop across one such membrane is

$$-dp = \frac{1}{R\lambda^2}\frac{d}{d\lambda} W(\lambda,\lambda,\lambda^{-2})dR . \qquad (1.6.2)$$

Hence, the total inflation pressure is

$$p_o - p_1 = \int_{R_o}^{R_1} \frac{1}{R\lambda^2}\frac{dW}{d\lambda}\, dR . \qquad (1.6.3)$$

By using $r = \lambda R$ in (1.6.1), we obtain

$$R^3(\lambda^3 - 1) = R_o^3(\lambda_o^3 - 1). \qquad (1.6.4)$$

By using this relation to express R in terms of λ in the integral, we obtain

$$p_o - p_1 = P(\lambda_o) - P(\lambda_1) , \qquad (1.6.5)$$

where

$$P(\lambda) = \int_1^{\lambda} \frac{1}{\lambda^3 - 1}\frac{d}{d\lambda} W(\lambda,\lambda,\lambda^{-2})d\lambda . \qquad (1.6.6)$$

The function $P(\lambda)$ is the excess pressure required to stretch the internal radius of a spherical cavity in an infinite body by the amount λ .

The function $P(\lambda)$ is the same for all spheres of a given material. If it is determined empirically by tests on one sphere, results for all other spheres (all values of R_0 and R_1) are given by (1.6.5). This is an example of modelling, which usually cannot be carried out purely by dimensional analysis if the theory is highly non-linear.

1.7 Tube Forming

As another elementary example, let us consider a particularly simple special case of the problem of flexure. A slab is bent around until opposite edges meet, whereupon they are joined so that a cylindrical tube is formed, and the tube is left in a state of self stress. The problem is to determine the dimensions of the tube from those of the slab.

Let the slab be bounded by the planes $X = \pm H$, with the ends that are to be joined at $Y = \pm L$. We will suppose that the slab is very long in the Z-direction, and, to make the example simpler, ignore that coordinate.

Symmetry suggests that the deformation should carry planes $Y =$ constant onto planes $\theta =$ constant, in cylindrical coordinates, and that planes $X =$ constant should be bent into cylinders $r =$ constant. We suppose that the Z-coordinate of each particle is unaltered. The relation between r and X is determined by the incompressibility condition; the volume of material initially between the planes X=0 and X=const. must be the same as that between cylinders of radii $r_0 = r(0)$ and $r = r(X)$ after the deformation. Thus, the final cylindrical coordinates of a particle are

given in terms of its original cartesian coordinates by

$$\pi(r^2-r_o^2) = 2LX,\ \theta = \pi Y/L,\ z = Z\ . \qquad (1.7.1)$$

Symmetry and the constraints (incompressibility and the stipulation of no elongation in the Z-direction) have determined the deformation to within an unknown parameter r_o. The parameter r_o is to be determined by minimizing the energy.

An elementary cube initially bounded by coordinate surfaces is stretched into a brick-shaped element bounded by the cylindrical coordinate surfaces. The stretch ratios are

$$\lambda_1 = dr/dX = L/\pi r,\ \lambda_2 = rd\theta/dY = \pi r/L,\ \lambda_3 = dz/dZ = 1. \qquad (1.7.2)$$

Let us write λ for λ_2. Then λ_1 is λ^{-1}, and the energy per unit of initial volume is $W(\lambda,\lambda^{-1},1)$. From (1.7.1), the variation of λ with X is given by

$$\lambda^2 = \lambda_o^2 + 2\pi X/L\ . \qquad (1.7.3)$$

The total energy, per unit of length in the Z-direction, is

$$E = 2L \int_{-H}^{H} W\, dX\ . \qquad (1.7.4)$$

This is to be minimized with respect to r_o, or $\lambda_o\ (=\pi r_o/L)$. By treating W as a function of λ^2, for convenience, we obtain

$$\int_{-H}^{H} \frac{dW}{d(\lambda^2)} \frac{\partial(\lambda^2)}{\partial(\lambda_o^2)}\, dX = 0\ . \qquad (1.7.5)$$

But $\partial(\lambda^2)/\partial(\lambda_o^2)$ is unity. Now, also,

$$\frac{\partial W}{\partial X} = \frac{dW}{d(\lambda^2)} \frac{2\pi}{L} . \tag{1.7.6}$$

Hence, the integrand in (1.7.5) is proportional to $\partial W/\partial X$. Integration gives

$$W|_{X=H} = W|_{X=-H} , \tag{1.7.7}$$

and this is the equation that must be solved for r_o (or λ_o).

Now, since $W(\lambda,\lambda^{-1},1)$ depends symmetrically on λ and λ^{-1} if the material is isotropic, (1.7.7) is satisfied if the value of λ at X=H is the same as the value of λ^{-1} at X=-H. Thus, with (1.7.3) we obtain

$$(\lambda_o^2+2\pi H/L)(\lambda_o^2-2\pi H/L) = 1. \tag{1.7.8}$$

Hence,

$$\lambda_o^4 = 1+(2\pi H/L)^2 . \tag{1.7.9}$$

Thus, the desired information is obtained without knowing anything at all about W, except its symmetry.

1.8 Tensor Stress-Strain Relations

In any deformation, homogeneous or not, through each particle there are three fibers that are perpendicular both before and after the deformation. We call them principal fibers, and their directions are called principal directions of strain. In isotropic elastic materials, the stresses on surface elements perpendicular to such directions are purely normal stresses (by symmetry).

In some simple problems, such as the inflation of a spherical shell, the principal directions can be located by inspection. More often this is not the case, even when the deformation is completely specified. To avoid the ingenuity or work required to compute principal directions, we use stress-strain relations in tensor form.

Let $\underline{u}^1, \underline{u}^2$, and $\underline{u}^3$ be unit vectors along the principal directions. With principal stresses σ_α given by (1.3.8), the total stress $\underline{\sigma}$ is

$$\underline{\sigma} = \Sigma \sigma_\alpha \underline{u}^\alpha \underline{u}^\alpha = 2W_1 \underline{g} - 2W_2 \underline{g}^{-1} - p\underline{I} \ , \tag{1.8.1}$$

where

$$\underline{g} = \Sigma \lambda_\alpha^2 \underline{u}^\alpha \underline{u}^\alpha \ , \ \underline{g}^{-1} = \Sigma \lambda_\alpha^{-2} \underline{u}^\alpha \underline{u}^\alpha, \ \underline{I} = \Sigma \underline{u}^\alpha \underline{u}^\alpha \ . \tag{1.8.2}$$

The usefulness of this relation lies in the fact that the strain, $\underline{g}$, can be computed directly, without first finding λ_α and $\underline{u}^\alpha$. Let $\underline{x}(\underline{X})$ be a deformation carrying the particle initially at $\underline{X}$ to the place $\underline{x}$. A fiber whose initial span is $d\underline{X}$ is mapped onto

$$d\underline{x} = \underline{F} \, d\underline{X} \ . \tag{1.8.3}$$

If the three principal fibers initially lie along the orthogonal directions $\underline{v}^\alpha$, then the deformation gradient $\underline{F}$ must have the form

$$\underline{F} = \Sigma \lambda_\alpha \underline{u}^\alpha \underline{v}^\alpha \ . \tag{1.8.4}$$

Hence, we see that

$$\underline{F}\underline{F}^T = \Sigma \lambda_\alpha \underline{u}^\alpha \underline{v}^\alpha \cdot \Sigma \lambda_\beta \underline{v}^\beta \underline{u}^\beta = \Sigma \lambda_\alpha^2 \underline{u}^\alpha \underline{u}^\alpha = \underline{g} \ . \tag{1.8.5}$$

Now, $\underline{F}$ can be computed directly; in cartesian coordinates, its components are

$$F_{iA} = x_{i,A} = \partial x_i / \partial X_A \ . \tag{1.8.6}$$

Thus, the components of $\underline{g}$ are

$$g_{ij} = F_{iA}F_{jA} = x_{i,A}x_{j,A} \ . \tag{1.8.7}$$

1.9. Shear

Even in such a trivial deformation as simple shear,

$$x = X + \kappa Y, \quad y = Y, \quad z = Z, \tag{1.9.1}$$

the principal directions in the x,y plane are not obvious. However, the deformation gradient $\underline{F}$ and strain $\underline{g}$ can be written down at sight:

$$\underline{F}^T = \nabla x = \underline{\nabla}(\underline{X} + \kappa Y \underline{i}) = \underline{I} + \kappa \underline{j}\underline{i} \ , \tag{1.9.2}$$

$$\begin{aligned} \underline{g} &= \underline{F}\underline{F}^T = (\underline{I} + \kappa \underline{i}\underline{j}) \cdot (\underline{I} + \kappa \underline{j}\underline{i}) \\ &= \underline{I} + \kappa(\underline{i}\underline{j} + \underline{j}\underline{i}) + \kappa^2 \underline{i}\underline{i} \ . \end{aligned} \tag{1.9.3}$$

The matrix of components of $\underline{g}^{-1}$ is the matrix of cofactors of $\underline{g}$, since $\underline{g}$ is symmetric and its determinant is unity:

$$\underline{g}^{-1} = \underline{I} - \kappa(\underline{i}\underline{j} + \underline{j}\underline{i}) + \kappa^2 \underline{j}\underline{j} \ . \tag{1.9.4}$$

Thus, from (1.8.1), the stress is

$$\underline{\sigma} = 2(W_1 + W_2)\kappa(\underline{i}\underline{j} + \underline{j}\underline{i}) + 2W_1\kappa^2 \underline{i}\underline{i} - 2W_2\kappa^2 \underline{j}\underline{j} - p'\underline{I} \ , \tag{1.9.5}$$

where $p' = p - 2(W_1 - W_2)$.

If W_1 and W_2 are positive, which is the case for rubber, a simple shearing deformation requires, in addition to the shearing stress, an extra tension along the direction of shear and an extra pressure on the slip surfaces y=const. We see that in fact,

$$\sigma_{11}-\sigma_{22} = \kappa\sigma_{12} , \tag{1.9.6}$$

so the stresses σ_{11} and σ_{22} cannot possibly be equal when κ and σ_{12} are different from zero. The value of W_2 is typically 20%, say, of the value of W_1 (in rubber), so the second normal stress difference, $\sigma_{22}-\sigma_{33} = -2W_2\kappa^2$, is not large in comparison to the first normal stress difference, $\sigma_{11}-\sigma_{22} = 2(W_1+W_2)\kappa^2$. Thus, the effect of normal stress differences is mainly an extra tension along the direction of shear. Similar normal stress differences arise in shearing flows of viscoelastic fluids, such as solutions of the high polymers that elasticity theory treats in bulk, cross-linked form. Of course, in fluids the stresses are not connected through a strain-energy function, but the state of stress produced by shearing is qualitatively like that for elastic shear.

Since the principal stretch in the z-direction is unity, the other two are reciprocal, say λ and $1/\lambda$. Thus I_1 and I_2 are equal:

$$I_1 = I_2 = 3+\kappa^2 = 1+\lambda^2+\lambda^{-2} . \tag{1.9.7}$$

This is true for any plane deformation, and it simplifies their analysis. The apparent shear modulus, σ_{12}/κ, is

$$\mu(\kappa^2) = \sigma_{12}/\kappa = 2(W_1+W_2) = 2\,\frac{dW}{d(\kappa^2)} . \qquad (1.9.8)$$

Since a large shear, say $\kappa=1$, corresponds to only a moderate stretch λ, the shear modulus can be considered roughly constant over a feasible range of κ.

1.10. Torsion

The torsion of a right circular cylindrical rod maps the particle initially at R,θ,Z to the place r,θ,z given by

$$r = R , \quad \theta = \theta+\tau Z , \quad z = Z . \qquad (1.10.1)$$

Let $\underline{\nabla}$ be the gradient with respect to R,θ,Z. Then

$$\begin{aligned} \underline{F}^T = \underline{\nabla x} &= \underline{\nabla}[r\underline{i}_r(\theta)+z\underline{i}_z] \\ &= \underline{\nabla} r\, \underline{i}_r + r\underline{\nabla}\theta\, \underline{i}_r'(\theta) + \underline{\nabla} z\, \underline{i}_z \\ &= \underline{i}_R\, \underline{i}_r + (\underline{i}_\theta+\tau R\underline{i}_Z)\underline{i}_\theta + \underline{i}_Z\underline{i}_z \end{aligned} \qquad (1.10.2)$$

and, writing $\kappa = \tau R$,

$$\underline{g} = \underline{F}\underline{F}^T = \underline{I}+\kappa(\underline{i}_\theta\underline{i}_z+\underline{i}_z\underline{i}_\theta) + \kappa^2\underline{i}_\theta\underline{i}_\theta . \qquad (1.10.3)$$

We see that the strain has the same general form as for simple shearing. The stress, then, like that for shear, is

$$\underline{\sigma} = 2(W_1+W_2)\kappa(\underline{i}_\theta\underline{i}_z+\underline{i}_z\underline{i}_\theta) + 2W_1\kappa^2\underline{i}_\theta\underline{i}_\theta - 2W_2\kappa^2\underline{i}_z\underline{i}_z - p\,\underline{I} . \qquad (1.10.4)$$

It is easy to see by symmetry that the axial and azimuthal components of the equilibrium equation $\mathrm{Div}\,\underline{\sigma} = \underline{0}$ are satisfied trivially if the reaction pressure p is constant over each membrane r=constant. The radial component is satisfied by a suitable

adjustment of $p(r)$.

Because of the extra tension along the direction of shearing, the θ-direction, azimuthal circles act as if they were stretched strings. The reaction pressure is largest at the axis of the rod because of this squeezing. If, instead of a solid rod, we twisted a hollow cylindrical shell, the cylinder would contract unless it were put under internal pressure. With rod or shell, compressive forces must be applied to the ends $Z=\text{const.}$ to prevent extension.

The torsional moment M is related to the twist per unit length, τ, through the work-energy relation

$$M d\tau = d \int_0^{R_0} 2\pi R W \, dR \, . \tag{1.10.5}$$

Here R_0 is the radius of the rod. W is a function of I_1 and I_2, both of which are equal to $3+\kappa^2$, where $\kappa = \tau R$. Hence,

$$M = 2\pi \int_0^{R_0} R \frac{dW}{d(\kappa^2)} 2\tau R^2 \, dR$$

$$= 2\pi \int_0^{\kappa_0^2} \frac{1}{\tau^3} \frac{dW}{d(\kappa^2)} \kappa^2 \, d(\kappa^2) \, . \tag{1.10.6}$$

Here $\kappa_0 = \tau R_0$. Introducing the shear modulus $\mu(\kappa^2)$, we obtain

$$\tau^3 M = \pi \int_0^{\kappa_0^2} \mu(\kappa^2) \kappa^2 \, d(\kappa^2) \, . \tag{1.10.7}$$

The modelling rule for this case shows that $\tau^3 M$ is expressible as a single function of $(\tau R_0)^2$ for all rods of the same material.

A. C. Pipkin

1.11. Small Superposed Deformations

The main analytical technique for solving problems that do not have a high degree of symmetry is linearization, the staple of the applied mathematician. A reasonably simple state of deformation that is probably close to the desired solution is guessed. It may be an equilibrium state, but need not be. The unknown deformation is then treated, hopefully, as a small perturbation on the guess. I don't want to get involved with the details of any such problem, but I would like to discuss a simple qualitative idea that is helpful in thinking about perturbations.

When a small additional deformation is superposed on some existing state of large deformation, stresses are changed because of the additional distortion. But also, the stress field is changed because the additional deformation rotates the stress that was already present. In some cases the former effect is entirely negligible in comparison to the latter. When this is true, the apparent moduli of the material are determined by the forces acting on it, rather than by anything that we would regard as a material property.

Small deflection of a stretched string is an example everyone is familiar with. The change in length of the string is quadratic in the deflection, and the resulting change of tension is negligible. The restoring force is due to the change in direction of the string tension when the string is deflected.

As another example, consider a small shear of a slab that has been mashed to λ times its initial thickness. The deformation is described by the mapping

$$x = \lambda^{-1/2}X+\kappa\lambda Z \ , \quad y = \lambda^{-1/2}Y \ , \quad z = \lambda Z \ . \tag{1.11.1}$$

A short calculation shows that the strain invariants are

$$I_1 = \lambda^2+2\lambda^{-1}+\kappa^2\lambda^2 \quad \text{and} \quad I_2 = \lambda^{-2}+2\lambda+\kappa^2\lambda \ . \tag{1.11.2}$$

Thus, a small amount of shear κ produces no first-order change in the invariants, or no first-order change in the principal stretches. The small shear merely rotates the stress field already present.

Let A be the initial area (before mashing) of a face z=const., let F be the total normal force on it, and let S be the total shearing force. Then,

$$F d\lambda + S(\lambda d\kappa) = A dW \ . \tag{1.11.3}$$

When $\kappa=0$, the force F is

$$F = A(\lambda-\lambda^{-2})(2W_1+\lambda^{-1}2W_2) \ . \tag{1.11.4}$$

For any κ , the shearing force S is

$$S = A(2W_1+\lambda^{-1}2W_2)\lambda\kappa \ . \tag{1.11.5}$$

Hence, the apparent shear modulus $S/A\kappa$ is given at $\kappa=0$ by

$$\left.\frac{S}{A\kappa}\right|_{\kappa=0} = \frac{F}{A(1-\lambda^{-3})} \ . \tag{1.11.6}$$

If the original force and deformation are known, the modulus is known, independent of any knowledge of the form of W.

The same effect governs the torsional modulus of an initially stretched rod. With a deformation of the form

$$r = \lambda^{-1/2}R, \ \theta = \Theta+\tau\lambda Z \ , \quad z = \lambda Z \ , \tag{1.11.7}$$

the invariants have the form (1.11.2), with $\kappa = \tau r = \tau\lambda^{-1/2}R$:

$$I_1 = \lambda^2+2\lambda^{-1}+\lambda(\tau R)^2, \quad I_2 = \lambda^{-2}+2\lambda+(\tau R)^2 \ . \tag{1.11.8}$$

The extending force F and twisting moment M satisfy

$$F d\lambda + M(\lambda d\tau) = d \int_0^{R_o} 2\pi WR \, dR \ . \tag{1.11.9}$$

Since W is independent of R when $\tau=0$, the force is

$$F = \pi R_o^2(\lambda-\lambda^{-2})(2W_1+\lambda^{-1}2W_2) \ , \tag{1.11.10}$$

when there is no twist. Similarly, the torsional modulus at $\tau=0$ is

$$\frac{\partial M}{\partial \tau} = \frac{2\pi}{\lambda} \int_0^{R_o} \frac{\partial^2 W}{\partial \tau^2} R dR = \frac{\pi}{2} R_o^4(2W_1+\lambda^{-1}2W_2) \ . \tag{1.11.11}$$

Thus,

$$\frac{\partial M}{\partial \tau} = \frac{R_o^2 F}{2(\lambda-\lambda^{-2})} \ . \tag{1.11.12}$$

2. Viscoelastic Flow

The solution of viscoelastic flow problems is complicated by the lack of any single, simple form of constitutive equation that will describe all of the material behavior that is easy to observe. The reason is not difficult to find. It is not entirely a matter of our ignorance, although there is much that we do not know about viscoelastic properties. The real difficulty is inherent in the idea of a viscoelastic fluid. We call a material a fluid if it will flow under easily accessible experimental conditions, and call it elastic if it shows solid-like behavior under other, equally accessible, conditions. There is a natural desire to describe all of the behavior that is easy to observe within one equation, and an understandable hope that this equation should turn out to be reasonably simple. These aims are contradictory when we can observe large elastic deformations and steady shearing flows in the same material.

There are various tractably simple forms of stress-deformation relations, each valid in some specified range of flow conditions. These narrow-range descriptions can be viewed as approximations to the ideal omnibus equation that would describe every aspect of the material's behavior. Determining which approximation is relevant in a given problem is an integral part of the process of solving the problem.

2.1. Simplest Properties of Viscoelastic Fluids

A fluid is called viscoelastic if it exhibits elasticity in shear. Suppose that a thin (inertialess) layer of fluid is sheared

instantaneously by the amount κ and then held motionless. According to the Navier-Stokes (Newtonian) approximation, the shearing stress σ required to do this would vary in time in proportion to $\kappa'(t) = \kappa\delta(t)$, being enormously large at the instant of shearing but relaxing to zero immediately. Under the elastic approximation, $\sigma(t)$ rises to some value depending on κ and then remains constant, never relaxing at all. In viscoelasticity theory we recognize that the stress actually has an intermediate kind of time behavior. For small κ, the stress has the form $\sigma(t) = \kappa\mu(t)$. The stress-relaxation modulus $\mu(t)$ may rise to a very large value at $t=0$, but it dies out to negligible values within some time of order T, say. If times as short as T are not easily accessible to observation, we are content with the Newtonian approximation; if T is a time too long to wait, we call the material elastic. We recognize viscoelastic behavior when T has a convenient size such as one second.

If the amount of shear is varied a little as time progresses, we can approximate the shearing stress by a superposition of the stresses due to each elementary step $d\kappa(t)$:

$$\sigma(t) = \int_{-\infty}^{t} \mu(t-t')d\kappa(t') \ . \qquad (2.1.1)$$

This is the one-dimensional form of the constitutive equation of linear viscoelasticity theory.

This approximation is applicable to steady simple shearing $\kappa(t) = \gamma t$, if the amount of shear in one relaxation time, γT, is small. The stress is $\sigma = \eta_0\gamma$; this is the Newtonian approximation.

The viscosity coefficient is

$$\eta_0 = \int_0^\infty \mu(t)dt \ . \qquad (2.1.2)$$

The mean relaxation time T can be defined by

$$\eta_0 T = \int_0^\infty \mu(t)t \ dt \ . \qquad (2.1.3)$$

In the case of a variable shearing, $\kappa'(t) = \gamma(t)$, the Newtonian approximation is valid when γT is small and relatively constant over intervals of the order of T.

In an oscillatory shearing $\kappa(t) = \kappa_0 \exp(i\omega t)$, the linear viscoelastic approximation is valid if κ_0 is small, regardless of how large ω may be. The stress is equal to $\eta^*(\omega)\kappa'(t)$ (real part), where the complex viscosity η^* is defined by

$$\eta^*(\omega) = \int_0^\infty \mu(t)e^{-i\omega t} \ dt = \eta_1(\omega) - i\eta_2(\omega) \ . \qquad (2.1.4)$$

If ωT is small, the viscosity is approximately

$$\eta^* = \eta_0 - i\omega T\eta_0 \ . \qquad (2.1.5)$$

On the other hand, when ωT is very large, the viscosity is

$$\eta^* \sim \mu(0)/i\omega \ , \qquad (2.1.6)$$

and the response is approximately that of an elastic material, $\sigma = \mu(0)\kappa(t)$.

In steady shearing motions for which γT is not small, it cannot be expected that the linear approximation will be valid.

The stress is some non-linear function of the shear rate. For a polymer melt or solution, the apparent viscosity $\eta(\gamma)$, equal to $\sigma(\gamma)/\gamma$, descends rapidly from its zero shear rate value η_0 toward a much lower limiting value for large γT. The mean relaxation time T can be estimated crudely from measurements of the apparent viscosity function; the viscosity is relatively constant when the rate of shear is below $1/T$.

At the opposite extreme from steady shearing motion, we can consider sudden, large-amplitude shearing. At times small in comparison to T, before any perceptible stress-relaxation has taken place, the stress is some function of the amount of shear: $\sigma = \kappa\mu_e(\kappa^2)$. The linear elastic shear modulus $\mu_e(0)$ is equal to the value of the linear relaxation modulus $\mu(t)$ at time zero (see(2.1.6)).

2.2. Flow Diagnosis

To visualize the relations among the material properties that have been mentioned and the regions of validity of various approximations, it is useful to characterize flows in an over-simplified way by the values of two dimensionless parameters, the shear amplitude A and the frequency ωT. Both are to be defined only loosely. We plot these parameters on distorted scales so that all values from zero to infinity lie in a square (Fig. 1).

In problems of forced vibration, the meaning of the frequency ω is obvious. More generally, we use high ωT to mean sudden motions and low ωT to mean smooth motions. The edge $\omega T = \infty$ on the

flow diagnosis diagram is the region of elastic response; interpreting ω as $1/t_o$, where t_o is the total time of observation, the limiting case $T/t_o = \infty$ is that in which there is no stress relaxation within the period of observation. In this domain, the parameter A stands for a typical amount of shear, as estimated from boundary or initial conditions.

The edge $A = 0$ is the domain of linear viscoelasticity theory. Its intersection with the edge $\omega T = \infty$ is the domain of linear elasticity theory, and its intersection with the edge $\omega T = 0$ is the domain of Newtonian (Navier-Stokes) fluid dynamics.

The edge $\omega T = 0$ corresponds to viscometric flows and other absolutely steady shearing motions. On this edge, A is equal to γT, the amount of shear in one relaxation time.

The interior region in which A is large and ωT is neither large nor small is an area of real ignorance. Flows corresponding to such values are usually handled by optomistic guess-work, based on extrapolations from the edges. However, there is hope for the future. To every pair of values of A and ωT there corresponds some materially steady motion, and some of these motions are accessible to observation in the orthogonal rheometer, a relatively recent invention (see Sec. 5).

2.3 Relative Strain Histories

Let's turn to some matters of notation that will occasionally be needed for three-dimensional problems. Let $\underline{x}$ be the position of a particle at a reference time t, and let

$\underline{x}(s,t,\underline{x})$ be the position of that particle at time s. A fiber of fluid along $d\underline{x}$ at time t lies along $d\underline{p}$ at time s; the mapping

$$d\underline{p} = \underline{F}\, d\underline{x} \tag{2.3.1}$$

defines the deformation gradient $\underline{F}(s,t)$. $\underline{F}$ is also a function of the particle $\underline{x}$, but, with the understanding that we are dealing with a Lagrangian description, we omit $\underline{x}$ to save writing.

$\underline{F}$ defines the local rotation and stretching of fibers. To get a measure of stretching alone, we observe that

$$d\underline{p}\cdot d\underline{p} = d\underline{x}\cdot\underline{F}^{T}\underline{F}\, d\underline{x} \tag{2.3.2}$$

and define the strain by

$$\underline{G}(s,t) = \underline{F}^{T}(s,t)\underline{F}(s,t). \tag{2.3.3}$$

In problems in which the strain does not change much over an interval of the order of T preceding the reference time t, it may be possible to approximate the strain history by a few terms of its expansion in powers of s-t:

$$\underline{G}(s,t) = \Sigma\frac{1}{n!}\,\underline{A}_n(t)(s-t)^n. \tag{2.3.4}$$

The derivatives,

$$\underline{A}_n(t) = \frac{D^n}{Ds^n}\,\underline{G}(s,t)\Big|_{s=t}, \tag{2.3.5}$$

are called the Rivlin-Ericksen tensors. $\underline{A}_0$ is $\underline{I}$, the identity, and $\underline{A}_1$ is twice the classical strain-rate tensor. The higher derivatives can be calculated, in terms of the velocity $\underline{u}$, from the recursion relation

$$A_{ij}^{(n+1)} = \frac{D}{Dt}\,A_{ij}^{(n)} + A_{ik}^{(n)}u_{k,j} + A_{jk}^{(n)}u_{k,i}. \tag{2.3.6}$$

The operation giving $\underline{A}_{n+1}$ from $\underline{A}_n$ is called Oldroyd differentiation.

2.4 Linear Viscoelastic Approximation

In the linear viscoelastic approximation, the stress has the form

$$\underline{\sigma} = -p\underline{I} + \int_{-\infty}^{t} \mu(t-s)\frac{D}{Ds}\,\underline{G}(s,t)ds. \qquad (2.4.1)$$

We assume that the fluid is incompressible; p is the reaction pressure. The strain history is non-linear in the velocity, so this approximation includes terms that are negligible when the approximation is valid at all. However, precisely which terms are negligible depends on the problem, and it is not always immediately obvious.

Walters has used this equation more than anyone else, I believe. His analysis of the balance rheometer [6], with Jones, is a recent application of it.

As an exercise, consider steady simple shearing motion, with velocity $\underline{u} = \gamma y\underline{i}$. Then, in quick order, we obtain

$$\underline{p}(s,t,\underline{x}) = \underline{x} + (s-t)\gamma\underline{i}y, \qquad (2.4.2)$$

$$\underline{F}(s,t,\underline{x}) = \underline{I} + (s-t)\gamma\underline{i}\underline{j}, \qquad (2.4.3)$$

$$\begin{aligned}\underline{G}(s,t,\underline{x}) &= \underline{I} + (s-t)\gamma(\underline{i}\underline{j}+\underline{j}\underline{i}) + (s-t)^2\gamma^2\underline{j}\underline{j} \\ &= \underline{I} + (s-t)\underline{A}_1 + \tfrac{1}{2}(s-t)^2\underline{A}_2, \qquad (2.4.4)\end{aligned}$$

$$\frac{D}{Ds}\,\underline{G}(s,t,\underline{x}) = \underline{A}_1 + (s-t)\underline{A}_2, \qquad (2.4.5)$$

and, with the definitions of η_o and T given earlier,

$$\underline{\sigma} = -p\underline{I} + \eta_o\underline{A}_1 - \eta_o T\underline{A}_2. \qquad (2.4.6)$$

Notice the peculiarity that although the acceleration is zero, the strain-acceleration $\underline{A}_2$ is not.

Now, the term $T\underline{A}_2$ is of order γT in comparison to $\underline{A}_1$, and so it is negligible when the linear approximation is valid. As we shall see later (Sec. 4), (2.4.6) does not correctly account for all second-order terms; a term proportional to $\underline{A}_1^2$ is also needed. Thus, to the extent that (2.4.6) is valid, the correct expression for the stress is the Navier-Stokes approximation:

$$\underline{\sigma} = -p\underline{I} + \eta_0\underline{A}_1. \tag{2.4.7}$$

As a second example, consider the plane steady motion with velocity $\underline{u} = \underline{N}\underline{x}$, where

$$\underline{N} = \varepsilon(\underline{i}\underline{i} - \underline{j}\underline{j}) + \omega(\underline{j}\underline{i} - \underline{i}\underline{j}). \tag{2.4.8}$$

Being unused to Lagrangian fluid dynamics, and assuming too much from the result of the preceding exercise, I would have guessed that the strain-rate ε should be small in comparison to $1/T$ in order for the linear viscoelastic approximation to be valid, and that it would reduce to the Navier-Stokes equation again. In fact, ε can be arbitrarily large provided that ω is still larger, and the stress is then nothing like what the Navier-Stokes approximation would give.

To see why, it is necessary (or rather, sufficient) to carry out the computation of $\underline{G}$. From

$$\dot{p}_{i,A} = u_{i,j}p_{j,A}, \quad \text{or} \quad \dot{\underline{F}}(s,t) = \underline{N}\underline{F}(s,t), \tag{2.4.9}$$

with the initial condition $\underline{F}(t,t) = \underline{1}$ we obtain

$$\underline{F}(s,t) = e^{(s-t)\underline{N}} = \Sigma\frac{1}{n!}(s-t)^n\underline{N}^n. \tag{2.4.10}$$

Since

$$\underline{N}^2 = -(\omega^2 - \varepsilon^2)(\underline{i}\underline{i} + \underline{j}\underline{j}) = -\Omega^2\underline{\Delta} \quad \text{(say)} \qquad (2.4.11)$$

it follows that

$$\underline{N}^{2n} = (-1)^n\Omega^{2n}\underline{\Delta} \quad \text{and} \quad \underline{N}^{2n+1} = (-1)^n\Omega^{2n+1}(\underline{N}/\Omega). \qquad (2.4.12)$$

The case $\underline{N}^0 = \underline{I}$ is an exception. Hence,

$$\underline{F}(s,t) = \underline{k}\underline{k} + \underline{\Delta}\cos\Omega(s-t) + (\underline{N}/\Omega)\sin\Omega(s-t). \qquad (2.4.13)$$

Thus,

$$\underline{G}(s,t) = \underline{k}\underline{k} + \underline{\Delta}\cos^2\Omega(s-t) + (\underline{N}^T\underline{N}/\Omega)\sin^2\Omega(s-t)$$

$$+ \Omega^{-1}(\underline{N}+\underline{N}^T)\sin\Omega(s-t)\cos\Omega(s-t). \qquad (2.4.14)$$

We see that $\underline{G}$ is close to $\underline{I}$ at all times, which is sufficient for the linear approximation to be valid, if ε/ω is small. In that case, by neglecting terms of order $(\varepsilon/\omega)^2$ we obtain $\Omega = \omega$ and

$$\underline{N}^T\underline{N}/\Omega^2 = \underline{\Delta} - (2\varepsilon/\omega)(\underline{i}\underline{j} + \underline{j}\underline{i}). \qquad (2.4.15)$$

Then,

$$\underline{G}(s,t) = \underline{I} - (\varepsilon/\omega)(\underline{i}\underline{j} + \underline{j}\underline{i})[1 - \cos 2\omega(s-t)]$$

and

$$+ (\varepsilon/\omega)(\underline{i}\underline{i} - \underline{j}\underline{j})\sin 2\omega(s-t), \qquad (2.4.16)$$

$$\frac{D}{Ds}\underline{G}(s,t) = 2\varepsilon(\underline{i}\underline{j} + \underline{j}\underline{i})\sin 2\omega(t-s)$$

$$+ 2\varepsilon(\underline{i}\underline{i} - \underline{j}\underline{j})\cos 2\omega(t-s). \qquad (2.4.17)$$

By using this expression in (2.4.1) and recalling the definition (2.1.4) of the complex viscosity, we obtain

$$\underline{\sigma} = -p\underline{I} + 2\varepsilon(\underline{i}\underline{j} + \underline{j}\underline{i})\eta_2(2\omega) + 2\varepsilon(\underline{i}\underline{i} - \underline{j}\underline{j})\eta_1(2\omega). \qquad (2.4.18)$$

The result is surprising in many ways. First, the linear approximation is valid even for arbitrarily large values

of the strain-rate, provided that it is small in comparison to the vorticity. Second, the Newtonian approximation is not necessarily valid even when εT is small; we recover the Navier-Stokes equation only if ωT is also small. Third, the stress depends on the vorticity, and in the oddest conceivable way, wrapped up inside some highly non-linear functions. Fourth, although the flow is absolutely steady in every usual sense, the relevant material properties are ones that we think of in connection with oscillatory motions. The explanation is simple enough. When ε/ω is small, the flow has elliptical, nearly circular streamlines. Radial fibers stretch and contract alternately, twice in each revolution. So far as such a fiber is concerned, the motion is indeed a small oscillation.

For any values of ε and ω, the stress components must be time-independent functions of ε and ω, since there is nothing else for the stress to depend upon. These motions are <u>materially steady</u> (Sec. 5).

They are also <u>controllable</u> flows, which satisfy the momentum equation no matter what the stress components may turn out to be. Here, this is true because the extra stress (produced by the deformation) is constant in space, so its divergence vanishes trivially; and, the inertial term $\rho(\underline{u}\cdot\underline{\nabla})\underline{u}$ is $\rho\underline{N}^2\underline{x}$, and $\underline{N}^2$ is symmetric, so this can be balanced by the gradient of a reaction pressure,

$$p = -(\rho/2)\underline{x}\cdot\underline{N}^2\underline{x}. \qquad (2.4.19)$$

3. Viscometric Flow

Flows in which each material element undergoes steady uniaxial shearing are called viscometric. They are particularly simple to discuss because there are no transient effects of stress-relaxation, the stress having reached a steady state. The stress response is completely characterized by the response in plane steady simple shearing motion, one of the simplest conceivable flows.

3.1 Slip Surfaces, Shear Axes, and Shear Rate

Let us first consider some kinematically admissible examples of viscometric flow, without regard to their dynamical admissibility. All viscometric flows can be visualized as the relative sliding motion of a sheaf of inextensible material surfaces, which we call slip surfaces. In the more important cases, each slip surface moves as if it were rigid. The simpler cases are steady motions with straight, circular, or helical streamlines. In these cases, each slip surface slides tangentially and always occupies the same locus in space.

In the case of steady parallel flows, $\underline{u} = u(x,y)\underline{k}$, the slip surfaces are the general cylinders $u(x,y) = \text{constant}$. The flow is a relative sliding motion of these cylinders. The direction of relative sliding, which we call $\underline{a}$, coincides with the direction of motion in these cases. The direction normal to the slip surfaces, which we call $\underline{b}$, is parallel to $\nabla\underline{u}$. We add a third unit vector $\underline{c}$ so that $\underline{a}$, $\underline{b}$, and $\underline{c}$ form an orthonormal system at each particle, and call these the shear axes. The shear rate γ is equal to $|\nabla\underline{u}|$. The velocity gradient has the

form $\underline{\nabla u} = \underline{\nabla u}\ \underline{k} = \gamma \underline{b}\underline{a}$. At any given particle, γ is constant in time, and the motion in the vicinity of that particle looks like a steady plane shearing motion.

Flows with velocity fields of the form

$$\underline{u} = u(z)\underline{i} + v(z)\underline{j} \tag{3.1.1}$$

have parallel plane slip surfaces z = constant moving in skew directions. In these flows $\underline{b} = \underline{k}$. The direction of relative sliding, $\underline{a}$, and the shear rate, γ, are given together by

$$\gamma \underline{a} = u'(z)\underline{i} + v'(z)\underline{j}. \tag{3.1.2}$$

Then $\underline{\nabla u} = \gamma \underline{b}\underline{a}$. Since γ is a function of z, it is constant in time at each particle. Notice that the direction of relative sliding is not the same as the direction of motion.

In steady flows with coaxial circular streamlines, for which $\underline{u} = r\omega(r,z)\underline{i}_\theta$, the slip surfaces are the surfaces of constant angular velocity ω. They form a nested set of surfaces of revolution, each one rotating as if it were rigid. The direction of relative sliding is the azimuthal direction, $\underline{a} = \underline{i}_\theta$, and $\underline{b}$ is parallel to $\underline{\nabla}\omega$. The velocity gradient is

$$\underline{\nabla u} = r\underline{\nabla}\omega\underline{i}_\theta + \omega(\underline{i}_r\underline{i}_\theta - \underline{i}_\theta\underline{i}_r). \tag{3.1.3}$$

The second term corresponds to rigid rotation. The first term has the form $\gamma \underline{b}\underline{a}$ if we identify γ as $r|\underline{\nabla}\omega|$. Then γ is constant in time at each particle, and the motion in the vicinity of that particle, as seen from the rotating axes $\underline{a}$, $\underline{b}$, $\underline{c}$, is a steady simple shearing motion.

Slip surfaces that are coaxial circular cylinders can translate in the axial direction and rotate about it simultaneously, producing a motion with helical streamlines:

$$\underline{u} = r\omega(r)\underline{i}_\theta + u(r)\underline{i}_z. \qquad (3.1.4)$$

The direction $\underline{b}$ is radial. Since the velocity gradient is

$$\underline{\nabla u} = \underline{i}_r(r\omega'\underline{i}_\theta + u'\underline{i}_z) + \omega(\underline{i}_r\underline{i}_\theta - \underline{i}_\theta\underline{i}_r), \qquad (3.1.5)$$

by identifying the first term as $\gamma\underline{b}\underline{a}$ we find that

$$\gamma\underline{a} = r\omega'(r)\underline{i}_\theta + u'(r)\underline{i}_z. \qquad (3.1.6)$$

There is another class of viscometric flows with helical streamlines, which we call helicoidal after the shapes of their slip surfaces. They have velocity fields of the form

$$\underline{u} = (r\underline{i}_\theta + c\underline{i}_z)\omega(r, z - c\theta). \qquad (3.1.7)$$

Pearson was the first person who ever mentioned such a flow to me. In these flows, all streamlines have the same rise per turn, $2\pi c$. The slip surfaces are the helicoids ω = constant, so $\underline{b}$ is parallel to $\underline{\nabla}\omega$. The direction of relative sliding is parallel to $\underline{u}$. To find the shear rate, we first write out the velocity gradient:

$$\underline{\nabla u} = \underline{\nabla}\omega(r\underline{i}_\theta + c\underline{i}_z) + \omega(\underline{i}_r\underline{i}_\theta - \underline{i}_\theta\underline{i}_r). \qquad (3.1.8)$$

The first term has the form $\gamma\underline{b}\underline{a}$ if we take γ^2 to be

$$\gamma^2 = (r^2 + c^2)\underline{\nabla}\omega\cdot\underline{\nabla}\omega\ . \qquad (3.1.9)$$

We verify that this is constant along streamlines.

The preceding examples by no means exhaust all of the kinematically admissible possibilities [7], but they include all of the physically more important cases.

3.2 Stress

In a steady simple shearing motion $\underline{u} = \gamma y \underline{i}$, the stress components are functions of γ because they have nothing else to depend upon. In saying that the motion is steady, I mean that the shear rate γ has been constant for so long that the stress has reached a steady state.

Assuming that the fluid is isotropic, it follows from symmetry that the zx and zy components of stress must be zero. The shearing stress σ_{xy} is some odd function of γ, and the normal stress differences are even functions of γ that vanish if there has never been any motion:

$$\sigma_{xy} = \gamma\eta(\gamma^2), \quad \sigma_{xx} - \sigma_{yy} = \gamma^2 N_1(\gamma^2), \quad \sigma_{yy} - \sigma_{zz} = -\gamma^2 N_2(\gamma^2). \tag{3.2.1}$$

The apparent viscosity η and the normal stress coefficients N_1 and N_2 are called the viscometric functions. In polymer solutions, η and N_1 have roughly the same form. If there were a complete analogy with elasticity theory, they would differ only by a constant of proportionality. Again by analogy with elasticity, I expect that N_2 is small in comparison to N_1 and that, with the sign convention shown, N_2 is positive. Thus, in addition to the shearing stress, there are normal stresses that are roughly equivalent to an extra tension in the direction of shearing

(the $\underline{a}$-direction). At high shear rates, this extra tension can be much larger than the shearing stress.

In simple shearing, the shear axes $\underline{a}$, $\underline{b}$, $\underline{c}$ are the coordinate axes $\underline{i}$, $\underline{j}$, $\underline{k}$. The stress can be written all together as

$$\underline{\sigma} = -p\underline{I} + \gamma\eta(\underline{ab} + \underline{ba}) + \gamma^2(N_1-N_2)\underline{aa} - \gamma^2 N_2\underline{bb}. \qquad (3.2.2)$$

Here p is a reaction pressure; we consider only incompressible fluids.

This expression for the stress is immediately applicable to all of the more complicated shearing motions discussed previously. For example, in the parallel motions $\underline{u} = u(x,y)\underline{k}$, for which $\underline{a} = \underline{k}$ and $\gamma\underline{b} = \underline{\nabla}u$, the stress is

$$\underline{\sigma} = -p\underline{I} + \eta(\underline{k\nabla}u + \underline{\nabla u k}) + \underline{\nabla}u\cdot\underline{\nabla}u(N_1-N_2)\underline{kk} - N_2\underline{\nabla u\nabla}u. \qquad (3.2.3)$$

In the circular motions $\underline{u} = r\omega\underline{i}_\theta$, for which $\underline{a} = \underline{i}_\theta$ and $\gamma\underline{b} = r\underline{\nabla}\omega$, the stress is

$$\underline{\sigma} = -p\underline{I} + r\eta(\underline{i}_\theta\underline{\nabla}\omega + \underline{\nabla}\omega\underline{i}_\theta) + r^2\underline{\nabla}\omega\cdot\underline{\nabla}\omega(N_1-N_2)\underline{i}_\theta\underline{i}_\theta$$

$$- N_2 r^2\underline{\nabla}\omega\underline{\nabla}\omega. \qquad (3.2.4)$$

The stress for all of the other cases can be written down just as easily.

3.3 Controllable Flows

Of course, whether or not such a stress field satisfies the momentum equation depends on exactly what the velocity is, and what forms the viscometric functions have. There are a

very few special cases in which the velocity field can be fully specified in advance, and the momentum equation is then satisfied trivially, no matter what forms the viscometric functions may have. The prototype $\underline{u} = \gamma y \underline{i}$ is such a case; the stress produced by this motion is constant in space, if the reaction pressure is constant, so the momentum equation is automatically satisfied. We call such flows completely controllable.

Completely controllable flows would be ideal for use in the experimental determination of the viscometric functions, since the analysis of data would not be complicated by a simultaneous determination of an unknown velocity field. They would be, if it were not for the fact that all completely controllable flows involve some practical impossibility such as infinite parallel plates.

All completely controllable flows are known [7,8,9]. One is the flow $\underline{u} = c\theta \underline{i}_z$ corresponding to shearing between non-parallel plates, one fixed and the other moving parallel to the line where they would intersect. Another is the circular flow with angular velocity $\gamma \, \mathrm{Ln}(r/r_o)$, for which the shear rate $r\omega'$ is uniform. This flow requires an azimuthal pressure gradient, so it cannot be maintained in a full circular annulus. And that's all there are, aside from motions obtained by superposing global rigid motions on the preceding ones.

There are a few flows that would be completely controllable if it were not for inertial effects. This is true of the special helicoidal flows that have the form

$$\underline{u} = (\omega_o/h)(z-c\theta)(r\underline{i}_\theta + c\underline{i}_z). \qquad (3.3.1)$$

The degenerate case c = 0 is the well known torsional flow produced by a rotating disc viscometer (approximately). The slip surfaces are the parallel planes z = constant, rotating with angular velocity $\omega_o z/h$ about the z-axis. Centrifugal force is suppressed by reducing the gap h between fixed plate and rotating disc.

The only other flow that is completely controllable with neglect of inertia is a wierd case with slip surfaces that curl up until they overlap; the motion is not steady with respect to any frame of reference, although the motion near each particle is steady with respect to the shear axes at that particle. This case was turned up by working from a strict, legalistic definition of controllability [7,8].

3.4 Partially Controllable Flows

The flows that are actually used for measuring the viscometric functions are partially controllable, either exactly or to a good approximation. Partially controllable flows are those in which normal stress differences do not influence the motion. The velocity field depends on the form of the apparent viscosity function, but the flow is so highly symmetrical that normal stresses are automatically equilibrated by the reaction pressure, whatever forms the normal stress functions may have.

As an example, consider plane Poiseuille flow, the flow in a channel with walls at $y = \pm L$, produced by a pressure

gradient $\partial\sigma_{xx}/\partial x = G$. With $\underline{u} = u(y)\underline{i}$, the momentum equation is satisfied if the reaction pressure has the form

$$p = -Gx - \gamma^2 N_2 + \text{const.}, \tag{3.4.1}$$

where $\gamma = u'(y)$, and if the shearing stress balances the pressure gradient:

$$\gamma\eta = -Gy. \tag{3.4.2}$$

The velocity is found from the latter equation; the former merely gives p when u is already known. Thus, the motion is partially controllable; the velocity field is not affected by normal stress differences.

Let's complete this problem. Let $\gamma = \Gamma(\sigma)$ be the inverse of $\sigma = \gamma\eta(\gamma^2)$; $\Gamma(\sigma)$ is odd in σ. Then (3.4.2) gives

$$u'(y) = \gamma = -\Gamma(Gy), \tag{3.4.3}$$

and an integration, with $u(-L) = 0$, gives $u(y)$. The flux is of interest:

$$Q = \int_{-L}^{L} u\,dy = yu\Big|_{-L}^{L} - \int_{-L}^{L} yu'\,dy$$

$$= \int_{-L}^{L} y\Gamma(Gy)dy = (2/G^2)\int_{0}^{GL} \sigma\Gamma(\sigma)d\sigma. \tag{3.4.4}$$

This is the analog, for channel flow, of the Weissenberg-Rabinowitsch-Mooney formula for pipe flow, which dates from around 1930. It shows that G^2Q is a function of GL, and plotting data in this way will show what function $\Gamma(\sigma)$, or ultimately η, may be.

People have been measuring apparent viscosities for a long time, without always being aware of the presence of

normal stress differences. It was possible to interpret experimental data correctly, without knowing about normal stresses, because the flows used in shear viscosity measurements are, ordinarily, partially controllable.

Poiseuille and Couette flows are partially controllable, and so are the helical flows with coaxial circular cylindrical slip surfaces that one might produce in the annular region between two cylinders, with or without an axial pressure gradient. The only other partially controllable flows are those with parallel plane slip surfaces moving in skew directions, which can be regarded as small-gap approximations to motions between cylinders. The flow in a cone-and-plate viscometer is partially controllable to the same degree of approximation that the shear rate is uniform, but not exactly.

3.5 Nearly Viscometric Flows

If a viscometric flow is not at least partially controllable, then it is not likely to be dynamically admissible at all. The form of the apparent viscosity function determines the velocity field; then either the normal stresses turn out to be equilibrated automatically, or, by some curious accident the forms of the normal stress functions might be such that equilibrium occurs. However, the latter possibility is remote.

Thus, if we were interested only in exact solutions, the theory of viscometric flow would be only a neat correlation of data from various kinds of viscometers. But, of course, we hope and expect that viscometric data can be used predictively

in the approximate analysis of flows that are not exactly viscometric. Let me discuss a few examples in qualitative terms.

3.6 Tube Flows

Parallel flows are kinematically admissible in tube flow problems, but they are partially controllable only if the tube cross-section is circular (or annular). Unequilibrated normal stresses should produce transverse flow. The normal stress difference that is relevant is $\gamma^2 N_2$. If this were zero, rectilinear flow would be dynamically admissible. The assumption $N_2 = 0$ is called the Weissenberg hypothesis. Now, in fact, no one has ever seen any transverse flow, so far as I know. We might regard this as a clear confirmation of the Weissenberg hypothesis, and thus be led to an essential simplification in the form of the viscometric constitutive equation. We might, if we believed too much in the relevance of exact solutions.

Let us suppose that the motion is rectilinear, $\underline{u} = u(x,y)\underline{k}$, and see where the inconsistency arises. The stress would have the form (3.2.3). The momentum equation, with no acceleration, becomes

$$\underline{\nabla} p = \underline{\nabla}\cdot(\eta \underline{\nabla} u)\underline{k} - \underline{\nabla}\cdot(N_2 \underline{\nabla} u \underline{\nabla} u). \qquad (3.6.1)$$

For the right-hand member to be irrotational, as the left-hand member is, the z component must be constant:

$$\underline{\nabla}\cdot(\eta \underline{\nabla} u) = -G. \qquad (3.6.2)$$

With u = 0 on the tube wall, this determines u. Then,

$$p = -Gz + P(x,y), \tag{3.6.3}$$

where P must satisfy

$$\underline{\nabla}P = -\underline{\nabla}\cdot(N_2\underline{\nabla}u\underline{\nabla}u)$$

$$= -\underline{\nabla}\cdot(N_2\underline{\nabla}u)\underline{\nabla}u - N_2\gamma\underline{\nabla}\gamma. \tag{3.6.4}$$

Thus,

$$p = -Gz - \int_0^\gamma N_2(\gamma^2)\gamma d\gamma + P_1(x,y), \tag{3.6.5}$$

where P_1 must satisfy

$$\underline{\nabla}P_1 = -\underline{\nabla}\cdot(N_2\underline{\nabla}u)\underline{\nabla}u. \tag{3.6.6}$$

Since the right-hand side is parallel to $\underline{\nabla}u$, the surfaces P_1 = constant must coincide with surfaces u = constant. Thus, $P_1 = P_1(u)$, and

$$P_1'(u) = -\underline{\nabla}\cdot(N_2\underline{\nabla}u). \tag{3.6.7}$$

Now, unless the right-hand side is also constant over surfaces u = constant, we have a contradiction. An easy way out is to suppose that $N_2 = 0$, but it is not the only way.

Suppose that N_2 is roughly proportional to n. Whether it is or not, we can write $N_2 = T_2 n(\gamma^2) + n(\gamma^2)$, and choose the time constant T_2 so as to make $n(\gamma^2)$ vanish at the largest shear rate involved. Then neglecting n, and recalling that u satisfies (3.6.2), we obtain

$$P_1'(u) = -T_2\underline{\nabla}\cdot(n\underline{\nabla}u) = GT_2. \tag{3.6.8}$$

Thus,

$$p = -Gz - \int_0^\gamma N_2(\gamma^2)\gamma\, d\gamma + GT_2u(x,y) + \text{const.} \tag{3.6.9}$$

Only the error $n(\gamma^2)$ is left over to produce transverse flow. This should be small in comparison to N_2, which, in turn, is probably small in comparison to N_1. The extra tension along streamlines, $\gamma^2 N_1$, will tend to hold them straight. Consequently, it should not be surprising if the transverse flow were too small to notice.

3.7 Flow in Tilted Troughs

As a variation on the preceding problem, consider the flow under gravity down a trough inclined at an angle α to the horizontal. In Newtonian flow, the free surface conditions are satisfied on a flat free surface, $y = 0$, say, and the flow is the same as in a tube formed from the trough and its mirror image in the free surface (as Stokes observed, in the first paper on viscous flow).

We can try the same trick for viscoelastic flow. We take the trough flow to be half of the flow in a tube symmetrical in the plane $y = 0$. The axial pressure gradient G is replaced by $\rho g \sin \alpha$, and the pressure (3.6.9) is modified by deleting $-Gz$ and adding a hydrostatic part $-\rho g y \cos \alpha$ (taking y negative in the fluid):

$$p = -\rho g y \cos \alpha + \rho g T_2 (\sin \alpha) u(x,y) - \int_0^\gamma N_2(\gamma^2)\gamma \, d\gamma + \text{const.} \qquad (3.7.1)$$

In the case of a trough of semi-circular cross-section, the full Poiseuille flow is partially controllable, and the approximate term involving T_2 can be replaced by an exact expression. The same is true for a deep channel with sides at $x = \pm L$, say. In the latter case, the pressure is

$$p = -\rho g y \cos \alpha - N_2 \gamma^2 + \text{const.} \qquad (3.7.2)$$

However, this is not exact for flow in a trough in any case, because the condition of constant pressure on the free surface is not met. On $y = 0$, the normal direction is the $\underline{c}$ - direction, so the normal stress σ_{yy} is equal to $-p$. This is constant over the surface $y = 0$ only if $N_2 = 0$.

The unequilibrated normal stress will cause the free surface to warp until the extra weight of fluid above $y = 0$ supplies the extra normal stress. A first approximation to the shape of the warped surface is obtained by setting p = constant in (3.7.2) (for deep channels):

$$y = -(N_2 \gamma^2 / \rho g \cos \alpha) + \text{const.} \qquad (3.7.3)$$

The solution is only approximate because on this new free surface, the tangential stress is not exactly zero.

R. I. Tanner told me about these results, and also showed me photographs of the warped free surface. For the fluid involved, polyisobutylene in cetane, there is a pronounced upward bulge in the free surface. With γ zero at the center and large at the walls, (3.7.3) shows that the free surface stands highest at the center if N_2 is positive. For that one fluid, at least, it is visibly evident that N_2 is positive.

3.8 Anti-Centrifugal Effects

Circular viscometric flows are kinematically admissible in problems involving steady rotation of an immersed body of

revolution, but usually not exactly dynamically admissible. Within Newtonian fluid dynamics, departures from circular flow are caused by centrifugal force. Viscoelastic fluids can show striking anti-centrifugal effects, in which the direction of transverse flow is directly opposite to the direction that would be produced by centrifugal force.

The extra tension along the direction of shearing causes the (nearly) circular streamlines to act as if they were stretched fibers, with tension highest where the shear rate is largest. This tension causes the fibers to contract, producing radially inward motion where the shear rate is largest.

For example, in the flow produced by a rotating sphere, centrifugal force would cause a transverse flow outward at the equator, where the speed is high, and inward at the poles. However, the shear rate is also highest at the equator, and thus so is the extra tension in azimuthal circles. The tendency for such fibers to contract can produce inward motion in the equatorial plane, with outward motion at the poles. Giesekus [10] has made beautiful pictures of this effect, for cones as well as spheres, by injecting dye with a hypodermic. A recent paper by Griffiths, Jones, and Walters [11] analyzes this effect in the case of a rotating disc.

The flow produced by a rotating cylinder would be a partially controllable viscometric flow in the absence of a free surface. With a free surface, and a vertical rotating cylinder, the extra tension forces fluid upward near the rod,

producing the well known climbing effect. This, and other effects that can be interpreted in terms of an extra tension in the direction of shearing, first became well known through the work of Weissenberg.

3.9 Laminar Drag Reduction

In flow in curved pipes, circular or helicoidal viscometric flows are kinematically admissible. Unlike the case with straight pipes, in which the streamlines are straight in the viscometric approximation, the extra tension along the direction of shearing produces a net sideways force when the pipe is curved. Whether as a centrifugal (inertial) effect or an anti-centrifugal (viscoelastic) effect, there is a pronounced transverse flow.

This produces a very interesting effect, which has been observed and explained by Barnes and Walters [12]. The transverse flow decreases the apparent viscosity by increasing the over-all shear rate. Hence, the pressure gradient that is required in order to produce a given flux is lower than it would be if there were no transverse flow; under the right circumstances it is easier to pump fluid through a curved pipe than through a straight pipe.

At first this sounds impossible, because the higher shear rate produced by the transverse flow necessarily implies larger energy dissipation, even if the viscosity is lowered. However, the rate of work is GQ, where Q is the flux and G is the mean axial pressure gradient, so an increase of Q at fixed G automatically supplies the extra work.

3.10 Boundary Layers

Although I have been speaking of global viscometric flows, the viscometric character of a strain history is a local property of the history of a material element, and it is possible for some particles in a flow to be undergoing viscometric motion even though others are not.

For example, in any steady flow with no slip at solid boundaries, the flow at each boundary point is exactly viscometric. For, the tangential components of the gradient of velocity are zero, and thus $\underline{\nabla u} = \underline{n}(\underline{n}\cdot\underline{\nabla})\underline{u}$ at the boundary, $\underline{n}$ being the normal to the boundary. This has the viscometric form $\gamma\underline{b}\underline{a}$ with $\underline{b} = \underline{n}$ and $\gamma\underline{a} = (\underline{n}\cdot\underline{\nabla})\underline{u}$. The shear rate at a boundary particle never changes because the particle stays put. Thus, whatever the flow may be like away from the boundary, provided that it is steady, the stress at the boundary is given exactly by the viscometric constitutive equation:

$$\underline{\sigma} = -p\underline{I} + \eta(\underline{n}\,\partial\underline{u}/\partial n + \partial\underline{u}/\partial n\,\underline{n}) + (N_1-N_2)(\partial\underline{u}/\partial n)(\partial\underline{u}/\partial n) - N_2|\partial\underline{u}/\partial n|^2\,\underline{n}\underline{n}. \quad (3.10.1)$$

Since the flow in a boundary layer is nearly a steady parallel flow, we can presume that the stress will be nearly of the viscometric form throughout the boundary layer, as it is, exactly, at the wall. Certain qualitative effects are easy to see. If the apparent viscosity goes down as the shear rate goes up, the boundary layer will be thinner than in the Newtonian approximation, which uses the apparent viscosity at

zero shear rate. The "pressure" that is constant through the thickness is $p + \gamma^2 N_2$, and the "pressure" whose variation along the boundary direction drives the flow is $p - \gamma^2(N_1-N_2)$. The former is the outside pressure p_o. The latter, which is the forcing term, is then $p_o - \gamma^2 N_1$, the outside pressure modified by the extra tension along the direction of shearing. Since $\gamma^2 N_1$ can be very large, its variation should be taken into account. The affect of an adverse pressure gradient is magnified by this extra tension, and it appears that separation should occur sooner than predicted by the Newtonian approximation.

3.11 Analytical Treatment of Nearly Viscometric Flows

Flows that are nearly viscometric can be treated by perturbation methods. However, there are some complications that should be mentioned.

The first is that if the velocity field is not exactly viscometric, the shear axes and shear rate are not well defined. To get around this difficulty, we first notice that in flows that are exactly viscometric, the first two Rivlin-Ericksen tensors are

$$\underline{A}_1 = \gamma(\underline{a}\underline{b} + \underline{b}\underline{a}) \quad \text{and} \quad \underline{A}_2 = 2\gamma^2 \underline{b}\underline{b} \qquad (3.11.1)$$

(and all others vanish). Then, also,

$$\underline{A}_1^2 = \gamma^2(\underline{a}\underline{a} + \underline{b}\underline{b}) \text{ and } \gamma^2 = \frac{1}{2} \operatorname{tr} \underline{A}_1^2. \qquad (3.11.2)$$

Consequently, the stress is

$$\underline{\sigma} = -p\underline{I} + \eta\underline{A}_1 - \frac{1}{2} N_1 \underline{A}_2 + (N_1-N_2)\underline{A}_1^2, \qquad (3.11.3)$$

with γ^2 equal to half of tr $\underline{A}_1^2$. For a flow that is not viscometric, we can still calculate $\underline{A}_1$ and $\underline{A}_2$ unambiguously, and then use this expression for the stress.

This procedure is rather arbitrary. Even if we stipulate that an approximation in terms of $\underline{A}_1$ and $\underline{A}_2$ alone will be tried, and that it must reduce to the exact stress when the flow is viscometric, (3.11.3) is far from being the only possible choice.

However, (3.11.3) has been used very widely, although under a different interpretation. Later we will consider the second-order slow-motion approximation for perturbations on Newtonian flow. It has the form (3.11.3), with constant coefficients (the values of the viscometric functions at zero shear rate). Solutions based on the second-order approximation often agree well with observations even when there is not the least doubt that the experimental conditions are far outside the range in which the second-order approximation could be valid.

Apparently, what happens is that if the flow is actually nearly viscometric, the relation (3.11.3) is a good approximation to the stress, and not bad even when the viscometric functions are treated as constants. Of course, the values used for these constants should be the values of the viscometric functions at a shear rate appropriate to the problem; the values at the highest shear rate involved would give good accuracy where accuracy is most important. Giesekus [10] has used this kind of extrapolation effectively to explain even rather complicated flows.

A. C. Pipkin

More accurate treatment of nearly viscometric flows involves the consideration of material properties not embodied in the viscometric functions, since transient effects of stress relaxation and steady effects of biaxial or triaxial shearing, or stretching motions, can come into play. For example, even in uniaxial shearing at a nearly constant rate $\gamma(t)$, the shearing stress is of the form

$$\sigma(t) = \gamma(t)\eta(\gamma^2(t)) + \int_{-\infty}^{t} \mu[\gamma^2(t);\ t-s][\gamma(t-s) - \gamma(t)]\, ds, \tag{3.11.4}$$

to first order in the difference history $\gamma(t-s) - \gamma(t)$. Notice that the perturbation stress relaxation modulus is a function of the shear rate for the main motion. Indeed, (3.11.4) is not even internally consistent unless

$$d(\gamma\eta)/d\gamma = \int_0^{\infty} \mu(\gamma^2;t)dt. \tag{3.11.5}$$

Since the left-hand member is a function of γ, so is μ. The fact that μ depends on γ is borne out by experimental evidence [13-17]. The generalization of (3.11.4) to three-dimensional form is fairly complicated [18,19]. So far, the relatively little work that has been done on this more accurate treatment of nearly viscometric flow problems has been directed toward experimental determination of the relaxation moduli, or has foundered on the lack of data.

4. Slow Viscoelastic Flow

In a motion with characteristic velocity U and characteristic length L, shear rates are of the order of U/L, and the amount of shear in time T is of the order of A = TU/L. To estimate ωT, suppose that it is appropriate to estimate the jerkiness of the motion in terms of a frequency $\omega \sim D/Dt$, and suppose that the material derivative of any function is of the order of U/L in comparison to that function. Then for a characteristic frequency ωT we obtain TU/L again.

We use the term slow viscoelastic flow for flows in which both A and ωT are small. Since the slowness is with respect to a time scale defined by the material, the motion might be extremely rapid by some other standard. In particular, the Reynolds number may be either large or small in a slow viscoelastic flow. Boundary-layer flow of a material with a very short relaxation time, such as air or water ($T \sim 10^{-8}$ sec) is a slow viscoelastic flow. The beginning of a creeping flow (zero Reynolds number) is not, because the difference between motion and no motion is too jerky.

4.1 Stress

The stress at a given time is mainly determined by the strain history over an immediately preceding interval of the order of the mean relaxation time T. Suppose that the motion is smooth enough that the strain at time t-s relative to the state at time t can be represented by a Taylor series in the lag s:

$$\underline{G}(t-s,t) = \Sigma \frac{1}{n!}(-s)^n \underline{A}_n(t). \qquad (4.1.1)$$

If $\underline{A}_n$ is of order $(U/L)^n$, then the n-th term of the series is of order $(TU/L)^n$, over the interval that matters. If TU/L is small, the strain history is approximately determined, over this interval, by $\underline{A}_1$ alone. It is determined a little more exactly by $\underline{A}_1$ and $\underline{A}_2$ together, and so on. Consequently, the same is true of the stress.

The stress, $\underline{\sigma} = -p\underline{I} + \underline{S}$, is in part a reaction to the constraint of incompressibility. The extra stress $\underline{S}$ is the part to be described by a constitutive equation. We assume that $\underline{S}$ can be expanded as

$$\underline{S} = \underline{S}_1 + \underline{S}_2 + \underline{S}_3 + \cdots , \qquad (4.1.2)$$

where $\underline{S}_n$ is of order $(TU/L)^n$ in the parameter. The terms of various orders that are smooth, isotropic functions of the tensors $\underline{A}_n$ are

$$\begin{aligned} &O(U/L): && \underline{A}_1, \\ &O(U^2/L^2): && \underline{A}_2,\ \underline{A}_1^2, \\ &O(U^3/L^3): && \underline{A}_3,\ \underline{A}_1 \operatorname{tr} \underline{A}_1^2,\ \underline{A}_1\underline{A}_2 + \underline{A}_2\underline{A}_1, \end{aligned} \qquad (4.1.3)$$

and so on, to list only those that cannot be eliminated by using algebraic identities or the constraint condition. Hence, for $\underline{S}_1$ we take

$$\underline{S}_1 = \eta_0 \underline{A}_1, \qquad (4.1.4)$$

the Navier-Stokes approximation. $\underline{S}_2$ has the form

$$\underline{S}_2 = \eta_0[-T\underline{A}_2 + (T + T^*)\underline{A}_1^2]. \qquad (4.1.5)$$

The coefficients η_0 and $\eta_0 T$ are the zeroth and first moments of the linear stress-relaxation modulus, as defined earlier. The coefficient of $\underline{A}_1^2$ is written in such a curious way because it is convenient later. The coefficients appearing so far are related to the viscometric functions by

$$\eta_0 = \eta(0), \quad 2\eta_0 T = N_1(0), \quad \eta_0(T+T^*) = N_1(0) - N_2(0). \qquad (4.1.6)$$

Since N_2 is probably positive and smaller than N_1, T^* is probably positive and a little smaller than T.

In the third order, the coefficient of $\underline{A}_3$ is proportional to the second moment of the linear stress-relaxation modulus. If the other two third-order terms are recombined as their sum and difference, the coefficient of the sum is proportional to the value of $d\eta(\gamma^2)/d(\gamma^2)$ at $\gamma = 0$. The coefficient of the difference is a material property that is not related either to the linear stress-relaxation modulus or to the viscometric functions.

4.2 Solution of Problems

The expression of the stress as, effectively, a power series in the parameter TU/L carries with it a method of solution of flow problems that is not only convenient but obligatory. Solutions must be obtained by ordinary perturbation methods, so far as the parameter TU/L is concerned, beginning with Newtonian flow as the lowest approximation. Higher-order

approximations are obtained by iteration or an equivalent power-series expansion method.

To simplify the exposition, let us restrict attention to creeping flow (zero Reynolds number). Then the equations of motion are

$$\underline{\nabla}\cdot\underline{u} = 0 \quad \text{and} \quad \underline{\nabla}p = \text{Div}\,\underline{S}. \tag{4.2.1}$$

Let us write the velocity and pressure as

$$\underline{u} = \underline{u}_1 + \underline{u}_2 + \ldots \quad \text{and} \quad p = p_1 + p_2 + \ldots, \tag{4.2.2}$$

where the n-th term is of order n in the parameter TU/L. Then $\underline{\nabla}\cdot\underline{u}_n = 0$, and the momentum equation yields

$$\begin{aligned}
&\underline{\nabla}p_1 - \eta_o\nabla^2\underline{u}_1 = \underline{0},\\
&\underline{\nabla}p_2 - \eta_o\nabla^2\underline{u}_2 = \text{Div}\,\underline{S}_2(\underline{u}_1),\\
&\underline{\nabla}p_3 - \eta_o\nabla^2\underline{u}_3 = \text{Div}\,\underline{S}_2(\underline{u}_1,\underline{u}_2) + \text{Div}\,\underline{S}_3(\underline{u}_1),
\end{aligned} \tag{4.2.3}$$

and so on. Here $\underline{S}_2(\underline{u}_1)$ means $\underline{S}_2$, evaluated with the velocity field $\underline{u}_1$, while $\underline{S}_2(\underline{u}_1,\underline{u}_2)$ stands for the third-order terms that arise when $\underline{S}_2$ is evaluated with $\underline{u}_1 + \underline{u}_2$.

It might appear more elegant to truncate the approximation to $\underline{S}$ at some point and then solve the resulting equation exactly, rather than by a boring perturbation procedure. Of course, nothing would be gained in the way of accuracy by doing this, but that is no objection. What *is* an objection to such a procedure is that spurious solutions which are not perturbations of Newtonian flow can arise. For motions that are not small perturbations of Newtonian flow, the constitutive equations used

here have no justification, and "solutions" that are far from Newtonian are meaningless.

In spite of the fact that we must not stray far from Newtonian flow, it is possible to find interesting results. Second or third-order effects can be interesting if they are qualitatively unlike Newtonian flow. For example, transverse flows due to normal stress effects are interesting by their existence, even when they are only small perturbations on the main motion.

4.3 Tanner's Theorem

What is more surprising is that there are interesting, even valuable, results to be obtained in cases in which the Newtonian flow $\underline{u}_1$ is still correct to second order, i.e. $\underline{u}_2 = \underline{0}$. From (4.2.3) we see that this can be the case only when Div $\underline{S}_2(\underline{u}_1)$ is the gradient of a scalar, f, say. If it is, the second-order momentum equation is satisfied by $\underline{u}_2 = \underline{0}$ and $p_2 = f$.

Let me simplify the notation by writing $\underline{u}_1 = \underline{u}$, $\underline{A}_1 = \underline{A}$, and $\underline{A}_2 = \underline{B}$, with the understanding that $\underline{A}$ and $\underline{B}$ are evaluated in terms of $\underline{u}$ ($=\underline{u}_1$). Then,

$$\underline{S}_2(\underline{u}_1) = -\eta_0 T(\underline{B} - \underline{A}^2) + \eta_0 T^* \underline{A}^2. \qquad (4.3.1)$$

Whether or not $\underline{S}_2(\underline{u}_1)$ is irrotational, the following modification of a result essentially due to Tanner [20] is useful:

Div $(\underline{B}-\underline{A}^2)$ is irrotational whenever Div $\underline{A}$ is irrotational and $\underline{\nabla}\cdot\underline{u} = 0$. If P is the potential for Div $\underline{A}$, so that

Div $\underline{A} = \underline{\nabla}P$, then

$$\text{Div}(\underline{B} - \underline{A}^2) = \underline{\nabla}(\frac{DP}{Dt} + \frac{1}{2}\gamma^2). \tag{4.3.2}$$

Here γ is the absolute shear rate, defined by

$$2\gamma^2 = \text{tr } \underline{A}^2. \tag{4.3.3}$$

The proof requires only a manipulation of the messy expression for Div $\underline{B}$. Since $\underline{A}$ and $\underline{B}$ are defined by

$$A_{ij} = u_{i,j} + u_{j,i} \tag{4.3.4}$$

and

$$B_{ij} = DA_{ij}/Dt + A_{ik}u_{k,j} + A_{jk}u_{k,i}\ , \tag{4.3.5}$$

we obtain

$$B_{ij,j} = D(A_{ij,j})/Dt + u_{k,j}A_{ij,k} + (A_{ik}u_{k,j})_{,j}$$

$$+ A_{jk,j}u_{k,i} + A_{jk}u_{k,ij}. \tag{4.3.6}$$

The second and third terms combine to yield Div $\underline{A}^2$, if $\underline{\nabla}\cdot\underline{u} = 0$ as assumed. The final term is equal to $\underline{\nabla}\gamma^2/2$. With Div $\underline{A} = \underline{\nabla}P$, the first and fourth terms combine to give $\underline{\nabla}$ (DP/Dt). Rearrangement yields (4.3.2).

In creeping flow, the potential P is p_1/η_o, and the second-order momentum equation takes the form

$$\underline{\nabla}(p_2 + T\frac{Dp_1}{Dt} + \frac{1}{2}\eta_o T\gamma^2) - \eta_o\nabla^2\underline{u}_2 = \eta_o T^* \text{ Div } \underline{A}^2. \tag{4.3.7}$$

There are several important cases in which Div $\underline{A}^2$ is also irrotational.

4.4 Plane Flow

Plane flow is the simplest and broadest case. In plane flow, $\underline{A}$ is essentially a 2 x 2 matrix, and since tr $\underline{A}$ = 0 ($\underline{\nabla}\cdot\underline{u} = 0$) it follows that $\underline{A}^2$ reduces to an isotropic pressure in the plane of flow:

$$\underline{A}^2 = \gamma^2(\underline{I} - \underline{kk}). \qquad (4.4.1)$$

Thus, Div $\underline{A}^2$ is equal to $\underline{\nabla}\gamma^2$. Consequently, the second-order momentum equation is satisfied if $\underline{u}_2 = \underline{0}$ and

$$p_2 = -T\frac{Dp_1}{Dt} + \eta_0(T^* - \tfrac{1}{2}T)\gamma^2. \qquad (4.4.2)$$

As an application, consider a shearing flow over a plane wall y = 0, and the disturbance that is produced in it by a deep, narrow slot in the wall. Let x = 0 be the slot centerline, and let x = $\pm$L, for y negative, be the walls of the slot. We specify that the velocity is asymptotic to $\gamma_0 y\underline{i}$ far from the slot mouth for y positive, and that the velocity approaches zero deep in the slot.

First consider the Newtonian approximation. If the Reynolds number defined by $\rho(\gamma_0 L)L/\eta_0$ is small, the creeping flow approximation is applicable to the analysis of the disturbance. Since creeping flow is reversible, the streamlines are symmetrical about the slot axis x = 0. Thus, the y-component of velocity, v, is an odd function of x, so v, v_{xx}, and v_{yy} all vanish on x = 0. From $dp_1/dy = \eta_0\nabla^2 v$, we conclude that p_1 is

constant along the slot axis. Thus, the pressure deep in the slot, $p_1(0,-\infty)$, is the same as the undisturbed pressure $p_\infty = p_1(0,\infty)$. It follows that the undisturbed pressure can be measured by measuring, instead, the pressure in the hole. The error, an inertial effect, can be made as small as desired by decreasing the hole width to such an extent that the creeping flow approximation is good.

But now consider the second-order approximation. The stress is

$$\underline{\sigma} = -[p_1 - T\frac{Dp_1}{Dt} + \eta_o(T^* - \tfrac{1}{2}T)\gamma^2]\underline{I} - \eta_o T\underline{B} + \eta_o(T+T^*)\underline{A}^2. \quad (4.4.3)$$

Far from the slot, where the shearing motion is undisturbed and $p_1 = p_\infty$, the stress component σ_{yy} is

$$-p_u = \sigma_{yy} = -p_\infty - \tfrac{1}{2}\eta_o T\gamma_o^2 . \quad (4.4.4)$$

We write p_u for the undisturbed value of $-\sigma_{yy}$, which is the undisturbed pressure of the fluid against the wall. The gauge pressure p_g is the value of $-\sigma_{yy}$ (or $-\sigma_{xx}$) deep in the slot. Since $\gamma = 0$ and $p_1 = p_\infty$ there, we obtain $p_g = p_\infty$. Hence,

$$p_g - p_u = -\tfrac{1}{2}\eta_o T\gamma_o^2. \quad (4.4.5)$$

The gauge reading is lower than the pressure that the fluid would exert against the wall if no hole were present, and the error is not diminished by making the hole smaller.

In terms of the normal stress coefficient N_1, the error is

$$p_g - p_u = -\tfrac{1}{4}N_1\gamma^2. \quad (4.4.6)$$

Although we have deduced this result only for γ so small that the limiting value $N_1(0)$ can be used, it is found experimentally that the error is still approximated well by $-N_1(\gamma^2)\gamma^2/4$ at large shear rates [21]. I mentioned earlier that the second-order equation often gives results that are better than they should be, if the flow is nearly viscometric; this is an example.

The connection of the pressure error with the extra tension along the direction of shearing, $N_1\gamma^2$, suggests a qualitative explanation of the error, which Tanner [21] has pointed out. Streamlines dip in slightly as they pass the mouth of the hole. The extra tension, roughly along streamlines, exerts an upward force that partly balances the downward thrust p_u.

Although the analysis was carried out for the simplest possible geometry, of course the geometry of the exterior flow has nothing to do with it. The disturbance is localized near the mouth of the hole, and the shear rate that determines the pressure error is the wall shear rate that would exist where the hole is, if there were no hole. The error is, so to speak, a property of the material, so it will be perfectly consistent from one viscometer to another.

Since normal stress measurements are often made with the use of pressure holes, and the measurements are always interpreted under the assumption that $p_g = p_u$ for a small enough hole, much of what we though we knew about normal stress differences is quantitatively inaccurate. The effect was only recently discovered; suspicions were first raised by Broadbent, Kaye, Lodge, and Vale [22].

4.5 Parallel Flows

In connection with viscometric flows it was pointed out that if N_2 were a constant multiple of η, flow in tubes could be rectilinear. In the second-order approximation both η and N_2 are constant, so N_2 is trivially a multiple of η. This is another case in which Div $\underline{A}^2$ is irrotational and the Newtonian velocity field is still accurate to second order. The approximation (3.6.9) for the pressure becomes

$$p = -Gz - \tfrac{1}{2}(T-T^*)(\eta_0\gamma^2 - 2Gu) + \text{const.} \qquad (4.5.1)$$

The normal thrust against the wall of the tube is

$$-\underline{n}\cdot\underline{\underline{\sigma}}\underline{n} = p - \eta_0\underline{n}\cdot\underline{\underline{A}}\underline{n} + \eta_0\underline{T}\underline{n}\cdot\underline{B}\underline{n} - \eta_0(T+T^*)\underline{n}\cdot\underline{A}^2\underline{n}, \qquad (4.5.2)$$

where $\underline{n}$ is the normal to the wall. This reduces to

$$-\underline{n}\cdot\underline{\underline{\sigma}}\underline{n} = -Gz + \tfrac{1}{2}(T - T^*)\eta_0\gamma^2 + \text{const.}, \qquad (4.5.3)$$

where γ is the Newtonian shear rate at the place considered.

The value of $T-T^*$ can be deduced from measurements of the thrust at two places where the shear rates are different. The difference of the undisturbed values is

$$p_{u1} - p_{u2} = \tfrac{1}{2}(T - T^*)\eta_0(\gamma_1^2 - \gamma_2^2), \qquad (4.5.4)$$

if z is the same at each place. If pressure holes are used, and the pressure error is evaluated by using the plane flow result, the different in gauge pressures will be

$$\begin{aligned} p_{g1} - p_{g2} &= p_{u1} - p_{u2} - \tfrac{1}{2}\eta_0 T(\gamma_1^2 - \gamma_2^2) \\ &= -\tfrac{1}{2}\eta_0 T^*(\gamma_1^2 - \gamma_2^2). \end{aligned} \qquad (4.5.5)$$

If the pressure-hole effect were not known, we would interpret T^* as T^*-T. The error in interpretation of data could be quite sizable. (See note below.)

4.6 Potential Flow

Bruce Caswell has pointed out to me that potential flow is another case in which Div $\underline{A}^2$ is irrotational. With $\underline{u}_1 = \underline{\nabla}\phi$, and $\nabla^2\phi = 0$, then Div $\underline{A}^2 = \underline{\nabla}\gamma^2$. Since Div $\underline{A}$ is zero, its potential P can be taken to be zero as well. Then the second-order pressure is

$$p_2 = -\frac{1}{2}\eta_0(T - 2T^*)\gamma^2. \qquad (4.6.1)$$

A reason for looking at this is to see whether or not the outside pressure used in boundary-layer theory should be modified. The first order pressure, given by Bernoulli's equation, is of order ρU^2. Then p_2/p_1 is of the order of

$$\frac{\eta_0 T(U/L)^2}{\rho U^2} = \frac{1}{Re}\frac{TU}{L}. \qquad (4.6.2)$$

Thus, even if we retain $0(TU/L)$ terms, we would omit p_2 as $0(1/Re)$.

Note added in proof: Kearsley [36] has considered the difference between the undisturbed pressure and the pressure deep in a slot that is parallel to the direction of flow. In this case, the tension across the mouth of the slot, which produces the pressure error, is caused by the second normal stress difference rather than the first. The pressure error has the form (4.5.4), with station 1 deep in the slot (whence $\gamma_1 = 0$) and station 2 at the slot mouth (where γ_2 is approximately the wall shear rate in the absence of a slot).

5. Materially Steady Motions

The history of motion of a material element is materially steady if the velocity gradient is constant in time at the particle considered. The axes with respect to which the velocity gradient appears constant need not be the axes of an inertial frame, nor need they be the same for different particles. Viscometric flows are materially steady. In them, the velocity gradient as viewed from the shear axis system has the form $\nabla \underline{u} = \gamma \underline{b}\underline{a}$, with γ constant in time. Couette flow is an example in which the shear axes rotate with respect an inertial frame, and the velocity gradient with respect to an inertial frame does not have constant components.

The steady, homogeneous, plane motions $\underline{u} = \underline{\underline{N}}\underline{x}$ that were considered in Sec. 2.4 are materially steady, trivially. The axes with respect to which the velocity gradient is steady are the same at all particles and do not rotate. In connection with those motions, recall that the material response was described in certain cases by the dynamic viscosity at arbitrary values of the frequency. The complete time-independence of the velocity gradient $\underline{\underline{N}}$ did not by any means prevent oscillatory motions of material fibers.

The particular simplicity of materially steady motions, with regard to material response, lies in the fact that if the velocity gradient $\underline{\underline{N}}$ (where $d\underline{u} = \underline{\underline{N}}\, d\underline{x}$) is constant in time, then the extra stress is also constant and simply a function of $\underline{\underline{N}}$. This function, $\underline{S}(\underline{\underline{N}})$, embodies not only the viscometric functions but also such information as the dynamic viscosity.

It appears that the idea of materially steady motions provides a framework within which a large amount of information can be organized. Indeed, there are materially steady motions with arbitrary values of the flow diagnosis parameters A and ωT. Complete knowledge of the stress response in all materially steady motions would by far exceed what we know about any fluid.

However, it is not certain that the stress response in all such cases can actually be observed. I am not referring to the infinite amount of work involved, but to the question of whether or not a flow with a specified value of $\underline{N}$ can be produced, physically. Homogeneous motions are not that easy to produce.

The task of the problem-solver at this stage is to discover materially steady flows with experimentally feasible configurations. So far, very few are known.

5.1 Stress: The Reiner-Rivlin Paradox

Let me say a little about the relation between stress and velocity gradient. To simplify matters, suppose that the velocity gradient is constant in space as well as time. Then, wherever a particle may have been in the past, it always experienced the same velocity gradient $\underline{N}$. Consequently, there is nothing else for the extra stress to depend upon, in the way of kinematical variables:

$$\underline{\sigma} + p\underline{I} = \underline{S}(\underline{N}). \qquad (5.1.1)$$

According to classical ideas, superposition of a rotation, with velocity gradient $\underline{\omega} = -\underline{\omega}^{T}$, cannot affect the

stress. For, whether or not such a superposed rotation is seen depends as much on the motion of the observer as on that of the fluid. Hence,

$$\underline{S}(\underline{N}) = \underline{S}(\underline{N} + \underline{\omega}), \tag{5.1.2}$$

for all antisymmetric $\underline{\omega}$. This holds in particular when $-\underline{\omega}$ is the antisymmetric part of $\underline{N}$, so

$$\underline{S}(\underline{N}) = \underline{S}(\underline{\varepsilon}), \tag{5.1.3}$$

where $\underline{\varepsilon}$ is the strain-rate, the symmetric part of $\underline{N}$. Conversely, it is also sufficient for the extra stress to be a function of the strain-rate, since every such function satisfies the assumption (5.1.2) identically, for all antisymmetric $\underline{\omega}$.

If the fluid is isotropic, $\underline{S}(\underline{\varepsilon})$ is an isotropic function, so the stress has the representation

$$\underline{\sigma} = -p\underline{I} + 2\eta\underline{\varepsilon} + f\underline{\varepsilon}^2. \tag{5.1.4}$$

This is the Reiner-Rivlin equation. It is well known that it does not agree with data. In fact, it yields the value zero for the first normal stress difference, $N_1\gamma^2$, in viscometric flow. There must be something wrong with the argument.

The original assumption, (5.1.1), was made air-tight by postulating a situation with nothing for the stress to depend upon except $\underline{N}$. Indeed, the relative deformation history is found immediately from the equation

$$\frac{D}{Ds}\,\underline{F}(s,t) = \underline{N}\,\underline{F}(s,t), \tag{5.1.5}$$

with the initial condition $\underline{F}(t,t) = \underline{I}$, to be

$$\underline{F}(s,t) = \exp[(s-t)\underline{N}], \tag{5.1.6}$$

and this is completely determined by $\underline{N}$.

The difficulty is with the innocent invariance assumption (5.1.2). That assumption would be correct if we were assuming that the stress is determined in every motion by the concurrent value of the velocity gradient. However, we are not now doing so. Instead, we are assuming that the stress depends on the deformation history, and restricting attention to cases in which the whole history is characterized by the constant tensor $\underline{N}$. In this context, (5.1.2) means that the stress is the same for the two deformation histories with constant velocity gradients $\underline{N} + \underline{\omega}$ and $\underline{N}$. A glance at (5.1.6) shows that these two histories need have little in common. There is no reason why they should produce the same stress.

Giesekus [23] has worked out a correct canonical form, replacing the Reiner-Rivlin equation, for the stress in materially steady motions. It does, of course, involve the rotation as well as the strain-rate. It is fairly complicated, so I won't reproduce it here.

5.2 Motions with Flakes

Let us return to the problem of discovering materially steady motions that can be produced experimentally. The condition of no slip at solid boundaries is a limiting factor. It can be avoided when the fluid is so stiff that boundaries are not needed to support it. However, it may be profitable to seek flows that can be compatible with a no-slip condition.

The velocity variation $d\underline{u} = \underline{N}\, d\underline{x}$ in the neighborhood of a point of a solid boundary must be such that the neighboring part of the boundary is mapped onto itself by the motion. Let

$\underline{a}$, $\underline{b}$, $\underline{c}$ be axes with $\underline{b}$ normal to the boundary. The neighboring flake of boundary can rotate about the $\underline{b}$-axis with constant angular velocity ω, say. We can specify that the $\underline{a}$-direction is parallel to $\underline{b}\cdot\nabla\underline{u}$. Then the velocity gradient in the $\underline{a}$, $\underline{b}$, $\underline{c}$ system has the form

$$\underline{N} = \omega(\underline{ac} - \underline{ca}) + \gamma\,\underline{ab}. \qquad (5.2.1)$$

Although it is not necessary for the velocity gradient to have this form away from the boundary, flows in which it does so should be particularly simple. Through each particle there would be a segment, or flake, of material that moves rigidly, and it seems probable that these flakes would have to fit together to form material surfaces that would move without stretching. This has been proved [7] for the special case of viscometric flows, for which $\omega = 0$; the surfaces in question are the slip surfaces.

No motions with a velocity gradient of the form (5.2.1) and $\omega \neq 0$ are known except homogeneous motions. Fortunately, there exists a device that can produce homogeneous motions of this kind.

5.3 The Maxwell-Chartoff Rheometer

Huigol [24] has shown that the flow in Maxwell and Chartoff's [25] orthogonal rheometer is materially steady and not viscometric. This rheometer involves a pair of parallel discs, rotating with equal angular velocities about axes that are parallel but not coincident. Let one disc be in the plane

$y = 0$, and let it rotate about the y-axis; let the other be in the plane $y = h$, rotating about the axis $x = 0$, $z = -k$. It is kinematically admissible for the fluid on a plane $y =$ constant to rotate about an axis $x = 0$, $z = -ky/h$, so that the velocity field is

$$u = \omega(z + \alpha y), \quad v = 0, \quad w = -\omega x \quad (\alpha = k/h). \tag{5.3.1}$$

The velocity gradient $u_{i,j} = N_{ij}$ is

$$\underline{N} = \omega(\underline{i}\underline{k} - \underline{k}\underline{i}) + \alpha\omega\underline{i}\underline{j}. \tag{5.3.2}$$

Since this is constant, so is the extra stress. The momentum equation is satisfied with neglect of centrifugal force, which can be suppressed by making the gap h small enough.

The extra stress has the form

$$\begin{aligned}\underline{S} = {} & \underline{i}\underline{i}\, S_{11}(\alpha,\omega) + \underline{j}\underline{j}\, S_{22}(\alpha,\omega) + \underline{k}\underline{k}\, S_{33}(\alpha,\omega) \\ & + (\underline{i}\underline{j} + \underline{j}\underline{i})S_{12}(\alpha,\omega) + (\underline{j}\underline{k} + \underline{k}\underline{j})S_{23}(\alpha,\omega) \\ & + (\underline{k}\underline{i} + \underline{i}\underline{k})S_{13}(\alpha,\omega).\end{aligned} \tag{5.3.3}$$

It is possible to learn a little about the functions S_{ij} by symmetry considerations. Any transformation that does not alter the strain history must not alter $\underline{S}$, either. Such transformations include those that do not alter $\underline{N}$.

Changing α to $-\alpha$ and $\underline{j}$ to $-\underline{j}$ does not alter $\underline{N}$, so the same transformation must leave $\underline{S}$ unaltered. It follows that S_{11}, S_{22}, S_{33}, and S_{13} must be even functions of α, while S_{12} and S_{23} must be odd. It is also evident that changing the signs of both ω and $\underline{i}$ does not change $\underline{N}$. For the same to be true of $\underline{S}$, S_{11}, S_{22}, S_{33}, and S_{23} must be even functions of ω, and S_{12} and S_{13} must be odd.

Since only normal stress differences are determinate if the fluid is incompressible, we can arbitrarily set S_{33} equal to zero.

There is no motion if $\omega = 0$, and only a pure rotation if $\alpha = 0$, so $\underline{S}$ must vanish in both of these cases.

To summarize these observations, we write

$$S_{11} = \alpha^2\omega^2 f_{11}, \quad S_{22} = \alpha^2\omega^2 f_{22}, \quad S_{33} = 0,$$
$$S_{12} = \alpha\omega f_{12}, \quad S_{23} = \alpha\omega^2 f_{23}, \quad S_{13} = \alpha^2\omega^3 f_{13}, \tag{5.3.4}$$

where the functions f_{ij} are functions of α^2 and ω^2. The coefficient of f_{13} is written as $\alpha^2\omega^3$ instead of $\alpha^2\omega$ for convenience later, to make f_{13} finite in the viscometric limit.

It is conceivable that the strain history, $\underline{G}$, could have more symmetry than $\underline{N}$, for example invariance under interchange of axes; further conclusions about $\underline{S}$ could then be drawn. However, it does not. By using the same methods that were employed in Sec. 2.4, we find that

$$\underline{G}(s,t) = \underline{I} + (\underline{ij} + \underline{ji})\,\alpha \sin \omega(s-t)$$
$$+ [\alpha(\underline{jk} + \underline{kj}) + 2\alpha^2\underline{jj}][1 - \cos \omega(s-t)]. \tag{5.3.5}$$

The functions in (5.3.4) can be related to the viscometric functions. In the limiting case in which $\omega \to 0$ and $\alpha \to \infty$ with $\omega\alpha$ fixed at the value γ, $\underline{N}$ takes the viscometric form $\gamma\underline{ab}$, with $\underline{a} = \underline{i}$ and $\underline{b} = \underline{j}$. Hence, in this limit, S_{11} must approach $\gamma^2(N_1-N_2)$, S_{22} must approach $-\gamma^2 N_2$, S_{12} must approach $\gamma\eta$, and both S_{13} and S_{23} must vanish. If we write the functions f_{ij} as functions of ω^2 and $\alpha^2\omega^2$, i.e.

$$f_{ij} = f_{ij}(\omega^2, \alpha^2\omega^2), \tag{5.3.6}$$

then the limiting values of these functions are

$$f_{11}(0,\gamma^2) = N_1(\gamma^2) - N_2(\gamma^2), \quad f_{22}(0,\gamma^2) = -N_2(\gamma^2),$$

$$f_{12}(0,\gamma^2) = \eta(\gamma^2). \tag{5.3.7}$$

It is sufficient for f_{13} and f_{23} to be finite.

The functions f_{ij} can also be connected with the complex viscosity $\eta^*(\omega)$ of the linear viscoelastic approximation. From (5.3.5) we find that $\underline{G}$ is always close to $\underline{I}$, and thus the linear approximation is valid, provided that α is small. Then on linearizing with respect to α, we obtain

$$\frac{D}{Ds}\,\underline{G}(s,t) = (\underline{ij} + \underline{ji})\alpha\omega \cos \omega(t-s) - (\underline{jk}+\underline{kj})\alpha\omega \sin \omega(t-s). \tag{5.3.8}$$

Then

$$\int_{-\infty}^{t} \mu(t-s)\frac{D}{Ds}\,\underline{G}(s,t)ds = (\underline{ij} \pm \underline{ji})\,\alpha\omega\eta_1(\omega)$$

$$- (\underline{jk} + \underline{kj})\alpha\omega\eta_2(\omega). \tag{5.3.9}$$

Thus, S_{11}, S_{22}, S_{33}, and S_{13} must vanish in comparison to α when α is small, while S_{12} and S_{23} must be asymptotic to $\alpha\omega\eta_1$ and $-\alpha\omega\eta_2$, respectively. For the former to be true, it is sufficient that f_{11}, f_{22}, and f_{13} remain finite in the limit. The latter conditions give

$$f_{12}(\omega^2,0) = \eta_1(\omega) \quad \text{and} \quad f_{23}(\omega^2,0) = -\eta_2(\omega)/\omega. \tag{5.3.10}$$

As a check, notice that η_1 and η_2 are by definition respectively even and odd functions of ω.

The fact that $f_{12}(0,\gamma^2)$ is $\eta(\gamma^2)$, while $f_{12}(\omega^2,0)$ is $\eta_1(\omega)$, is tantalizing since plots of η versus γ and η_1 versus

ω usually have pretty much the same shape.

Except in these limiting cases, there is no data yet, so far as I know.

5.4 Steady Simple Extension

Steady simple extension is a non-viscometric, materially steady motion. The velocity gradient is

$$\underline{N} = \varepsilon \underline{i}\underline{i} - (\varepsilon/2)(\underline{j}\underline{j} + \underline{k}\underline{k}). \qquad (5.4.1)$$

By symmetry, the xy, yz, and zx components of the extra stress are zero, and the yy and zz components are equal. The normal stress difference $\sigma_{xx} - \sigma_{yy}$, divided by ε, is called the Trouton viscosity, or extensional viscosity. Some data for polystyrene [26] indicates that the extensional viscosity can be several hundred times as large as η, because it remains relatively constant while η decreases at higher shear rates. Of course, at zero shear rate, where the Newtonian approximation is valid, the extensional viscosity is only three times as large as the shear viscosity.

This result suggests that fluids will avoid extensional motions when an alternative of the viscometric shearing kind is kinematically admissible. As an example, consider the flow into a pipe or channel from a reservoir. In the usual Newtonian picture, fluid in the reservoir would approach the outlet from all directions. However, this requires squeezing in the direction perpendicular to the direction of flow, and extension along the flow direction. Less extension would occur if the fluid were relatively motionless except in a slender cone approaching

the outlet. There would be shearing between this upstream jet and the surrounding fluid, rather than extension throughout the whole fluid. Metzner, Uebler, and Chan Man Fong [27] have observed this effect. I am not sure that they would agree with my explanation of it, but I believe that it is essentially the same as theirs.

6. Controllability

In order to determine the strain energy function of an elastic material, or, equivalently, its stress-deformation relation, one can subject the material to known deformations and measure the forces required to produce them. A theoretical difficulty in this procedure is in producing deformations that can be regarded as completely known. An assumed deformation, compatible with the observed surface displacements, may very well fail to produce an equilibrium stress field. One needs some assurance that the assumed deformation is actually the one produced.

In most problems of elastic deformation with prescribed surface displacements, there is no way to determine what the interior deformation may be until the stress-deformation relation has been specified. However, there are some problems with such a high degree of symmetry that only one state of deformation has all of the symmetries of the data. In such cases, the symmetrical deformation is the solution, unless the problem has more than one solution.

Of course, homogeneous deformations (translation-invariant in all directions) are suitable in homogeneous materials (properties translation-invariant), because the resulting stress field will be homogeneous, and thus an equilibrium field regardless of the specific values of the stresses. However, homogeneous deformations are not always experimentally convenient.

We say that a deformation is controllable for all materials in a given class if it can be supported in equilibrium with surface tractions alone, in every material in that class. Thus, for example, any given homogeneous deformation is controllable in homogeneous, elastic materials. Conversely, no deformation is controllable for this very broad class of materials except homogeneous ones.

An inhomogeneous deformation can be controllable only for some more narrowly defined class of materials. For example, expansion of a spherical shell and torsion of a circular cylindrical rod (Sec. 1) are controllable in homogeneous, isotropic, incompressible elastic materials. In these cases, and others, two kinds of restrictions on the class of materials are useful: symmetries, and constraints. Constraints reduce the number of kinematically admissible deformations, and symmetries reduce the number that have all of the symmetry of the data.

The idea of controllability is useful in other areas as well. Controllable viscometric flows [7,8,9] have been mentioned in Sec. 3. In such flows, a velocity field is prescribed in advance, and it is found to satisfy the momentum equations for any homogeneous, isotropic, incompressible, viscoelastic fluid. In problems of heat flux, we seek temperature fields that will produce a solenoidal field of heat flux in every material in some specified class. In any area, we call a field of some sort controllable if it leads to an exact solution of all relevant equations, for all materials in a prescribed class.

A. C. Pipkin

6.1 An Example of the Use of Constraints

To begin with an extreme example of the usefulness of constraints, let us consider homogeneous isotropic elastic materials that are both incompressible and reinforced with a family of parallel, inextensible fibers. Every plane deformation of such a material is controllable.

We consider only plane deformations, $x = x(X,Y)$, $y = y(X,Y)$, (and $z = Z$). Let the X-direction be the initial direction of the inextensible fibers. To pose a definite problem, consider a block of material with two of its edges initially along the X-axis and the Y-axis, and suppose that these two edges are mapped onto intersecting curves C_1 and C_2. Curiously enough, the problem is well-set, or at least it has no more than one solution. The constraints require that a fiber initially along $Y = Y_o$ map onto a curve that is parallel to C_1, at a distance Y_o from it. The position of the fiber along this curve is determined by the condition that the end $X = 0$ lies on C_2.

Because there is only one deformation that is even kinematically admissible, it doesn't matter what the relation between extra stress and deformation may be. The stress field that is produced will automatically be an equilibrium field. The reactions to the constraints will adjust themselves to equilibrate the extra stress produced by the deformation, whatever it may be.

To see how this happens within the mathematics, consider the expression for the stress:

$$\underline{\sigma} = -p\underline{I} + T\underline{t}\underline{t} + \underline{S}. \qquad (6.1.1)$$

Here $-p\underline{I}$ is the reaction to incompressibility, $T\underline{t}\underline{t}$ is the reaction to inextensibility, and $\underline{S}$ is the extra stress. The unit vector $\underline{t}$ is tangent to the fiber direction, and $T\underline{t}\underline{t}$ is a tensile stress along the fiber direction. The equilibrium equation is

$$\underline{\nabla}p = \underline{t}\,\underline{\nabla}\cdot(T\underline{t}) + T(\underline{t}\cdot\underline{\nabla})\underline{t} + \mathrm{Div}\,\underline{S}. \qquad (6.1.2)$$

By separating this into components in the fiber direction, the normal direction $\underline{n}$, and the z-direction $\underline{k}$, we obtain

$$\begin{aligned} \underline{t}\cdot\underline{\nabla}p &= \underline{t}\cdot\underline{\nabla}T + T\underline{\nabla}\cdot\underline{t} + (\mathrm{Div}\,\underline{S})\cdot\underline{t}, \\ \underline{n}\cdot\underline{\nabla}p &= T\underline{n}\cdot(\underline{t}\cdot\underline{\nabla})\underline{t} + (\mathrm{Div}\,\underline{S})\cdot\underline{n}, \\ \underline{k}\cdot\underline{\nabla}p &= 0. \end{aligned} \qquad (6.1.3)$$

(The z-component of Div $\underline{S}$ is zero since S_{zx} and S_{zy} must vanish, by symmetry.) No matter what function of x and y Div $\underline{S}$ may be, there are functions p and T that satisfy these equations, and we can determine them by integrating along characteristics.

6.2. Controllable Heat Conduction

It is possible to seek out controllable states systematically, rather than waiting to discover them as solutions of definitely prescribed problems. To illustrate the methods that are used, we consider problems of controllable heat conduction, which have been discussed by Petroski and Carlson [28] and Laws [29].

The materials to be considered are isotropic, and it is assumed that the heat flux $\underline{q}$ is determined by the temperature T and the temperature gradient, $\underline{\nabla}T$. In isotropic materials, $\underline{q}$ and $\underline{\nabla}T$ must be (anti-) parallel, so

$$\underline{q} = -K(T,\underline{\nabla}T\cdot\underline{\nabla}T)\underline{\nabla}T. \tag{6.2.1}$$

In a steady state, $\underline{\nabla}\cdot\underline{q} = 0$, and thus

$$K\nabla^2T + K_1\underline{\nabla}T\cdot\underline{\nabla}T + K_2\underline{\nabla}(\underline{\nabla}T\cdot\underline{\nabla}T)\cdot\underline{\nabla}T = 0. \tag{6.2.2}$$

Here K_1 is the derivative of K with respect to its i-th argument.

We say that a temperature field is controllable if it satisfies (6.2.2) no matter what form the function K has. At a given point, the values of K, K_1, and K_2 can be assigned arbitrarily by a suitable choice of the form of K. Hence, the coefficients of K, K_1, and K_2 must all vanish at that point if T is controllable. But the point is arbitrary, so a controllable field T must satisfy

$$\nabla^2T = \underline{\nabla}T\cdot\underline{\nabla}T = \underline{\nabla}(\underline{\nabla}T\cdot\underline{\nabla}T)\cdot\underline{\nabla}T = 0, \tag{6.2.3}$$

everywhere. Since $|\underline{\nabla}T|$ vanishes, T must be uniform. Thus, in the class of materials described by constitutive equations of the form (6.2.1), there are no controllable temperature fields except constant fields [28].

This does not mean that all controllable states are trivial, as Laws [29] has pointed out. Suppose that $\underline{q}$, rather than T, is to be controlled. In that case it is more convenient to use a constitutive equation of the form

$$-\underline{\nabla}T = R(T,\underline{q}\cdot\underline{q})\underline{q}. \tag{6.2.4}$$

Now, the right-hand side must be irrotational, since the left-hand side is:

$$R\underline{\nabla} \times \underline{q} + R_1 \underline{\nabla} T \times \underline{q} + R_2 \underline{\nabla}(\underline{q}\cdot\underline{q}) \times \underline{q} = \underline{0}. \tag{6.2.5}$$

For this to be true no matter what function R is, $\underline{q}$ must satisfy

$$\underline{\nabla} \times \underline{q} = \underline{\nabla} T \times \underline{q} = \underline{\nabla}(\underline{q}\cdot\underline{q}) \times \underline{q} = \underline{0}, \tag{6.2.6}$$

as well as $\underline{\nabla}\cdot\underline{q} = 0$.

The first equation implies that $\underline{q}$ has a potential, $\underline{q} = -\underline{\nabla}\theta$, say, and thus $\underline{q}$ is normal to a family of surfaces θ = const. Let us also write $\underline{q} = q\underline{n}$, where $\underline{n}$ is the unit normal to a θ-surface. The third equation states that $\underline{\nabla} q$ is parallel to $\underline{q}$, so q is constant over each θ-surface, i.e. $\underline{q} = q(\theta)\underline{n}$. The condition $\underline{\nabla}\cdot\underline{q} = 0$ yields

$$q\underline{\nabla}\cdot\underline{n} = -\underline{n}\cdot\underline{\nabla} q(\theta) = -q'(\theta)\underline{n}\cdot\underline{\nabla}\theta = qq'. \tag{6.2.7}$$

Thus, for $q \neq 0$, $\underline{\nabla}\cdot\underline{n}$ is also a function of θ, $q'(\theta)$. But $\underline{\nabla}\cdot\underline{n}$ is the mean curvature of the θ-surface. Thus, each θ-surface has constant mean curvature. Now, we recall that $|\underline{\nabla}\theta|$ is constant over each θ-surface, so they are parallel surfaces. We now save work by quoting a theorem: parallel surfaces of constant mean curvature are parallel planes, coaxial circular cylinders, or concentric spheres. Thus, θ is a function of x in an appropriate cartesian system, or of radius in a cylindrical or spherical system. The flux $\underline{q}$ is in parallel straight lines or is radial, and its magnitude depends only on x, or r, and it is solenoidal, so it has one of the forms

$$\begin{aligned} &\underline{q} = \text{const.}, \quad \underline{q} = \frac{c}{r}\,\underline{i}_r \quad \text{(cylindrical)}, \\ &\qquad\qquad \underline{q} = \frac{c}{r^2}\,\underline{i}_r \quad \text{(spherical)}. \end{aligned} \tag{6.2.8}$$

From the second equation in (6.2.6), T is a function of x, or r, as the case may be. It is determined by solving (6.2.4). For example, with $\underline{q}$ constant, in which case $T = T(x)$, we obtain

$$-T'(x) = R(T,q^2)q, \tag{6.2.9}$$

which yields

$$x = x_o - \int_{T_o}^{T} \frac{dT}{qR(T,q^2)} . \tag{6.2.10}$$

6.3 Controllable States of Dielectrics

The equations governing the electric field strength $\underline{E}$ and the dielectric displacement field $\underline{D}$ in electrostatics have the same form as those governing $-\underline{\nabla}T$ and $\underline{q}$ in heat conduction. $\underline{E}$ is irrotational, and $\underline{D}$ is solenoidal:

$$\underline{E} = -\underline{\nabla}V, \qquad \underline{\nabla}\cdot\underline{D} = 0. \tag{6.3.1}$$

However, in an isotropic dielectric the constitutive equation would have the form

$$\underline{D} = f(\underline{E}\cdot\underline{E})\underline{E} \qquad \text{or} \qquad \underline{E} = g(\underline{D}\cdot\underline{D})\underline{D}, \tag{6.3.2}$$

with no explicit V-dependence of f or g to match the T-dependence of the analogous functions in heat conduction.

If $\underline{D}$ is to be controlled, the possibilities are the same as those listed in (6.2.8) for $\underline{q}$; the T-dependence of the resistivity was not used in obtaining (6.2.8). Thus, the dielectric displacement is controllable in parallel-plate, cylindrical and spherical condensers.

If $\underline{E}$, or rather V, is to be controlled, in place of the three controllability conditions (6.2.3) for the analogous

heat conduction case, there are only two:

$$\nabla^2 V = \underline{\nabla} V \cdot \underline{\nabla}(\underline{\nabla} V \cdot \underline{\nabla} V). \tag{6.3.3}$$

It is much too hard for me to prove, but these equations imply that V can be represented in an appropriate cylindrical system by

$$V = -a\theta - bz. \tag{6.3.4}$$

Then

$$\underline{E} = \frac{a}{r}\,\underline{i}_\theta + b\underline{i}_z \tag{6.3.5}$$

and

$$\underline{D} = f\left(\frac{a^2}{r^2} + b^2\right)\left(\frac{a}{r}\,\underline{i}_\theta + b\underline{i}_z\right). \tag{6.3.6}$$

The lines of flux are helical, in general, and the magnitude of the flux is constant over each of the cylinders on which these helices lie, so the flux is obviously divergence-free.

6.4 Electrical Conduction, Diffusion through Porous Media, Magnetism

There are several theories that involve the steady-state flux of some conserved quantity under the action of a conservative driving force. The forces and fluxes for a few cases are

	Force	Flux
Heat conduction	$-\underline{\nabla} T$	$\underline{q}$
Electrostatics	$\underline{E} = -\underline{\nabla} V$	$\underline{D}$
Magnetostatics	$\underline{H}$	$\underline{B}$
Electrical conduction	$\underline{E} = -\underline{\nabla} V$	$\underline{J}$
Diffusion through porous media	$-\underline{\nabla} p$	$\underline{u}$

In each case, the force $\underline{F}$ is irrotational, $\underline{\nabla} \times \underline{F} = \underline{0}$, and the flux $\underline{f}$ is solenoidal, $\underline{\nabla} \cdot \underline{f} = 0$, in appropriate problems. In isotropic materials, the flux is controllable if it is a radial field, in one, two, or three dimensions, of the form (6.2.8), and the force is controllable if it is a helical field of the form (6.3.5) (except in heat conduction with temperature-sensitive conductivity).

6.5 Controllable Elastic Deformations

The preceding problems illustrate the methods of analysis used in seeking controllable solutions. Instead of boundary conditions and well-set problems, there are additional equations. The over-determined system is solved by geometrical methods.

In the theory of finite deformations of homogeneous, isotropic, incompressible, elastic materials, most of the known controllable deformations were first obtained as solutions of definite problems of technological interest. The remainder have been found by systematic search. The search has not been completed because the analysis is much too hard in certain degenerate cases.

The known examples are the following:

0. Homogeneous (isochoric) deformations.
1. Expansion of a spherical shell, or a shell first turned inside out:

$$r = \pm(R^3 - R_0^3 \pm r_0^3)^{1/3}, \quad \theta = \Theta, \quad \phi = \pm\Phi. \tag{6.5.1}$$

2. Certain deformations of right circular cylindrical rods and tubes:

$$r^2 = AR^2 + B, \quad \theta = C\Theta + DZ, \quad z = E\Theta + FZ. \tag{6.5.2}$$

3. Bending and associated deformations:

$$r^2 = 2AX + B, \quad \theta = CY + DZ, \quad z = EY + FZ. \tag{6.5.3}$$

4. Straightening and associated deformations:

$$x = \tfrac{1}{2} AR^2 + B, \quad y = C\Theta + DZ, \quad z = E\Theta + FZ. \tag{6.5.4}$$

5. Azimuthal shearing and associated deformations:

$$r = AR, \quad \theta = B \log R + C\Theta, \quad z = Z/A^2C. \tag{6.5.5}$$

In families 2, 3, and 4, the condition of no volume change requires that

$$A(CF - DE) = 1. \tag{6.5.6}$$

The coordinates are cartesian or cylindrical, except in family 1. Big letters are initial coordinates, and little letters are final coordinates.

There are many classes of materials that include homogeneous, isotropic, incompressible, elastic materials as special cases. A deformation can be controllable in one of these wider classes only if it is controllable in each special case, so the families listed above turn up over and over again.

To understand how changes in the class of materials may affect the list of controllable deformations, let us consider, first, some geometrical features of the deformations previously listed.

One of the three fields of principal directions of strain is normal to a family of surfaces, which we call

principal membranes. Accordingly, in isotropic materials there is no shearing stress between principal membranes. Next, the three principal extension ratios are constant over each principal membrane. For the class of materials considered, this implies that the principal stresses (apart from the reaction pressure) are also constant over each principal membrane.

In Family 1 the principal membranes are spherical. The analysis of this deformation in Sec. 1.6 made explicit use of the principal membranes.

In Family 2, they are cylindrical both before and after the deformation, but possibly with different radii. Families 3 and 4 are limiting cases of Family 2, with principal membranes that are parallel planes, cylinders with infinite radii, either before the deformation or after it. In all of these cases, the two fields of principal directions tangential to the principal membranes form two orthogonal families of geodesics on them, and in consequence, the field of extra stress is so highly symmetrical that it is self-equilibrating in the tangential directions. Equilibrium in the normal direction is produced by a suitable reaction pressure, which is constant over each principal membrane.

In Families 0 and 5 the principal stretches are the same everywhere, so there are three orthogonal families of principal membranes. In Family 0, each is a set of parallel planes. In Family 5, the planes z = constant are the obvious

set. The fields of principal directions tangential to these planes have constant components in polar coordinates, so the same is true of the extra stress, and this is enough symmetry to guarantee equilibrium.

6.6 Fiber-reinforced Materials

A deformation can be controllable in some class of anisotropic materials only if it is controllable in isotropic materials, because a material with specified point symmetry might accidentally have more symmetry than is demanded.

Consider, for example, transversely isotropic materials (that are homogeneous, incompressible, and elastic). For a definite physical picture, think of an isotropic matrix reinforced with parallel elastic fibers. Now, to find controllable deformations of such a material, we examine the list in Sec. 6.5. We seek cases in which the fiber direction does not spoil the symmetry of the problem. For this to be true, the fiber direction must be either parallel or perpendicular to the principal membrane at each point.

Family 1, the spherical case, is ruled out because straight, parallel fibers cannot be normal to the spheres everywhere, or tangential everywhere.

In Families 2 and 4, the principal membranes are coaxial cylinders initially. No symmetry is lost if the fibers lie along the axial direction initially, so, if they do, these deformations are all controllable.

In Family 3, the principal membranes are parallel planes initially, and they are bent into coaxial circular cylinders. If the preferred direction is initially perpendicular to the principal membranes, or in any tangential direction, no relevant symmetry is lost, and these deformations are controllable.

Homogeneous deformations, Family 0, are all controllable in any homogeneous material, so the fiber direction is irrelevant.

The deformations in Family 5 remain controllable in fiber-reinforced materials if the fibers lie along the axial direction.

Since these lectures are about problem-solving, rather than the theory of anything, I should claim that a large number of problems involving fiber-reinforced materials have just been solved. Of course, it remains to write down the extra stress and determine the reaction pressure, and to look for features of the solutions that may be particularly interesting or valuable. All of this requires a great deal of work, but the work will not involve any worries about solving partial differential equations.

6.7 Layered Materials

Instead of relaxing the requirement that the material be isotropic, as in the preceding section, consider the possibility of relaxing the homogeneity requirement. The picture in terms of principal membranes makes it clear that only layered materials can be considered, with properties constant over principal membranes. Conversely, it is easy to see that Families 0 to 4 remain controllable if the lamination coincides with the principal membranes. Family 5 doesn't work. The trouble is that in this family, unlike

the others, the reaction pressure is not constant over the principal membranes $z = \text{const.}$, so letting properties be z - dependent spoils the symmetry.

6.8 Viscoelastic Materials

Now consider relaxing the requirement that the material be elastic. When we take into account the stress relaxation that occurs in real materials, we must be concerned with controllable motions, rather than single deformations.

Since elastic materials are a special case of viscoelastic materials, a motion that is controllable in viscoelastic materials must at each instant be a controllable elastic deformation. However, this is not sufficient.

In all cases except homogeneous deformations, controllability depends upon the fact that there is no shearing stress between principal membranes. This requires that the normal direction be a principal direction of stress. In isotropic elastic materials, this requirement is met because the principal directions of stress coincide with the principal directions of strain, and the direction normal to a principal membrane is a principal direction of strain by definition.

However, in isotropic viscoelastic materials it is no longer true that the principal directions of stress and strain always coincide. The stress depends on past deformations as well as on the present state, and the principal directions of stress might well be closer to fibers that were principal fibers previously, than to those that are principal fibers presently.

Nevertheless, there is no real difficulty. For, if a given fiber is a principal fiber of strain at all times, then by symmetry it must also be a principal direction of stress at all times. Thus, to assure that there is no shearing stress between principal membranes, it is necessary and sufficient that the material surfaces which are principal membranes at one instant are also principal membranes at every other instant. This requirement is automatically satisfied if the motion is generated by varying the parameters within one of the listed families of elastic deformations. In general it rules out changing from one family to another during the course of the motion, although there are exceptions.

Thus, for example, time-dependent radial expansion of a spherical shell is controllable in viscoelastic materials because the same spherical membranes are principal membranes at all times. Eversion is not controllable, because intermediate states in the process of turning the shell inside out are not controllable elastic deformations.

Families 2 and 4 should be considered together. The principal membranes are generally cylindrical, with time-varying radii. Family 4 describes the deformation at instants when the cylinders happen to be completely flattened out into parallel planes. It is considered as distinct from Family 2 only because of notational difficulty.

In Family 3, the principal membranes are parallel planes initially, and by varying the parameters we obtain motions that include flexing the membranes variably in time. Motions generated by varying the parameters in Family 5 have

planes z = constant as principal membranes at all times.

In all of these cases, the principal directions of strain tangential to the membranes generally involve different fibers at different times, so there is no way to tell where the tangential principal directions of stress may be. However, controllability involves only the symmetry of these fields of principal stresses, and does not rely on knowing their directions exactly. Time-dependent variation of parameters does not spoil the symmetry.

Quasi-static motions of the kinds described are, accordingly, controllable. A closer look would be required to see which cases might remain controllable if inertia is taken into account.

These motions have been examined in more detail by Carroll [30,31] and Wineman [32]. Carroll has also considered transversely isotropic materials [31] and orthotropic materials [33].

6.9 Flux Through Highly Deformable Materials

Theories of flux, such as those listed in Sec. 6.4, can be combined with the theory of large elastic deformation. For example, one might consider heat conduction in deformed materials, as Petroski and Carlson [34] and Laws [29] have done. That theory is formally identical to the theory of electrical conduction in deformable materials, aside from the temperature-dependence of the conductivity or resistivity and the stress.

A. C. Pipkin

Rubber loaded with carbon black is a good electrical conductor, and the conductivity is strongly dependent on the state of deformation, so let us consider that case. When the material is deformed, the relation between current density $\underline{J}$ and field strength $\underline{E}$ is like that for an undeformed orthotropic material. The three principal planes of strain at a point act as planes of reflectional symmetry for the conductivity. If $\underline{E}$ lies along a principal direction, so does $\underline{J}$, by symmetry. If $\underline{E}$ is orthogonal to a principal direction, so is $\underline{J}$, although it need not be parallel to $\underline{E}$. When $\underline{E}$ does not lie in any principal plane, symmetry gives no conclusion about the direction of $\underline{J}$.

Just as the deformation affects the conductivity, the electric field may affect the stress. However, let us leave such electrostrictive effects aside for the time being.

If a state of deformation and flux is to be controllable, then in particular, the flux must be controllable even when the deformation has no effect on it. Hence, controllable states must involve either a (generally) helical electric field, of the form (6.3.5), or a field of current density that is uniform, or radial in two or three dimensions, of the form (6.2.8).

For a given field of one of these forms to remain controllable when the deformation does affect the conductivity, it is necessary that the field and the deformation can be combined without spoiling the symmetry of either one. Let us consider each family of controllable elastic deformations in turn.

A uniform field, $\underline{E}$ or $\underline{J}$, can of course be combined with any homogeneous deformation to give a controllable state.

In the case of radial expansion of a spherical shell, a radial field of current can be fitted in without spoiling the symmetry.

In families 2 and 3, principal membranes are coaxial cylinders in the deformed state. Radial current, along a principal direction of strain, gives a controllable state, and so do helical electric fields with lines of force lying tangential to the principal membranes.

In Family 4, the principal membranes are parallel planes in the deformed state. Uniform currents perpendicular to these planes and uniform electric fields in any tangential direction produce controllable states.

Family 5 can be combined with uniform electric field and uniform current in the axial direction.

6.10 Electrostriction

In problems involving polarization and electrostriction in deformed materials, the dielectric displacement $\underline{D}$ plays the role that $\underline{J}$ did in electrical conduction. However, it is necessary to take into account the effect of the fields on the stress, which we omitted to do in discussing conduction.

When $\underline{E}$, or $\underline{D}$, lies along a principal direction of strain, then the principal directions of stress must still coincide with the principal directions of strain. If the field lies orthogonal to a principal direction of strain, that particular principal

direction must be a principal direction of stress, although the symmetry is spoiled within the principal plane in which the field lies.

With these observations, examination of the controllable states listed in Sec. 6.9 (reading $\underline{D}$ for $\underline{J}$) shows that each one is still controllable under electrostrictive stresses.

For this particular theory, it has been proved [35] that the preceding list includes all states that are controllable. Thus, although we do not know that the list of controllable elastic deformations listed in Sec. 6.5 is complete, we do know that there are no others that remain controllable under electrostriction. The proof was not direct; it involved systematically searching out all possible controllable cases, using methods similar to those employed in Secs. 6.2 and 6.3. The curious Family 5 was turned up in the course of this search.

Acknowledgement

Preparation of this review was carried out under a grant from the U. S. National Science Foundation to Brown University. Their support is gratefully acknowledged.

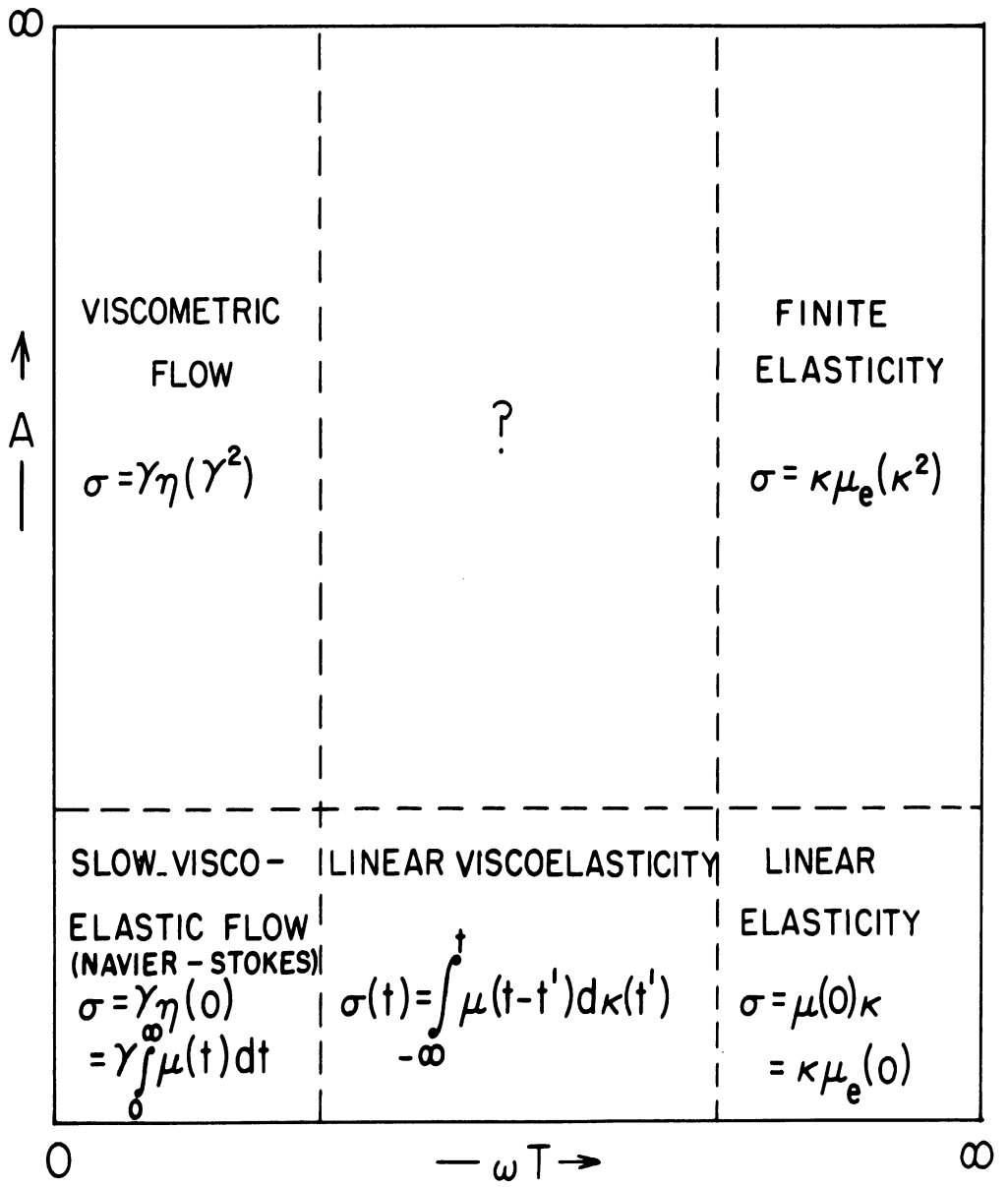


FIG. 1 FLOW DIAGNOSIS DIAGRAM (SEE SEC. 2.2)

References A. C. Pipkin

[1] Green, A. E., and W. Zerna: Theoretical Elasticity. Oxford: Clarendon Press, 1954.

[2] Green, A. E., and J. E. Adkins: Large Elastic Deformations and Non-Linear Continuum Mechanics. Oxford: Clarendon Press, 1960.

[3] Truesdell, C., and W. Noll: The Nonlinear Field Theories of Mechanics. Encyclopedia of Physics III/3. New York: Springer-Verlag, 1965.

[4] Jaunzemis, W.: Continuum Mechanics. New York: MacMillan, 1967.

[5] Treloar, L. R. G.: The Physics of Rubber Elasticity. Second Edition. Oxford: Clarendon Press, 1967.

[6] Jones, T. E. R., and K. Walters: A Theory for the Balance Rheometer. Brit. J. Appl. Phys. (J. Phys. D) 2, 815 (1969).

[7] Yin, W.-L., and A. C. Pipkin: Kinematics of Viscometric Flow. ARMA (Forthcoming).

[8] Yin, W.-L.: Ph.D. thesis, Brown University, Division of Applied Mathematics, 1969. University Microfilms, Ann Arbor, Michigan.

[9] Pipkin, A. C.: Controllable Viscometric Flows. Quart. Appl. Math. 26, 87 (1968).

[10] Giesekus, H.: Some Secondary Flow Phenomena in General Viscoelastic Fluids. Proc. Fourth Int. Cong. on Rheology, Part 1, 249 (1965).

[11] Griffiths, D. F., D. T. Jones, and K. Walters: A Flow Reversal Due to Edge Effects. J. Fluid Mech. 36, 161 (1969).

[12] Barnes, H. A., and K. Walters: Dynamic Similarity and Drag Reduction in Flow of Elastic Liquids through Curved Pipes. Nature 219, 57 (1968).

[13] Osaki, K., M. Tamura, M. Kurata, and T. Kotaka: Complex Modulus of Concentrated Polymer Solutions in Steady Shear. J. Phys. Chem. 69, 4183 (1965).

[14] Booij, H. C.: Influence of Superimposed Steady Shear Flow on the Dynamic Properties of Non-Newtonian Fluids. Rheologica Acta 5, 215 (1966).

[15] MacDonald, I. F., and R. B. Bird: Complex Modulus of Concentrated Polymer Solutions in Steady Shear. J. Phys. Chem. 70, 2068 (1966).

[16] Simmons, J. M.: Dynamic Modulus of Polyisobutylene Solutions in Superposed Steady Shear Flow. Rheologica Acta 7, 184 (1968).

[17] Tanner, R. I., and J. M. Simmons: Combined Simple and Sinusoidal Shearing in Elastic Liquids. Chem. Eng. Sci. 22, 1803 (1967).

[18] Pipkin, A. C. and D. R. Owen: Nearly Viscometric Flows. Phys. Fluids 10, 836 (1967).

[19] Pipkin, A. C.: Small Displacements Superposed on Viscometric Flow. Trans. Soc. Rheology 12, 397 (1968).

[20] Tanner, R. I.: Plane Creeping Flows of Incompressible Second-order Fluids. Phys. Fluids 9, 1246 (1966).

[21] Tanner, R. I., and A. C. Pipkin: Intrinsic Errors in Pressure-Hole Measurements. Trans. Soc. Rheology. Forthcoming.

[22] Broadbent, J. M., A. Kaye, A. S. Lodge, and D. G. Vale: Possible Systemátic Error in the Measurement of Normal Stress Differences in Polymer Solutions in Steady Shear Flow. Nature 217, 55 (1968).

[23] Giesekus, H.: Die Rheologische Zustandsgleichung elasto-viskoser Flüssigkeiten – insbesondere von Weissenberg-Flüssigkeiten – für allgemeine und stationäre Fliessvorgänge. ZAMM 42, 32 (1962).

[24] Huilgol, R.: On the Properties of the Motion with Constant Stretch History Occurring in the Maxwell Rheometer. Math. Res. Center, U. of Wisc., Madison (1969).

[25] Maxwell, B., and R. P. Chartoff: Studies of a Polymer Melt in an Orthogonal Rheometer. Trans. Soc. Rheology 9, 41 (1965).

[26] Ballman, R. L.: Extensional Flow of Polystyrene Melt. Rheologica Acta 4, 137 (1965).

[27] Metzner, A. R., E. A. Uebler, and C. F. Chan Man Fong: Converging Flows of Viscoelastic Fluids. U. of Delaware, Newark (1968).

[28] Petroski, H. J., and D. E. Carlson: Controllable States of Rigid Heat Conductors. ZAMP 19, 372 (1968).

[29] Laws, N.: Controllable States in Thermoelasticity.

[30] Carroll, M. M.: Controllable Deformations of Incompressible Simple Materials. I.J.E.S. 5, 515 (1967).

[31] Carroll, M. M.: Finite Deformations of Incompressible Simple Solids. I. Isotropic Solids. II. Transversely Isotropic Solids. Quart. J. Mech. and Appl. Math. 21, 147-170, 269-278 (1968).

[32] Wineman, A. S. : Motions Possible in Every Incompressible Isotropic Simple Solid and Fluid. Dep't Eng. Mech., U. of Mich., Ann Arbor (1967).

[33] Carroll, M. M.: Finite Bending, Stretching, and Shearing of a Block of Orthotropic, Incompressible Simple Solid. J. Appl. Mech.

[34] Petroski, H. J., and D. E. Carlson: Controllable States of Elastic Heat Conductors. ARMA 31, 127 (1968).

[35] Singh, M., and A. C. Pipkin: Controllable States of Elastic Dielectrics. ARMA 21, 169 (1966).

[36] Kearsley, E.: Intrinsic Errors for Pressure Measurements in a Slot Along a Flow. National Bureau of Standards, Washington, D. C. (1969).

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C. I. M. E.)

R. S. RIVLIN

AN INTRODUCTION TO NON-LINEAR CONTINUUM MECHANICS

Corso tenuto a Bressanone dal 3 all'11 settembre 1969

CONTENTS

Chapter 4. Thermodynamic Considerations

4.1 Description of the thermal state

4.2 The energy balance equation

4.3 The heat flux vector

4.4 The equations of motion and the dissipation equation

4.5 The second law of thermodynamics

4.6 Implications of the Clausius inequality

4.7 Elastic materials

4.8 Constraints

4.9 Cauchy elastic materials

Chapter 5. Finite Elasticity Theory

5.1 Introduction

5.2 Restrictions due to effect of a superposed rotation -

(i) Fromulation of the problem
(ii) Method I
(iii) Method 2
(iv) The stress

5.3 Restrictions imposed by material symmetry

5.4 Symmetry groups

5.5 The invariant-theoretical problem posed by material symmetry

5.6 Constitutive equations of finite elasticity theory for isotropic materials

5.7 Constitutive equations of finite elasticity for incompressible isotropic materials

Chapter 6. Some Results in the Theory of Invariants

6.1 Definitions

6.2 Integrity bases for vectors under the full and proper orthogonal groups

6.3 Integrity basis for symmetric second-order tensors under the full or proper orthogonal group

6.4 Invariant tensor-valued functions

6.5 Invariants of functionals

6.6 Invariant tensor-valued functionals

Chapter 7. Viscoelastic Materials

7.1 Introduction

7.2 Restrictions due to effect of a superposed rotation -
(i) Formulation of the problem
(ii) Method 1
(iii) Method 2
(iv) The Cauchy and Kirchoff-Piola stresses

7.3 Hereditary materials

7.4 Constitutive equations of the differential type

7.5 The Rivlin-Ericksen tensors

7.6 Restrictions imposed by material symmetry on constitutive equations of the functional type

7.7 Restrictions imposed by material symmetry on constitutive equations of the differential type

7.8 Constraints

7.9 Viscoelastic fluids.

R. S. Rivlin

Chapter 1

INTRODUCTION

Since the Second World War there has been a considerable interest in the development of phenomenological theories describing the relation between force and deformation in bodies of material which do not obey either the linear laws of the classical theories of elasticity and the hydrodynamics of viscous fluids, or the simple and explicit non-linear laws of gas dynamics.

During the first half of the present century much effort has been devoted to the elaboration of these classical theories and their application to specific problems of importance in - or at any rate suggested by - some area of science and technology. During this period the development of non-linear theories in the same rational spirit as inspired the development of the classical theories in the nineteenth century was, in large measure, neglected. This was undoubtedly due, at any rate in part, to a preoccupation with the more developed theories which were nearer the point at which they might provide some insight into technological or scientific problems.

In addition, advances may well have been inhibited by a conviction that no progress could be made in developing non-linear theories unless some completely explicit constitutive equation could be written down for the material. This was usually chosen on the basis of an alleged simplicity - the "simplest" constitutive equation which could yield the type of phenomena of interest - or on the basis of some microscopic or molecular model of a particular material. One of the difficulties with the former approach lay in the fact that simplicity is very much a subjective matter, depending considerably on the choice of the variables in terms of which the relation is expressed. Moreover, there is usually no special reason why any particular material should obey the simplest law. The main difficulty with the second approach is that even if the microscopic or molecular structure is well understood, the passage to the constitutive equation, expressed in phenomenological terms, is usually very difficult and cannot be made at all without so many idealizations of both the model and the mathematics as to leave the significance of the result in serious question.

The more modern approach stems largely from the realization that it is possible to write down rather general constitutive equations in canonical form from purely phenomenological considerations. This realization is already involved

in the theory of finite elasticity as it existed at the end of the last century. There, the constitutive equation is given by a statement that the strain-energy function must depend on the nine deformation gradients and it is shown (§5.2) that this dependence must be through the six independent components of the strain-energy function, further restriction of form resulting if the material has some symmetry. In the case when the material is isotropic, the latter dependence is through only three specified functions of these strain components (§5.6). Much of the development of non-linear continuum mechanics has consisted of the extension of this principle to dissipative materials in which the constitutive equation takes the form of an expression for the stress tensor in terms of one or more kinematic tensors (see Chapter 7 below). In this development the main considerations which are introduced are twofold. One is that the superposition on the deformation of the body of a rigid rotation results in the rotation of the stress tensor by an equal amount. The other is that in physical and engineering problems the material with which one is concerned has some symmetry and this enables one, in all cases, to write the constitutive equation in canonical form. The mathematical technique for doing this stems from the classical theory of invariants, but considerable development of the theory has been necessary for this purpose. This development is discussed in Chapter 6.

The question of the application of the theories developed here to specific problems is taken up in the lectures of Professor Pipkin. The technique he has used for presenting the essential content of the various problems is very largely that used in papers by Rivlin, and Green and Rivlin*, although not that by which the problems were originally solved - the direct application of the constitutive equations and equations of motion developed here. The latter procedure is, of course, of much wider applicability although it may present less feel for the physical content of the calculations.

* R.S. Rivlin, Large Elastic Deformations, Chapter 10, in "Rheology", Vol.1, ed. F.R. Eirich, (publ. Academic Press, New York, 1956); R.S. Rivlin, Proc. First International Conference on Hemorheology, Reykjavik, Iceland, 1966, p.157 (publ. Pergamon, Oxford, 1967); W.A. Green and R.S. Rivlin, Acta Mechanica, 5, 254 (1968).

R. S. Rivlin

Chapter 2

KINEMATICS OF DEFORMATION

1. Description of the deformation

As a body undergoes deformation, the vector positions of the particles of the body with respect to a fixed origin change. We denote by $\underset{\sim}{x}$ the vector position at time t of a generic particle with respect to a fixed origin O. Let $\underset{\sim}{X}$ be the vector position of this particle with respect to O at some reference time T (say). Then, we can use $\underset{\sim}{X}$ as a means of identifying the particles of the body. If now we specify the dependence of the vector $\underset{\sim}{x}$ on $\underset{\sim}{X}$ and t, thus

$$\underset{\sim}{x} = \underset{\sim}{x}(\underset{\sim}{X},t) \ , \tag{2.1.1}$$

then we have a complete description of the deformation of the body.

We now consider that at each instant t, $\underset{\sim}{x}$ is a continuous, differentiable function of $\underset{\sim}{X}$ except possibly at a finite number of points or on a finite number of lines and surfaces. We also assume that the relation (2.1.1) between $\underset{\sim}{x}$ and $\underset{\sim}{X}$ is uniquely invertible for all values of t except possibly at a finite number of points or on a finite number of lines and surfaces.

We shall use a rectangular cartesian reference system x in which the components of $\underset{\sim}{X}$ are denoted X_A and the components of $\underset{\sim}{x}$ are denoted x_i. We shall use an analogous notation for the components in this system of other vectors. In this notation we can rewrite (2.1.1) as

$$x_i = x_i(X_A, t) \ , \tag{2.1.2}$$

in which x_i is differentiable with respect to X_A, except possibly at a finite number of points, or on a finite number of lines and surfaces. The nine spatial derivatives $\partial x_i/\partial X_A$ are called the deformation gradients.

2. The Cauchy strain

Consider two neighboring particles P_1 and P_2 whose vector positions at the reference time T are $\underset{\sim}{X}$ and $\underset{\sim}{X} + d\underset{\sim}{X}$. Let their vector positions be $\underset{\sim}{x}$ and $\underset{\sim}{x} + d\underset{\sim}{x}$ at time t. Let dL and dℓ be the distances between the particles at times T and t respectively. Then,

$$(dL)^2 = d\underset{\sim}{X}.d\underset{\sim}{X} = dX_A dX_A$$

and (2.2.1)

$$(d\ell)^2 = d\underset{\sim}{x}.d\underset{\sim}{x} = dx_i dx_i .$$

Now*,

$$dx_i = x_{i,A} dX_A \ . \tag{2.2.2}$$

Introducing this into $(2.2.1)_2$, we obtain

$$(d\ell)^2 = C_{AB} dX_A dX_B \ , \tag{2.2.3}$$

where

$$C_{AB} = x_{i,A} x_{i,B} \ . \tag{2.2.4}$$

If we take a second rectangular cartesian system $\bar{x}$ in which the components of the vectors $\underset{\sim}{X}$ and $\underset{\sim}{x}$ are $\bar{X}_A$ and $\bar{x}_i$ respectively, we have

$$(d\ell)^2 = \bar{C}_{PQ} d\bar{X}_P d\bar{X}_Q \ , \tag{2.2.5}$$

---

*Throughout we shall use the notation $_{,A}$ to denote $\partial/\partial X_A$ and the notation $_{,i}$ to denote $\partial/\partial x_i$.

where

$$\bar{C}_{PQ} = \frac{\partial \bar{x}_i}{\partial \bar{X}_P} \frac{\partial \bar{x}_i}{\partial \bar{X}_Q} . \tag{2.2.6}$$

It is easy to show that

$$\bar{C}_{PQ} = \bar{X}_{P,A} \bar{X}_{Q,B} C_{AB} . \tag{2.2.7}$$

C_{AB} and $\bar{C}_{AB}$ are thus the components in the systems x and $\bar{x}$ respectively of a second-order cartesian tensor. This is called the Cauchy strain tensor. We note from its definition by (2.2.4) that it is a symmetric tensor.

Let $\underset{\sim}{L}$ and $\underset{\sim}{\ell}$ be unit vectors in the direction of the linear element from P_1 to P_2 at times T and t respectively. Then,

$$\underset{\sim}{L} = d\underset{\sim}{X}/dL \quad \text{and} \quad \underset{\sim}{\ell} = d\underset{\sim}{x}/d\ell . \tag{2.2.8}$$

In cartesian notation, these expressions may be written as

$$L_A = dX_A/dL \quad , \quad \ell_i = dx_i/d\ell . \tag{2.2.9}$$

Introducing (2.2.9) into (2.2.2), we obtain

$$\ell_i = \frac{dL}{d\ell} x_{i,A} L_A . \tag{2.2.10}$$

Since

$$\underset{\sim}{\ell} . \underset{\sim}{\ell} = \ell_i \ell_i = 1 , \tag{2.2.11}$$

we have, from (2.2.9) and (2.2.10),

$$\left(\frac{d\ell}{dL}\right)^2 = C_{AB} L_A L_B . \tag{2.2.12}$$

This result can also be obtained directly from (2.2.3) by dividing throughout by $(dL)^2$.

3. The volume change

Consider an elementary tetrahedron in the body which at time T has a vertex located at the point P, with vector position $\underset{\sim}{X}$ relative to the fixed origin O, and three edges PQ_1, PQ_2, PQ_3 parallel to the axes of the rectangular cartesian system x. Let us suppose that the vector positions of Q_1, Q_2, Q_3 relative to P are $d\underset{\sim}{X}^{(1)}, d\underset{\sim}{X}^{(2)}, d\underset{\sim}{X}^{(3)}$. We denote the components in the system x of these three vectors by $(dX_1,0,0)$, $(0,dX_2,0)$, $(0,0,dX_3)$ respectively. The volume dV of the tetrahedron $PQ_1Q_2Q_3$ is given by

$$dV = \frac{1}{6}[d\underset{\sim}{X}^{(1)}, d\underset{\sim}{X}^{(2)}, d\underset{\sim}{X}^{(3)}] = \frac{1}{6}\, dX_1 dX_2 dX_3 , \qquad (2.3.1)$$

where the square brackets denote the scalar triple product of the three vectors.

In the deformation P, Q_1, Q_2, Q_3 move to, say, p, q_1, q_2, q_3 respectively. The vector positions at time t of q_1, q_2, q_3 with respect to p are given by $\underset{\sim}{x}_{,1}dX_1$, $\underset{\sim}{x}_{,2}dX_2$, $\underset{\sim}{x}_{,3}dX_3$ respectively. The components of these three vectors in the system x are

$$x_{i,1}dX_1 , \quad x_{i,2}dX_2 , \quad x_{i,3}dX_3$$

respectively. The volume dv of the tetrahedron $pq_1q_2q_3$ is given by

$$dv = \frac{1}{6}[\underset{\sim}{x}_{,1}dX_1, \underset{\sim}{x}_{,2}dX_2, \underset{\sim}{x}_{,3}dX_3]$$

$$= \frac{1}{6}\, \varepsilon_{ijk}\, x_{i,1}x_{j,2}x_{k,3}\, dX_1 dX_2 dX_3 . \qquad (2.3.2)$$

From (2.3.2) and (2.3.1) we have

$$\frac{dv}{dV} = \varepsilon_{ijk} x_{i,1} x_{j,2} x_{k,3} \cdot \tag{2.3.3}$$

This provides a formula for the ratio between the volume of an element of the material at times t and T respectively. It may also be written as the Jacobian determinant of $\underset{\sim}{x}$ with respect to $\underset{\sim}{X}$, thus:

$$\frac{dv}{dV} = \frac{\partial \underset{\sim}{x}}{\partial \underset{\sim}{X}} = |x_{i,A}| \cdot \tag{2.3.4}$$

We can express dv/dV in terms of the Cauchy strain tensor. We see from the definition of C_{AB} given in (2.2.4) that

$$|C_{AB}| = |x_{i,A}|^2 \cdot \tag{2.3.5}$$

With (2.3.4), we obtain

$$\frac{dv}{dV} = |C_{AB}|^{1/2} \cdot \tag{2.3.6}$$

4. The unit normal to a surface

We consider a surface drawn in a material and deforming with it; i.e. particles which are on the surface at time T remain on it as it deforms. Let the equation of the surface in the system x at time T be

$$\Phi(X_A) = 0 \ . \tag{2.4.1}$$

Then its equation at time t is

$$\phi(x_i) = 0 \ , \tag{2.4.2}$$

where $\phi(x_i)$ is defined by

$$\phi(x_i) = \Phi[X_A(x_i)] \ . \tag{2.4.3}$$

Conversely, we may consider $\Phi(X_A)$ to be defined in terms of $\phi(x_i)$ by

$$\Phi(X_A) = \phi[x_i(X_A)] \ . \tag{2.4.4}$$

Let $\underset{\sim}{N}$ be the outward-drawn normal to the surface (2.4.1) at the particle P, i.e. at $\underset{\sim}{X}$, and let $\underset{\sim}{n}$ be the outward-drawn normal to the surface (2.4.2) at the same particle P, i.e. at $\underset{\sim}{x}$. Then the components N_A and n_i, in the system x, of $\underset{\sim}{N}$ and $\underset{\sim}{n}$ respectively are given by

$$N_A = \Lambda\Phi_{,A} \quad \text{and} \quad n_i = \lambda\phi_{,i} \ , \tag{2.4.5}$$

where*

$$\Lambda = (\Phi_{,C}\Phi_{,C})^{-1/2} \quad \text{and} \quad \lambda = (\phi_{,j}\phi_{,j})^{-1/2} . \tag{2.4.6}$$

From (2.4.4), we obtain

$$\Phi_{,A} = \phi_{,i}x_{i,A} . \tag{2.4.7}$$

Using (2.4.5) we obtain from (2.4.7)

$$N_A = \frac{\Lambda}{\lambda} n_i x_{i,A} . \tag{2.4.8}$$

*Our taking Λ and λ as the inverses of the positive square roots, defines essentially what we mean by the outward-drawn normal in each case.

5. The Finger strain

We now consider a second particle $\overline{P}$ neighboring the particle P, but not lying on the surface. Let $\underset{\sim}{X} - d\underset{\sim}{X}$ and $\underset{\sim}{x} - d\underset{\sim}{x}$ be the position vectors of $\overline{P}$ at times T and t respectively. Let dN and dn denote the perpendicular distances from $\overline{P}$ to the tangent planes at P at times T and t respectively. Then, dN is the distance from $\underset{\sim}{X} - d\underset{\sim}{X}$ to the tangent plane to (2.4.1) at $\underset{\sim}{X}$. We thus have

$$dN = N_A dX_A = \Lambda \Phi_{,A} dX_A \ . \tag{2.5.1}$$

Similarly dn is the distance from $\underset{\sim}{x} - d\underset{\sim}{x}$ to the tangent plane to (2.4.2) at $\underset{\sim}{x}$. Thus,

$$dn = n_i dx_i = \lambda \phi_{,i} dx_i \ . \tag{2.5.2}$$

With (2.4.7), we obtain from (2.5.1) and (2.5.2)

$$\frac{dn}{dN} = \frac{\lambda}{\Lambda} \ . \tag{2.5.3}$$

From (2.5.3) and (2.4.8) we have

$$N_A \frac{dn}{dN} = n_i x_{i,A} \ . \tag{2.5.4}$$

Since $N_A N_A = 1$, (2.5.4) yields

$$\left(\frac{dn}{dN}\right)^2 = c_{ij} n_i n_j \ , \tag{2.5.5}$$

where

$$c_{ij} = x_{i,A} x_{j,A} \ . \tag{2.5.6}$$

If we take a second rectangular cartesian system $\bar{x}$ in which the components of the vectors $\underset{\sim}{X}$ and $\underset{\sim}{x}$ are $\bar{X}_A$ and $\bar{x}_i$ respectively, we have

$$\left(\frac{dn}{dN}\right)^2 = \bar{c}_{ij}\bar{n}_i\bar{n}_j \ , \tag{2.5.7}$$

where

$$\bar{c}_{ij} = \frac{\partial \bar{x}_i}{\partial \bar{X}_A}\frac{\partial \bar{x}_j}{\partial \bar{X}_A} \tag{2.5.8}$$

It is easily seen that

$$\bar{c}_{ij} = \bar{x}_{i,p}\bar{x}_{j,q}c_{pq} \ . \tag{2.5.9}$$

c_{ij} and $\bar{c}_{ij}$ are thus the components in the system x and $\bar{x}$ respectively of a second-order cartesian tensor. This is called the Finger strain tensor. We note from its definition by (2.5.6) that it is a symmetric tensor.

6. The element of area

We define the vector element of area of the surface in the usual manner as the product of the (scalar) area and the unit vector in the direction of the outward-drawn normal to the element. Let $d\underset{\sim}{A}$ and $d\underset{\sim}{a}$ be corresponding vector elements of area of the surface at the particle P at times T and t respectively. Let dA and da be the corresponding scalar areas of these elements. Then,

$$d\underset{\sim}{A} = \underset{\sim}{N}dA \quad \text{and} \quad d\underset{\sim}{a} = \underset{\sim}{n}da \ . \tag{2.6.1}$$

We consider a conical element of volume of the material which has the particle P as vertex and the elements of area $d\underset{\sim}{A}$ and $d\underset{\sim}{a}$ as bases at times T and t respectively. Let dV and dv be the volumes of these elements at times T and t respectively. Then,

$$dV = \frac{1}{3}\, dN dA \quad \text{and} \quad dv = \frac{1}{3}\, dn da \ . \tag{2.6.2}$$

From (2.3.4) we have

$$\frac{dv}{dV} = \frac{\partial \underset{\sim}{x}}{\partial \underset{\sim}{X}} = \frac{dn\ da}{dN\ dA} \ . \tag{2.6.3}$$

Introducing this result into (2.5.4) and (2.5.5), we obtain

$$N_A \frac{dA}{da} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} n_i x_{i,A} \tag{2.6.4}$$

and

$$\left(\frac{dA}{da}\right)^2 = \frac{1}{(\partial \underset{\sim}{x}/\partial \underset{\sim}{X})^2} c_{ij} n_i n_j \ . \tag{2.6.5}$$

R. S. Rivlin

Chapter 3

FORCES ACTING ON A BODY

1. Description of the forces

We consider a body to be acted upon by forces of two kinds--forces distributed throughout the body, which are called body forces, and forces acting on the surface of the body, which are called surface forces.

We consider an element of the body which has mass dm, say. We denote by $\underset{\sim}{\phi}dm$ the force acting on this mass and we assume that the force system acting is such that $\underset{\sim}{\phi}$ is everywhere finite. $\underset{\sim}{\phi}$ is called the body force per unit mass.

The forces acting on an element of surface of the body are specified per unit area of the surface element. Since the area of a surface element changes during the deformation, it is necessary to be explicit as to the instant at which this area is measured. Two methods of specifying surface forces are commonly used. In one of these the force is specified per unit area of the surface measured at the reference time T. In the other the force is specified per unit area of the surface measured at time t. We note that in both cases the force is that existing at time t.

Thus, let dA and da be the areas of a surface element measured at times T and t respectively. We denote by

$\underset{\sim}{F}dA$ or by $\underset{\sim}{f}da$ the force acting on this surface element at time t, so that, by definition,

$$\underset{\sim}{F}dA = \underset{\sim}{f}da. \tag{3.1.1}$$

$\underset{\sim}{F}$ is called the surface force at time t per unit area of surface measured at time T, and $\underset{\sim}{f}$ is called the surface force at time t per unit area of surface measured at time t.

2. Newton's second law

We consider a body to undergo a deformation described by (2.1.1). We assume that the body forces $\underset{\sim}{\phi}$ per unit mass and surface forces $\underset{\sim}{F}$ per unit area measured at the reference time T are acting on the body.

Newton's second law, expressed in the form that the resultant force on the body is equal to the rate of change of linear momentum of the body, implies that

$$\int_m \underset{\sim}{\phi} dm + \int_A \underset{\sim}{F} dA = \frac{d}{dt} \int_m \dot{\underset{\sim}{x}} dm \ , \tag{3.2.1}$$

where the dot denotes differentiation with respect to time. The first and last integrals in (3.2.1) are taken over the whole mass of the body and represent respectively the resultants of the body forces acting on the body and of the linear momentum. The second integral, which is taken over the surface of the body measured at time T, is the resultant of the surface forces acting on the body at time t.

Let ρ_0 denote the density of the material measured at time T and let dV be the volume at time T of the material element of mass dm. Then,

$$dm = \rho_0 dV \ . \tag{3.2.2}$$

Introducing this into (3.2.1), we obtain

$$\int_V \rho_0 \underset{\sim}{\phi} dV + \int_A \underset{\sim}{F} dA = \int_V \rho_0 \ddot{\underset{\sim}{x}} dV \ . \tag{3.2.3}$$

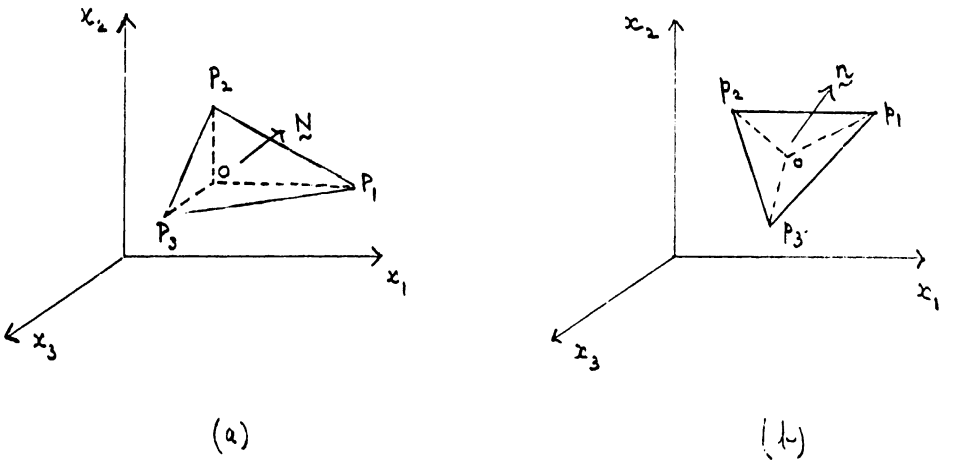


FIG. 1.

3. The definition of stress

We apply the formula (3.2.3) to an infinitesimal element of the body which has, at time T, the form of a tetrahedron $OP_1P_2P_3$ with its edges OP_1, OP_2, OP_3 parallel to the axes x_1, x_2, x_3 respectively of the rectangular cartesian reference system x, as shown in Fig.1(a). We consider that O is at point $\underset{\sim}{X}$ and that the outward-drawn normal to the face $P_1P_2P_3$ of the tetrahedron is in the direction of the unit normal $\underset{\sim}{N}$. At time t, the edges op_1, op_2, op_3 of the tetrahedron $op_1p_2p_3$ into which $OP_1P_2P_3$ is deformed will, in general, no longer be parallel to the axes of the reference system x, as shown in Fig.1(b). We denote by $\underset{\sim}{n}$ the unit normal to the face $p_1p_2p_3$ of this tetrahedron.

Let α be the area of the face $P_1P_2P_3$ of the tetrahedron $OP_1P_2P_3$ and let α_A be the area of the face which is perpendicular to the x_A-axis, i.e. of the face OP_2P_3. Then

$$\alpha_A = \alpha N_A \, . \tag{3.3.1}$$

Also, let Δ be the volume of the tetrahedron $OP_1P_2P_3$.

We shall apply the formula (3.2.3) to a body consisting of the infinitesimal material tetrahedron which occupies the region $OP_1P_2P_3$ at time T and $op_1p_2p_3$ at time t. We denote by $\underset{\sim}{F}$ the force acting at time t on the face $p_1p_2p_3$ of the tetrahedron, measured per unit area of the face $P_1P_2P_3$.

Thus, $\alpha\underset{\sim}{F}$ is the resultant surface force acting on $p_1p_2p_3$. We denote by $\underset{\sim}{\pi}_1$ the force acting at time t on the face op_2p_3 of the tetrahedron, measured per unit area of the face OP_2P_3. Thus, $\alpha_1\underset{\sim}{\pi}_1$ is the resultant surface force acting on op_2p_3 at time t. We attach analogous meanings to $\underset{\sim}{\pi}_2$ and $\underset{\sim}{\pi}_3$.

Applying (3.2.3) to the elementary material tetrahedron, we obtain, with this notation,

$$\rho_o\underset{\sim}{\phi}\Delta + (\alpha\underset{\sim}{F} - \alpha_A\underset{\sim}{\pi}_A) = \rho_o\underset{\sim}{\ddot{x}}\Delta \ .$$

Introducing (3.3.1), this becomes

$$(\underset{\sim}{F} - N_A\underset{\sim}{\pi}_A) + \rho_o(\underset{\sim}{\phi} - \underset{\sim}{\ddot{x}})\frac{\Delta}{\alpha} = 0 \ . \tag{3.3.2}$$

We now consider the limit as the linear dimensions of the tetrahedron tend to zero, the orientations of the faces of the tetrahedron remaining fixed. Then $\Delta/\alpha \to 0$ and equation (3.3.2) becomes

$$\underset{\sim}{F} = N_A\underset{\sim}{\pi}_A \ . \tag{3.3.3}$$

The three vectors $\underset{\sim}{\pi}_A$ have nine components in the reference system x. We denote the components of $\underset{\sim}{\pi}_A$ by π_{Ai}. In terms of the components of $\underset{\sim}{F}$ and $\underset{\sim}{\pi}_A$ in the cartesian reference system x, the relation (3.3.3) may be written

$$F_i = N_A\pi_{Ai} \ . \tag{3.3.4}$$

We now define three vectors $\bar{\pi}_A$ and their nine components $\bar{\pi}_{Ai}$ in an analogous manner with respect to another rectangular cartesian reference system $\bar{x}$. Let $\underset{\sim}{N}_1$ be the unit vector in the direction of the axis $\bar{x}_1$ of this system and let $\underset{\sim}{N}_2$ and $\underset{\sim}{N}_3$ be unit vectors in the directions of $\bar{x}_2$ and $\bar{x}_3$. Let N_{1A} be the components in the system x of the unit vector $\underset{\sim}{N}_1$ and, more generally, let N_{BA} be the components in the system x of the unit vector $\underset{\sim}{N}_A$.

We note that $\underset{\sim}{\bar{\pi}}_1$ is the force at time t acting on an element of area which was normal to $\underset{\sim}{N}_1$ at time T, measured per unit area at time T. Thus, from (3.3.3), we obtain

$$\underset{\sim}{\bar{\pi}}_1 = N_{1A}\underset{\sim}{\pi}_A \ . \tag{3.3.5}$$

More generally, we have

$$\underset{\sim}{\bar{\pi}}_B = N_{BA}\underset{\sim}{\pi}_A \ . \tag{3.3.6}$$

We have already introduced the notation $\bar{\pi}_{Bi}$ for the components of the vector $\underset{\sim}{\bar{\pi}}_B$ in the system $\bar{x}$. We now introduce the notation $\bar{\pi}^*_{Bi}$ for the components of this vector in the system x. Then,

$$\bar{\pi}_{Bi} = N_{ij}\bar{\pi}^*_{Bj} \ . \tag{3.3.7}$$

From (3.3.6) we have

$$\bar{\pi}^*_{Bj} = N_{BA}\pi_{Aj} \ . \tag{3.3.8}$$

Introducing this result into (3.3.7), we obtain

$$\bar{\pi}_{Bi} = N_{BA}N_{ij}\pi_{Aj} \cdot \tag{3.3.9}$$

We note that $\bar{\pi}_{Ai}$ and π_{Ai} are the components of a cartesian tensor in the coordinate systems $\bar{x}$ and x respectively. This is called the Kirchoff-Piola stress tensor.

4. Equations of motion

We now return to the equation (3.2.3) which expresses Newton's second law for a body of arbitrary magnitude on which body forces $\underset{\sim}{\phi}$ and surface forces $\underset{\sim}{F}$ are acting. We have seen that the relation (3.3.3) must be valid at every point of the body and consequently at its surface. Substituting for $\underset{\sim}{F}$ in (3.2.3), we obtain

$$\int_V \rho_o \underset{\sim}{\phi} dV + \int_A N_B \underset{\sim}{\pi}_B dA = \int_V \rho_o \ddot{\underset{\sim}{x}} dV \ . \tag{3.4.1}$$

We use the Divergence Theorem to convert the surface integral in this equation into a volume integral and thus obtain

$$\int_V (\rho_o \underset{\sim}{\phi} + \underset{\sim}{\pi}_{B,B} - \rho_o \ddot{\underset{\sim}{x}}) dV = 0 \ . \tag{3.4.2}$$

This equation may now be applied to an infinitesimal material element, yielding

$$\underset{\sim}{\pi}_{B,B} + \rho_o \underset{\sim}{\phi} - \rho_o \ddot{\underset{\sim}{x}} = 0 \ . \tag{3.4.3}$$

This is the Kirchoff-Piola equation of motion. In terms of components in the system x, we may write it as

$$\pi_{Bi,B} + \rho_o \phi_i - \rho_o \ddot{x}_i = 0 \ . \tag{3.4.4}$$

5. The moment equation

We now apply the law that the moment about a point of the forces acting on a body is equal to the rate of change of angular momentum about that point. We obtain

$$\int_V (\underset{\sim}{x}\times\rho_o\underset{\sim}{\phi})\,dV + \int_A (\underset{\sim}{x}\times\underset{\sim}{F})\,dA = \frac{d}{dt}\int_V (\underset{\sim}{x}\times\rho_o\dot{\underset{\sim}{x}})\,dV \ . \tag{3.5.1}$$

Introducing (3.3.3) into the second integral in (3.5.1) and using the divergence theorem, we obtain

$$\int_A (\underset{\sim}{x}\times\underset{\sim}{F})\,dA = \int_A (\underset{\sim}{x}\times\underset{\sim}{\pi}_B)N_B\,dA = \int_V (\underset{\sim}{x}\times\underset{\sim}{\pi}_B)_{,B}\,dV$$

$$= \int_V (\underset{\sim}{x}_{,B}\times\underset{\sim}{\pi}_B + \underset{\sim}{x}\times\underset{\sim}{\pi}_{B,B})\,dV \ . \tag{3.5.2}$$

We also have

$$\frac{d}{dt}\int_V (\underset{\sim}{x}\times\rho_o\dot{\underset{\sim}{x}})\,dV = \int_V (\underset{\sim}{x}\times\rho_o\ddot{\underset{\sim}{x}})\,dV. \tag{3.5.3}$$

Introducing these results into (3.5.1) and applying the resulting equation to an infinitesimal material element, so that we may omit the integral signs, we obtain

$$\underset{\sim}{x}\times(\underset{\sim}{\pi}_{B,B} + \rho_o\underset{\sim}{\phi} - \rho_o\ddot{\underset{\sim}{x}}) + \underset{\sim}{x}_{,B}\times\underset{\sim}{\pi}_B = 0 \ . \tag{3.5.4}$$

Using the result (3.4.3), we obtain

$$\underset{\sim}{x}_{,B}\times\underset{\sim}{\pi}_B = 0 \ . \tag{3.5.5}$$

In terms of the components in the coordinate system x, this result may be written

$$\epsilon_{ijk} x_{j,B} \tau_{Bk} = 0 \ . \tag{3.5.6}$$

To clarify the significance of this result we write down one of the components in this vector equation, say that with i = 1. We have

$$x_{2,B} \tau_{B3} = x_{3,B} \tau_{B2} \ . \tag{3.5.7}$$

Thus, (3.5.5), or (3.5.6), is a statement that $x_{j,B} \tau_{Bk}$ is symmetric with respect to interchange of j and k.

6. The Cauchy equations

It may be noted that so far T is an arbitrary reference time, not necessarily fixed, which may be chosen in any manner we please. As a particular choice, we may take it to coincide with the instant t at which the forces are measured.

In this section we explore the consequences of making this choice. Firstly, in (3.1.1) $\underset{\sim}{F}$ becomes $\underset{\sim}{f}$ and A becomes a, so that the equation is satisfied identically. Also the position $\underset{\sim}{X}$ to which the equations apply becomes the same as $\underset{\sim}{x}$. Equation (3.2.3) then becomes

$$\int_v \rho \underset{\sim}{\phi} dv + \int_a \underset{\sim}{f} da = \int_v \rho \ddot{\underset{\sim}{x}} dv \ , \qquad (3.6.1)$$

since the volume V now becomes the volume v of the body at time t and ρ_o becomes the material density ρ at time t. In §3 the tetrahedron $OP_1P_2P_3$ becomes identical with the tetrahedron $op_1p_2p_3$ and the unit vector normal to its slant surface becomes $\underset{\sim}{n}$. Thus, the vector $\underset{\sim}{\pi}_1$ is now the force at time t acting on the face op_2p_3 measured per unit area of op_2p_3. We shall denote it by $\underset{\sim}{\sigma}_1$. Analogous meanings attach to $\underset{\sim}{\pi}_2$ and $\underset{\sim}{\pi}_3$ and we shall denote them by $\underset{\sim}{\sigma}_2$ and $\underset{\sim}{\sigma}_3$. We shall denote these collectively by $\underset{\sim}{\sigma}_j$ and their components in the cartesian reference system x by σ_{ji}. The nine quantities σ_{ji} are the components of a tensor which is called the Cauchy stress tensor.

Making these substitutions in (3.3.3), we obtain

$$\underset{\sim}{f} = n_j \underset{\sim}{\sigma}_j \ . \tag{3.6.2}$$

In terms of the components of $\underset{\sim}{f}$ and $\underset{\sim}{\sigma}_j$ in the cartesian reference system this equation may be written

$$f_i = n_j \sigma_{ji} \ . \tag{3.6.3}$$

In the same way the equations of motion (3.4.3) yield

$$\underset{\sim}{\sigma}_{j,j} + \rho \underset{\sim}{\phi} - \rho \ddot{\underset{\sim}{x}} = 0 \ . \tag{3.6.4}$$

These are called the Cauchy equations of motion. They may be written in alternative form as

$$\sigma_{ji,j} + \rho \phi_i - \rho \ddot{x}_i = 0 \ . \tag{3.6.5}$$

Again applying the same procedure to equation (3.5.7) we obtain

$$\sigma_{ij} = \sigma_{ji} \ . \tag{3.6.6}$$

This expresses the fact that the Cauchy stress tensor is symmetric.

7. Relation between the Kirchoff-Piola and Cauchy stress

From (3.1.1) we note that the Kirchoff-Piola stress vector $\underset{\sim}{F}$ and Cauchy stress vector $\underset{\sim}{f}$ are related by

$$\underset{\sim}{F}dA = \underset{\sim}{f}da \ , \tag{3.7.1}$$

where dA and da are the areas at time T and t respectively of an element of a surface drawn in the body and deforming with it. Introducing (3.3.4) and (3.6.3) into (3.7.1), we obtain

$$N_A \pi_{Ai} dA = n_j \sigma_{ji} da \ . \tag{3.7.2}$$

Using (2.6.4), we obtain

$$n_j \sigma_{ji} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} n_j x_{j,A} \pi_{Ai} \ . \tag{3.7.3}$$

This result is valid for arbitrary orientation of the surface element considered, i.e. for arbitrary orientation of the unit vector $\underset{\sim}{n}$. Taking $n_j = \delta_{j1}$, i.e. the unit vector $\underset{\sim}{n}$ is parallel to the x-axis, we obtain

$$\sigma_{1i} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{1,A} \pi_{Ai} \ . \tag{3.7.4}$$

We obtain analogous results by taking $n_j = \delta_{j2}$ and δ_{j3}. The three results may be written

$$\sigma_{ji} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{j,A} \pi_{Ai} \ . \tag{3.7.5}$$

The converse relation

$$\pi_{Ai} = (\partial \underset{\sim}{x}/\partial \underset{\sim}{X})X_{A,j}\sigma_{ji} \tag{3.7.6}$$

follows immediately.

8. The Piola stress

We now define yet another stress tensor, the Piola stress tensor. We denote its components in the rectangular cartesian system x by P_{AB} and define it by

$$P_{AB} = X_{B,i}\pi_{Ai} \ . \tag{3.8.1}$$

With (3.7.6), we have

$$P_{AB} = (\partial \underset{\sim}{x}/\partial \underset{\sim}{X})X_{A,j}X_{B,i}\sigma_{ji} \ . \tag{3.8.2}$$

From (3.8.1) and (3.8.2), we obtain the inverse relations

$$\pi_{Ai} = x_{i,B}P_{AB} \ , \tag{3.8.3}$$

and

$$\sigma_{ji} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{i,A}x_{j,B}P_{AB} \ . \tag{3.8.4}$$

R. S. Rivlin

9. Constitutive equations

In continuum mechanics we usually wish to solve problems in which a body of material is acted on by specified forces, or certain parts of its boundary are subjected to specified displacements, or in which combinations of these conditions are imposed.

A complete solution of the problem posed then involves a complete description of the resulting deformation and a complete determination of the stress components in the body. (We note that if the deformation is determined, then any of the stresses--whether Cauchy, Kirchoff-Piola, or Piola--suffices for the determination of the remaining components.) Let us consider the Cauchy equations (3.6.3), (3.6.5) and (3.6.6). Equation (3.6.6) tells us that of the nine components of the Cauchy stress, only six are independent. Equation (3.6.5) then provides three differential equations for the determination of these six independent stress components, even if we consider the body force ϕ_i and the deformation x_i known.

Plainly, we cannot determine the stress in the body without some further information. This is provided by the constitutive equations for the material of the body. Throughout these lectures we shall consider the materials, with which we are concerned, to be such that the stress at a particle of the body is determined by the deformation gradients at the

particle, although the stress at time t may depend on the deformation gradients at the particle at times prior to t as well as on those at time t. Materials for which the stress at a particle depends on the deformation gradients at the particle (and not on their values at other particles or on their spatial derivatives) are sometimes called simple materials*.

When expressions for the stress in terms of the deformation gradients are introduced into (3.6.5) we obtain three equations for the three components x_i. These may, in particular cases, be differential equations, or differential-functional equations.

*The nomenclature appears to have originated with Noll (Arch. Rat'l. Mech. Anal. 2, 197 (1958)).

Chapter 4

THERMODYNAMIC CONSIDERATIONS

1. Description of the thermal state

So far we have been concerned only with the action of forces on the body. In so doing it has not been necessary to discuss the thermal conditions which prevail in the body under consideration.

In this chapter, however, we shall explore the implications of the first and second laws of thermodynamics when they are applied to a body which undergoes deformation*. In order to do this, we cannot avoid introducing the conditions of temperature and heat flux existing in the body.

We consider a body to undergo a deformation described by equation (2.1.1). We consider that at the generic particle P, which is at $\underset{\sim}{X}$ at the reference time T, the temperature at an arbitrary time t is θ. If the dependence of θ on $\underset{\sim}{X}$ and t is specified, thus

$$\theta = \theta(\underset{\sim}{X},t) \ , \tag{4.1.1}$$

---

*The approach adopted here is essentially that of Green and Rivlin, which was originally used by them in somewhat wider contexts (Arch. Rat'l. Mech. Anal. 16, 325 (1964); ibid. 17, 113 (1964)). It was used in the present context in a later note (Z.A.M.P. 15, 290 (1964)). The presentation here differs from those previously given in its use of the Clausius inequality instead of the much more doubtful Clausius-Duhem inequality.

then the temperature distribution is given throughout the body at all times. We denote the temperature at the reference time T by Θ, so that

$$\Theta(\underset{\sim}{X}) = \theta(\underset{\sim}{X},T) \ . \tag{4.1.2}$$

The flow of heat into the body is described in a manner which to some extent parallels the description in §3.1 of the system of forces applied to the body. We consider that heat enters the body in two ways--by generation throughout the volume of the body and by flow across the surface of the body.

We denote by χdm the rate at which heat is generated in an element of the body of mass dm at time t. χ is called the rate of heat generation per unit mass.

The rate at which heat enters the body across its surface is specified per unit area of the surface. Since the area changes during the deformation, it is necessary to be explicit as to the instant at which this area is measured. As in the case of the description of surface forces, we use two methods of specifying the rate of heat flow across the surface. In each of these we specify, for historical reasons, the rate of heat flow out of the body rather than into it. The rate of heat flow into the body is then given by the negative of this quantity.

We denote by QdA, or qda, the rate at which heat flows out of the body at time t across an element of surface

whose area at time T was dA and whose area at time t is da. By this definition

$$QdA = qda \ . \tag{4.1.3}$$

With (2.6.5), this equation yields

$$q = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}}(C_{ij}n_i n_j)^{1/2}Q \ . \tag{4.1.4}$$

We note that with these definitions, the total rate Q (say) at which heat enters the body is given by

$$Q = \int_m \chi dm - \int_A QdA \ . \tag{4.1.5}$$

In equation (4.1.5) we may replace the integration over the mass of the body by integration over its volume V at time T. The equation then becomes

$$Q = \int_m \rho_o \chi dV - \int_A QdA \ . \tag{4.1.6}$$

2. The energy balance equation

The first law of thermodynamics tells us that when a body undergoes deformation, the sum of the work done by the applied forces and the heat flowing into the body (measured in energy units) is equal to the total change of energy in the body. It is assumed that the latter may be written as the sum of the change in kinetic energy and a term called the change in internal energy which is unaltered by the superposition on the assumed deformation of a rigid body motion.

Applying this law to the changes occurring in an infinitesimal time interval we see that the sum of the rate R (say) at which work is done by the forces applied to the body and the rate Q at which heat is entering the body is equal to the sum of the rate of change of the kinetic energy and the rate of change of the internal energy. We denote the kinetic energy and internal energy by T and U respectively and their rates of change are, therefore, $\dot{T}$ and $\dot{U}$ respectively. We thus have

$$R + Q = \dot{T} + \dot{U} \ . \tag{4.2.1}$$

At time t, the rate at which work is done by the applied forces is given by

$$R = \int_V \rho_0 \underset{\sim}{\phi} . \dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F} . \dot{\underset{\sim}{x}} dA \ . \tag{4.2.2}$$

The first term represents the rate at which work is done by the body forces and the second that at which it is done by the surface forces.

The kinetic energy T is given by

$$T = \int_V \frac{1}{2} \rho_o \dot{\underset{\sim}{x}} . \dot{\underset{\sim}{x}} dV \ . \tag{4.2.3}$$

We now assume that the internal energy is a distributive property of the body, i.e. to each element of mass dm we can assign an internal energy Udm such that the total internal energy of the body U is given by

$$U = \int_m U dm = \int_V \rho_o U dV \ . \tag{4.2.4}$$

U is then the internal energy per unit mass at time t.

Introducing the expressions (4.2.2), (4.2.3), (4.2.4) and (4.1.6) into (4.2.1), we obtain

$$\int_V \rho_o \underset{\sim}{\phi} . \dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F} . \dot{\underset{\sim}{x}} dA + \int_V \rho_o \chi dV - \int_A Q dA$$

$$= \int_V \rho_o \ddot{\underset{\sim}{x}} . \dot{\underset{\sim}{x}} dV + \int_V \rho_o \dot{U} dV \ . \tag{4.2.5}$$

This is the energy balance equation.

3. The heat flux vector

If we examine the various quantities occurring in equation (4.2.5) we note that $\underset{\sim}{\phi}$, $\underset{\sim}{F}$, χ, Q and $\dot{U}$ are unaltered if we superpose on the assumed deformation a rigid motion. This fact can be used to derive from (4.2.5) the fundamental mechanical equations (3.3.3), (3.4.3) and (3.5.5) together with additional equations which must be satisfied by the heat flux.

We first apply the energy balance equation (4.2.5) to a body consisting of the material tetrahedron described in §3.3, which occupies the regions $OP_1P_2P_3$ and $op_1p_2p_3$ at times T and t respectively (see Fig.1). We adopt the same notation as that introduced in §3.3 and we introduce the further notation Q_1 for the rate at which heat leaves the tetrahedron at time t across the face op_2p_3, per unit area of OP_2P_3. Analogous meanings are attached to Q_2 and Q_3.

Taking the volume V in (4.2.5) to be the elementary tetrahedron $OP_1P_2P_3$, we obtain

$$\rho_0\Delta(\underset{\sim}{\phi}.\dot{\underset{\sim}{x}} + \chi - \ddot{\underset{\sim}{x}}.\dot{\underset{\sim}{x}} - \dot{U}) + (\alpha\underset{\sim}{F} - \alpha_A\underset{\sim}{\pi}_A).\dot{\underset{\sim}{x}} - (\alpha Q - \alpha_A Q_A) = 0 . \qquad (4.3.1)$$

Here we have preserved the notation Q for the rate at which heat leaves the tetrahedron at time t across the face $p_1p_2p_3$, measured per unit area of $P_1P_2P_3$. Introducing (3.3.1) and

allowing $\Delta/\alpha \to 0$, as in §3.3, we obtain

$$(\underset{\sim}{F}-N_A\underset{\sim}{\pi}_A).\dot{\underset{\sim}{x}} - (Q-N_AQ_A) = 0 . \tag{4.3.2}$$

We now consider the effect of superposing on the assumed deformation a uniform translational velocity of arbitrary magnitude, i.e. the effect of changing $\dot{\underset{\sim}{x}}$ by a constant vector $\underset{\sim}{c}$ (say). Evidently, $(\underset{\sim}{F}_A-N_A\underset{\sim}{\pi}_A)$ and $(Q-N_AQ_A)$ are left unchanged and equation (4.3.2) becomes

$$(\underset{\sim}{F}-N_A\underset{\sim}{\pi}_A).(\dot{\underset{\sim}{x}}+\underset{\sim}{c}) - (Q-N_AQ_A) = 0 . \tag{4.3.3}$$

Subtracting (4.3.2) from (4.3.3), we obtain

$$(\underset{\sim}{F}-N_A\underset{\sim}{\pi}_A).\underset{\sim}{c} = 0 . \tag{4.3.4}$$

Since equation (4.3.4) is valid for an arbitrary value of vector $\underset{\sim}{c}$, we obtain equation (3.3.3), viz.

$$\underset{\sim}{F}-N_A\underset{\sim}{\pi}_A = 0 . \tag{4.3.5}$$

Introducing this result into (4.3.2), we obtain

$$Q-N_AQ_A = 0 . \tag{4.3.6}$$

We bear in mind that equations (4.3.5) and (4.3.6) are valid at each point of the body.

4. The equations of motion and the dissipation equation

We now return to equation (4.2.5). We note, using (4.3.5) and applying the divergence theorem, that

$$\int_A \underset{\sim}{F}.\dot{\underset{\sim}{x}}dA = \int_A N_B \underset{\sim}{\pi}_B.\dot{\underset{\sim}{x}}dA = \int_V (\underset{\sim}{\pi}_B.\dot{\underset{\sim}{x}})_{,B}dV$$

$$= \int_V (\underset{\sim}{\pi}_{B,B}.\dot{\underset{\sim}{x}} + \underset{\sim}{\pi}_B.\dot{\underset{\sim}{x}}_{,B})dV \ . \qquad (4.4.1)$$

Similarly using (4.3.6) and applying the divergence theorem, we obtain

$$\int_A QdA = \int_A N_B Q_B dA = \int_V Q_{B,B}dV \ . \qquad (4.4.2)$$

Introducing these results into (4.2.5) and applying the resulting equation to an infinitesimal material element, we obtain

$$(\underset{\sim}{\pi}_{B,B} + \rho_o\underset{\sim}{\phi} - \rho_o\underset{\sim}{\ddot{x}}).\dot{\underset{\sim}{x}} + \rho_o\chi - Q_{B,B} - \rho_o\dot{U}$$

$$+ \underset{\sim}{\pi}_B.\dot{\underset{\sim}{x}}_{,B} = 0 \ . \qquad (4.4.3)$$

We again consider the effect of superposing on the assumed deformation a uniform translational velocity of arbitrary magnitude $\underset{\sim}{c}$, say. We obtain the Kirchoff-Piola equations of motion (cf. equation (3.4.3))

$$\underset{\sim}{\pi}_{B,B} + \rho_o \underset{\sim}{\phi} - \rho_o \ddot{\underset{\sim}{x}} = 0 \tag{4.4.4}$$

and the equation

$$-\rho_o \dot{U} + \rho_o \chi - Q_{B,B} + \underset{\sim}{\pi}_B \cdot \dot{\underset{\sim}{x}}_{,B} = 0 , \tag{4.4.5}$$

which is called the dissipation equation.

Equation (3.5.5) can be derived from (4.4.5) by considering the effect of superposing on the assumed deformation a time-dependent rigid rotation. We assume this to be such that the positions of the various particles of the body at time t are unaltered, but the angular velocity of the body at time t is changed by $\underset{\sim}{\Omega}$, say. The velocity at time t of the particle which is at $\underset{\sim}{x}$ is thereby changed to $\dot{\underset{\sim}{x}} + \underset{\sim}{x} \times \underset{\sim}{\Omega}$.

Replacing $\dot{\underset{\sim}{x}}$ by this vector in (4.4.5), we obtain

$$-\rho_o \dot{U} + \rho_o \chi - Q_{B,B} + \underset{\sim}{\pi}_B \cdot (\dot{\underset{\sim}{x}} + \underset{\sim}{x} \times \underset{\sim}{\Omega})_{,B} = 0 . \tag{4.4.6}$$

Subtracting (4.4.5) from (4.4.6), we obtain, bearing in mind that $\underset{\sim}{\Omega}$ is constant throughout the body,

$$\underset{\sim}{\pi}_B \cdot (\underset{\sim}{x} \times \underset{\sim}{\Omega})_{,B} = \underset{\sim}{\pi}_B \cdot (\underset{\sim}{x}_{,B} \times \underset{\sim}{\Omega}) = -\underset{\sim}{\Omega} \cdot (\underset{\sim}{x}_{,B} \times \underset{\sim}{\pi}_B) = 0 . \tag{4.4.7}$$

This relation is valid for arbitrary $\underset{\sim}{\Omega}$ and consequently the relation (3.5.5) follows, i.e.

$$\underset{\sim}{x}_{,B} \times \underset{\sim}{\pi}_B = 0 . \tag{4.4.8}$$

R. S. Rivlin

5. The second law of thermodynamics

In classical thermodynamics the difference in the entropy of a body in two states of equilibrium A and B, say, is defined in the following way. Let S_A and S_B denote the entropies in the states of equilibrium A and B respectively. It is considered that the body is taken from state B to state A by a reversible path. The path is assumed to be such that at each point of it the absolute temperature θ throughout the body is independent of position. θ may, however, vary as the body is taken from state B to state A. Let dH denote the amount of heat added to the body at the point on the path at which the temperature is θ. Then $S_A - S_B$ is defined by

$$S_A - S_B = \int_B^A \frac{dH}{\theta} \,. \tag{4.5.1}$$

We now consider that the body is taken from the state of equilibrium B to the state of equilibrium A by some other path, which is not necessarily reversible. At each point of this path the temperature is not necessarily constant throughout the body. We now form the ratio of the amount of heat entering an element of the body to the absolute temperature of the element at the instant at which it enters. We then integrate this over the whole body and over the path from state B to state A.

Let $\underset{\sim}{x}$ and θ be the vector position and absolute temperature respectively at time t at a particle of the body which was at vector position $\underset{\sim}{X}$ at some reference time T. We adopt the notation χ for the rate per unit mass at which heat enters the body throughout its volume and Q for the rate at which heat leaves the body across its surface, per unit area measured at time T. Then the ratio of the amount of heat entering the body in time dt to the absolute temperature at the point of entry, integrated over the body, is

$$\left(\int_m \frac{\chi}{\theta}\, dm - \int_A \frac{Q}{\theta}\, dA\right)dt \ . \tag{4.5.2}$$

Using $dm = \rho_o dV$, we see that the integral of this quantity over the path from state B to state A is

$$\int_{t_B}^{t_A}\left[\int_V \frac{\rho_o \chi}{\theta}\, dV - \int_A \frac{Q}{\theta}\, dA\right] dt \ . \tag{4.5.3}$$

t_A and t_B are the times at which the body reaches the states B and A respectively.

The second law of thermodynamics tells us that the quantity (4.5.3) is less than or equal to the increase in entropy, ΔS, say, in passing from the state of equilibrium B to the state of equilibrium A, thus:

$$\Delta S \geq \int_{t_B}^{t_A} \left[\int_V \frac{\rho_o \chi}{\theta} \, dV - \int_\Lambda \frac{Q}{\theta} \, d\Lambda \right] dt \, . \qquad (4.5.4)$$

The inequality (4.5.4) is called the Clausius inequality.

6. Implications of the Clausius inequality

We introduce into the Clausius inequality expressed by (4.5.4) the relation (4.3.6). Using the divergence theorem, we note that

$$\int_A \frac{Q}{\theta}\, dA = \int_A \frac{N_B Q_B}{\theta}\, dA = \int_V \left(\frac{Q_B}{\theta}\right)_{,B} dV$$

$$= \int_V \left(\frac{1}{\theta}\, Q_{B,B} - \frac{Q_B}{\theta^2}\, \theta_{,B}\right) dV \ . \qquad (4.6.1)$$

We now assume that the entropy of the body in any state of equilibrium is an extensive property, so that S_A and S_B and hence ΔS may be expressed as integrals over the body, thus:

$$\Delta S = \int_m \Delta S dm = \int_V \rho_o \Delta S dV \ , \qquad (4.6.2)$$

where ΔS is the change of entropy per unit mass. Equations (4.6.1) and (4.6.2) are introduced into (4.5.4) and the body considered is taken to be an infinitesimal element, so that the integral signs over V may be omitted. We thus obtain

$$\rho_o \Delta S \geq \int_{t_B}^{t_A} \left[\frac{1}{\theta}(\rho_o \chi - Q_{B,B}) + \frac{Q_B}{\theta^2}\, \theta_{,B}\right] dt \ . \qquad (4.6.3)$$

The dissipation equation (4.4.5) is now used to substitute for $(\rho_o \chi - Q_{B,B})$ in (4.6.2). This yields

$$\rho_o \Delta S \geq \int_{t_B}^{t_A} \left[\frac{1}{\theta} (\rho_o \dot{U} - \underset{\sim}{\pi}_B \cdot \dot{\underset{\sim}{x}}_{,B}) + \frac{Q_B}{\theta^2} \theta_{,B} \right] dt \ . \tag{4.6.4}$$

We now consider the special case when the temperature is independent of time. In this case the inequality (4.6.4) yields

$$\rho_o (\Delta U - \theta \Delta S) \leq \int_{t_B}^{t_A} \left(\underset{\sim}{\pi}_B \cdot \dot{\underset{\sim}{x}}_B - \frac{Q_B}{\theta} \theta_{,B} \right) dt \ , \tag{4.6.5}$$

where ΔU denotes the increase in internal energy per unit mass in passing from state B to state A. The quantity $U-\theta S$ is usually called the Helmholtz free energy. We shall use the symbol W to denote the Helmholtz free energy per unit mass, thus:

$$W = U - \theta S \ . \tag{4.6.6}$$

The increase ΔW in the Helmholtz free energy in passing from state B to state A, when the temperature of the material element is the same in the two states, is given by

$$\Delta W = \Delta U - \theta \Delta S \ . \tag{4.6.7}$$

Then, (4.6.5) becomes

$$\rho_o \Delta W \leq \int_{t_B}^{t_A} \left(\underset{\sim}{\pi}_B \cdot \dot{\underset{\sim}{x}}_{,B} - \frac{Q_B}{\theta} \theta_{,B} \right) dt \ . \tag{4.6.8}$$

7. Elastic materials

We now apply the inequality (4.6.4) to the situation in which the states B and A are neighboring states, each of which is an equilibrium state, attained at times t and t + dt respectively. The increases in S, U and W in the passage from state B to state A are infinitesimal and will be denoted dS, dU and dW respectively. The inequality (4.6.4) yields, when applied to this situation,

$$\rho_o(dU-\theta dS) \leq \underset{\sim}{\pi}_B.d\underset{\sim}{x}_{,B} - \frac{Q_B}{\theta}\,\theta_{,B}\,dt, \tag{4.7.1}$$

where $\underset{\sim}{x}$ and $\underset{\sim}{x} + d\underset{\sim}{x}$ are the vector positions of the particle considered at times t and t + dt respectively. Equation (4.7.1) may be rewritten in terms of the components of $\underset{\sim}{\pi}_B$ and $d\underset{\sim}{x}$ in the rectangular cartesian system x, thus:

$$\rho_o(dU-\theta dS) \leq \pi_{Bi}dx_{i,B} - \frac{Q_B}{\theta}\,\theta_{,B}\,dt\ . \tag{4.7.2}$$

From (4.6.6) we have

$$dU - \theta dS = dW + Sd\theta\ . \tag{4.7.3}$$

With (4.7.2), this yields

$$\rho_o(dW+Sd\theta) \leq \pi_{Bi}dx_{i,B} - \frac{Q_B}{\theta}\,\theta_{,B}\,dt\ . \tag{4.7.4}$$

We define an elastic material as one for which the Helmholtz free energy and the Kirchoff-Piola stress depend

on the deformation gradients only, i.e.

$$W = W(x_{j,P},\theta) \quad \text{and} \quad \pi_{Bi} = \pi_{Bi}(x_{j,P},\theta) \ . \tag{4.7.5}$$

From (4.7.5) we have

$$dW = \frac{\partial W}{\partial x_{i,B}}\, dx_{i,B} + \frac{\partial W}{\partial \theta}\, d\theta \ . \tag{4.7.6}$$

Introducing (4.7.6) into (4.7.4), we obtain

$$\left(\rho_o \frac{\partial W}{\partial x_{i,B}} - \pi_{Bi}\right) dx_{i,B} + \rho_o \left(S + \frac{\partial W}{\partial \theta}\right) d\theta \leq - \frac{Q_B}{\theta}\, \theta_{,B}\, dt \ . \tag{4.7.7}$$

Since the coefficients of $dx_{i,B}$, $d\theta$ and dt in (4.7.7) are independent of $dx_{i,B}$, $d\theta$ and dt, with the assumption that the inequality (4.7.7) is valid for all infinitesimal values of these quantities, we obtain

$$\pi_{Bi} = \rho_o \left(\frac{\partial W}{\partial x_{i,B}}\right)_\theta , \quad S = -\left(\frac{\partial W}{\partial \theta}\right)_{x_{i,B}} , \quad \theta_{,B} = 0 \ . \tag{4.7.8}$$

We have written the first two of these results in forms which emphasize the independent variables which are kept constant in the derivation. The last of the relations (4.7.8) expresses the fact that a temperature gradient cannot exist in a body in thermodynamic equilibrium.

We may also express the Kirchoff-Piola stress in terms of the internal energy U per unit mass. U and W may be regarded either as functions of $x_{i,B}$ and θ, or of $x_{i,B}$ and S. Thus,

$$dW = \left(\frac{\partial W}{\partial x_{i,B}}\right)_{\theta} dx_{i,B} + \left(\frac{\partial W}{\partial \theta}\right)_{x_{i,B}} d\theta$$

$$= \left(\frac{\partial W}{\partial x_{i,B}}\right)_{S} dx_{i,B} + \left(\frac{\partial W}{\partial S}\right)_{x_{i,B}} dS \ . \qquad (4.7.9)$$

From this it follows immediately, with (4.7.8), that

$$\left(\frac{\partial W}{\partial x_{i,B}}\right)_{\theta} = \left(\frac{\partial W}{\partial x_{i,B}}\right)_{S} + S\left(\frac{\partial \theta}{\partial x_{i,B}}\right)_{S} \ . \qquad (4.7.10)$$

From (4.6.6), it follows that

$$\left(\frac{\partial W}{\partial x_{i,B}}\right)_{S} = \left(\frac{\partial U}{\partial x_{i,B}}\right)_{S} - S\left(\frac{\partial \theta}{\partial x_{i,B}}\right)_{S} \qquad (4.7.11)$$

From (4.7.10), (4.7.11) and (4.7.8), we obtain

$$\pi_{Bi} = \rho_o \left(\frac{\partial W}{\partial x_{i,B}}\right)_{\theta} = \rho_o \left(\frac{\partial U}{\partial x_{i,B}}\right)_{S} \qquad (4.7.12)$$

Using (3.7.5), we obtain immediately an expression for the Cauchy stress

$$\sigma_{ji} = \frac{\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{j,B}\left(\frac{\partial W}{\partial x_{i,B}}\right)_{\theta} = \frac{\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{j,B}\left(\frac{\partial U}{\partial x_{i,B}}\right)_{S} \ . \qquad (4.7.13)$$

If we are considering isothermal deformations of a body, it is preferable to regard the thermodynamic potentials as functions of $x_{i,B}$ and θ. We then characterize the material by expressing the Helmholtz free energy as a function of $x_{i,B}$ and θ and use the first expression in (4.7.12) or (4.7.13) to determine the stress. On the other hand, if we are considering isentropic deformations, it is preferable to regard the thermodynamic potentials as functions of $x_{i,B}$ and S. The material is then characterized by expressing the internal energy as a function of $x_{i,B}$ and S and we use the last expression in (4.7.12) or (4.7.13) to determine the stress.

Care must be exercised in interpreting equation (4.7.12), for it might at first sight appear that the stress associated with specified deformation gradients is independent of whether these are attained under isothermal or isentropic conditions. That this is not correct can be seen in the following way. Let θ_o and S_o denote the values of θ and S respectively at the reference time T. Then, for an isothermal deformation

$$\pi_{Bi} = \left(\frac{\partial W(x_{p,Q}, \theta_o)}{\partial x_{i,B}} \right)_{\theta} , \tag{4.7.14}$$

while for an isentropic deformation

$$\pi_{Bi} = \left(\frac{\partial U(x_{p,Q}, S_o)}{\partial x_{i,B}} \right)_{S} . \tag{4.7.15}$$

The values of π_{Bi} given by (4.7.14) and (4.7.15) are not the same notwithstanding equation (4.7.12). The latter, applied to an isothermal deformation, states that

$$\left(\frac{\partial W(x_{p,Q},\theta_o)}{\partial x_{i,B}}\right)_\theta = \left(\frac{\partial U(x_{p,Q},S)}{\partial x_{i,B}}\right)_S , \qquad (4.7.16)$$

where S is the entropy which corresponds to the state in which the deformation gradients are $x_{i,B}$ and the temperature is θ_o. On the other hand, when applied to an isentropic deformation, equation (4.7.12) states that

$$\left(\frac{\partial W(x_{p,Q},\theta)}{\partial x_{i,B}}\right)_\theta = \left(\frac{\partial U(x_{p,Q},S_o)}{\partial x_{i,B}}\right)_S , \qquad (4.7.17)$$

where θ is the temperature which corresponds to the state in which the deformation gradients are $x_{i,B}$ and the entropy is S_o.

The expressions for the stress in an elastic material in terms of thermodynamic potential functions, given by (4.7.12) and (4.7.13), depend on the assumption that the body is in equilibrium. The formulae may, however, be applied when the body is not in equilibrium if we make a further assumption regarding the material and are careful to define appropriately what we mean by entropy in a non-equilibrium situation. We shall call a material for which these further assumptions are valid a perfectly elastic material.

We define the entropy and internal energy at any instant in a perfectly elastic material which is not in equilibrium as the entropy and internal energy the body would have if held in equilibrium in precisely the same state of deformation and temperature as it instantaneously occupies. With this definition we assert the validity of the second law of thermodynamics in the form (4.5.4), and the argument leading to the results (4.7.12) and (4.7.13) remains unchanged. Whether or not the argument can be applied to a particular material is a matter which can be decided either by experiment or by an understanding of the physical structure of the material and of the physical significance of entropy and internal energy. It should be borne in mind, also, that the nature of the departures from equilibrium for which the stress may still be derived from thermodynamic potential functions may vary from material to material.

8. Constraints

In passing from (4.7.7) to (4.7.8) it was assumed that the variations $dx_{i,B}$, $d\theta$ and dt may be chosen independently. If the material is such that the deformations to which it can be subjected are restricted by some constraint, this assumption is not necessarily correct.

For example, if the material is incompressible, then only isochoric deformations (i.e. deformations in which each material element remains unchanged in volume) are possible. It follows from (2.3.3) that the deformations to which the material can be subjected are restricted by the constraint

$$\partial \underset{\sim}{x}/\partial \underset{\sim}{X} = \varepsilon_{ik\ell} x_{i,1} x_{k,2} x_{\ell,3} = 1 \ . \tag{4.8.1}$$

It follows that the quantities $dx_{i,B}$ in (4.7.7) are subject to the constraint

$$\varepsilon_{ik\ell}(dx_{i,1}x_{k,2}x_{\ell,3} + dx_{k,2}x_{i,1}x_{\ell,3} + dx_{\ell,3}x_{i,1}x_{k,2}) = 0 \ . \tag{4.8.2}$$

The method of Lagrange undetermined multipliers can be used to relax this constraint in (4.7.7). From (4.7.7) we have

$$\left(\rho_o \frac{\partial W}{\partial x_{i,B}} - \pi_{Bi}\right) dx_{i,B} - p\varepsilon_{ik\ell}(dx_{i,1}x_{k,2}x_{\ell,3} + dx_{i,2}x_{k,3}x_{\ell,1} + dx_{i,3}x_{k,1}x_{\ell,2}) + \rho_o\left(S + \frac{\partial W}{\partial \theta}\right)d\theta \leq -\frac{Q_B}{\theta}\theta_{,B}\, dt \ , \tag{4.8.3}$$

where p is arbitrary. This relation is now valid for all infinitesimal values of $dx_{i,B}$, $d\theta$ and dt. We therefore obtain, in place of $(4.7.8)_1$,

$$\pi_{1i} = \rho_o\left(\frac{\partial W}{\partial x_{i,1}}\right)_\theta - p\varepsilon_{ik\ell}x_{k,2}x_{\ell,3}\ , \tag{4.8.4}$$

with analogous equations for π_{2i} and π_{3i}. We also obtain, as before, equations $(4.7.8)_2$ and $(4.7.8)_3$. It follows that

$$\pi_{Bi} = \rho_o\left(\frac{\partial W}{\partial x_{i,B}}\right)_\theta - \frac{1}{2}\,p\varepsilon_{ik\ell}\varepsilon_{BPQ}x_{k,P}x_{\ell,Q}\ . \tag{4.8.5}$$

Introducing (4.8.5) into (3.7.5), and bearing in mind (4.8.1), we obtain an expression for the Cauchy stress

$$\sigma_{ji} = \rho_o x_{j,B}\left(\frac{\partial W}{\partial x_{i,B}}\right)_\theta - p\delta_{ij}\ . \tag{4.8.6}$$

This expression replaces, for an incompressible material, the expression (4.7.13) for the Cauchy stress which is valid when there is no constraint on the deformations which the material can undergo. We notice that in the latter case, if the expression for the Helmholtz free energy W as a function of the deformation gradients is known, the Cauchy stress can be calculated explicitly. However, in the case when the material is incompressible and the Cauchy stress is given by (4.8.6), if W is known as a function of the deformation gradients, the Cauchy stress is still undetermined to the extent of a term $p\delta_{ij}$; this has the nature of an arbitrary hydrostatic pressure.

As another example of a constraint, we have the possibility that there is, in the material, a direction of inextensibility, say linear elements of the material which are parallel to the 1-axis at time T cannot change in length. From (2.2.12), this implies that

$$C_{11} = x_{i,1}x_{i,1} = 1 \ . \tag{4.8.7}$$

From (4.8.7) we have

$$x_{i,1}dx_{i,1} = x_{i,1}dx_{i,B}\delta_{B1} = 0 \ . \tag{4.8.8}$$

Introducing this constraint on $dx_{i,B}$ into (4.7.7) by means of a Lagrange undetermined multiplier τ, say, we obtain

$$\pi_{Bi} = \rho_0 \frac{\partial W}{\partial x_{i,B}} + \tau x_{i,1}\delta_{B1} \ . \tag{4.8.9}$$

If the possible deformations are subject to more than one constraint, then more than one undetermined multiplier must be introduced, one for each constraint. Suppose there are α independent constraints on the deformation gradients

$$f_\beta(x_{i,B}) = 0 \quad (\beta = 1,\ldots,\alpha) \ . \tag{4.8.10}$$

We obtain

$$\frac{df_\beta}{dx_{i,B}} dx_{i,B} = 0 \ . \tag{4.8.11}$$

Let $\tau_\beta (\beta = 1,\ldots,\alpha)$ be α undetermined multipliers. Then, we

obtain from (4.7.7),

$$\pi_{Bi} = \rho_o \frac{\partial W}{\partial x_{i,B}} + \sum_{\beta=1}^{\alpha} \tau_\beta \frac{df_\beta}{\partial x_{i,B}} . \qquad (4.8.12)$$

9. Cauchy elastic materials

In deriving, from (4.7.7), the expression $(4.7.8)_1$, we make the assumption that the Helmholtz free energy W, the entropy S and the Kirchoff-Piola stress π_{Ai} depend only on the deformation gradients $x_{i,A}$ and the absolute temperature θ. A material for which this assumption is valid is called an elastic material*.

In contrast, a Cauchy elastic material is defined as a material for which the stress at any instant depends only on the deformation gradients and temperature existing at that instant, while the Helmholtz free energy and internal energy do not depend solely on these quantities.

We thus have

$$\pi_{Ai} = \pi_{Ai}(x_{p,Q}, \theta) \ . \tag{4.9.1}$$

It can be shown that, from a physical standpoint, the concept of a Cauchy elastic material is somewhat artificial.

We first note that the rate at which work is done by the body force $\underset{\sim}{\phi}$ per unit mass and surface forces $\underset{\sim}{F}$, per unit area measured at time T, acting on a body at time t is given by

$$\int_V \rho_0 \underset{\sim}{\phi}.\dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F}.\dot{\underset{\sim}{x}} dA \ , \tag{4.9.2}$$

*The term hyperelastic material has been used for such a material to distinguish it from a Cauchy elastic material. In view of the implications of the discussion in this section, it does not seem worthwhile to introduce a new term.

where V is the domain occupied by the body at time T and A is the surface of the body at time T. Introducing (3.4.3) and (3.3.3) and using the divergence theorem, we obtain

$$\int_V \rho_o \underset{\sim}{\phi}.\dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F}.\dot{\underset{\sim}{x}} dA ,$$

$$= \int_V (\rho_o \phi_i + \pi_{Bi,B})\dot{x}_i dV + \int_V \pi_{Bi}\dot{x}_{i,B} dV$$

$$= -\int_V \rho_o \dot{x}_i \ddot{x}_i dV + \int_V \pi_{Bi}\dot{x}_{i,B} dV . \qquad (4.9.3)$$

The work done by the forces acting on the body in a deformation taking place in the time interval t_1 to t_2 is thus

$$\int_{t_1}^{t_2} dt\left[\int_V \rho_o \underset{\sim}{\phi}.\dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F}.\dot{\underset{\sim}{x}} dA\right]$$

$$= -\left[\frac{1}{2}\rho_o \dot{x}_i \dot{x}_i\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \int_V \pi_{Bi}\dot{x}_{i,B} dV . \qquad (4.9.4)$$

In the case when this deformation is cyclic, in the sense that

$$\underset{\sim}{x}(t_2+\tau) = \underset{\sim}{x}(t_1+\tau),\ \theta(t_2+\tau) = \theta(t_1+\tau),\ \dot{\underset{\sim}{x}}(t_2+\tau) = \dot{\underset{\sim}{x}}(t_1+\tau), \qquad (4.9.5)$$

equation (4.9.4) becomes

$$\int_{t_1}^{t_2} dt\left[\int_V \rho_o \underset{\sim}{\phi}.\dot{\underset{\sim}{x}} dV + \int_A \underset{\sim}{F}.\dot{\underset{\sim}{x}} dA\right]$$

$$= \int_{t_1}^{t_2} dt \int_V \pi_{Bi}\dot{x}_{i,B} dV ; \qquad (4.9.6)$$

i.e. the stress work done in a cyclic deformation is equal to the work done by the external forces.

Now, for any cycle of deformation, the work done by the external forces is positive, negative, or zero. If it is positive, then we consider the inverse cycle of deformation in which a particle which is at X_A at the reference time T moves to $\bar{x}_i(t)$ at time t, the temperature at this time being $\bar{\theta}(t)$, where

$$\bar{x}_i(t) = x_i(t_1+t_2-t) \ , \quad \bar{\theta}(t) = \theta(t_1+t_2-t) \ . \tag{4.9.7}$$

Using the dissipation equation (4.4.5) together with (4.9.6) and (4.4.2), we see that the work done by the body in the time interval t_1 to t_2 is

$$-\left[\int_V \rho_o U dV\right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \int_V \rho_o \chi dV - \int_{t_1}^{t_2} dt \int_A Q dA \ ;$$

i.e. in the cycle considered internal energy and heat are converted into useful work with an efficiency of 100 percent. If we exclude this unlikely, if not impossible, class of materials, we must conclude that the work done by the external forces in <u>all</u> cyclic deformations is zero. From this it follows that the stress is derivable from a potential function.

R. S. Rivlin

Chapter 5

FINITE ELASTICITY THEORY

1. Introduction

It has been seen in §4.7 that the Piola-Kirchoff stress π_{Ai}, referred to a rectangular cartesian system x, at time t is given in terms of W, the Helmholtz free energy per unit mass of the material, by

$$\pi_{Bi} = \rho_o \frac{\partial W}{\partial x_{i,B}} , \tag{5.1.1}$$

where ρ_o is the material density at the reference time T. W is regarded as a function of the deformation gradients $x_{i,B}$ and the absolute temperature θ. The Cauchy stress is then obtained by using the formula (3.7.5) as

$$\sigma_{ji} = \frac{\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{j,B} \frac{\partial W}{\partial x_{i,B}} . \tag{5.1.2}$$

Since, as was seen in §2.3, $\partial \underset{\sim}{x}/\partial \underset{\sim}{X}$ is the ratio between the volumes of a material element at times t and T, we have

$$\rho = \rho_o/(\partial \underset{\sim}{x}/\partial \underset{\sim}{X}) , \tag{5.1.3}$$

where ρ is the material density at time t. Equation (5.1.2) may then be rewritten as

$$\sigma_{ji} = \rho x_{j,B} \frac{\partial W}{\partial x_{i,B}} \quad . \tag{5.1.4}$$

The form of W as a function of the deformation gradients is not entirely arbitrary, but is subject to certain restrictions. These arise in two ways. Firstly, the superposition on the assumed deformation of a rigid rotation leaves the Helmholtz free energy W unchanged and this restricts the manner in which W can depend on the deformation gradients. Secondly, the material may possess some symmetry and this provides a further restriction on the form of W. It will be seen that the problem of determining explicitly the resulting limitations on the form of W can be reduced to a problem in the theory of invariants.

2. Restrictions due to effect of a superposed rotation

(i) Formulation of the problem

Consider a deformation in which a particle located at $\underset{\sim}{X}$ at the reference time T moves to $\underset{\sim}{x}$ at time t. The deformation is then described in a rectangular cartesian coordinate system x by

$$x_i = x_i(x_B, t) \ . \tag{5.2.1}$$

We call this deformation A. We suppose further that the Helmholtz free energy W per unit mass of a body which undergoes deformation is a function of the deformation gradients, thus:

$$W = F(x_{i,B}) \ . \tag{5.2.2}$$

We now consider a second deformation--deformation B--which differs from A by a rigid rotation. We suppose that in this deformation the particle which is at $\underset{\sim}{X}$ at time T moves to $\underset{\sim}{\bar{x}}$ at time t. In the reference system x, this deformation is described by

$$\bar{x}_i = \bar{x}_i(X_B, t) \ . \tag{5.2.3}$$

Since deformation B differs from deformation A by a rigid rotation, we may write

$$\bar{x}_i = s_{ij} x_j \ , \tag{5.2.4}$$

where s_{ij} is a proper orthogonal transformation, i.e.

$$s_{ij}s_{ik} = s_{ji}s_{ki} = \delta_{jk} , \quad |s_{ij}| = 1 . \tag{5.2.5}$$

The Helmholtz free energy $\overline{W}$ associated with this deformation is, from (5.2.2), given by

$$\overline{W} = F(\overline{x}_{i,B}) . \tag{5.2.6}$$

We assume that the conditions of the system are such that

$$\overline{W} = W . \tag{5.2.7}$$

Then, from (5.2.2), (5.2.6) and (5.2.4),

$$F(s_{ij}x_{j,B}) = F(x_{i,B}) , \tag{5.2.8}$$

and this relation must be satisfied for all proper orthogonal s_{ij}, i.e. for all s_{ij} satisfying (5.2.5).

Equation (5.2.8) implies a restriction on the manner in which F can depend on the deformation gradients $x_{i,B}$. This restriction is, however, given by (5.2.8) in implicit form and our next problem is to make the restriction explicit. We shall give two different methods by which this can be done.*

---

*Although the main results here, contained in equation (5.2.9), are classical, the first method of deriving them stems from that used by Green and Rivlin (Arch. Rat'l. Mech. Anal. 1, 1 (1957)) in a more general context. The second method stems from that used by Noll (Arch. Rat'l. Mech. Anal. 2, 197 (1958)) in this same more general context (see also §7.2).

(ii) Method 1

We recognize that equation (5.2.8) expresses the fact that W is a scalar invariant under the proper orthogonal group of transformations of the three vectors $x_{i,1}$, $x_{i,2}$, $x_{i,3}$.

We now make use of a theorem due to Cauchy which states that any scalar invariant under the proper orthogonal group of transformations of μ vectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$ may be expressed as a function of the $\frac{1}{2}\mu(\mu+1)$ inner products $\underset{\sim}{v}_\alpha \cdot \underset{\sim}{v}_\beta (\alpha,\beta = 1,\ldots,\mu)$ and $\frac{1}{2}\mu(\mu-1)(\mu-2)$ scalar triple products $[\underset{\sim}{v}_\alpha, \underset{\sim}{v}_\beta, \underset{\sim}{v}_\gamma]$ $(\alpha,\beta,\gamma = 1,2,\ldots,\mu;\ \alpha,\beta,\gamma$ all unequal).

This theorem implies that W must be a function of the six inner products $x_{i,A}\, x_{i,B}$ $(A,B = 1,2,3)$ and the scalar triple product of the three vectors $x_{i,1}$, $x_{i,2}$, $x_{i,3}$. This may be expressed as $\varepsilon_{ijk} x_{i,1} x_{j,2} x_{k,3}$. We recognize that $x_{i,A} x_{i,B}$ are the components in the system x of the Cauchy strain and we have already seen in §2.3 that for deformations possible in a real body this scalar triple product may be expressed as $(\det \underset{\sim}{C})^{1/2}$, where $\underset{\sim}{C}$ is the Cauchy strain matrix referred to the system x. It follows that W must be expressible as a function of the components of the Cauchy strain tensor, i.e.

$$W = W(C_{AB}) \ . \qquad (5.2.9)$$

(iii) Method 2

Another method by which the restrictions on W implied by (5.2.8) may be made explicit is the following. Since (5.2.8) must be valid for all proper orthogonal s_{ij}, i.e. for all s_{ij} satisfying (5.2.5), it must be valid for*

$$s_{ij} = \left(\underset{\sim}{C}^{-1/2}\right)_{iP} x_{j,P} \ . \tag{5.2.10}$$

Introducing this expression for s_{ij} into (5.2.8), we see, with (2.2.4), that

$$\begin{aligned} F(x_{i,B}) &= F\left\{\left(\underset{\sim}{C}^{-1/2}\right)_{iP} x_{j,P} x_{j,B}\right\} \\ &= F\left\{\left(\underset{\sim}{C}^{-1/2}\right)_{iP} C_{PB}\right\} \\ &= F\left\{\left(\underset{\sim}{C}^{-1/2}\right)_{iB}\right\} \ . \end{aligned} \tag{5.2.11}$$

Thus W must be expressible in the form (5.2.9). Since C_{AB}, regarded as a function of $x_{i,P}$, satisfies the condition (5.2.8) for <u>all</u> proper orthogonal s_{ij}, it is evident that any W of the form (5.2.9) must do so.

---

*We have already seen that s_{ij} defined in this way is proper orthogonal.

(iv) The stress

Introducing the expression (5.2.9) for W into (5.1.1), we obtain

$$\pi_{Bi} = \rho_o x_{i,P} \left(\frac{\partial W}{\partial C_{BP}} + \frac{\partial W}{\partial C_{PB}} \right) . \tag{5.2.12}$$

Similarly, from (5.1.2), we obtain

$$\sigma_{ji} = \frac{\rho_o}{\partial \underset{\sim}{x} / \partial \underset{\sim}{X}} x_{j,B} x_{i,P} \left(\frac{\partial W}{\partial C_{BP}} + \frac{\partial W}{\partial C_{PB}} \right) . \tag{5.2.13}$$

These are the restrictions on the forms of the Kirchoff-Piola and Cauchy stress tensors, which result from the fact that W is unchanged by the superposition on the assumed deformation of a rigid rotation.

Employing the definition (3.8.1) for the Piola stress P_{AB}, we see from (5.2.12) that

$$P_{AB} = \rho_o \left(\frac{\partial W}{\partial C_{BA}} + \frac{\partial W}{\partial C_{AB}} \right) . \tag{5.2.14}$$

In the case when the material is incompressible these equations have to be modified slightly, since the expressions for the Kirchoff-Piola and Cauchy stresses are given by (4.8.5) and (4.8.6) respectively. Introducing (5.2.9) into these equations, we obtain for an incompressible elastic material

$$\pi_{Bi} = \rho_o x_{i,P}\left(\frac{\partial W}{\partial C_{BP}} + \frac{\partial W}{\partial C_{PB}}\right) - \frac{1}{2}\, p\varepsilon_{ik\ell}\varepsilon_{BPQ}x_{R,P}x_{\ell,Q} \tag{5.2.15}$$

and

$$\sigma_{ji} = \rho_o x_{j,B}x_{i,P}\left(\frac{\partial W}{\partial C_{BP}} + \frac{\partial W}{\partial C_{PB}}\right) - p\delta_{ij}\ . \tag{5.2.16}$$

From the definition (3.8.1) for the Piola stress P_{AB} and (5.2.15), we see that

$$P_{AB} = \rho_o\left(\frac{\partial W}{\partial C_{BA}} + \frac{\partial W}{\partial C_{AB}}\right) - pX_{A,i}X_{B,i}\ . \tag{5.2.17}$$

We note that

$$X_{A,i}X_{B,i} = \left(\underset{\sim}{C}^{-1}\right)_{AB}\ , \tag{5.2.18}$$

so that (5.2.17) may be rewritten

$$P_{AB} = \rho_o\left(\frac{\partial W}{\partial C_{BA}} + \frac{\partial W}{\partial C_{AB}}\right) - p\left(\underset{\sim}{C}^{-1}\right)_{AB}\ . \tag{5.2.19}$$

3. Restrictions imposed by material symmetry

Most materials of practical interest possess some symmetry. Most commonly they are isotropic. Otherwise they may possess fiber symmetry, or have the symmetry of one or other of the crystal classes.

It is convenient to describe the symmetry of a material in terms of equivalent coordinate systems. Let x and $\bar{x}$ be two fixed rectangular cartesian coordinate systems with a common origin. We consider two deformations A and B. A is described by

$$x_i = f_i(X_B,t) \ , \tag{5.3.1}$$

where X_B and x_i are the coordinates in the system x of a particle at times T and t respectively. B is described by

$$\bar{x}_i = f_i(\bar{X}_B,t) \ , \tag{5.3.2}$$

where $\bar{X}_B$ and $\bar{x}_i$ are the coordinates in the system $\bar{x}$ of a particle at times T and t respectively. We note that we have taken the functions to be the same in both cases. This means that the deformation B viewed from the coordinate system $\bar{x}$ looks precisely the same as the deformation A viewed from the coordinate system x.

The Helmholtz free energy per unit mass W associated with the deformation A is, of course, a function of the deformation gradients $\partial x_i/\partial X_A$ and, as we have seen in §5.2, depends

depends on these through the six independent components C_{AB} in the reference system x of the Cauchy strain; thus:

$$W = F(C_{AB}) \ , \tag{5.3.3}$$

where

$$C_{AB} = \frac{\partial x_i}{\partial X_A} \frac{\partial x_i}{\partial X_B} \ . \tag{5.3.4}$$

The Helmholtz free energy per unit mass $\overline{W}$ associated with the deformation $\mathcal{B}$ is, of course, a function of the deformation gradients $\partial \overline{x}_i / \partial \overline{X}_A$, and again it follows from §5.2 that $\overline{W}$ depends on these through the six components $\overline{C}_{AB}$ in the reference system $\overline{x}$ of the Cauchy strain associated with the deformation $\mathcal{B}$, thus:

$$W = \overline{F}(\overline{C}_{AB}) \ , \tag{5.3.5}$$

where

$$\overline{C}_{AB} = \frac{\partial \overline{x}_i}{\partial \overline{X}_A} \frac{\partial \overline{x}_i}{\partial \overline{X}_B} \ . \tag{5.3.6}$$

We note that the functions F and $\overline{F}$ which occur in (5.3.3) and (5.3.5) are not, in general, the same. If they are then the two coordinate systems x and $\overline{x}$ are said to be equivalent*. In this case W is given by (5.3.3) and we may write (5.3.5) as

$$\overline{W} = F(\overline{C}_{AB}) \ . \tag{5.3.7}$$

---

*Here we imply equivalence only from the point of view of finite elasticity theory.

We note that if we take

$$\bar{x}_i = x_i \quad \text{and} \quad \bar{X}_A = X_A \ , \tag{5.3.8}$$

i.e. we choose the deformation B so that it is related to the reference system $\bar{x}$ in precisely the same way as the deformation A is related to the reference system x, then

$$\bar{W} = W \ . \tag{5.3.9}$$

Since the reference systems x and $\bar{x}$ are both rectangular cartesian systems, they are related by an orthogonal transformation, which we shall denote by s_{ij}. Let points with coordinates x_i and X_A in the system $\bar{x}$ have coordinates x_i' and X_A' in the system x, where

$$x_i' = s_{ij}x_j \ , \quad X_A' = s_{AB}X_B \ . \tag{5.3.10}$$

From (5.3.10)

$$\frac{\partial x_i'}{\partial X_A'} = s_{ij}s_{AB}\frac{\partial x_i}{\partial X_B} \ . \tag{5.3.11}$$

We denote by C_{AB}' the components in the system x of the Cauchy strain tensor associated with the deformation B. Then,

$$C_{AB}' = \frac{\partial x_i'}{\partial X_A'}\frac{\partial x_i'}{\partial X_B'} \ . \tag{5.3.12}$$

Introducing (5.3.11) into (5.3.12) and recalling that the transformation s_{ij} is a proper orthogonal transformation, we

obtain

$$C'_{AB} = s_{AP}s_{BQ}C_{PQ} \ . \tag{5.3.13}$$

From (5.3.3) we see that the Helmholtz free energy per unit mass associated with the deformation B is $F(C'_{AB})$ and from (5.3.6) we see that it is given by $F(\bar{C}_{AB})$. Thus

$$F(C'_{AB}) = F(\bar{C}_{AB}) \ . \tag{5.3.14}$$

Now, from (5.3.8), (5.3.4) and (5.3.6), we have

$$\bar{C}_{AB} = C_{AB} \ . \tag{5.3.15}$$

Introducing (5.3.13) and (5.3.15) into (5.3.14), we obtain

$$F(s_{AP}s_{BQ}C_{PQ}) = F(C_{AB}) \ . \tag{5.3.16}$$

Equation (5.3.16) is valid for all transformations $\underset{\sim}{s}$ relating equivalent coordinate systems and represents, in implicit form, a restriction on the manner in which F (and hence W) can depend on C_{AB}. The manner in which this implicit restriction can be made explicit will be explained in the following two sections. It will be seen that the details of the problem and the explicit form which can be given to the dependence of W on C_{AB} depend very much on the precise nature of the material symmetry.

R. S. Rivlin

4. Symmetry groups

The symmetry of a material can be described by specifying all the rectangular cartesian coordinate systems which are equivalent to a given rectangular cartesian system. The set of transformations relating every pair of these equivalent coordinate systems forms a group of transformations. This is the symmetry group of the material. Since each transformation of the group relates two rectangular cartesian coordinate systems, it is an orthogonal transformation. The symmetry group for the material is, therefore, either the full orthogonal group or a subgroup of it.

If the symmetry group is the full orthogonal group then every pair of rectangular cartesian coordinate systems, whether right or left handed (i.e. which can be brought into parallelism by a rigid rotation or rotation-reflection), are equivalent. The material is then isotropic and centrosymmetric. If the symmetry group is the proper orthogonal group, only rectangular cartesian coordinate systems of the same hand (i.e. which can be brought into parallelism by a rigid rotation) are equivalent. The material is then a non-centrosymmetric isotropic material.

Other symmetry groups which have particular interest in mechanics are the transverse isotropy groups and the crystallographic groups.

R S Rivlin

In a transversely isotropic material there is a single direction of rotational symmetry. Let this be the direction of the x_3-axis of a rectangular cartesian coordinate system x. Then, every rectangular cartesian coordinate system of the same hand which has its 3-axis in this direction is equivalent to x. We denote by θ the angle by which the system x has to be rotated about the x_3-axis in order to bring it into coincidence with an equivalent system $\bar{x}$. Then,

$$\bar{x}_1 = x_1\cos\theta + x_2\sin\theta ,$$
$$\bar{x}_2 = -x_1\sin\theta + x_2\cos\theta ,$$
$$\bar{x}_3 = x_3 ,$$

where x_i and $\bar{x}_i$ are the coordinates of a generic point, fixed in space, in the systems x and $\bar{x}$ respectively. We note that as far as the 1- and 2-coordinates are concerned, the transformation group is the two-dimensional proper orthogonal group. As far as the 3-coordinate is concerned, the transformation group is the identity group. In addition, the transversely isotropic group may or may not possess one or both of the following symmetries--reflection symmetry in a plane normal to the axis of rotational symmetry (i.e. the x_3-axis) and reflection symmetry in a plane containing the axis of rotational symmetry. There are thus four possible types of transversely isotropic material. Unfortunately there does not appear to be any generally accepted nomenclature for distinguishing between them.

There are thirty-two possible crystal symmetries--the so-called crystal classes, or crystallographic point groups. Each of these crystal classes can be characterized by a group of transformations which is a subgroup of the full orthogonal group. In each case it is a finite group. In order to define the group, it is not generally necessary to specify all the transformations of the group. Certain of these, called the generating transformations, suffice. The generating transformations of a group is a set of transformations of the group which has the property that each of the transformations of the group can be obtained by forming products from the generating transformations. For a specified group the generating transformations are not necessarily unique. We take an example. Consider the group

$$\underset{\sim}{I} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad \underset{\sim}{D}_1 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{Vmatrix},$$

$$\underset{\sim}{D}_2 = \begin{Vmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{Vmatrix} \quad \underset{\sim}{D}_3 = \begin{Vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

We note that $\underset{\sim}{D}_1$ and $\underset{\sim}{D}_2$ are generating transformations, since

$$\underset{\sim}{I} = \underset{\sim}{D}_1^2 \quad \text{and} \quad \underset{\sim}{D}_3 = \underset{\sim}{D}_1\underset{\sim}{D}_2.$$

Alternatively $\underset{\sim}{D}_2$ and $\underset{\sim}{D}_3$ can be taken as the generating transformations, or $\underset{\sim}{D}_3$ and $\underset{\sim}{D}_1$.

The following transformations are adequate for the description of all thirty-two crystal classes*:

$$\underset{\sim}{I} = (1,1,1)\ , \qquad \underset{\sim}{C} = (-1,-1,-1)\ ,$$

$$\underset{\sim}{R}_1 = (-1,1,1,)\ , \quad \underset{\sim}{R}_2 = (1,-1,1)\ , \quad \underset{\sim}{R}_3 = (1,1,-1)\ ,$$

$$\underset{\sim}{D}_1 = (1,-1,-1)\ , \quad \underset{\sim}{D}_2 = (-1,1,-1)\ , \quad \underset{\sim}{D}_3 = (-1,-1,1)\ ,$$

$$\underset{\sim}{T}_1 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix}, \quad \underset{\sim}{T}_2 = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{Vmatrix}, \quad \underset{\sim}{T}_3 = \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$\underset{\sim}{M}_1 = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{Vmatrix}, \quad \underset{\sim}{M}_2 = \begin{Vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix},$$

$$\underset{\sim}{S}_1 = \begin{Vmatrix} -1/2 & \sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \quad \underset{\sim}{S}_2 = \begin{Vmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

We note that $\underset{\sim}{I}$ is the identity transformation, $\underset{\sim}{C}$ denotes a central inversion and $\underset{\sim}{R}_1$, $\underset{\sim}{R}_2$, $\underset{\sim}{R}_3$ denote reflections in the 23, 31 and 12 planes respectively. $\underset{\sim}{D}_1$, $\underset{\sim}{D}_2$ and $\underset{\sim}{D}_3$ denote rotations through 180° about the 1, 2 and 3 directions respectively. $\underset{\sim}{T}_1$ denotes a reflection in the 13-plane followed

---

*The notation (a,b,c) is used to denote the diagonal matrix

$$\begin{Vmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{Vmatrix}$$

by a rotation through 90^{o} about the 1-axis; $\underset{\sim}{T}_2$ denotes a reflection in the 21-plane followed by a rotation through 90^{o} about the 2-axis; $\underset{\sim}{T}_3$ denotes a reflection in the 32-plane followed by a rotation through 90^{o} about the 3-axis. $\underset{\sim}{M}_1$ and $\underset{\sim}{M}_2$ denote rotations of 120^{o} and 240^{o} respectively about the tetrahedral direction (i.e. the direction making equal angles with the positive senses of the 1, 2 and 3 directions). $\underset{\sim}{S}_1$ and $\underset{\sim}{S}_2$ denote rotations of 120^{o} and 240^{o} respectively about the 3-direction.

5. The invariant-theoretical problem posed by material symmetry

We have seen in §5.3 that for an elastic material the Helmholtz free energy must depend on the deformation gradients through the six components of the Cauchy strain C_{AB}. This results from the fact that the superposition on the assumed deformation of a rigid rotation leaves the Helmholtz free energy unaltered. We thus have (cf. (5.3.3))

$$W = F(C_{AB}) \ . \tag{5.5.1}$$

We have also seen that if the material has some symmetry, this may be characterized by a group of transformations which relate a rectangular cartesian coordinate system to equivalent rectangular cartesian coordinate systems. Let $\underset{\sim}{s}$ be a transformation of this group, then we have seen (cf. (5.3.16)) that F must satisfy the relation

$$F(s_{AP}s_{BQ}C_{PQ}) = F(C_{AB}) \ . \tag{5.5.2}$$

This relation must be satisfied for every transformation of the symmetry group of the material, $\{\underset{\sim}{s}\}$ say.

Equation (5.5.2) states that F is a scalar invariant of the symmetric second-order tensor under the group of transformations $\{\underset{\sim}{s}\}$. It follows from results in the theory of invariants, which will be discussed later in §6.1, that F must be expressible as a function of a finite set of

invariants which depends on the group $\{\underset{\sim}{s}\}$. This set of invariants is called a function basis for invariants of the symmetric second-order tensor C_{AB} under the group $\{\underset{\sim}{s}\}$. If F is a polynomial function of C_{AB} then it must be expressible as a polynomial function of a finite set of polynomial invariants which again depends on the group $\{\underset{\sim}{s}\}$. This set of polynomial invariants is called an integrity basis for invariants of the symmetric second-order tensor C_{AB} under the group $\{\underset{\sim}{s}\}$.

If $F_1,F_2,\ldots,F_N$ are the elements of a function basis for invariants of the symmetric second-order tensor C_{AB} under the group $\{\underset{\sim}{s}\}$, then any function F satisfying (5.5.2) must be expressible as a function of $F_1,F_2,\ldots,F_N$, thus:

$$F = F(F_1,F_2,\ldots,F_N) \ . \tag{5.5.3}$$

We note that the function F in (5.5.3) is understood in a generic sense and is not, in general, the same function F as occurs in (5.5.2). Now, if one of the elements of the function basis, F_N say, is expressible as a function of the others, we can express F as a function of the elements of the smaller function basis $F_1,F_2,\ldots,F_{N-1}$. When all the redundant elements of the function basis have been eliminated in this way, we arrive at a minimal or irreducible function basis.

Analogously, if $I_1,I_2,\ldots,I_N$ are the elements of an integrity basis for invariants of the symmetric second-order tensor C_{AB} under the group $\{\underset{\sim}{s}\}$, then any polynomial

function F satisfying (5.5.2) must be expressible as a polynomial function of $I_1, I_2, \ldots, I_N$. If one of the elements of this integrity basis, I_N say, is expressible as a polynomial in the others, then F may be expressed as a polynomial in $I_1, I_2, \ldots I_{N-1}$ only. $I_1, I_2, \ldots, I_{N-1}$ then constitutes an integrity basis consisting of fewer elements than the original one. When all the redundant elements of the integrity basis have been eliminated in this way, we arrive at a minimal or irreducible integrity basis.

It can be shown fairly readily that an integrity basis necessarily forms a function basis, but it is not necessarily true that an irreducible integrity basis forms an irreducible function basis.

An irreducible integrity basis for the second-order symmetric tensor C_{AB} under the proper orthogonal group is formed by the three quantities

$$\operatorname{tr} \underset{\sim}{C} , \quad \operatorname{tr} \underset{\sim}{C}^2 , \quad \operatorname{tr} \underset{\sim}{C}^3 , \tag{5.5.4}$$

where $\underset{\sim}{C}$ is the matrix formed by the components C_{AB} of the Cauchy strain tensor. Alternatively, the set of three quantities I_1, I_2 and I_3 defined by

$$I_1 = \operatorname{tr} \underset{\sim}{C} , \quad I_2 = \frac{1}{2}[(\operatorname{tr} \underset{\sim}{C})^2 - \operatorname{tr} \underset{\sim}{C}^2] , \quad I_3 = \det \underset{\sim}{C} , \tag{5.5.5}$$

also forms an irreducible integrity basis. The relation between the integrity bases (5.5.4) and (5.5.5) can be easily appreciated

from the identity

$$\det \underset{\sim}{C} \equiv \frac{1}{6}[(\operatorname{tr} \underset{\sim}{C})^3 - 3 \operatorname{tr} \underset{\sim}{C} \operatorname{tr} \underset{\sim}{C}^2 + 2 \operatorname{tr} \underset{\sim}{C}^3] , \qquad (5.5.6)$$

which can be easily established in the following way. We first note that

$$\det \underset{\sim}{C} = \frac{1}{6} \varepsilon_{KLM}\varepsilon_{PQR}C_{KP}C_{LQ}C_{MR} , \qquad (5.5.7)$$

where ε_{KLM} is the alternating symbol. Now, we use the result

$$\varepsilon_{KLM}\varepsilon_{PQR} = \begin{vmatrix} \delta_{KP} & \delta_{KQ} & \delta_{KR} \\ \delta_{LP} & \delta_{LQ} & \delta_{LR} \\ \delta_{MP} & \delta_{MQ} & \delta_{MR} \end{vmatrix} . \qquad (5.5.8)$$

Expanding the determinant on the right-hand side of (5.5.8) and using the expression for $\varepsilon_{KLM}\varepsilon_{PQR}$ so obtained in (5.5.7), we readily obtain the identity (5.5.6).

We can use the relations (5.5.5) and (5.5.6) to obtain expressions for $\operatorname{tr} \underset{\sim}{C}$, $\operatorname{tr} \underset{\sim}{C}^2$, $\operatorname{tr} \underset{\sim}{C}^3$ in terms of I_1, I_2, I_3 thus:

$$\begin{aligned} &\operatorname{tr} \underset{\sim}{C} = I_1, \ \operatorname{tr} \underset{\sim}{C}^2 = I_1^2 - 2I_2 , \\ &\operatorname{tr} \underset{\sim}{C}^3 = 3I_3 - 3I_1I_2 + I_1^3 . \end{aligned} \qquad (5.5.9)$$

It is evident that (5.5.5) may be used instead of (5.5.4) as an irreducible integrity basis since if $\operatorname{tr} \underset{\sim}{C}$, $\operatorname{tr} \underset{\sim}{C}^2$ and $\operatorname{tr} \underset{\sim}{C}^3$ are given, I_1, I_2 and I_3 are determined uniquely by the polynomial relations (5.5.5) with (5.5.6). Conversely, if I_1, I_2

and I_3 are given, tr $\underset{\sim}{C}$, tr $\underset{\sim}{C}^2$ and tr $\underset{\sim}{C}^3$ are determined uniquely by the polynomial relations (5.5.9).

We note that tr $\underset{\sim}{C}$, tr $\underset{\sim}{C}^2$ and tr $\underset{\sim}{C}^3$--and hence I_1, I_2 and I_3--are unaltered by a central inversion. Since any improper orthogonal transformation is the resultant of a central inversion and a proper orthogonal transformation, it follows that tr $\underset{\sim}{C}$, tr $\underset{\sim}{C}^2$ and tr $\underset{\sim}{C}^3$--and hence I_1, I_2 and I_3--forms an irreducible integrity basis for the symmetric second-order tensor C_{AB} under the full orthogonal group.

Since tr $\underset{\sim}{C}$, tr $\underset{\sim}{C}^2$, tr $\underset{\sim}{C}^3$ and I_1, I_2, I_3 are integrity bases for the symmetric second-order tensor under the full, or proper, orthogonal group, they are also function bases. It can, in fact, be shown in each case that they are irreducible function bases.

For transversely isotropic materials*, for which the x_3-axis is the axis of rotational symmetry, an irreducible integrity basis for the symmetric second-order tensor C_{AB} is formed by the three invariants I_1, I_2, I_3, defined by (5.5.5) and also C_{33} and $C_{3B}C_{B3}$. These five quantities form an irreducible integrity basis for each of the four types of transverse isotropy described in §5.4. Again, these five quantities also form a function basis and it can be shown that this function basis is irreducible.

*The result, specifically in its application to elasticity theory, is due to a development of Ericksen and Rivlin (J. Rat'l. Mech. Anal. 3, 281 (1954)) and Adkins (Arch. Rat'l. Mech. Anal. 4, 193 (1960)). As a result in invariant theory it was probably known long before.

Irreducible integrity bases have been found for the second-order symmetric tensor C_{AB} for each of the crystal classes, by Smith and Rivlin*. These are given below. In each case the Schoenflies and Hermann-Mauguin nomenclatures for the crystal class are given, together with the generating transformations for the group. The notation for these transformations is given in §5.4.

Crystal class		Generating transform-tions	Integrity basis
Schoen-flies	Hermann-Mauguin		
Triclinic system			
C_1	1	$\underset{\sim}{I}$	$C_{AB}(A,B = 1,2,3)$
S_2	$\bar{1}$	$\underset{\sim}{C}$	
Monoclinic system			
C_2	2	$\underset{\sim}{D}_1$	$C_{11},C_{22},C_{33},C_{23},C_{31}^2,C_{12}^2,$
C_{1h}	m	$\underset{\sim}{R}_1$	$C_{12}C_{31}$
C_{2h}	2/m	$\underset{\sim}{C},\underset{\sim}{R}_1$	
Orthorhombic system			
V	222	$\underset{\sim}{D}_1,\underset{\sim}{D}_2$	$C_{11},C_{22},C_{33},C_{23}^2,C_{31}^2,C_{12}^2,$
C_{2v}	2mm	$\underset{\sim}{D}_1,\underset{\sim}{R}_2$	$C_{23}C_{31}C_{12}$
V_h	mmm	$\underset{\sim}{C},\underset{\sim}{R}_1,\underset{\sim}{R}_2$	
Tetragonal system			
S_4	$\bar{4}$	$\underset{\sim}{D}_1\underset{\sim}{T}_3$	$C_{11}+C_{22},C_{33},C_{23}^2+C_{31}^2,C_{12}^2,$
C_4	4	$\underset{\sim}{R}_1\underset{\sim}{T}_3$	$C_{11}C_{22},C_{12}(C_{11}-C_{22}),$
C_{4h}	4/m	$\underset{\sim}{C},\underset{\sim}{R}_1\underset{\sim}{T}_3$	$C_{23}C_{31}(C_{11}-C_{22}),C_{23}C_{31}C_{12},$ $C_{12}(C_{31}^2-C_{23}^2),C_{11}C_{23}^2+C_{22}C_{31}^2,$ $C_{23}C_{31}(C_{31}^2-C_{23}^2),C_{23}^2C_{31}^2$

*G. F. Smith and R. S. Rivlin, Trans. Amer. Math. Soc. 88, 175 (1958).

D_4	422	$\underset{\sim}{R}_1\underset{\sim}{T}_3,\underset{\sim}{D}_1$	$C_{11} + C_{22}, C_{33}, C_{23}^2 + C_{31}^2, C_{12}^2,$
V_d	$\bar{4}2m$	$\underset{\sim}{D}_1\underset{\sim}{T}_3,\underset{\sim}{D}_1$	$C_{11}C_{22}, C_{23}C_{31}C_{12},$
C_{4v}	4mm	$\underset{\sim}{R}_1\underset{\sim}{T}_3,\underset{\sim}{R}_1$	$C_{11}C_{23}^2 + C_{22}C_{31}^2, C_{23}^2C_{31}^2$
D_{4h}	4/mmm	$\underset{\sim}{C},\underset{\sim}{R}_1,\underset{\sim}{R}_1\underset{\sim}{T}_3$	
Hexagonal system			
C_3	3	$\underset{\sim}{S}_1$	$C_{33}, C_{11} + C_{22}, C_{11}C_{22} - C_{12}^2,$
C_{3i}	$\bar{3}$	$\underset{\sim}{C}\underset{\sim}{S}_1$	$C_{11}[C_{11} + 3C_{22})^2 - 12C_{12}^2], C_{31}^2 + C_{23}^2,$
			$C_{31}(C_{31}^2 - 3C_{23}^2), (C_{11} - C_{22})C_{31} -$
			$2C_{12}C_{23}, (C_{22} - C_{11})C_{23} - 2C_{12}C_{31},$
			$3C_{12}(C_{11} - C_{22})^2 - 4C_{12}^3, C_{23}(C_{23}^2 -$
			$3C_{31}^2), C_{22}C_{31}^2 + C_{11}C_{23}^2 - 2C_{23}C_{31}C_{12},$
			$C_{31}[(C_{11} + C_{22})^2 + 4(C_{12}^2 - C_{22}^2)] -$
			$8C_{11}C_{12}C_{23}, C_{23}[(C_{11} + C_{22})^2 +$
			$4(C_{12}^2 - C_{22}^2)] + 8C_{11}C_{12}C_{31},$
			$(C_{11} - C_{22})C_{23}C_{31} + C_{12}(C_{23}^2 - C_{31}^2)$
D_3	32	$\underset{\sim}{S}_1,\underset{\sim}{D}_1$	$C_{33}, C_{11} + C_{22}, C_{11}C_{22} - C_{12}^2, C_{31}^2 + C_{23}^2,$
C_{3v}	3m	$\underset{\sim}{S}_1,\underset{\sim}{R}_1$	$C_{11}[(C_{11} + 3C_{22})^2 - 12C_{12}^2], C_{23}(C_{23}^2 -$
D_{3d}	$\bar{3}2/m$	$\underset{\sim}{R}_1,\underset{\sim}{C}\underset{\sim}{S}_1$	$3C_{31}^2), (C_{11} - C_{22})C_{23} + 2C_{12}C_{31}, C_{11}C_{31}^2$
			$+ C_{22}C_{23}^2 + 2C_{23}C_{31}C_{12}, C_{23}[(C_{11} + C_{22})^2$
			$-4(C_{22}^2 - C_{12}^2)] + 8C_{11}C_{12}C_{31}$
C_{3h}	$\bar{6}$	$\underset{\sim}{R}_3\underset{\sim}{S}_1$	$C_{33}, C_{11} + C_{22}, C_{11}C_{22} - C_{12}^2, C_{31}^2 + C_{23}^2,$
C_6	6	$\underset{\sim}{D}_3\underset{\sim}{S}_1$	$C_{11}[(C_{11} + 3C_{22})^2 - 12C_{12}^2],$
C_{6h}	6/m	$\underset{\sim}{C}\underset{\sim}{S}_1,\underset{\sim}{R}_3$	$C_{31}(C_{31} - 3C_{23}^2)^2, C_{11}C_{23}^2 +$
			$C_{22}C_{31}^2 - 2C_{23}C_{31}C_{12}, C_{12}(C_{31}^2 - C_{23}^2) +$
			$(C_{22} - C_{11})C_{31}C_{23}, 3C_{12}(C_{11} - C_{22})^2 -$
			$4C_{12}^3, C_{31}C_{23}[3(C_{31}^2 - C_{23}^2) -$

		Contd.	$4C_{31}^2C_{23}^2], C_{11}(C_{31}^4 + 3C_{23}^4) + 2C_{22}C_{31}^2(C_{31}^2 + 3C_{23}^2) - 8C_{12}C_{23}C_{31}^3, C_{31}^2[(C_{11} + C_{22})^2 - 4(C_{22}^2 - C_{12}^2)] - 2C_{11}[(C_{11} + 3C_{22})(C_{31}^2 + C_{23}^2) - 4C_{23}C_{31}C_{12}], C_{23}C_{31}[(C_{11} + C_{22})^2 - 4(C_{22}^2 - C_{12}^2)] + 4C_{11}C_{12}(C_{23}^2 - C_{31}^2), C_{12}[(C_{31}^2 + C_{23}^2)^2 + 4C_{23}^2(C_{31}^2 - C_{23}^2)] - 4C_{31}^3C_{23}(C_{11} - C_{22})$
D_{3h}	$\bar{6}$m2	$\underset{\sim}{R}_3\underset{\sim}{S}_1, \underset{\sim}{R}_1$	$C_{33}, C_{11} + C_{22}, C_{11}C_{22} - C_{12}^2, C_{31}^2 + C_{23}^2,$
D_6	622	$\underset{\sim}{D}_3\underset{\sim}{S}_1, \underset{\sim}{D}_1$	$C_{11}[(C_{11} + 3C_{22})^2 - 12C_{12}^2],$
C_{6v}	6mm	$\underset{\sim}{D}_3\underset{\sim}{S}_1, \underset{\sim}{R}_1$	$C_{31}^2(C_{31}^2 - 3C_{23}^2)^2, C_{11}C_{23}^2 + C_{22}C_{31}^2$
D_{6h}	6/mmm	$\underset{\sim}{D}_3\underset{\sim}{S}_1, \underset{\sim}{R}_1, \underset{\sim}{C}$	$- 2C_{23}C_{31}C_{12}, C_{11}(C_{31}^4 + 3C_{23}^4) + 2C_{22}C_{31}^2(C_{31}^2 + 3C_{23}^2) - 8C_{12}C_{23}C_{31}^3, C_{31}^2[(C_{11} + C_{22})^2 - 4(C_{22}^2 - C_{12}^2)] - 2C_{11}[(C_{11} + 3C_{22})(C_{31}^2 + C_{23}^2) - 4C_{23}C_{31}C_{12}]$
Cubic system			
T	23	$\underset{\sim}{D}_1, \underset{\sim}{M}_1$	$\Sigma C_{11}, \Sigma C_{22}C_{33}, \Sigma C_{23}^2, C_{11}C_{22}C_{33},$
T_h	m3	$\underset{\sim}{C}, \underset{\sim}{D}_1, \underset{\sim}{M}_1$	$C_{23}C_{31}C_{12}, \Sigma C_{31}^2C_{12}^2, \Sigma C_{22}C_{12}^2, \Sigma C_{33}C_{31}^2, \Sigma C_{33}C_{22}^2, \Sigma C_{12}^2C_{31}^4, \Sigma C_{11}C_{31}^2C_{12}^2, \Sigma C_{23}^2C_{22}C_{33}, \Sigma C_{23}^2C_{33}C_{11}, \Sigma C_{23}^2C_{31}^2C_{22}$
T_d	$\bar{4}$3m	$\underset{\sim}{D}_1, \underset{\sim}{M}_1, \underset{\sim}{T}_1$	$\Sigma C_{11}, \Sigma C_{22}C_{33}, \Sigma C_{23}^2, C_{11}C_{22}C_{33},$
O	432	$\underset{\sim}{R}_3\underset{\sim}{T}_1, \underset{\sim}{M}_2$	$C_{23}C_{31}C_{12}, \Sigma C_{31}^2C_{12}^2, \Sigma C_{33}C_{12}^2,$
O_h	m3m	$\underset{\sim}{R}_3\underset{\sim}{T}_1, \underset{\sim}{M}_2, \underset{\sim}{C}$	$\Sigma C_{11}C_{31}^2C_{12}^2, \Sigma C_{23}^2C_{22}C_{33}$

6. Constitutive equations of finite elasticity theory for isotropic materials

We have seen in §5.5 that the manner in which W, the Helmholtz free energy for an elastic material which has some symmetry, can depend on the components C_{AB} of the Cauchy strain tensor is restricted. Thus, if

$$W = F(C_{AB}) \ , \tag{5.6.1}$$

then the function F must satisfy the relation

$$F(s_{AP}s_{BQ}C_{PQ}) = F(C_{AB}) \tag{5.6.2}$$

for every transformation of the symmetry group of the material $\{\underset{\sim}{s}\}$ say. This states that F is a scalar invariant of the symmetric second-order tensor C_{AB} under the group $\{\underset{\sim}{s}\}$. F, and hence W, may therefore be expressed as a function of the elements of an irreducible integrity basis*.

In the case when the material is isotropic, whether centrosymmetric or not, it has been seen in §5.5 that such an integrity basis is formed by the quantities I_1, I_2 and I_3 defined by (5.5.5). We therefore write

$$W = W(I_1,I_2,I_3) \ , \tag{5.6.3}$$

where (cf. equations (5.5.5) and (5.5.6))

---

*We have already pointed out that an integrity basis is necessarily a function basis.

$$I_1 = \text{tr } \underset{\sim}{C} , \quad I_2 = \frac{1}{2}[(\text{tr } \underset{\sim}{C})^2 - \text{tr } \underset{\sim}{C}^2] ,$$

$$I_3 = \det \underset{\sim}{C} = \frac{1}{6}[(\text{tr } \underset{\sim}{C})^3 - 3 \text{ tr } \underset{\sim}{C} \text{ tr } \underset{\sim}{C}^2 + 2 \text{ tr } \underset{\sim}{C}^3] . \qquad (5.6.4)$$

Introducing (5.6.3) into (5.2.13), we obtain

$$\sigma_{ji} = \frac{\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} x_{j,B} x_{i,P} \sum_{\alpha=1}^{3} W_\alpha \left(\frac{\partial I_\alpha}{\partial C_{BP}} + \frac{\partial I_\alpha}{\partial C_{PB}} \right) , \qquad (5.6.5)$$

where the notation

$$W_\alpha = \partial W / \partial I_\alpha \qquad (5.6.6)$$

is used.

We note from (5.6.4) that

$$\frac{\partial I_1}{\partial C_{AB}} = \delta_{AB} , \quad \frac{\partial I_2}{\partial C_{AB}} = I_1 \delta_{AB} - C_{AB} ,$$

$$\frac{\partial I_3}{\partial C_{AB}} = I_2 \delta_{AB} - I_1 C_{AB} + C_{AQ} C_{QB} . \qquad (5.6.7)$$

Introducing (5.6.7) into (5.6.5) and bearing in mind the definition of the Cauchy strain C_{AB}, we obtain

$$\sigma_{ji} = \frac{2\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} \{(W_1 + I_1 W_2 + I_2 W_3) c_{ij} - (W_2 + I_1 W_3) c_{ik} c_{kj} + W_3 c_{ik} c_{k\ell} c_{\ell j}\} , \qquad (5.6.8)$$

where c_{ij} denotes the Finger strain defined by (cf. equation (2.5.6))

$$c_{ij} = x_{i,A} x_{j,A} . \qquad (5.6.9)$$

Equation (5.6.8) may be written slightly more succintly in matrix notation thus:

$$\underset{\sim}{\sigma} = \frac{2\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} \{(W_1+I_1W_2+I_2W_3)\underset{\sim}{c} - (W_2+I_1W_3)\underset{\sim}{c}^2 + W_3\underset{\sim}{c}^3\} . \qquad (5.6.10)$$

We can easily verify that

$$\text{tr } \underset{\sim}{C} = \text{tr } \underset{\sim}{c} , \quad \text{tr } \underset{\sim}{C}^2 = \text{tr } \underset{\sim}{c}^2 \quad \text{and} \quad \text{tr } \underset{\sim}{C}^3 = \text{tr } \underset{\sim}{c}^3, \qquad (5.6.11)$$

so that I_1, I_2, I_3 may be expressed in terms of $\text{tr } \underset{\sim}{c}, \text{tr } \underset{\sim}{c}^2, \text{tr } \underset{\sim}{c}^3$ by (cf. equations (5.6.4))

$$\begin{aligned} I_1 &= \text{tr } \underset{\sim}{c} , \quad I_2 = \frac{1}{2}[(\text{tr } \underset{\sim}{c})^2 - \text{tr } \underset{\sim}{c}^2] , \\ I_3 &= \det \underset{\sim}{c} = \frac{1}{6}[(\text{tr } \underset{\sim}{c})^3 - 3 \text{ tr } \underset{\sim}{c} \text{ tr } \underset{\sim}{c}^2 + 2 \text{ tr } \underset{\sim}{c}^3] . \end{aligned} \qquad (5.6.12)$$

If we apply the well known Hamilton-Cayley theorem to the matrix $\underset{\sim}{c}$, we obtain, with (5.6.12)

$$\underset{\sim}{c}^3 - I_1\underset{\sim}{c}^2 + I_2\underset{\sim}{c} - I_3\underset{\sim}{I} = 0 . \qquad (5.6.13)$$

Using this result to substitute for $\underset{\sim}{c}^3$ in (5.6.10), we obtain*

$$\underset{\sim}{\sigma} = \frac{2\rho_o}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} \{(W_1+I_1W_2)\underset{\sim}{c} - W_2\underset{\sim}{c}^2 + I_3W_3\underset{\sim}{I}\} . \qquad (5.6.14)$$

Yet another form for $\underset{\sim}{\sigma}$ can be obtained in the following way. Dividing (5.6.13) throughout by $\underset{\sim}{c}$, we obtain

$$\underset{\sim}{c}^2 - I_1\underset{\sim}{c} + I_2\underset{\sim}{I} - I_3\underset{\sim}{c}^{-1} = 0 . \qquad (5.6.15)$$

---

*This result was first obtained by Finger (Sitzungsber. Akad. Wiss. Wien 103, 1073 (1894)).

Using this result to substitute for $\underset{\sim}{C}^2$ in (5.6.14), we obtain

$$\underset{\sim}{\sigma} = \frac{2\rho_o}{\partial\underset{\sim}{x}/\partial\underset{\sim}{X}} \{W_1\underset{\sim}{C} - I_3W_2\underset{\sim}{C}^{-1} + (I_2W_2+I_3W_3)\underset{\sim}{I}\} . \qquad (5.6.16)$$

Equations (5.6.14) and (5.6.16) are alternative forms for the Cauchy stress in an isotropic elastic material.

An expression for the Piola stress can be obtained by introducing (5.6.3) into (5.2.14). We obtain

$$P_{AB} = \rho_o \sum_{\alpha=1}^{3} \frac{\partial W}{\partial I_\alpha}\left(\frac{\partial I_\alpha}{\partial C_{BA}} + \frac{\partial I_\alpha}{\partial C_{AB}}\right) . \qquad (5.6.17)$$

Employing the expressions (5.6.7), this yields

$$\underset{\sim}{P} = 2\rho_o\{(W_1+I_1W_2+I_2W_3)\underset{\sim}{I} - (W_2+I_1W_3)\underset{\sim}{C} + W_3\underset{\sim}{C}^2\} , \qquad (5.6.18)$$

where $\underset{\sim}{P}$ is the matrix $\|P_{AB}\|$. An alternative expression for $\underset{\sim}{P}$ is obtained by using an expression for $\underset{\sim}{C}$ analogous to (5.6.15), derived from the Hamilton-Cayley theorem for the matrix $\underset{\sim}{C}$. We have

$$\underset{\sim}{C}^2 - I_1\underset{\sim}{C} + I_2\underset{\sim}{I} - I_3\underset{\sim}{C}^{-1} = 0 . \qquad (5.6.19)$$

Employing (5.6.19) in (5.6.18), we obtain

$$\underset{\sim}{P} = 2\rho_o\{(W_1+I_1W_2)\underset{\sim}{I} - W_2\underset{\sim}{C} + I_3W_3\underset{\sim}{C}^{-1}\} . \qquad (5.6.20)$$

The corresponding expression for the Kirchoff-Piola stress can be obtained by using (3.8.3), viz.

$$\pi_{Ai} = x_{i,B}P_{AB} , \qquad (5.6.21)$$

or (3.7.6), viz.

$$\pi_{Ai} = (\partial\underset{\sim}{x}/\partial\underset{\sim}{X})X_{A,j}\sigma_{ji} . \qquad (5.6.22)$$

7. Constitutive equations of finite elasticity for incompressible isotropic materials

If the material is incompressible, the volume of each material element remains unchanged in the deformation so that (cf. (2.3.4) and (2.3.5))

$$\partial \underset{\sim}{x}/\partial \underset{\sim}{X} = (\det \underset{\sim}{C})^{1/2} = 1 \ . \tag{5.7.1}$$

It has been seen that in this case the expression for the Cauchy stress is given by (5.2.16) and that for the Kirchoff-Piola stress by (5.2.15).

We note from (5.6.4) that the constraint condition (5.7.1) may be expressed in the form

$$I_3 = 1 \ . \tag{5.7.2}$$

It follows from (5.6.3) that for an incompressible isotropic material the Helmholtz free energy is expressible as a function of the two variables I_1 and I_2 thus:

$$W = W(I_1, I_2) \ . \tag{5.7.3}$$

Introducing this form for W into (5.2.16) and (5.2.17), we obtain, following a procedure similar to that used in deriving (5.6.14), (5.6.16) and (5.6.20)*,

---

*R. S. Rivlin, Phil. Trans. A 241, 379 (1948).

$$\underset{\sim}{\sigma} = 2\rho_o\{(W_1+I_1W_2)\underset{\sim}{c} - W_2\underset{\sim}{c}^2\} - p\underset{\sim}{I}$$

$$= 2\rho_o\{W_1\underset{\sim}{c} - W_2\underset{\sim}{c}^{-1}\} - p\underset{\sim}{I} \qquad (5.7.4)$$

and

$$\underset{\sim}{P} = 2\rho_o\{(W_1+I_1W_2)\underset{\sim}{I} - W_2\underset{\sim}{C}\} - p\underset{\sim}{C}^{-1} .$$

Chapter 6

SOME RESULTS IN THE THEORY OF INVARIANTS

1. Definitions

Let $v^{(1)}_{i\ldots j}$, $v^{(2)}_{k\ldots \ell}$, ..., $v^{(N)}_{m\ldots n}$ be the components in a rectangular cartesian coordinate system x of the N tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$, which are not necessarily of the same order. As a special case, one or more of the N tensors may be a tensor of zero order (i.e. a scalar) or of first order (i.e. a vector). Moreover, they may be absolute tensors or relative tensors. Let $\bar{v}^{(1)}_{i\ldots j}$, $\bar{v}^{(2)}_{k\ldots \ell}$, ..., $\bar{v}^{(N)}_{m\ldots n}$ be the components of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ in the rectangular cartesian coordinate system $\bar{x}$.

Now, suppose that y is a scalar function of the components in the system x and that $\bar{y}$ is the same function of the components in the system $\bar{x}$, thus:

$$y = F(v^{(1)}_{i\ldots j}, v^{(2)}_{k\ldots \ell}, \ldots, v^{(N)}_{m\ldots n})$$

and (6.1.1)

$$\bar{y} = F(\bar{v}^{(1)}_{i\ldots j}, \bar{v}^{(2)}_{k\ldots \ell}, \ldots, \bar{v}^{(N)}_{m\ldots n}) .$$

Let the coordinate systems x and $\bar{x}$ be related by the orthogonal transformation $\underset{\sim}{s}$, so that, in vector notation,

$$\underset{\sim}{\bar{x}} = \underset{\sim}{s}\underset{\sim}{x} , \qquad (6.1.2)$$

or, in indicial notation,

$$\bar{x}_i = s_{ij}x_j \ . \tag{6.1.3}$$

Then,

$$\begin{aligned}
\bar{v}^{(1)}_{i\ldots j} &= s_{ip}\ldots s_{jq}v^{(1)}_{p\ldots q} \ , \\
\bar{v}^{(2)}_{k\ldots \ell} &= s_{kr}\ldots s_{\ell s}v^{(2)}_{r\ldots s} \ , \\
&\ldots\ldots \\
\bar{v}^{(N)}_{m\ldots n} &= s_{mt}\ldots s_{nu}v^{(N)}_{t\ldots u} \ ,
\end{aligned} \tag{6.1.4}$$

provided that the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ are absolute tensors. If one of the tensors $\underset{\sim}{v}_1$, say, is a relative tensor, then the appropriate transformation relating its components in the systems $\bar{x}$ and x is

$$\bar{v}^{(1)}_{i\ldots j} = (\det \underset{\sim}{v}_1)s_{ip}\ldots s_{jq}v^{(1)}_{p\ldots q} \ , \tag{6.1.5}$$

instead of $(6.1.4)_1$.

If y and $\bar{y}$, defined by (6.1.1), are equal for all transformations $\underset{\sim}{s}$ belonging to a group $\{\underset{\sim}{s}\}$, then F is said to be a <u>scalar invariant</u> of the tensors $\underset{\sim}{v}_1,\ldots\underset{\sim}{v}_N$ under the group $\{\underset{\sim}{s}\}$. If F is a polynomial function of its arguments, then it follows from a theorem of Hilbert (the First Main Theorem of the Theory of Invariants) that there must exist a <u>finite</u> set of polynomial invariants $I_1,I_2,\ldots,I_\mu$, say, of $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_n$, in terms of which F may be expressed as a polynomial, thus:

$$F = P(I_1,\ldots,I_\mu) \ . \tag{6.1.6}$$

Such a set of polynomial invariants is called an integrity basis. Thus, the First Main Theorem asserts the existence of a finite integrity basis. If we omit all elements of an integrity basis which may be expressed as polynomials in the remaining ones, the remaining set of elements forms an irreducible integrity basis. We shall consider that $I_1,I_2,\ldots,I_\mu$ form an irreducible integrity basis.

We now consider the case when F is a single-valued function of its arguments rather than a polynomial function. It has been shown by Pipkin and Wineman (1963)* that, in this case, F may be expressed as a single-valued function of $I_1,I_2,\ldots,I_\mu$.

More generally suppose that $F_1,F_2,\ldots,F_\nu$ is a set of single-valued functions of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$, which are invariant under the group $\{\underset{\sim}{s}\}$. Suppose further that an arbitrary single-valued invariant function of $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ is expressible as a single-valued function of $F_1,F_2,\ldots,F_\nu$. Then $F_1,F_2,\ldots,F_\nu$ is called a function basis for invariants of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ under the group $\{\underset{\sim}{s}\}$. The result of Pipkin and Wineman then implies that an integrity basis is also a function basis. It does not, however, follow that an irreducible integrity basis is also an irreducible function basis, since one element of the irreducible integrity basis $I_1,I_2,\ldots,I_\nu$

* A.C. Pipkin and A.S. Wineman, Arch. Rat'l. Mech. Anal. 12, 420 (1963)

may be a single-valued function, but not a polynomial function, of the others. This may arise in the following manner.

If $I_1,\ldots,I_\nu$ is an irreducible integrity basis, although none of the elements is expressible as a polynomial function of the remaining ones, there may nevertheless exist one or more implicit polynomial relations between the invariants $I_1,\ldots,I_\nu$. Such a relation is called a syzygy. Suppose one of these syzygies takes the form

$$I_\alpha Q(I_1,\ldots,I_{\alpha-1},I_{\alpha+1},\ldots,I_\nu) = R(I_1,\ldots,I_{\alpha-1},I_{\alpha+1},\ldots,I_\nu) \ , \qquad (6.1.7)$$

where Q and R are polynomials in the indicated arguments. Then this relation may be used to express I_α as a single-valued function of the remaining I's, thus

$$I_\alpha = R/Q \ , \qquad (6.1.8)$$

provided $Q \neq 0$. If $Q = 0$ for some values of $I_1,\ldots,I_{\alpha-1}$, $I_{\alpha+1},\ldots,I_\nu$ then, for these values, the expression (6.1.8) must be replaced by some other expression for I_α as a function of $I_1,\ldots,I_{\alpha-1},I_{\alpha+1},\ldots,I_\nu$. Thus, in principle, I_α may be omitted from the function basis.

However, the resulting expression for an invariant function must be used with considerable caution, since its continuity and differentiability properties, when it is regarded

as a function of $I_1,\ldots,I_{\alpha-1},I_{\alpha+1},\ldots,I_\nu$, may be very different from those when it is regarded as a function of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$. For example, suppose F is an invariant function, so that

$$F = F(I_1,\ldots,I_{\alpha-1},I_{\alpha+1},\ldots,I_\nu) \ , \tag{6.1.9}$$

and we wish to calculate $\partial F/\partial v^{(1)}_{i\ldots j}$ by means of the expression

$$\frac{\partial F}{\partial v^{(1)}_{i\ldots j}} = \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\nu} \frac{\partial F}{\partial I_\beta}\frac{\partial I_\beta}{\partial v^{(1)}_{i\ldots j}} \ . \tag{6.1.10}$$

It may well be that although F is differentiable with respect to $v^{(1)}_{i\ldots j}$, the derivatives $\partial F/\partial I_\beta$ may not exist.

2. Integrity bases for vectors under the full and proper orthogonal groups

Irreducible integrity bases have been found for a number of choices of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ and of the group $\{\underset{\sim}{s}\}$. For example, in the case when $\{\underset{\sim}{s}\}$ is the full orthogonal group and $\underset{\sim}{v}_1,\ldots\underset{\sim}{v}_N$ are polar (i.e. absolute) vectors, an irreducible integrity basis is formed by the $\frac{1}{2}N(N+1)$ inner products

$$\underset{\sim}{v}_P\cdot\underset{\sim}{v}_P (P = 1,\ldots,N) \text{ and } \underset{\sim}{v}_P\cdot\underset{\sim}{v}_Q (P,Q = 1,\ldots,N;\ P < Q) \ . \quad (6.2.1)$$

Again, when $\{\underset{\sim}{s}\}$ is the proper orthogonal group and $\underset{\sim}{v}_1,\ldots\underset{\sim}{v}_N$ are either polar or axial (i.e. relative) vectors, an irreducible integrity basis is formed by the $\frac{1}{2}N(N+1)$ terms (6.2.1) and, in addition, the $\frac{1}{6}N(N-1)(N-2)$ scalar triple products

$$[\underset{\sim}{v}_P,\underset{\sim}{v}_Q,\underset{\sim}{v}_R] \ (P,Q,R = 1,\ldots N;\ P < Q < R) \ . \quad (6.2.2)$$

The integrity basis (6.2.1) can be presented in a slightly different form. Let $\underset{\sim}{a}$ and $\underset{\sim}{b}$ be two polar vectors. Then, (6.2.1) is equivalent to the set of quantities we obtain from

$$\underset{\sim}{a}.\underset{\sim}{a} \quad \text{and} \quad \underset{\sim}{a}.\underset{\sim}{b} \quad (6.2.3)$$

by making the following replacements:

(i) in $\underset{\sim}{a}.\underset{\sim}{a}$ replace $\underset{\sim}{a}$ by $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ in turn;

(ii) in $\underset{\sim}{a}.\underset{\sim}{b}$ replace $\underset{\sim}{a}$ by $\underset{\sim}{v}_1$ and $\underset{\sim}{b}$ by $\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ in turn,

replace $\underset{\sim}{a}$ by $\underset{\sim}{v}_2$ and $\underset{\sim}{b}$ by $\underset{\sim}{v}_3,\ldots,\underset{\sim}{v}_N$ in turn,

.....

replace $\underset{\sim}{a}$ by $\underset{\sim}{v}_{N-1}$ and $\underset{\sim}{b}$ by $\underset{\sim}{v}_N$.

Consequently (6.2.3) is called a table of typical invariants for a set of N vectors under the full orthogonal group. We note that if N = 1 (i.e. we are considering the invariants of only one vector), the table of typical invariants (6.2.3) reduces to

$$\underset{\sim}{a}.\underset{\sim}{a} \tag{6.2.4}$$

and the integrity basis is $\underset{\sim}{v}_1.\underset{\sim}{v}_1$ (i.e. the square of the magnitude of the vector $\underset{\sim}{v}_1$). If N = 2, the table of typical invariants is (6.2.3) and the irreducible integrity basis consists of the terms

$$\underset{\sim}{v}_1.\underset{\sim}{v}_1,\ \underset{\sim}{v}_2.\underset{\sim}{v}_2,\ \underset{\sim}{v}_1.\underset{\sim}{v}_2 \tag{6.2.5}$$

(i.e. the squares of the magnitudes of the vectors and their inner product). We note from (6.2.1) that for N(>2) vectors, no further types of terms are introduced into the integrity basis, which consists simply of the squares of the magnitudes of the vectors and the inner product of each pair which can be chosen from the N vectors. Indeed, we see that the integrity basis for N(>2) polar vectors under the full orthogonal group consists of the integrity bases for every pair of vectors which can be selected from the N vectors.

Again, a table of typical invariants for an irreducible integrity basis of N vectors under the proper orthogonal group is

$$\underset{\sim}{a}.\underset{\sim}{a}\ ,\quad \underset{\sim}{a}.\underset{\sim}{b}\ ,\quad [\underset{\sim}{a},\underset{\sim}{b},\underset{\sim}{c}]\ , \tag{6.2.6}$$

where $\underset{\sim}{a}$, $\underset{\sim}{b}$, and $\underset{\sim}{c}$ are vectors. The integrity basis for one vector $\underset{\sim}{v}_1$ is then $\underset{\sim}{v}_1.\underset{\sim}{v}_1$, as before. The integrity basis for two vectors $\underset{\sim}{v}_1$ and $\underset{\sim}{v}_2$ consists of the three terms (6.2.5). It is only when we come to the integrity basis for three vectors $\underset{\sim}{v}_1$, $\underset{\sim}{v}_2$ and $\underset{\sim}{v}_3$, say, that a new type of term is introduced into the integrity basis. The integrity basis for N(>3) vectors involves no new types of term and consists of the integrity bases for each set of three vectors which can be selected from the four vectors $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$.

3. Integrity basis for symmetric second-order tensors under the full or proper orthogonal group

Finally the table of typical invariants for an irreducible integrity basis under the full orthogonal (or proper orthogonal) group of an arbitrary number of symmetric second-order tensors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_N$ may be written down in terms of the six symmetric second-order tensors $\underset{\sim}{a}$, $\underset{\sim}{b}$, $\underset{\sim}{c}$, $\underset{\sim}{d}$, $\underset{\sim}{e}$, $\underset{\sim}{f}$. It consists of the traces of the products formed from $\underset{\sim}{a}$, $\underset{\sim}{b}$, $\underset{\sim}{c}$, $\underset{\sim}{d}$, $\underset{\sim}{e}$, $\underset{\sim}{f}$ given in the following table*:

(1) $\underset{\sim}{a}$, $\underset{\sim}{a}^2$, $\underset{\sim}{a}^3$,

(2) $\underset{\sim}{a}\underset{\sim}{b}$, $\underset{\sim}{a}\underset{\sim}{b}^2$, $\underset{\sim}{a}^2\underset{\sim}{b}$, $\underset{\sim}{a}^2\underset{\sim}{b}^2$,

(3) $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}$, $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}^2$, $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{a}^2$, $\underset{\sim}{c}\underset{\sim}{a}\underset{\sim}{b}^2$,
$\underset{\sim}{a}\underset{\sim}{b}^2\underset{\sim}{c}^2$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{a}^2$, $\underset{\sim}{c}\underset{\sim}{a}^2\underset{\sim}{b}^2$

(4) $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}$, $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}$,
$\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2$, $\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}^2$, $\underset{\sim}{d}\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}^2$, $\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{a}\underset{\sim}{c}^2$,
$\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{a}\underset{\sim}{b}^2$, $\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{a}^2$, $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{a}^2$, $\underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}^2$,
$\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{d}^2$, $\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{b}^2\underset{\sim}{d}^2$, $\underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{b}^2\underset{\sim}{c}^2$,
$\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{a}^2\underset{\sim}{d}^2$, $\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{a}^2\underset{\sim}{c}^2$, $\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{a}^2\underset{\sim}{b}^2$,
$\underset{\sim}{b}\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{a}^2$, $\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{a}\underset{\sim}{b}^2$, $\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}^2$, $\underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2$, (6.3.1)

(5) $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}$, $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{e}\underset{\sim}{c}$, $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{d}$,
$\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{e}$, $\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}\underset{\sim}{d}$, $\underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{e}$,
$\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}^2$, $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{e}^2$, $\underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}^2$, $\underset{\sim}{b}\underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}^2$ and terms obtained from these by cyclically permuting $\underset{\sim}{a}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}$

*This result results from a development in a number of papers by Rivlin, Spencer and G. F. Smith. (R. S. Rivlin, J. Rat'l. Mech. Anal. 4, 681 (1955); A. J. M. Spencer and R. S. Rivlin, Arch. Rat'l. Mech. Anal. 2, 309 (1959), 2 435 (1959), 4, 214 (1959); A. J. M. Spencer, ibid. 7, 64 (1961); G. F. Smith, ibid. 18, 282 (1965)). The results of these papers have been summarized and their derivation streamlined by R. S. Rivlin and G. F. Smith, Contributions to Mechanics, ed. D. Abir, publ. Pergamon (1969).

$$
(6)\quad \begin{array}{l} \underset{\sim}{a}\underset{\sim}{c}\underset{\sim}{f}\underset{\sim}{e}\underset{\sim}{b}\underset{\sim}{d}, \ \underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{f}\underset{\sim}{e}, \ \underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{f}\underset{\sim}{b}\underset{\sim}{e}, \\ \underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{f}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{e}, \ \underset{\sim}{a}\underset{\sim}{d}\underset{\sim}{f}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}, \ \underset{\sim}{a}\underset{\sim}{e}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{f}, \\ \underset{\sim}{a}\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{f}, \ \underset{\sim}{a}\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{f}, \ \underset{\sim}{a}\underset{\sim}{e}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{f}, \\ \underset{\sim}{a}\underset{\sim}{e}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{f} \end{array}
$$

We note that the traces of the elements in the above table of typical invariants, listed under (1), involve one tensor only, those listed under (2) involve two tensors, those listed under (3) involve three tensors, and so on. None of them involves more than six tensors.

An irreducible integrity basis for one symmetric second-order tensor, $\underset{\sim}{y}_1$ say, is obtained by taking the traces of the terms (1), with $\underset{\sim}{a}$ replaced by $\underset{\sim}{y}_1$. We thus obtain

$$
\operatorname{tr} \underset{\sim}{y}_1, \quad \operatorname{tr} \underset{\sim}{y}_1^2, \quad \operatorname{tr} \underset{\sim}{y}_1^3, \tag{6.3.2}
$$

in agreement with the result given in (5.5.4).

An irreducible integrity basis for two symmetric second-order tensors, $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$ say, is obtained by taking the traces of the terms (1), with $\underset{\sim}{a}$ replaced successively by $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$ and the traces of the terms (2) with $\underset{\sim}{a}$ and $\underset{\sim}{b}$ replaced by $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$ respectively. The elements obtained from the terms (1) form, of course, irreducible integrity bases for $\underset{\sim}{y}_1$ and for $\underset{\sim}{y}_2$.

An irreducible integrity basis for three symmetric second-order tensors $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, $\underset{\sim}{y}_3$ say, may be analogously obtained by taking the traces of the products listed in (3) with $\underset{\sim}{a},\underset{\sim}{b},\underset{\sim}{c}$ replaced by $\underset{\sim}{y}_1,\underset{\sim}{y}_2,\underset{\sim}{y}_3$ respectively and irreducible integrity bases for the pairs of tensors $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$, $\underset{\sim}{y}_2$ and $\underset{\sim}{y}_3$, $\underset{\sim}{y}_3$ and $\underset{\sim}{y}_1$.

Irreducible integrity bases for four, five and six symmetric second-order tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_6$ may be analogously obtained. Thus for six symmetric second-order tensors an irreducible integrity basis is formed by taking the traces of the products listed in (6) with $\underset{\sim}{a},\underset{\sim}{b},\ldots,\underset{\sim}{f}$ replaced by $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_6$ respectively and irreducible integrity bases for all possible selections of five matrices from the six matrices $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_6$.

An irreducible integrity basis for N(>6) symmetric second-order matrices $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ is formed by the irreducible bases for all possible selections of six matrices from the N matrices $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots\underset{\sim}{v}_N$.

4. Invariant tensor-valued functions

We suppose now that each of the components of a tensor, $\underset{\sim}{t}$ say, in a rectangular cartesian coordinate system x is a function of the components of N tensors $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ in the same coordinate system. Then, in any other coordinate system, the components of $\underset{\sim}{t}$ are also functions (but not necessarily the same functions) of the components of $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$. $\underset{\sim}{t}$ is said to be a tensor-valued (or tensor) function of the tensors $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$. We denote this by

$$\underset{\sim}{t} = \underset{\sim}{F}(\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N) \ . \tag{6.4.1}$$

If $t_{i\ldots j}$, $v^{(1)}_{k\ldots \ell}$, $v^{(2)}_{m\ldots n}$, $\ldots$, $v^{(N)}_{p\ldots q}$ are the components of $\underset{\sim}{t}$, $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ respectively in the system x, then we can write the relation (6.4.1) in the system x as

$$t_{i\ldots j} = F_{i\ldots j}\left(v^{(1)}_{k\ldots \ell},\ v^{(2)}_{m\ldots n},\ \ldots,\ v^{(N)}_{p\ldots q}\right) . \tag{6.4.2}$$

We note that each of the components in the system x of $\underset{\sim}{t}$ is a function of all of the components of $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_N$ in the system x.

If $\bar{x}$ is another rectangular cartesian coordinate system related to x by the transformation $\underset{\sim}{s}$, and $\bar{t}^{(1)}_{i\ldots j}$, $\bar{v}^{(1)}_{k\ldots \ell}$, $\ldots$ are the components of $\underset{\sim}{t},\underset{\sim}{v}_1,\ldots$ in this system, then the relation (6.4.1) may be expressed in the system $\bar{x}$ as

$$\bar{t}_{i\ldots j} = \bar{F}_{i\ldots j}\left(\bar{v}^{(1)}_{k\ldots \ell},\ \bar{v}^{(2)}_{m\ldots n},\ \ldots,\ \bar{v}^{(N)}_{p\ldots q}\right) . \tag{6.4.3}$$

In general the functions $F_{i\ldots j}$ in (6.4.2) and $\bar{F}_{i\ldots j}$ in (6.4.3) are different. However, if they are the same, then $\underset{\sim}{t}$ is said to be an invariant tensor-valued function of $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_N$ under the transformation $\underset{\sim}{s}$ and if they are the same for a group of transformations $\{\underset{\sim}{s}\}$, then $\underset{\sim}{t}$ is said to be an invariant tensor-valued function of $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_N$ under the group $\{\underset{\sim}{s}\}$. In this case (6.4.3) becomes

$$\bar{t}_{i\ldots j} = F_{i\ldots j}\left(\bar{v}^{(1)}_{k\ldots \ell}, \bar{v}^{(2)}_{m\ldots n}, \ldots, \bar{v}^{(N)}_{p\ldots q}\right), \tag{6.4.4}$$

and this is valid for all choices of $\bar{x}$ related to x by a transformation of the group $\{\underset{\sim}{s}\}$. We bear in mind that the components of the tensors in the systems $\bar{x}$ and x are related by the equations

$$\begin{aligned} \bar{t}_{i\ldots j} &= s_{ir}\ldots s_{js} t_{r\ldots s}\,, \\ \bar{v}^{(1)}_{k\ldots \ell} &= s_{ku}\ldots s_{\ell v} v^{(1)}_{u\ldots v}\,, \text{ etc.} \end{aligned} \tag{6.4.5}$$

From (6.4.2), (6.4.4) and $(6.4.5)_1$, we obtain

$$F_{i\ldots j}\left(\bar{v}^{(1)}_{k\ldots \ell}, \bar{v}^{(2)}_{m\ldots n}, \ldots, \bar{v}^{(N)}_{p\ldots q}\right) = s_{ir}\ldots s_{js} F_{r\ldots s}\left(v^{(1)}_{k\ldots \ell}, v^{(2)}_{m\ldots n}, \ldots, v^{(N)}_{p\ldots q}\right). \tag{6.4.}$$

This relation is, of course, valid for all $\underset{\sim}{s}$ belonging to the group $\{\underset{\sim}{s}\}$. It expresses, in implicit form, the restrictions on the tensor-valued function $F_{i\ldots j}$ which result from its invariance under the group $\{\underset{\sim}{s}\}$. Various procedures are available for obtaining these restrictions in explicit form.

Let $\underset{\sim}{\psi}$ be an arbitrary tensor of the same order as $\underset{\sim}{t}$. Its components $\psi_{i\ldots j}$ in the system x and $\bar{\psi}_{i\ldots j}$ in the system $\bar{x}$ are related by

$$\psi_{r\ldots s} = s_{ir}\cdots s_{js}\bar{\psi}_{i\ldots j} \,. \tag{6.4.7}$$

Multiplying equation (6.4.6) throughout by $\bar{\psi}_{i\ldots j}$ and using (6.4.7), we obtain

$$\begin{aligned}\bar{\psi}_{i\ldots j}\, F_{i\ldots j}\left(\bar{v}^{(1)}_{k\ldots \ell},\ \bar{v}^{(2)}_{m\ldots n},\ \ldots,\ \bar{v}^{(N)}_{p\ldots q}\right)&\\ = \psi_{i\ldots j}\, F_{i\ldots j}\left(v^{(1)}_{k\ldots \ell},\ v^{(2)}_{m\ldots n},\ \ldots,\ v^{(N)}_{p\ldots q}\right)&\\ = F \text{ (say)}.&\end{aligned} \tag{6.4.8}$$

We see that by introducing the arbitrary tensor $\underset{\sim}{\psi}$, we have changed the tensor equation (6.4.6) into a scalar equation (6.4.8). This equation expresses the fact that F is an invariant scalar function under the group $\{\underset{\sim}{s}\}$ of the tensors $\underset{\sim}{\psi},\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$, linear in $\underset{\sim}{\psi}$. For the moment, we will consider that $F_{i\ldots j}$ is a polynomial function of its arguments. Then, F must be expressible as a polynomial in the elements of an irreducible integrity basis for the tensors $\underset{\sim}{\psi},\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ under the group $\{\underset{\sim}{s}\}$. Furthermore, since F is linear in $\underset{\sim}{\psi}$, it must, when so expressed, be linear in those elements of the integrity basis which are themselves linear in $\underset{\sim}{\psi}$. Let $K_1,K_2,\ldots,K_\lambda$ be the elements of the irreducible integrity basis which are linear in $\underset{\sim}{\psi}$ and let $I_1,I_2,\ldots,I_\nu$ be the

elements which are independent of $\underset{\sim}{\psi}$. (The latter are, of course, the elements of an irreducible integrity basis for $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_N$ only.) Then

$$F = \sum_{\alpha=1}^{\lambda} A_\alpha K_\alpha \ , \tag{6.4.9}$$

where the A's are polynomials in $I_1, \ldots, I_\nu$ only.

Using (6.4.8), we can remove the tensor $\underset{\sim}{\psi}$, which is really extraneous to our problem, by differentiation, thus*:

$$F_{i\ldots j} = \frac{\partial F}{\partial \psi_{i\ldots j}} = \sum_{\alpha=1}^{\lambda} A_\alpha \frac{\partial K_\alpha}{\partial \psi_{i\ldots j}} \ . \tag{6.4.10}$$

So far we have assumed, in deriving the result (6.4.10), that $F_{i\ldots j}$ is a polynomial function of its arguments. It follows from a result of Pipkin and Wineman** that if $F_{i\ldots j}$ is a single-valued function of its arguments, the same result applies but now the A's in (6.4.9) and (6.4.10) are single-valued, rather than polynomial functions of the I's.

---

*The device used here of converting a tensor equation into a scalar equation and then using the integrity basis to obtain a canonical form appears to have been first used, at any rate in connection with the formulation of constitutive equations by G. F. Smith and R. S. Rivlin (Arch. Rat'l. Mech. Anal. 1, 107 (1957)). The same device was used in a more general context by A. C. Pipkin and R. S. Rivlin (Arch. Rat'l. Mech. Anal. 4, 129 (1959)).

**A. C. Pipkin and A. S. Wineman, Arch. Rat'l. Mech. Anal. 12, 420 (1963).

As an example, we take the case when the tensors $\underset{\sim}{t},\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ are symmetric second-order tensors. Then, $\underset{\sim}{\psi}$ is also a symmetric second-order tensor. In equation (6.4.9), the K's are now the elements in the irreducible integrity basis for N+1 symmetric second-order tensors $\underset{\sim}{\psi},\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$ which are linear in $\underset{\sim}{\psi}$ and the A's are functions of those elements of this integrity basis which are independent of $\underset{\sim}{\psi}$. It is then a simple matter to obtain a canonical form for $t_{ij}(=F_{ij})$ from the relation (6.4.10), which now becomes

$$F_{ij} = \frac{\partial F}{\partial \psi_{ij}} = \sum_{\alpha=1}^{\lambda} A_\alpha \frac{\partial K_\alpha}{\partial \psi_{ij}} . \qquad (6.4.11)$$

However, as a result of the symmetry of the tensors, the expression obtained will be formally ambiguous. It will, for example, not be symmetric for interchange of the subscripts i and j unless we write the expression F_{ij} in a form which is symmetric for such interchange. (This can always be done by making use of the symmetry of the tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_N$.) We can ensure that we obtain an expression for t_{ij} which is symmetric under interchange of i and j, by using the expression

$$F_{ij} = \frac{1}{2}\left(\frac{\partial F}{\partial \psi_{ij}} + \frac{\partial F}{\partial \psi_{ji}}\right) = \frac{1}{2}\sum_{\alpha=1}^{\lambda} A_\alpha \left(\frac{\partial K_\alpha}{\partial \psi_{ij}} + \frac{\partial K_\alpha}{\partial \psi_{ji}}\right) \qquad (6.4.12)$$

instead of (6.4.11).

As a special case, we assume that the symmetric second-order tensor $\underset{\sim}{t}$ depends on only two symmetric second-order tensors, $\underset{\sim}{v}_1$ and $\underset{\sim}{v}_2$ say, and the group $\{\underset{\sim}{s}\}$ is the full or proper orthogonal group. We can read off from the table of typical invariants (6.3.1) the elements of an irreducible integrity basis for $\underset{\sim}{v}_1$ and $\underset{\sim}{v}_2$. These are:

$$\begin{aligned} &\operatorname{tr} \underset{\sim}{v}_1,\ \operatorname{tr} \underset{\sim}{v}_1^2,\ \operatorname{tr} \underset{\sim}{v}_1^3, \\ &\operatorname{tr} \underset{\sim}{v}_2,\ \operatorname{tr} \underset{\sim}{v}_2^2,\ \operatorname{tr} \underset{\sim}{v}_2^3, \\ &\operatorname{tr} \underset{\sim}{v}_1\underset{\sim}{v}_2,\ \operatorname{tr} \underset{\sim}{v}_1^2\underset{\sim}{v}_2,\ \operatorname{tr} \underset{\sim}{v}_1\underset{\sim}{v}_2^2,\ \operatorname{tr} \underset{\sim}{v}_1^2\underset{\sim}{v}_2^2. \end{aligned} \tag{6.4.13}$$

Again, we can read off from the same table the elements of an irreducible integrity basis for $\underset{\sim}{v}_1, \underset{\sim}{v}_2$ and $\underset{\sim}{\psi}$, where $\underset{\sim}{\psi}$ is a symmetric second-order tensor, which are linear in $\underset{\sim}{\psi}$. These are:

$$\begin{aligned} &\operatorname{tr} \underset{\sim}{\psi} \\ &\operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_1,\ \operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_1^2, \\ &\operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_2,\ \operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_2^2, \\ &\operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_1\underset{\sim}{v}_2,\ \operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_1\underset{\sim}{v}_2^2,\ \operatorname{tr} \underset{\sim}{v}_2\underset{\sim}{\psi}\underset{\sim}{v}_1^2,\ \operatorname{tr} \underset{\sim}{\psi}\underset{\sim}{v}_1^2\underset{\sim}{v}_2^2. \end{aligned} \tag{6.4.14}$$

The element in the first line of (6.4.14) is obtained by taking $\underset{\sim}{a} = \underset{\sim}{\psi}$ in the elements (1) of (6.3.1) which are linear in $\underset{\sim}{a}$. The elements in the second line of (6.4.14) are obtained by taking $\underset{\sim}{a} = \underset{\sim}{\psi}$ and $\underset{\sim}{b} = \underset{\sim}{v}_1$ in the elements (2) in (6.3.1) which are linear in $\underset{\sim}{a}$. Again, the elements in the third line of (6.4.14) are obtained by taking $\underset{\sim}{a} = \underset{\sim}{\psi}$ and $\underset{\sim}{b} = \underset{\sim}{v}_2$ in the same elements. Finally,

the elements in the last line of (6.4.14) are obtained by taking $\underset{\sim}{a} = \underset{\sim}{\psi}$, $\underset{\sim}{b} = \underset{\sim}{v}_1$, $\underset{\sim}{c} = \underset{\sim}{v}_2$ in the elements (3) in (6.4.14) which are linear in $\underset{\sim}{a}$. The elements (6.4.14) are the elements $K_\alpha(\alpha=1,\ldots,\lambda)$ in (6.4.12). Introducing them and carrying out the differentiations, we obtain*

$$\underset{\sim}{t} = \underset{\sim}{F} = \|F_{ij}\| = A_1\underset{\sim}{I} + A_2\underset{\sim}{v}_1 + A_3\underset{\sim}{v}_1^2 + A_4\underset{\sim}{v}_2 + A_5\underset{\sim}{v}_2^2 + \frac{1}{2}\{A_6(\underset{\sim}{v}_1\underset{\sim}{v}_2+\underset{\sim}{v}_2\underset{\sim}{v}_1) + A_7(\underset{\sim}{v}_1\underset{\sim}{v}_2^2+\underset{\sim}{v}_2^2\underset{\sim}{v}_1) + A_8(\underset{\sim}{v}_2\underset{\sim}{v}_1^2+\underset{\sim}{v}_1^2\underset{\sim}{v}_2) + A_9(\underset{\sim}{v}_1^2\underset{\sim}{v}_2^2+\underset{\sim}{v}_2^2\underset{\sim}{v}_1^2)\}, \tag{6.4.15}$$

where the α's are single-valued functions of the invariants (6.4.14). If we take as our initial assumption that F_{ij} is a polynomial function of its argument tensors, then the A's in (6.4.15) are polynomial functions of the invariants (6.4.13).

In the particular case when $\underset{\sim}{t}$ depends only on one symmetric second-order tensor, $\underset{\sim}{v}_1$ say, we obtain, by taking $\underset{\sim}{v}_2 = 0$ in (6.4.15) and (6.4.13),

$$\underset{\sim}{t} = \underset{\sim}{F} = A_1\underset{\sim}{I} + A_2\underset{\sim}{v}_1 + A_3\underset{\sim}{v}_1^2, \tag{6.4.16}$$

where A_1, A_2 and A_3 are single-valued functions of tr $\underset{\sim}{v}_1$, tr $\underset{\sim}{v}_1^2$ and tr $\underset{\sim}{v}_1^3$.

More generally, if $\underset{\sim}{t}$ depends on the μ tensors $\underset{\sim}{v}_1,\ldots,\underset{\sim}{v}_\mu$, then the appropriate expressions for the K's in (6.4.12) are the elements in the integrity basis under the

---

*R. S. Rivlin, J. Rat'l. Mech. Anal. 4, 681 (1955).

orthogonal group for the $\mu+1$ symmetric second-order tensors $\underset{\sim}{\psi}, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_\mu$ which are linear in $\underset{\sim}{\psi}$. These are obtained from (6.3.1) in the following manner. The tensor $\underset{\sim}{a}$ is replaced by $\underset{\sim}{\psi}$. Then, in the elements (2), $\underset{\sim}{b}$ is replaced by $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$ in turn; in the elements (3), $\underset{\sim}{b}$, $\underset{\sim}{c}$ are replaced in turn by every selection of two tensors which can be made from $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$; in the elements (4), $\underset{\sim}{b}$, $\underset{\sim}{c}$, $\underset{\sim}{d}$ are replaced in turn by every selection of three tensors which can be made from $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$, and so on. The traces of the products so obtained, which are linear in $\underset{\sim}{\psi}$, are the K's in (6.4.12). It is evident that the terms $\partial K_\alpha / \partial \psi_{ij}$ in (6.4.12) may be obtained by differentiating with respect to $\underset{\sim}{a}$ the traces of the products in (6.3.1) which are linear in $\underset{\sim}{a}$ and replacing $\underset{\sim}{b}, \underset{\sim}{c}, \ldots, \underset{\sim}{f}$ in the products so obtained by selections from $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_\mu$. The table of products obtained is

(1) $\underset{\sim}{I}$

(2) $\underset{\sim}{b}$, $\underset{\sim}{b}^2$

(3) $\underset{\sim}{b}\underset{\sim}{c}$, $\underset{\sim}{b}\underset{\sim}{c}^2$, $\underset{\sim}{b}^2\underset{\sim}{c}$, $\underset{\sim}{b}^2\underset{\sim}{c}^2$

(4) $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}$, $\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}$, $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2$, $\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}^2$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{d}$, $\underset{\sim}{c}^2\underset{\sim}{d}\underset{\sim}{b}$, $\underset{\sim}{b}^2\underset{\sim}{c}\underset{\sim}{d}$, $\underset{\sim}{d}\underset{\sim}{b}^2\underset{\sim}{c}$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{d}^2$, $\underset{\sim}{c}\underset{\sim}{b}^2\underset{\sim}{d}^2$, $\underset{\sim}{d}\underset{\sim}{b}^2\underset{\sim}{c}^2$, $\underset{\sim}{b}^2\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{d}\underset{\sim}{c}$, $\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2$ (6.4.17)

(5) $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}$, $\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{e}\underset{\sim}{c}$, $\underset{\sim}{b}\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{d}$, $\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{e}$, $\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}\underset{\sim}{d}$, $\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{e}$, $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}^2$, $\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{e}^2$, $\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}^2$, $\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{e}^2\underset{\sim}{b}$, $\underset{\sim}{b}^2\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{e}$, $\underset{\sim}{e}\underset{\sim}{b}^2\underset{\sim}{c}\underset{\sim}{d}$, $\underset{\sim}{e}\underset{\sim}{d}\underset{\sim}{b}^2\underset{\sim}{c}$, $\underset{\sim}{b}^2\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{e}$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{d}\underset{\sim}{e}$, $\underset{\sim}{c}^2\underset{\sim}{d}\underset{\sim}{e}\underset{\sim}{b}$, $\underset{\sim}{e}\underset{\sim}{c}^2\underset{\sim}{d}\underset{\sim}{b}$, $\underset{\sim}{b}\underset{\sim}{c}^2\underset{\sim}{e}\underset{\sim}{d}$, $\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2\underset{\sim}{e}$, $\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}^2\underset{\sim}{e}$, $\underset{\sim}{d}^2\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{b}$, $\underset{\sim}{e}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{d}^2$

$$
\begin{array}{l}
\underset{\sim}{c}\underset{\sim}{f}\underset{\sim}{e}\underset{\sim}{b}\underset{\sim}{d}, \ \underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{f}\underset{\sim}{e}, \ \underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{f}\underset{\sim}{b}\underset{\sim}{e}, \\
\underset{\sim}{d}\underset{\sim}{f}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{e}, \ \underset{\sim}{d}\underset{\sim}{f}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{e}, \ \underset{\sim}{e}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{f}, \\
\underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{d}\underset{\sim}{f}, \ \underset{\sim}{e}\underset{\sim}{c}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{f}, \ \underset{\sim}{e}\underset{\sim}{d}\underset{\sim}{b}\underset{\sim}{c}\underset{\sim}{f}, \\
\underset{\sim}{e}\underset{\sim}{d}\underset{\sim}{c}\underset{\sim}{b}\underset{\sim}{f}
\end{array}
\qquad (6)
$$

Equation (6.4.12) may, of course, be rewritten as

$$
\underset{\sim}{F} = \frac{1}{2}\sum_{\alpha=1}^{\lambda} A_\alpha(\underset{\sim}{\pi}_\alpha + \underset{\sim}{\pi}_\alpha^t) \ , \qquad (6.4.18)
$$

where

$$
\underset{\sim}{\pi}_\alpha = \left\| \partial K_\alpha / \partial \psi_{ij} \right\|
$$

and $\underset{\sim}{\pi}_\alpha^t$ is the transpose of $\underset{\sim}{\pi}_\alpha$. Then the products $\underset{\sim}{\pi}_\alpha$ may be obtained from the terms (6.4.17) in the following way. We first take the term in line (1). Then, in the terms (2), $\underset{\sim}{b}$ is replaced by $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$ in turn; in the terms (3), $\underset{\sim}{b}$, $\underset{\sim}{c}$ are replaced in turn by every selection of two tensors which can be made from $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$; in the terms (4), $\underset{\sim}{b}$, $\underset{\sim}{c}$, $\underset{\sim}{d}$ are replaced in turn by every selection of three tensors which can be made from $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$, and so on. We note that since each of the tensors $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_\mu$ is symmetric, $\underset{\sim}{\pi}_\alpha^t$ may be obtained from $\underset{\sim}{\pi}_\alpha$ by writing its factors in reverse order, e.g. if $\underset{\sim}{\pi}_\alpha = \underset{\sim}{v}_1\underset{\sim}{v}_2\underset{\sim}{v}_3$, it is easily seen that $\underset{\sim}{\pi}_\alpha^t = \underset{\sim}{v}_3\underset{\sim}{v}_2\underset{\sim}{v}_1$.

The α's in (6.4.18) are functions of the elements of an irreducible integrity basis under the orthogonal group for the μ tensors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_\mu$. These are, of course, obtained from (6.3.1) in the manner described in §6.3.

5. Invariants of functionals

Let $\underset{\sim}{y}(\tau)$ be a tensor-valued function of a variable τ. Let x and $\bar{x}$ be two fixed rectangular cartesian coordinate systems related by the transformation $\underset{\sim}{s}$, thus:

$$\bar{\underset{\sim}{x}} = \underset{\sim}{\underset{\sim}{s}}\underset{\sim}{x} \ . \tag{6.5.1}$$

The components of $\underset{\sim}{y}(\tau)$ in these systems are denoted $v_{i\ldots j}(\tau)$ and $\bar{v}_{i\ldots j}(\tau)$ respectively and are related by

$$\bar{v}_{i\ldots j}(\tau) = s_{ip}\ldots s_{jq}v_{p\ldots q}(\tau) \ . \tag{6.5.2}$$

Let y now be a scalar <u>functional</u> of the components in the system x of the tensor-valued function $\underset{\sim}{y}(\tau)$ over the range $\tau = \tau', \tau''$ thus:

$$y = \overset{\tau''}{\underset{\tau'}{F}}[v_{i\ldots j}(\tau)] \ . \tag{6.5.3}$$

Let $\bar{y}$ be the <u>same</u> functional of the components of $\underset{\sim}{y}(\tau)$ in the system $\bar{x}$. Then,

$$\bar{y} = \overset{\tau''}{\underset{\tau'}{F}}[\bar{v}_{i\ldots j}(\tau)] \ . \tag{6.5.4}$$

If y and $\bar{y}$, defined by (6.5.3) and (6.5.4), are equal for all transformations $\underset{\sim}{s}$ of a group $\{\underset{\sim}{s}\}$, then y is said to be a scalar functional of $\underset{\sim}{y}(\tau)$ over $[\tau', \tau'']$, invariant under the group $\{\underset{\sim}{s}\}$. We may write

$$y = \overset{\tau''}{\underset{\tau'}{F}}[\underset{\sim}{y}(\tau)] \ . \tag{6.5.5}$$

In a loose sense we may regard y as a scalar invariant function of an infinite number of tensors--the values of $\underset{\sim}{y}(\tau)$ for all values of τ between τ' and τ''.

It follows from a development due primarily to Green and Rivlin, and Wineman and Pipkin that y may be expressed in canonical form in the following manner*.

We form the table of typical invariants for the integrity basis for the N tensors $\underset{\sim}{y}(\tau_1),\underset{\sim}{y}(\tau_2),\ldots,\underset{\sim}{y}(\tau_N)$ under the group $\{\underset{\sim}{s}\}$. We write this down in terms of as many of the tensors $\underset{\sim}{y}(\tau_1),\underset{\sim}{y}(\tau_2),\ldots$ as may be required. We denote the elements of the table so obtained by $I_p(P = 1,2,\ldots)$. We note that some of the elements I_p involve $\underset{\sim}{y}(\tau_1)$ only, others involve $\underset{\sim}{y}(\tau_1)$ and $\underset{\sim}{y}(\tau_2)$ only, yet others involve $\underset{\sim}{y}(\tau_1)$, $\underset{\sim}{y}(\tau_2)$ and $\underset{\sim}{y}(\tau_3)$ only, and so on. $I_p(P = 1,2,\ldots)$ are called the <u>basic invariants</u> of the history $\underset{\sim}{y}(\tau)$ under the group $\{\underset{\sim}{s}\}$. Then, y may be expressed as a function of linear functionals of I_1, of linear functionals of I_2, and so on.

<u>Example 1</u>. Suppose $\underset{\sim}{y}(\tau)$ is a polar vector and the group $\{\underset{\sim}{s}\}$ is the full orthogonal group. Then, the basic invariants of the history $\underset{\sim}{y}(\tau)$ are (cf. §6.2)

$$\underset{\sim}{y}(\tau_1).\underset{\sim}{y}(\tau_1) \ , \qquad \underset{\sim}{y}(\tau_1).\underset{\sim}{y}(\tau_2) \ . \tag{6.5.6}$$

*A. E. Green and R. S. Rivlin; Arch Rat'l. Mech. Anal. 1, 1 1957); A. S. Wineman and A. C. Pipkin, ibid. 17, 184 (1964). (See also, A. E. Green, R. S. Rivlin and A. J. M. Spencer, Arch. Rat'l. Mech. Anal. 3, 82 (1959); A. E. Green and R. S. Rivlin, ibid. 4, 387 (1960); R. S. Rivlin, ibid. 4, 262 (1960))

y is then a function of linear functionals of $\underset{\sim}{v}(\tau_1).\underset{\sim}{v}(\tau_1)$ and of linear functionals of $\underset{\sim}{v}(\tau_1).\underset{\sim}{v}(\tau_2)$. We note that $\underset{\sim}{v}(\tau_1).\underset{\sim}{v}(\tau_1)$ is itself a linear functional of $\underset{\sim}{v}(\tau_1).\underset{\sim}{v}(\tau_2)$. It follows that y must be expressible as a function of linear functionals of $\underset{\sim}{v}(\tau_1).\underset{\sim}{v}(\tau_2)$.

Example 2. Suppose that $\underset{\sim}{v}(\tau)$ is a symmetric second-order tensor and $\{\underset{\sim}{s}\}$ is again the full orthogonal group. The required table of typical invariants is obtained from (6.3.1) by replacing $\underset{\sim}{a},\underset{\sim}{b},\ldots,\underset{\sim}{f}$ by $\underset{\sim}{v}(\tau_1),\underset{\sim}{v}(\tau_2),\ldots,\underset{\sim}{v}(\tau_6)$ and taking the traces of the products so obtained. We note that any of the elements so obtained may be expressed as a linear functional of one of the multilinear elements. For example, $\mathrm{tr}\{\underset{\sim}{v}(\tau_1)\}^2$ is a linear functional of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)$ and $\mathrm{tr}\left[\underset{\sim}{v}(\tau_1)\{\underset{\sim}{v}(\tau_2)\}^2\right]$ is a linear functional of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\underset{\sim}{v}(\tau_3)$. It follows that y may be expressed as a function of functionals of each of the elements in the table discussed which is multilinear in one or more of the tensors $\underset{\sim}{v}(\tau_1),\underset{\sim}{v}(\tau_2),\ldots,\underset{\sim}{v}(\tau_6)$. Also, we note that if $\underset{\sim}{\pi}$ is a multilinear product in the first R of these tensors, then any functional of $\mathrm{tr}\ \underset{\sim}{\pi}$ may be expressed as a functional of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\ldots\underset{\sim}{v}(\tau_R)$. It follows that y may be expressed as a function of functionals of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)$, functionals of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)$, functionals of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\underset{\sim}{v}(\tau_3)$, ..., and functionals of $\mathrm{tr}\ \underset{\sim}{v}(\tau_1)\ldots\underset{\sim}{v}(\tau_6)$.

6. Invariant tensor-valued functionals*

We now suppose that each of the components of a tensor $\underset{\sim}{t}$, say, in a rectangular cartesian coordinate system x is a functional of the components in the system x of the tensor-valued function $\underset{\sim}{v}(\tau)$ over the range $\tau = \tau', \tau''$, thus

$$t_{i\ldots j} = \overset{\tau''}{\underset{\tau'}{F}}_{i\ldots j}[v_{p\ldots q}(\tau)] \ . \tag{6.6.1}$$

If $\bar{x}$ is another rectangular cartesian coordinate system related to x by the transformation $\underset{\sim}{s}$ and $\bar{t}_{i\ldots j}$ and $\bar{v}_{p\ldots q}(\tau)$ are the components of $\underset{\sim}{t}$ and $\underset{\sim}{v}(\tau)$ in the system $\bar{x}$, then

$$\bar{t}_{i\ldots j} = \overset{\tau''}{\underset{\tau'}{\bar{F}}}_{i\ldots j}[\bar{v}_{p\ldots q}(\tau)] \ , \tag{6.6.2}$$

where, in general, the functionals $\bar{F}_{i\ldots j}$ and $F_{i\ldots j}$ are different. However, if they are the same, then $\underset{\sim}{t}$ is said to be an invariant tensor-valued functional of the tensor-valued function $\underset{\sim}{v}(\tau)$ under the transformation $\underset{\sim}{s}$. In this case,

$$\bar{t}_{i\ldots j} = \overset{\tau''}{\underset{\tau'}{F}}_{i\ldots j}[\bar{v}_{p\ldots q}(\tau)] \ , \tag{6.6.3}$$

where

$$\bar{v}_{p\ldots q}(\tau) = s_{pm}\cdots s_{qn} v_{m\ldots n}(\tau)$$

and

$$\bar{t}_{i\ldots j} = s_{ik}\cdots s_{j\ell} t_{k\ldots \ell} \ . \tag{6.6.4}$$

---

*See references given in previous section.

If (6.6.3) holds for all transformations of a group $\{\underset{\sim}{s}\}$, then $\underset{\sim}{t}$ is said to be an invariant tensor-valued functional under the group $\{\underset{\sim}{s}\}$. From (6.6.1), (6.6.3) and $(6.6.4)_2$, we obtain

$$\overset{\tau''}{\underset{\tau'}{F}}_{i\ldots j}[\bar{v}_{p\ldots q}(\tau)] = s_{im}\ldots s_{jn}\,\overset{\tau''}{\underset{\tau'}{F}}_{m\ldots n}[v_{p\ldots q}(\tau)]\,. \tag{6.6.5}$$

This relation is, of course, valid for all $\underset{\sim}{s}$ belonging to the group $\{\underset{\sim}{s}\}$. It expresses, in implicit form, the restrictions on the tensor-valued functional $F_{i\ldots j}$ which result from its invariance under the group $\{\underset{\sim}{s}\}$. These restrictions may be put in explicit form in a manner similar to that adopted in §6.4 in discussing the corresponding problem for invariant tensor functions.

We introduce an arbitrary tensor $\underset{\sim}{\psi}$ of the same order as $\underset{\sim}{t}$, which has components $\psi_{i\ldots j}$ and $\bar{\psi}_{i\ldots j}$ respectively in the systems x and $\bar{x}$. Then,

$$\psi_{m\ldots n} = s_{im}\ldots s_{jn}\bar{\psi}_{i\ldots j}\,. \tag{6.6.6}$$

Multiplying (6.6.5) throughout by $\bar{\psi}_{i\ldots j}$, we obtain with (6.6.6)

$$\begin{aligned}\bar{\psi}_{i\ldots j}\,\overset{\tau''}{\underset{\tau'}{F}}_{i\ldots j}[\bar{v}_{p\ldots q}(\tau)] &= \psi_{i\ldots j}\,\overset{\tau''}{\underset{\tau'}{F}}_{i\ldots j}[v_{p\ldots q}(\tau)] \\ &= F, \text{ say.}\end{aligned} \tag{6.6.7}$$

From (6.6.7), we have

$$F_{i\ldots j}[v_{p\ldots q}(\tau)] = \frac{\partial F}{\partial \psi_{i\ldots j}}\,. \tag{6.6.8}$$

Equation (6.6.7) expresses the fact that the functional F is a scalar invariant functional under the group $\{\underset{\sim}{s}\}$ of the tensor-valued function $\underset{\sim}{v}(\tau)$ and the tensor $\underset{\sim}{\psi}$, its dependence on the latter being linear. The manner in which F can be expressed in canonical form follows from the development of Green and Rivlin (1957) and Wineman and Pipkin (1964).

We form the table of typical invariants for a tensor $\underset{\sim}{\psi}$ and an arbitrary number of tensors $\underset{\sim}{v}(\tau_1),\underset{\sim}{v}(\tau_2),\ldots,\underset{\sim}{v}(\tau_N)$ under the group $\{\underset{\sim}{s}\}$. We write this down in terms of $\underset{\sim}{\psi}$ and as many of the tensors $\underset{\sim}{v}(\tau_1),v(\tau_2),\ldots$ as may be required. We denote the elements of the table which are linear in $\underset{\sim}{\psi}$ by $K_P(P = 1,2,\ldots)$. We note that some of these elements involve $\underset{\sim}{\psi}$ and $\underset{\sim}{v}(\tau_1)$, others involve $\underset{\sim}{\psi}$, $\underset{\sim}{v}(\tau_1)$ and $\underset{\sim}{v}(\tau_2)$, others involve $\underset{\sim}{\psi}$, $\underset{\sim}{v}(\tau_1)$, $\underset{\sim}{v}(\tau_2)$ and $\underset{\sim}{v}(\tau_3)$, and so on. Then, F, defined by (6.6.7), may be expressed as the sum of linear functionals $L^{(1)},L^{(2)},\ldots$ of $K_1,K_2,\ldots$ respectively, each of which is a scalar functional of $\underset{\sim}{v}(\tau)$ invariant under the group $\{\underset{\sim}{s}\}$, thus:

$$F = \sum_{\mu} L^{(\mu)}[K_\mu] \ . \tag{6.6.9}$$

Example 1. Suppose $\underset{\sim}{v}(\tau)$ and $\underset{\sim}{\psi}$ are polar vectors and the group $\{\underset{\sim}{s}\}$ is the full orthogonal group. In the table of typical invariants (6.2.3) for an arbitrary number of polar vectors under the full orthogonal group, we take $\underset{\sim}{a} = \underset{\sim}{\psi}$, $\underset{\sim}{b} = \underset{\sim}{v}(\tau)$.

Then, there is only one element linear in $\underset{\sim}{\psi}$, i.e., there is only one K in (6.6.9). This is $\underset{\sim}{\psi}.\underset{\sim}{v}(\tau)$.
It follows that

$$F = L[\underset{\sim}{\psi}.\underset{\sim}{v}(\tau)] \ , \tag{6.6.10}$$

where L is a linear scalar functional of its indicated argument function and, from example 1 in §6.5, a scalar functional of $\underset{\sim}{v}(\tau_1)$, $\underset{\sim}{v}(\tau_2)$. Introducing (6.6.10) into (6.6.8), we obtain

$$F_i = \frac{\partial F}{\partial \psi_i} = L_i[\underset{\sim}{v}(\tau)], \text{ say} \ . \tag{6.6.11}$$

Example 2. Suppose that $\underset{\sim}{v}(\tau)$ and $\underset{\sim}{\psi}$ are symmetric second-order tensors and the group $\{\underset{\sim}{s}\}$ is the full orthogonal group. The table of typical invariants for an arbitrary number of symmetric second-order tensors $\underset{\sim}{\psi},\underset{\sim}{v}(\tau_1),\underset{\sim}{v}(\tau_2),\ldots,\underset{\sim}{v}(\tau_N)$ under the full orthogonal group is obtained by taking the traces of the products listed in (6.3.1). In this table we take $\underset{\sim}{a} = \underset{\sim}{\psi}$, $\underset{\sim}{b} = \underset{\sim}{v}(\tau_1)$, $\underset{\sim}{c} = \underset{\sim}{v}(\tau_2),\ldots,$ $\underset{\sim}{e} = v(\tau_5)$. We write down the elements in this table of typical invariants which are linear in $\underset{\sim}{\psi}$, omitting from the list certain elements, each of which has the property that it can be expressed as a linear functional of some element retained in the list. For example,

$\mathrm{tr}\ \underset{\sim}{\psi}\{\underset{\sim}{v}(\tau_1)\}^2$ is a linear functional of $\mathrm{tr}\ \underset{\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)$ and

$\mathrm{tr}\ \underset{\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_3)\underset{\sim}{v}(\tau_2)$ is equal to $\mathrm{tr}\ \underset{\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\underset{\sim}{v}(\tau_3)$, so that

any linear functional of tr $\underset{\sim}{\psi}\{\underset{\sim}{v}(\tau_1)\}^2$ can be expressed as a linear functional of tr $\underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)$ and any linear functional of tr $\underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_3)\underset{\sim}{v}(\tau_2)$ can be expressed as a linear functional of tr $\underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\underset{\sim}{v}(\tau_3)$. It then follows, using (6.3.1) in the way described, that the K's in (6.6.9) are, for the present case,

$$\text{tr } \underset{\sim}{\psi},\ \text{tr } \underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1),\ \text{tr } \underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\ , \qquad (6.6.12)$$

$$\ldots,\ \text{tr } \underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)\ldots\underset{\sim}{v}(\tau_5)\ .$$

Introducing these into (6.6.9), we obtain the canonical form for F as

$$F = L^{(0)}[\text{tr } \underset{\sim}{\psi}] + \sum_{\mu=1}^{5} L^{(\mu)}[\text{tr } \underset{\sim\sim}{\psi}\underset{\sim}{v}(\tau_1)\ldots\underset{\sim}{v}(\tau_\mu)], \qquad (6.6.13)$$

where the L's are linear scalar functionals of the indicated argument functions and scalar functionals of $\underset{\sim}{v}(\tau)$ invariant under the full orthogonal group. We note that we may, without loss of generality, take $L^{(0)}[\text{tr } \psi] = \alpha_0 \text{ tr } \underset{\sim}{\psi}$, where α_0 is a scalar functional of $\underset{\sim}{v}(\tau)$ invariant under the full orthogonal group. Also, from example 2 of §6.5 it follows that any scalar functional of $\underset{\sim}{v}(\tau)$ invariant under the full orthogonal group may be expressed as a scalar function of linear scalar functionals of tr $\underset{\sim}{v}(\tau_1)$, linear scalar functionals of tr $\underset{\sim}{v}(\tau_1)\underset{\sim}{v}(\tau_2)$,...,

and linear scalar functionals of tr $\underset{\sim}{v}(\tau_1)\ldots\underset{\sim}{v}(\tau_6)$. Hence α_o and the L's in (6.6.13) are of this form, in addition to being linear scalar functionals of the indicated arguments. Introducing (6.6.13) into (6.6.8) we see that

$$F_{ij} = \frac{\partial F}{\partial \psi_{ij}} = \alpha_o \delta_{ij} + \sum_{\mu=1}^{5} L_{ij}^{(\mu)}[\underset{\sim}{v}(\tau)\ldots\underset{\sim}{v}(\tau_\mu)], \text{ (say)} \qquad (6.6.14)$$

where δ_{ij} is the Kronecker delta and $L_{ij}^{(\mu)} = \partial L^{(\mu)}/\partial \psi_{ij}$.

In order to preserve formally, as well as factually, the symmetric character of F_{ij}, we can use, instead of (6.6.8),

$$F_{ij} = \frac{1}{2}\left(\frac{\partial F}{\partial \psi_{ij}} + \frac{\partial F}{\partial \psi_{ji}}\right). \qquad (6.6.15)$$

Introducing (6.6.13) into (6.6.15), we obtain

$$F_{ij}[\underset{\sim}{v}(\tau)] = \alpha_o \delta_{ij} + \sum_{\mu=1}^{5} \frac{1}{2} L_{ij}^{(\mu)}[\underset{\sim}{v}(\tau_1)\ldots\underset{\sim}{v}(\tau_\mu) + \underset{\sim}{v}(\tau_\mu)\ldots\underset{\sim}{v}(\tau_1)]. \qquad (6.6.16)$$

Chapter 7

VISCOELASTIC MATERIALS

1. Introduction

We now pass to a much more general class of materials than the elastic materials discussed in Chapter 5. These are materials for which the stress may depend not only on the values of the deformation gradients at the instant t, say, at which the stress is measured, but also on their values at all times prior to t, from the infinite past onwards. It is of little significance whether we make this assumption with respect to the Cauchy stress, the Kirchoff-Piola stress, or the Piola stress. Since they are related by the formulae (3.7.5), (3.8.3) and (3.8.4), if the assumption is valid for any one of them, it is valid for all of them.

In mathematical language we express our assumption regarding the kinematic variables on which is assumed that the stress may depend by saying that the Piola stress tensor P_{AB}, say, at time t is a tensor functional of the deformation gradients $x_{k,Q}(\tau)$ for all values of τ in the interval $-\infty<\tau\leq t$: i.e.

$$P_{AB} = P_{AB}\left[\overset{t}{\underset{\tau=-\infty}{x_{k,Q}(\tau)}}\right] . \tag{7.1.1}$$

From (3.8.4), (3.8.3) and (7.1.1), it follows that

$$\sigma_{ji} = \frac{1}{|\partial \underset{\sim}{x}/\partial \underset{\sim}{X}|} x_{i,A} x_{j,B} P_{AB}\left[x_{k,Q}^{t}(\tau)\right]_{\tau=-\infty} \tag{7.1.2}$$

and

$$\pi_{Ai} = x_{i,B} P_{AB}\left[x_{k,Q}^{t}(\tau)\right]_{\tau=-\infty} . \tag{7.1.3}$$

For succinctness, we have used σ_{ji}, P_{AB} and $x_{i,A}$ to denote the values of $\sigma_{ji}(\tau)$, $P_{AB}(\tau)$ and $x_{i,A}(\tau)$ at time t and we shall continue to use this notation. (We note, in connection with equation (7.1.2), that $x_{i,A}$ is a functional of $x_{k,Q}(\tau)$ defined over the range $-\infty<\tau\leq t$, since if the argument functions $x_{k,Q}(\tau)$ are given over this range, $x_{i,A}$ is surely uniquely determined, in conformity with the definition of a functional.)

2. Restrictions due to effect of a superposed rotation

(i) Formulation of the problem

In §5.2 we placed restrictions on the manner in which the Helmholtz free-energy for an elastic material can depend on the deformation gradients, which result from the assumption that a rigid rotation superposed on the assumed deformation of the body leaves the Helmholtz free energy unchanged. The physical consideration underlying this assumption was the absence of external fields, e.g. gravitational, electric, etc., which might lead to a change in the Helmholtz free energy when a rigid rotation is superposed on the assumed deformation.

A similar underlying physical consideration leads us to analyze the restrictions which can be placed on the form of the tensor functional P_{AB} in (7.1.1), as a result of the assumption that the superposition on the assumed deformation of an arbitrary time-dependent rigid rotation is associated with a corresponding rigid rotation of the applied force system.

Thus, suppose the body is subjected to a deformation

$$x_i(\tau) = x_i(X_A,\tau) \ , \tag{7.2.1}$$

and suppose that at time τ the Kirchoff-Piola stress associated with this deformation is $\pi_{Bi}(\tau)$. We now consider a new deformation

$$\bar{x}_i(\tau) = \bar{x}_i(X_A,\tau) \ , \tag{7.2.2}$$

which results from the superposition on the deformation (7.2.1) of an arbitrary rigid rotation, so that

$$\bar{x}_i(\tau) = a_{ij}(\tau)x_j(\tau) \ , \tag{7.2.3}$$

where $a_{ij}(\tau)$ is a proper orthogonal transformation and accordingly satisfies a relation of the form

$$a_{ij}(\tau)a_{ik}(\tau) = a_{ji}(\tau)a_{ki}(\tau) = \delta_{jk}, \ |a_{ij}(\tau)| = 1 \ . \tag{7.2.4}$$

From our assumption that the force system associated with the second deformation is rotated from that associated with the first deformation by the proper orthogonal transformation $a_{ij}(\tau)$, it follows that the Kirchoff-Piola stress $\bar{\pi}_{Bi}(\tau)$ associated with deformation (7.2.2) is given by

$$\bar{\pi}_{Bi}(\tau) = a_{ij}(\tau)\pi_{Bj}(\tau) \ . \tag{7.2.5}$$

The Piola stresses $\bar{P}_{AB}(\tau)$ and $P_{AB}(\tau)$ corresponding to the Kirchoff-Piola stresses $\bar{\pi}_{Ai}(\tau)$ and $\pi_{Ai}(\tau)$ respectively are given, from (3.8.1), by

$$\bar{P}_{AB}(\tau) = \frac{\partial X_B}{\partial \bar{x}_i(\tau)} \bar{\pi}_{Ai}(\tau)$$

and (7.2.6)

$$P_{AB}(\tau) = \frac{\partial X_B}{\partial x_i(\tau)} \pi_{Ai}(\tau) \ .$$

It follows from (7.2.3), (7.2.5) and (7.2.6) that

$$\bar{P}_{AB}(\tau) = P_{AB}(\tau) \ ; \tag{7.2.7}$$

the Piola stresses associated with the deformations (7.2.1) and (7.2.2) are equal. In accordance with our convention (cf. §7.1) we use the notation $\bar{P}_{AB} = \bar{P}_{AB}(t)$ and $P_{AB} = P_{AB}(t)$. Then, from (7.1.1),

$$P_{AB} = P_{AB}[\bar{x}_{k,Q}(\tau)] \ ,$$

and

$$P_{AB} = P_{AB}[x_{k,Q}(\tau)] \ . \tag{7.2.8}$$

We note that the form of the functional dependence on the argument functions is the same in both cases. Introducing (7.2.8) into (7.2.7), and using (7.2.3), we obtain

$$P_{AB}[a_{k\ell}(\tau)x_{\ell,Q}(\tau)] = P_{AB}[x_{k,Q}(\tau)] \ , \tag{7.2.9}$$

and this relation must be satisfied for all proper orthogonal $a_{ij}(\tau)$, i.e. for all $a_{ij}(\tau)$ satisfying (7.2.4).

Equation (7.2.9) implies a restriction on the manner in which P_{AB} can depend on the nine functions $s_{k,Q}(\tau)$. This restriction is, however, given by (7.2.9) in implicit form. Our next problem is to make this restriction explicit. We shall give two different methods by which this can be done.*

---

*Method 1 given below is essentially that due to A. E. Green and R. S. Rivlin (Arch. Rat'l. Mech. Anal. 1, 1 (1957)). Method 2 is essentially that due to W. Noll (Arch. Rat'l. Mech. Anal. 2, 197 (1958)). In the latter paper, equation (7.2.12) below is regarded as being valid for all orthogonal $a_{ij}(s)$ instead of merely proper orthogonal $a_{ij}(s)$. Although this procedure leads to the same result in the present instance, it has been criticized by R. S. Rivlin (Inelastic Behavior of Solids, e.d. W. F. Adler, R. I. Jaffee and M. F. Kanninen), publ. McGraw-Hill (in the press)).

These parallel the two methods used in discussing the corresponding problem for an elastic material in §5.2. In order to avoid any difficulty which may arise from the fact that the functions $x_{k,P}(\tau)$ are defined over an infinite range, we shall change the variable so that this range is mapped onto the finite range $0 \leqslant s \leqslant 1$. We do this by defining the new time-like variable s by

$$s = \frac{1}{t-\tau+1} . \tag{7.2.10}$$

Then,

$$P_{AB} = P_{AB}\Big[\overset{1}{\underset{s=0}{x_{k,Q}(s),t}}\Big] , \tag{7.2.11}$$

the functionals P_{AB} in (7.2.11) not, in general, being the same as those in (7.1.1). P_{AB} is, of course, the value of $P_{AB}(s)$ at s=1. Now, the restriction on P_{AB} in (7.2.9) becomes

$$P_{AB}\left[a_{k\ell}(s)x_{\ell,Q}(s),t\right] = P_{AB}\left[x_{k,Q}(s),t\right] \tag{7.2.12}$$

and this must be valid for all $a_{k\ell}(s)$ satisfying the orthogonality condition

$$a_{ij}(s)a_{ik}(s) = a_{ji}(s)a_{ki}(s) = \delta_{jk} , \quad |a_{ij}(s)| = 1. \tag{7.2.13}$$

(ii) Method 1

We recognize that equation (7.2.12) expresses the fact that for each value of AB, P_{AB} is an invariant scalar functional of the three argument vector functions of

s—$x_{k,1}(s)$, $x_{k,2}(s)$, $x_{k,3}(s)$. Since (7.2.12) is valid for arbitrary proper orthogonal $a_{ij}(s)$, we may choose

$$a_{ij}(s) = \delta_{ij}, \text{ unless } s = \bar{s} \text{ (some fixed value).}$$

The functional equation then becomes a statement that P_{AB} is an invariant function under the proper orthogonal group of the three vectors $x_{k,1}(\bar{s})$, $x_{k,2}(\bar{s})$, $x_{k,3}(\bar{s})$. Bearing in mind the fact that $|x_{k,A}(\bar{s})|$ is positive for a real deformation, we see, by an argument similar to that of §5.2 in which we discussed the form of the strain-energy function for an elastic material, that the dependence of P_{AB} on the deformation gradients at "time" $\bar{s}$ must be through the six components of the Cauchy strain at "time" $\bar{s}$, i.e. on $C_{PQ}(\bar{s})$. We repeat this argument for each value of s in the interval $0 \leqslant s \leqslant 1$, and arrive at the result that P_{AB} must depend on the nine deformation gradient functions $x_{i,A}(s)$ through the six components of the Cauchy strain $C_{AB}(s)$, over the interval $0 \leqslant s \leqslant 1$. Thus,

$$P_{AB} = P_{AB}[C_{PQ}(s), t] \ . \qquad (7.2.14)$$

(iii) Method 2

Another method by which the restrictions on P_{AB} implied by (7.2.12) may be made explicit is the following. Since (7.2.12) must be valid for all time-dependent proper orthogonal $a_{ij}(s)$, i.e. for all $a_{ij}(s)$ satisfying (7.2.13), it must be

valid for the particular choice

$$a_{ij}(s) = \{\underset{\sim}{C}(s)\}^{-\frac{1}{2}}{}_{iP} x_{j,P}(s) \ . \tag{7.2.15}$$

It is easily seen that $a_{ij}(s)$, defined in this way, satisfies the conditions (7.2.13) and is therefore a proper orthogonal transformation. Introducing (7.2.15) into (7.2.12), we obtain

$$\begin{aligned} P_{AB}[x_{k,Q}(s)] &= P_{AB}\left[\left\{C(s)\right\}^{-\frac{1}{2}}{}_{kP} C_{PQ}(s)\right] \\ &= P_{AB}\left[\left\{C(s)\right\}^{\frac{1}{2}}{}_{kQ}\right]. \end{aligned} \tag{7.2.16}$$

It follows that P_{AB} must be expressible as a functional of $\{\underset{\sim}{C}(s)\}^{\frac{1}{2}}$ and hence of $\underset{\sim}{C}(s)$, in accord with the result (7.2.14). As the final step in our argument, we note that if P_{AB} is expressible in the form (7.2.14), it automatically satisfies the relation (7.2.12) for all proper orthogonal $a_{ij}(s)$, not only for $a_{ij}(s)$ given by (7.2.15).

(iv) The Cauchy and Kirchoff-Piola stresses

Using the formulae (7.1.2) and (7.1.3) and making use of the fact that $x_{i,A}$ is itself a functional of $x_{i,A}(s)$, we obtain

$$\sigma_{ji} = x_{i,A} x_{j,B} P_{AB}[C_{PQ}(s),t] \tag{7.2.17}$$

and

$$\pi_{Ai} = x_{i,B} P_{AB}[C_{PQ}(s),t] \ . \tag{7.2.18}$$

We note that the functionals P_{AB} in (7.2.17) and (7.2.18) are not necessarily the same.

We can, of course, rewrite equations (7.2.14), (7.2.18) and (7.2.17) in the forms

$$P_{AB} = P_{AB}[C_{PQ}(\tau)] \ , \tag{7.2.19}$$

$$\pi_{Ai} = x_{i,B} P_{AB}[C_{PQ}(\tau)] \tag{7.2.20}$$

and

$$\sigma_{ji} = x_{i,A} x_{j,B} P_{AB}[C_{PQ}(\tau)] \ . \tag{7.2.21}$$

3. Hereditary materials

We may consider an important sub-class of the visco-elastic materials so far discussed. We see, from (7.2.11), that for a specified deformation gradient history, i.e. for specified functions $x_{k,Q}(s)$, the stress may depend on the actual time at which the history is carried out. For example, if we take two identical test-pieces and perform identical experiments on them on two different days, the results of the experiments may be different. We define a hereditary material as one for which the stress depends on the deformation gradient history only and not on the time at which it is carried out. This means that for a hereditary material, P_{AB} in equation (7.2.11) is independent of t and, consequently, so are P_{AB}, σ_{ji} and π_{Ai} in equations (7.2.14), (7.2.17) and (7.2.18) respectively, so that

$$P_{AB} = P_{AB}[C_{PQ}(s)] \ , \tag{7.3.1}$$

$$\sigma_{ji} = x_{i,A}x_{j,B}P_{AB}[C_{PQ}(s)] \ , \tag{7.3.2}$$

and

$$\pi_{Ai} = x_{i,B}P_{AB}[C_{PQ}(s)] \ , \tag{7.3.3}$$

the functionals P_{AB} in these three equations not necessarily being the same.

Essentially the assumption that a material is a hereditary material is an assumption that the material does not change its properties "on the shelf".

4. Constitutive equations of the differential type*

We now suppose that the deformation is such that $C_{PQ}(\tau)$ is expressible as a Taylor series about t, i.e. about the time at which the stress is measured. We then have

$$C_{PQ}(\tau) = \sum_{\alpha=0}^{\mu} \frac{1}{\alpha!} (\tau-t)^{\alpha} \frac{d^{\alpha}C_{PQ}}{dt^{\alpha}} + R_{\mu} , \tag{7.4.1}$$

where R_{μ} denotes the remainder after $\mu+1$ terms and the notation

$$\frac{d^{\alpha}C_{PQ}}{dt^{\alpha}} = \left.\frac{d^{\alpha}C_{PQ}(\tau)}{d\tau^{\alpha}}\right|_{\tau=t} \tag{7.4.2}$$

is used.

For certain deformations, $R_{\mu} = 0$. In this case, substitution of (7.4.1) in (7.2.19) yields

$$P_{AB} = F_{AB}(C_{PQ}, dC_{PQ}/dt, \ldots, d^{\mu}C_{PQ}/dt^{\mu}) ; \tag{7.4.3}$$

i.e. the functional dependence on the tensor-valued function $C_{PQ}(\tau)$ may be replaced by function dependence on the $\mu+1$ tensors $C_{PQ},\ldots,d^{\mu}C_{PQ}/dt^{\mu}$. This replacement is, of course, valid irrespective of the form of the functional P_{AB}.

For certain other classes of deformation, the replacement of the functional in (7.2.19) by a function may be made approximately, usually with some limitation on the nature of the functional P_{AB}.

*The ideas contained in this section are, in essence, contained in a paper by A. E. Green and R. S. Rivlin, Arch. Rat'l. Mech. Anal. 1, 1 (1957).

The result (7.4.3) can also be obtained from a constitutive assumption that the Kirchoff-Piola stress P_{AB} at time t, in the material considered, is a function of the deformation gradients $x_{p,Q}$, the velocity gradients $\dot{x}_{p,Q}$, the acceleration gradients $\ddot{x}_{p,Q}$, the second acceleration gradients $\dddot{x}_{p,Q}, \ldots$, the $(\mu-1)$th acceleration gradients $\overset{(\mu)}{x}_{p,Q}$, all of these being measured at time t, thus:

$$P_{AB} = F_{AB}(x_{p,Q}, \dot{x}_{p,Q}, \ldots, \overset{(\mu)}{x}_{p,Q}) \ , \tag{7.4.4}$$

the function dependence in (7.4.4) not necessarily being the same as that in (7.4.3). In order to obtain (7.4.3) from (7.4.4), we have, of course, to make use of the consideration that if we superpose on the assumed deformation a rigid time-dependent rotation P_{AB} is unaltered.

5. The Rivlin-Ericksen tensors

In this section, we define a set of tensors which are related to the tensors C_{PQ}, dC_{PQ}/dt, ..., $d^{\mu}C_{PQ}/dt^{\mu}$ and are useful in formulating constitutive equations of the differential type, particularly when the materials considered are isotropic. We denote these tensors by $A^{(0)}_{ij}$, $A^{(1)}_{ij}$, ..., $A^{(\mu)}_{ij}$ and define them by the relations

$$\begin{aligned} A^{(0)}_{ij} &= \frac{1}{2}\,\delta_{ij}\ , \\ A^{(\alpha)}_{ij} &= \frac{1}{2}\,X_{P,i}X_{Q,j}\,d^{\alpha}C_{PQ}/dt^{\alpha}\ (\alpha=1,\ldots,\mu)\ . \end{aligned} \tag{7.5.1}$$

The tensors $A^{(0)}_{ij},\ldots,A^{(\mu)}_{ij}$ defined in this way are usually called the Rivlin-Ericksen tensors. We note that the relation (7.5.1) may be written in inverse form as

$$\frac{d^{\alpha}C_{PQ}}{dt^{\alpha}} = 2x_{i,P}x_{j,Q}A^{(\alpha)}_{ij}\ (\alpha=0,1,\ldots,\mu)\ , \tag{7.5.2}$$

where $d^{\alpha}C_{PQ}/dt^{\alpha}$ is interpreted as C_{PQ} when $\alpha=0$.

We can easily obtain a formula by means of which $A^{(\alpha+1)}_{ij}$ can be calculated in terms of $A^{(\alpha)}_{ij}$. Thus, differentiating equation (7.5.2) with respect to t, we obtain

$$\begin{aligned} \frac{d^{\alpha+1}C_{PQ}}{dt^{\alpha+1}} = 2\Bigg(&x_{i,P}x_{j,Q}\,\frac{dA^{(\alpha)}_{ij}}{dt} + v_{i,P}x_{j,Q}A^{(\alpha)}_{ij} \\ &+ x_{i,P}v_{j,Q}A^{(\alpha)}_{ij}\Bigg)\ , \end{aligned} \tag{7.5.3}$$

where $v_i = dx_i/dt$. Multiplying (7.5.3) throughout by $\frac{1}{2} X_{P,m} X_{Q,n}$ and using $(7.5.1)_2$ with α replaced by $\alpha+1$, we obtain the desired result*

$$A_{mn}^{(\alpha+1)} = \frac{dA_{mn}^{(\alpha)}}{dt} + v_{i,m}A_{in}^{(\alpha)} + v_{i,n}A_{mi}^{(\alpha)} \quad (\alpha=0,1,\ldots,\mu) \ . \qquad (7.5.4)$$

Taking $\alpha=0$ in (7.5.4) and using (7.5.1), we note that

$$A_{mn}^{(1)} = \frac{1}{2} (v_{n,m}+v_{m,n}) \ ; \qquad (7.5.5)$$

i.e. $A_{mn}^{(1)}$ is the usual strain-velocity tensor.

---

*R. S. Rivlin and J. L. Ericksen, J. Rat'l. Mech. Anal. 4, 323 (1955).

6. Restrictions imposed by material symmetry on constitutive equations of the functional type

If a viscoelastic material possesses some symmetry, restrictions can be imposed on the form of the constitutive equation for the material. As in the case of an elastic material (cf. §5.3) the symmetry of the material may be described by a group of transformations relating equivalent coordinate systems.

We consider two rectangular cartesian coordinate systems x and $\bar{x}$ with a common origin. We consider two deformations A and B described by (cf. equations (5.3.1) and (5.3.2))

$$x_i(\tau) = f_i(X_A,\tau) \quad \text{and} \quad \bar{x}_i(\tau) = f_i(\bar{X}_A,\tau) \tag{7.6.1}$$

respectively.

Let P_{AB} and $\bar{P}_{AB}$ be the components at time t of the Piola stress in the systems x and $\bar{x}$ respectively associated with the deformations A and B respectively. Then, from (7.2.19), P_{AB} may be expressed in the form

$$P_{AB} = P_{AB}[C_{PQ}(\tau)] \ , \tag{7.6.2}$$

where $C_{PQ}(\tau)$ are the components in the system x of the Cauchy strain at time τ associated with the deformation A; i.e.

$$C_{PQ}(\tau) = \frac{\partial x_i(\tau)}{\partial X_P} \frac{\partial x_i(\tau)}{\partial X_Q} \ . \tag{7.6.3}$$

Similarly $\bar{P}_{AB}$ may be expressed in the form

$$\bar{P}_{AB} = \bar{P}_{AB}[\bar{C}_{PQ}(\tau)] \ , \tag{7.6.4}$$

where $\bar{C}_{PQ}(\tau)$ are the components in the system $\bar{x}$ of the Cauchy strain at time τ associated with the deformation $\mathcal{B}$; i.e.,

$$\bar{C}_{PQ}(\tau) = \frac{\partial \bar{x}_i(\tau)}{\partial \bar{X}_P} \ \frac{\partial \bar{x}_i(\tau)}{\partial \bar{X}_Q} \ . \tag{7.6.5}$$

In general the functionals P_{AB} and $\bar{P}_{AB}$ are different. If they are the same then the coordinate systems x and $\bar{x}$ are said to be <u>equivalent</u> and equation (7.6.4) may be written as

$$\bar{P}_{AB} = P_{AB}[\bar{C}_{PQ}(\tau)] \ . \tag{7.6.6}$$

The fact that (7.6.2) and (7.6.6) are simultaneously valid for all coordinate systems $\bar{x}$ equivalent to x implies (cf. §6.6) that the Piola stress tensor P_{AB} is a symmetric tensor-valued functional of the symmetric tensor function $C_{PQ}(\tau)$, invariant under the group $\{\underset{\sim}{s}\}$ describing the symmetry of the material. The results of §6.6 can now be used to write P_{AB} in canonical form. For example*, if the material considered is isotropic, whether or not it possesses a center of symmetry, replacing $\underset{\sim}{v}(\tau)$ by $\underset{\sim}{C}(\tau)$ $\left(= C_{PQ}(\tau)\right)$ and $F_{ij}[\underset{\sim}{v}(\tau)]$ by $P_{AB}\underset{\tau=-\infty}{\overset{t}{[}}\underset{\sim}{C}(\tau)]$ in (6.6.16), we obtain the canonical form for P_{AB}:

* See example 2 of §6.6.

$$P_{AB} = P_{AB}\,[\underset{\sim}{C}(\tau)]$$

$$= \alpha_o \delta_{AB} + \sum_{\mu=1}^{5} \frac{1}{2} \left\{ L_{AB}^{(\mu)} [\underset{\sim}{C}(\tau_1) \ldots \underset{\sim}{C}(\tau_\mu) + \underset{\sim}{C}(\tau_\mu) \ldots \underset{\sim}{C}(\tau_1)] \right\}, \qquad (7.6.7)$$

where $L_{AB}^{(\mu)}$ is an isotropic symmetric second-order tensor-valued functional, linear in its indicated argument and dependent also on scalar functions of linear functionals of the elements tr $\underset{\sim}{C}(\tau_1)$, tr $\underset{\sim}{C}(\tau_1)\underset{\sim}{C}(\tau_2)$,..., tr $\underset{\sim}{C}(\tau_1)\ldots\underset{\sim}{C}(\tau_6)$. α_o is, of course, also a scalar function of such linear functionals, since it is a scalar functional of $\underset{\sim}{C}(\tau)$ invariant under the full orthogonal group. From (7.6.7) we can easily obtain the Piola-Kirchoff and Cauchy stresses by using the formulae (3.8.3) and (3.8.4).

7. Restrictions imposed by material symmetry on constitutive equations of the differential type*

Instead of being a functional of the history of the Cauchy strain, we now consider the Piola stress to be a function of the Cauchy strain and its time derivatives at the instant at which the Piola stress is measured, as expressed by equation (7.4.3). In this case we have, instead of (7.6.2) and (7.6.6),

$$P_{AB} = F_{AB}(C_{PQ}, dC_{PQ}/dt, \ldots, d^{\mu}C_{PQ}/dt^{\mu})$$

and (7.7.1)

$$\overline{P}_{AB} = F_{AB}(\overline{C}_{PQ}, d\overline{C}_{PQ}/dt, \ldots, d^{\mu}\overline{C}_{PQ}/dt^{\mu}) \ ,$$

where C_{PQ} and $\overline{C}_{PQ}$, defined by (7.6.3) and (7.6.5) respectively, are the components of the Cauchy strain tensor $\underset{\sim}{C}$ and P_{AB} and $\overline{P}_{AB}$ are the components of the Piola stress tensor $\underset{\sim}{P}$ in the equivalent coordinate systems x and $\overline{x}$ respectively. We note that $dC_{PQ}/dt, \ldots, d^{\mu}C_{PQ}/dt^{\mu}$ and $d\overline{C}_{PQ}/dt, \ldots, d^{\mu}\overline{C}_{PQ}/dt^{\mu}$ are the components in the systems x and $\overline{x}$ respectively of the tensors $d\underset{\sim}{C}/dt, \ldots, d\underset{\sim}{C}^{\mu}/dt^{\mu}$. Equations (7.7.1) are, of course, valid for all choices of $\overline{x}$ equivalent to the system x. This expresses the fact that $\underset{\sim}{P}$ is a tensor-valued function of $\underset{\sim}{C}, d\underset{\sim}{C}/dt, \ldots, d\underset{\sim}{C}^{\mu}/dt^{\mu}$, invariant under the symmetry group $\{\underset{\sim}{s}\}$ for the material.

For example, if the material considered is isotropic, whether or not it possesses a center of symmetry, $\underset{\sim}{P}$ must be a

---

*See also R. S. Rivlin and J. L. Ericksen, J. Rat'l. Mech. Anal. 4, 323 (1955).

tensor-valued function of $\underset{\sim}{C}, d\underset{\sim}{C}/dt, \ldots, d\underset{\sim}{C}^{\mu}/dt^{\mu}$ invariant under the full (or proper) orthogonal group. (We note that it does not matter whether the full or proper orthogonal group is used, i.e. whether we consider the material to possess a center of symmetry or not, since all the tensors involved are symmetric tensors and therefore are unchanged by a central inversion.) The canonical form for $\underset{\sim}{P}$ is then obtained from (6.4.18) as

$$\underset{\sim}{P} = \sum_{\alpha=1}^{\lambda} B_{\alpha}(\underset{\sim}{\pi}_{\alpha}+\underset{\sim}{\pi}_{\alpha}^{(t)}) \ , \tag{7.7.2}$$

where the products $\underset{\sim}{\pi}_{\alpha}$ are obtained in the manner described in §6.4, with $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_{\mu}$ replaced by $\underset{\sim}{C}, d\underset{\sim}{C}/dt, \ldots, d\underset{\sim}{C}^{\mu}/dt^{\mu}$ and the B's are functions of the elements of an irreducible integrity basis under the orthogonal group for the tensors $\underset{\sim}{C}, d\underset{\sim}{C}/dt, \ldots, d^{\mu}\underset{\sim}{C}/dt^{\mu}$.

From (7.7.2), we obtain, with (3.8.4),

$$\underset{\sim}{\sigma} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} \sum_{\alpha=1}^{\lambda} \overline{B}_{\alpha}(\overline{\underset{\sim}{\pi}}_{\alpha}+\overline{\underset{\sim}{\pi}}_{\alpha}^{(t)}) \ , \tag{7.7.3}$$

where

$$\overline{\underset{\sim}{\pi}}_{\alpha} = \left\| x_{i,A} x_{j,B} \pi_{AB}^{(\alpha)} \right\| \ , \quad \underset{\sim}{\pi}_{\alpha} = \left\| \pi_{AB}^{(\alpha)} \right\| \ . \tag{7.7.4}$$

Bearing in mind that $\underset{\sim}{\pi}_{\alpha}$ is a product formed from $\underset{\sim}{C}, d\underset{\sim}{C}/dt, \ldots, d^{\mu}\underset{\sim}{C}/dt^{\mu}$, it is easily seen that $\overline{\underset{\sim}{\pi}}_{\alpha}$ is a product formed from the Finger strain tensor $\underset{\sim}{c}$ and the Rivlin-Ericksen tensors $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_{\mu}$.

This follows immediately from the definitions of $\underset{\sim}{c}$ and $\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_\mu$ in (2.5.6) and (7.5.2) respectively. For example, suppose $\underset{\sim}{\pi}_\alpha = \underset{\sim}{C}\, d\underset{\sim}{C}/dt$. We then have, from (7.7.4),

$$\bar{\pi}_{ij}^{(\alpha)} = x_{i,A}x_{j,B}\pi_{AB}^{(\alpha)} = x_{i,A}x_{j,B}C_{AP}\, dC_{PB}/dt\ . \tag{7.7.5}$$

With (7.5.2) and (2.5.6), equation (7.7.5) becomes

$$\bar{\pi}_{ij}^{(\alpha)} = 2c_{ik}c_{k\ell}A_{\ell m}^{(1)}c_{mj}\ , \tag{7.7.6}$$

which, with the notation

$$\underset{\sim}{A}_\alpha = \left\|A_{ij}^{(\alpha)}\right\|,\quad \underset{\sim}{c} = \left\|c_{ij}\right\|,\quad \underset{\sim}{\bar{\pi}}_\alpha = \left\|\bar{\pi}_{ij}^{(\alpha)}\right\|\ , \tag{7.7.7}$$

may be rewritten as

$$\underset{\sim}{\bar{\pi}}_\alpha = 2\underset{\sim}{c}^2\underset{\sim}{A}_1\underset{\sim}{c}\ . \tag{7.7.8}$$

We note that $\underset{\sim}{\bar{\pi}}_\alpha$ is of higher degree in $\underset{\sim}{c}$ than is $\underset{\sim}{\pi}_\alpha$ in $\underset{\sim}{C}$. Consequently, although in (7.7.2) $\underset{\sim}{\pi}_\alpha$ is of degree 5 or lower in $\underset{\sim}{C},d\underset{\sim}{C}/dt,\ldots,d^\mu\underset{\sim}{C}/dt^\mu$, $\underset{\sim}{\bar{\pi}}_\alpha$ in (7.7.3) may be of degree higher than 5 in $\underset{\sim}{c},\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_\mu$. We can, however, express each of the terms $\underset{\sim}{\bar{\pi}}_\alpha$ and hence $\underset{\sim}{\sigma}$ in canonical form by recognizing that each $\underset{\sim}{\bar{\pi}}_\alpha$ and hence $\underset{\sim}{\sigma}$ is a tensor-valued function of $\underset{\sim}{c},\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_\mu$, invariant under the orthogonal group*. It follows that $\underset{\sim}{\sigma}$ may be expressed in the form (7.7.3), where now the products $\bar{\pi}_\alpha$ are obtained in the manner described in §6.4, with $\underset{\sim}{v}_1$, $\underset{\sim}{v}_2,\ldots,\underset{\sim}{v}_\mu$

---

*We note that $\partial\underset{\sim}{x}/\partial\underset{\sim}{X} = (\det\underset{\sim}{c})^{1/2}$.

replaced by $\underset{\sim}{C}, \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_\mu$. The $\bar{B}$'s in (7.7.3) are functions of the elements of an irreducible integrity basis under the orthogonal group of the tensors $\underset{\sim}{C}, \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_\mu$. We note that $\partial \underset{\sim}{x}/\partial \underset{\sim}{X} = (\det \underset{\sim}{C})^{1/2}$ and we may therefore absorb the factor $1/(\partial \underset{\sim}{x}/\partial \underset{\sim}{X})$ in (7.7.3) into the B's, without loss of generality.

For example, in the case when $\mu=1$ in (7.7.1), so that our starting point is the relation

$$P_{AB} = F_{AB}(C_{PQ}, dC_{PQ}/dt) , \tag{7.7.9}$$

we find (cf. equation (6.4.15)) that equation (7.7.2) becomes

$$\begin{aligned} \underset{\sim}{P} = 2 \Bigg[& B_1 \underset{\sim}{I} + B_2 \underset{\sim}{C} + B_3 \underset{\sim}{C}^2 + B_4 \frac{d\underset{\sim}{C}}{dt} + B_5 \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 \Bigg] \\ & + B_6 \left(\underset{\sim}{C} \frac{d\underset{\sim}{C}}{dt} + \frac{d\underset{\sim}{C}}{dt} \underset{\sim}{C} \right) + B_7 \left\{ \underset{\sim}{C} \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 + \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 \underset{\sim}{C} \right\} \\ & + B_8 \left(\frac{d\underset{\sim}{C}}{dt} \underset{\sim}{C}^2 + \underset{\sim}{C}^2 \frac{d\underset{\sim}{C}}{dt} \right) + B_9 \left\{ \underset{\sim}{C}^2 \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 + \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 \underset{\sim}{C}^2 \right\}, \end{aligned} \tag{7.7.10}$$

where the B's are functions of the following quantities

$$\begin{aligned} & \operatorname{tr} \underset{\sim}{C}, \ \operatorname{tr} \underset{\sim}{C}^2, \ \operatorname{tr} \underset{\sim}{C}^3, \\ & \operatorname{tr} \frac{d\underset{\sim}{C}}{dt}, \ \operatorname{tr} \left(\frac{d\underset{\sim}{C}}{dt}\right)^2, \ \operatorname{tr} \left(\frac{d\underset{\sim}{C}}{dt}\right)^3 , \\ & \operatorname{tr} \underset{\sim}{C} \frac{d\underset{\sim}{C}}{dt}, \ \operatorname{tr} \underset{\sim}{C}^2 \frac{d\underset{\sim}{C}}{dt}, \ \operatorname{tr} \underset{\sim}{C} \left(\frac{d\underset{\sim}{C}}{dt}\right)^2, \\ & \operatorname{tr} \underset{\sim}{C}^2 \left(\frac{d\underset{\sim}{C}}{dt}\right)^2 . \end{aligned} \tag{7.7.11}$$

In the same case, equation (7.7.3) becomes

$$\begin{aligned}\underset{\sim}{\sigma} = \frac{1}{\partial \underset{\sim}{x}/\partial \underset{\sim}{X}} \Big\{ & 2\Big[\bar{B}_1 \underset{\sim}{I} + \bar{B}_2 \underset{\sim}{c} + \bar{B}_3 \underset{\sim}{c}^2 + \bar{B}_4 \underset{\sim}{A}_1 + \bar{B}_5 \underset{\sim}{A}_1^{\,2}\Big] \\ & + \bar{B}_6(\underset{\sim}{c}\underset{\sim}{A}_1 + \underset{\sim}{A}_1\underset{\sim}{c}) + \bar{B}_7(\underset{\sim}{c}\underset{\sim}{A}_1^{\,2} + \underset{\sim}{A}_1^{\,2}\underset{\sim}{c}) \\ & + \bar{B}_8(\underset{\sim}{A}_1\underset{\sim}{c}^2 + \underset{\sim}{c}^2\underset{\sim}{A}_1) + \bar{B}_9(\underset{\sim}{c}^2\underset{\sim}{A}_1^{\,2} + \underset{\sim}{A}_1^{\,2}\underset{\sim}{c}^2) \ , \end{aligned} \tag{7.7.12}$$

where the $\bar{B}$'s are functions of

$$\begin{aligned} & \operatorname{tr} \underset{\sim}{c}, \quad \operatorname{tr} \underset{\sim}{c}^2, \quad \operatorname{tr} \underset{\sim}{c}^3, \\ & \operatorname{tr} \underset{\sim}{A}_1, \quad \operatorname{tr} \underset{\sim}{A}_1^{\,2}, \quad \operatorname{tr} \underset{\sim}{A}_1^{\,3}, \\ & \operatorname{tr} \underset{\sim}{c}\underset{\sim}{A}_1, \quad \operatorname{tr} \underset{\sim}{c}^2\underset{\sim}{A}_1, \quad \operatorname{tr} \underset{\sim}{c}\underset{\sim}{A}_1^{\,2}, \\ & \operatorname{tr} \underset{\sim}{c}^2\underset{\sim}{A}_1^{\,2}. \end{aligned} \tag{7.7.13}$$

As has already been remarked we may, without loss of generality, absorb the factor $1/(\partial \underset{\sim}{x}/\partial \underset{\sim}{X})$ into the $\bar{B}$'s.

8. Constraints

We have already seen in the discussion of elastic materials in §4.8 that if the deformations to which an elastic material can be subjected are restricted by some constraint, then the stress is no longer determined by the deformation, but is subject to an indeterminacy, the exact nature of which depends on the character of the constraint.

For example, if the material is incompressible, then the constraint takes the form, cf. (4.8.1),

$$\partial \underset{\sim}{x}/\partial \underset{\sim}{X} = 1 \ . \tag{7.8.1}$$

From this it follows, cf. (4.8.2), that

$$\varepsilon_{ik\ell}(\dot{x}_{i,1}x_{k,2}x_{\ell,3} + \dot{x}_{k,2}x_{i,1}x_{\ell,3} + \dot{x}_{\ell,3}x_{i,1}x_{k,2})$$

$$= \frac{1}{2}\,\varepsilon_{ik\ell}\varepsilon_{ABC}\dot{x}_{i,A}x_{k,B}x_{\ell,c} = 0 \ . \tag{7.8.2}$$

Now from (4.4.5) we see that the dissipation equation may be written

$$-\rho_o\dot{U} + \rho_o\chi - Q_{B,B} + \pi_{Bi}\dot{x}_{i,B} = 0 \ . \tag{7.8.3}$$

From (7.8.2) and (7.8.3), we obtain

$$-\rho_o\dot{U} + \rho_o\chi - Q_{B,B} + (\pi_{Ai} + \frac{1}{2}\,p\varepsilon_{ik\ell}\varepsilon_{ABC}x_{k,B}x_{\ell,c})\dot{x}_{i,A} = 0 \ , \tag{7.8.4}$$

where p is an arbitrary quantity. Thus the addition of the term $\frac{1}{2}\,p\varepsilon_{ik\ell}\varepsilon_{ABC}x_{k,B}x_{\ell,C}$ to the Kirchoff-Piola stress leaves the

stress-power unaltered. For an incompressible material, we accordingly make a constitutive assumption for $(\pi_{Ai} + \frac{1}{2} p\epsilon_{ik\ell} \epsilon_{ABC} x_{k,B} x_{\ell,C})$ instead of for π_{Ai}. Alternatively, bearing in mind the relation (3.7.5), we may make a constitutive assumption for $\sigma_{ij} + p\delta_{ij}$ instead of for σ_{ji}. Thus, for an incompressible material with memory, we may make a constitutive assumption in the form

$$\pi_{Ai} + \frac{1}{2} p\epsilon_{ik\ell}\epsilon_{ABC} x_{k,B} x_{\ell,C} = F_{Ai}[x_{k,Q}(\tau)]$$

or (7.8.5)

$$\sigma_{ij} + p\delta_{ij} = F_{ij}[x_{k,Q}(\tau)] ,$$

the tensor-valued functionals F_{Ai} and F_{ij} being different in the two cases.

Similarly, for an incompressible material, bearing in mind the relation (3.8.4), we may replace the constitutive assumption (7.4.4) by the assumption

$$\sigma_{ij} + p\delta_{ij} = F_{ij}(x_{p,Q}, \dot{x}_{p,Q}, \ldots, \overset{(\mu)}{x}_{p,Q}) . \qquad (7.8.6)$$

Other constraints than incompressibility can be introduced in an analogous manner.

9. Viscoelastic fluids

While there is no clear general criterion which differentiates a solid from a fluid, we have certain loose criteria by which we recognize fluid-like behavior. One of these is that in a fluid the stress does not necessarily increase as the amount of deformation increases; the magnitude of the stress is associated with the rate at which the deformation changes rather than with the magnitude of the deformation itself, apart possibly for a dependence on the density of the fluid.

In accord with these concepts we may make the assumption that the components of the Cauchy stress in some rectangular cartesian system x depend on the density of the fluid and on the gradients of the velocity, acceleration, second acceleration and so on, all of these gradients being taken with respect to the configuration of the body at the time t at which the stress is measured. Thus,

$$\sigma_{ij} = F_{ij}(\rho, v_{p,q}, v^{(1)}_{p,q}, v^{(2)}_{p,q}, \ldots) , \qquad (7.9.1)$$

where $\underset{\sim}{v} = \dot{\underset{\sim}{x}}$ is the velocity of a particle situated at $\underset{\sim}{x}$ at time t, $\underset{\sim}{v}^{(1)} = \ddot{\underset{\sim}{x}}$ is its acceleration, $\underset{\sim}{v}^{(2)}$ is its second acceleration, and so on, and ρ denotes the density of the fluid. For an incompressible fluid the density is constant and the constitutive assumption (7.9.1) is replaced by*

---

*This is the assumption made by R. S. Rivlin and J. L. Ericksen, J. Rat'l. Mech. Anal. 4, 323 (1955).

$$\sigma_{ij} = F_{ij}(v_{p,q}, v^{(1)}_{p,q}, \ldots) - p\delta_{ij} , \qquad (7.9.2)$$

where p is undetermined if the deformation is prescribed. We notice that the constitutive assumption (7.9.2) is equivalent to the assumption (7.8.6) if in this relation the reference configuration is taken to be the configuration at time t (i.e. T = t). We note that in this case the Piola and Cauchy stresses become identical.

We have already remarked that (7.4.3) may be obtained from (7.4.4) by making use of the consideration that if we superpose on the assumed deformation a rigid time-dependent rotation, the Piola stress remains unchanged. In a similar fashion we see, with the relation (7.5.2), that the constitutive assumption (7.9.1) leads to

$$\sigma_{ij} = F_{ij}(\rho, A^{(1)}_{pq}, A^{(2)}_{pq}, \ldots) \qquad (7.9.3)$$

and the constitutive assumption (7.9.2) leads to

$$\sigma_{ij} = F_{ij}(A^{(1)}_{pq}, A^{(2)}_{pq}, \ldots) - p\delta_{ij} . \qquad (7.9.4)$$

We note that the functions F_{ij} in (7.9.1) and (7.9.3) are not generally the same. Also the functions F_{ij} in (7.9.2) and (7.9.4) are not generally the same.

We reiterate that the assumption (7.9.1) is made in a particular rectangular cartesian coordinate system x. In general, the functions F_{ij} in (7.9.1) depend on the rectangular

cartesian coordinate system in which (7.9.1) is written. If the fluid is isotropic, then the F_{ij} in (7.9.1) are the same for all choices of the rectangular cartesian reference system. It follows that F_{ij} in (7.9.3) are also the same for all choices of the rectangular cartesian reference system. Then $\underset{\sim}{\sigma}$ may be expressed in canonical form in terms of $\underset{\sim}{A}_1, \underset{\sim}{A}_2, \ldots$ in a manner similar to that employed in §7.7. Similar considerations apply to the incompressible case when the relevant equations are (7.9.2) and (7.9.4).

An analogous procedure was adopted by Noll* in specializing the constitutive equation (7.2.19) to the case when the material is a fluid.

Noll defined a <u>simple material</u> as one for which the Cauchy stress depends on the history of the deformation gradients; i.e. as a material satisfying the constitutive assumption

$$\sigma_{ij} = F_{ij}[x_{p,Q}(\tau)] \ . \tag{7.9.5}$$

He then defined a <u>simple fluid</u> as a simple material for which σ_{ij} is independent of the reference configuration with respect to which the gradients are measured, except insofar as the volume (or density) of a material element is changed in the passage between these configurations.+ Noting that we may rewrite (7.9.5) as

$$\sigma_{ij} = F_{ij}[x_{p,q}(\tau)x_{q,Q}] \ , \tag{7.9.6}$$

*W. Noll, Arch. Rat'l. Mech. Anal. <u>2</u>, 197 (1958).

+ Substantially the same idea for differentiating between a solid and a fluid had been previously expressed by J.G. Oldroyd, (Proc. Roy. Soc. A <u>200</u>, 523 (1950))

it follows that for a simple fluid we may write, in accordance with Noll's concept,

$$\sigma_{ij} = F_{ij}[x_{p,q}(\tau),\ |x_{q,Q}|] \tag{7.9.7}$$

or

$$\sigma_{ij} = F_{ij}[x_{p,q}(\tau),\ \rho_o/\rho] \ . \tag{7.9.8}$$

We note here that $|x_{q,Q}| = \rho_o/\rho$ where ρ and ρ_o are the material densities at times t and T respectively.

If the constitutive assumption (7.9.8) is made instead of (7.9.5), it is easy to see that σ_{ij} must be expressible in the form

$$\sigma_{ij} = F_{ij}^{*}[C_{pq}^{(t)}(\tau),\rho] \tag{7.9.9}$$

where $C_{pq}^{(t)}(\tau)$ is the Cauchy strain at time τ with respect to the configuration at time t, i.e.

$$C_{pq}^{(t)}(\tau) = x_{k,p}(\tau)x_{k,q}(\tau) \ . \tag{7.9.10}$$

This result is easily obtained from (7.2.21) by noting that now we may replace $\underset{\sim}{X}$ by $\underset{\sim}{x}$ and $x_{i,A}$ by δ_{iA}. For the incompressible case, we have analogously

$$\sigma_{ij} = F_{ij}[C_{pq}^{(t)}(\tau)] - p\delta_{ij} \ . \tag{7.9.11}$$

In a certain sense, merely writing down the expression (7.9.8) instead of (7.9.5) does not effect any change, for we note that

$$X_i = \int_{-\infty}^{t} \delta(\tau - T) x_i(\tau) d\tau \ , \qquad (7.9.12)$$

where $\delta(\tau-T)$ is a Dirac delta function. Here we bear in mind that $T < t$. Also,

$$|X_{Q,q}| x_{i,A} = \frac{1}{2} \varepsilon_{ijk} \varepsilon_{ABC} X_{B,j} X_{C,k} \ . \qquad (7.9.13)$$

Combining (7.9.12) and (7.9.13), we see that $x_{i,A}$ may be regarded as a functional of $x_{i,j}(\tau)$. Thus, making use of the relation

$$x_{i,A}(\tau) = x_{i,j}(\tau) x_{j,A} \ , \qquad (7.9.14)$$

we may consider any functional of $x_{i,A}(\tau)$ to be a functional of $x_{i,j}(\tau)$. The converse statement that we may consider any functional of $x_{i,j}(\tau)$ to be a functional of $x_{i,A}(\tau)$ is evidently also valid.

In order to embody in (7.9.9) the concept that σ_{ij} is independent of any initial state (apart, possibly, through its dependence on the density at time t) we must add the proviso that F_{ij} does not involve any fixed reference time. This is difficult to present in explicit form while F_{ij} is left as a general functional. We can, however, see the meaning of this more clearly by taking an artificial example. Suppose

$$\sigma_{ij} = \int_{-\infty}^{t} \alpha_{ijpq}(\tau - T) C_{pq}^{(t)}(\tau) d\tau \ . \qquad (7.9.15)$$

If this is to represent a simple fluid, then α_{ijpq} must be independent of any reference time T (other, of course, than the time t at which the stress is measured). As an example of what may occur if this is not the case, suppose

$$\alpha_{ijpq}(\tau-T) = \beta_{ijpq}\delta(\tau-T) \tag{7.9.16}$$

where β_{ijpq} is constant. Then,

$$\sigma_{ij} = \beta_{ijpq}C^{(t)}_{pq}(T) \tag{7.9.17}$$

and we have a material which behaves as an elastic solid rather than as one having fluid-like behavior.

Returning to (7.9.9) and assuming the deformation to be such that $C^{(t)}_{pq}(\tau)$ may be expanded as a Taylor series about the time t, we have (cf. §7.4)

$$C^{(t)}_{pq}(\tau) = \delta_{pq} + \sum_{\alpha=1}^{\mu}\frac{1}{\alpha!}(\tau-t)^{\alpha}\frac{d^{\alpha}C^{(t)}_{pq}}{dt^{\alpha}} + R_{\mu}\ , \tag{7.9.18}$$

where

$$d^{\alpha}C^{(t)}_{pq}/dt^{\alpha} = d^{\alpha}C^{(t)}_{pq}(\tau)/d\tau^{\alpha}\big|_{\tau=t}\ , \tag{7.9.19}$$

R_{μ} is the usual Taylor remainder and we have used the result from (7.9.10) that $C^{(t)}_{pq}(t) = \delta_{pq}$. Subject to the condition that R_{μ} be small enough, we see by introducing (7.9.17) into (7.9.9) that σ_{ij} may be expressed in the form

$$\sigma_{ij} = F_{ij}(\rho, dC^{(t)}_{pq}/dt, d^2C^{(t)}_{pq}/dt, \ldots, d^{\mu}C^{(t)}_{pq}/dt)\ . \tag{7.9.20}$$

In (7.5.2), we now take the reference configuration to be the configuration at time t; i.e. we take $T = t$ and $X_A = x_A$. We then obtain

$$\frac{d^{\alpha}C_{PQ}}{dt^{\alpha}} = 2A_{ij}^{(\alpha)} \quad . \tag{7.9.21}$$

Introducing this relation into (7.9.19) we see that σ_{ij} is expressible in the form (cf. (7.9.3))

$$\sigma_{ij} = F_{ij}(\rho, A_{ij}^{(1)}, A_{ij}^{(2)}, \ldots, A_{ij}^{(\mu)}) \quad . \tag{7.9.22}$$

Analogously, for an incompressible fluid, we obtain from (7.9.11), the relation (7.9.4).

Noll's concept of a simple fluid as a simple material for which σ_{ij} is independent of the reference configuration with respect to which the gradients are measured (apart from possible dependence on the density) implies, of course, that the material is isotropic. This is immediately evident from the consideration that, as a special case, we can change the reference configuration by a rigid rotation, without altering the stress. This does not, however, imply, as is sometimes supposed, that a simple material which has generally properties which we recognize as fluid-like is necessarily a simple fluid. This fact has been pointed out by Green*. In this respect the nomenclature introduced by Noll is highly misleading. It has, for example, led to the erroneous notion that a simple material which is a fluid cannot be anisotropic.

---

*A. E. Green, Proc. Roy. Soc. A 279, 437 (1964).

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

G.F. SMITH

THE GENERATION OF INTEGRITY BASES

Corso tenuto a Bressanone dal 3 all'11 settembre 1969

THE GENERATION OF INTEGRITY BASES

by

G. F. Smith
(Lehigh - University)

1. Introduction

We consider constitutive relations of the form

$$W = P(\underset{\sim}{\beta},\ldots,\underset{\sim}{\gamma}) \tag{1.1}$$

where W is a scalar and where P is a polynomial function. The symmetry properties of the material considered impose restrictions on the form of P. Thus, let $\underset{\sim}{e}_i$ denote the base vectors associated with a rectangular cartesian coordinate system $\underset{\sim}{x}$ whose orientation relative to the preferred directions (if any) in the material is specified. Let

$$\underset{\sim}{\bar{e}}_i^{(k)} = A_{ij}^{(k)} \underset{\sim}{e}_j \quad (i,j=1,2,3) \tag{1.2}$$

denote the vectors into which the $\underset{\sim}{e}_i$ are carried by the various symmetry transformations of the material. The $\underset{\sim}{\bar{e}}_i^{(k)}$ form the base vectors for the rectangular cartesian coordinate systems

$$\underset{\sim}{A_1 x},\ \underset{\sim}{A_2 x},\ \underset{\sim}{A_3 x},\ldots \tag{1.3}$$

which are referred to as the set of equivalent coordinate systems associated with the material considered. The reference frames (1.3) are specified by listing the set of orthogonal

matrices

$$\underset{\sim}{A}_k = \left\| A_{ij}^{(k)} \right\| \tag{1.4}$$

where $\underset{\sim}{A}_1 = \underset{\sim}{I}$. The set of matrices (1.4) forms a group which we denote by $\{\underset{\sim}{A}\}$ and refer to as the symmetry group of the material. Let β_i and $(A_k\ \beta)_i, \ldots,$ γ_j and $(A_k\ \gamma)_j$ denote the independent components of $\underset{\sim}{\beta}, \ldots,$ $\underset{\sim}{\gamma}$ when referred to the equivalent reference frames $\underset{\sim}{x}$ and $\underset{\sim}{A}_k\underset{\sim}{x}$ respectively. These quantities are related by the equations

$$(A_k\beta)_i = T_{ij}^{(k)}\beta_j, \ldots, \quad (A_k\gamma)_i = S_{ij}^{(k)}\gamma_j. \tag{1.5}$$

We require that the response function (1.1) shall have the same functional form when referred to any of the set of equivalent reference frames $\underset{\sim}{A}_k\underset{\sim}{x}$ (k=1,2,...) associated with the material. Thus, the function P must satisfy the equations

$$P(\beta_i, \ldots, \gamma_j) = P\left[T_{in}^{(k)}\beta_n, \ldots, S_{jm}^{(k)}\gamma_m\right] \quad (k=1,2,\ldots) \tag{1.6}$$

or

$$P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma}) = P(\underset{\sim}{T}_k\underset{\sim}{\beta}, \ldots, \underset{\sim}{S}_k\underset{\sim}{\gamma}) \quad (k=1,2,\ldots). \tag{1.7}$$

The function P is then said to be invariant under the group of transformations $\{\underset{\sim}{A}\} = \underset{\sim}{A}_1, \underset{\sim}{A}_2, \ldots$. The problem of concern here is to find the general form of the polynomial function

$P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ consistent with the restrictions (1.6) or (1.7). The solution to this problem is given by determining a minimal integrity basis for polynomial functions of $\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma}$ which are invariant under $\{\underset{\sim}{A}\}$. A minimal integrity basis is comprised of the smallest set of polynomial functions $I_1(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma}), \ldots, I_p(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$, each of which is invariant under $\{\underset{\sim}{A}\}$, such that any polynomial function $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ which is invariant under $\{\underset{\sim}{A}\}$ is expressible as a polynomial in the $I_1, \ldots, I_p$.

2. The generation of integrity bases

We consider the problem of determining the elements of the integrity basis for polynomial functions of N absolute vectors $y_1, y_2, \ldots, y_N$ which are invariant under the group $\{\underset{\sim}{A}\}$ associated with the crystal class D_2. This crystal class is characterized by the presence of three mutually orthogonal two-fold axes of rotation. If we take the x_1, x_2 and x_3 axes to lie along these two-fold axes of rotation, the symmetry transformations are the rotations through 180 degrees about each of the coordinate axes. The set of matrices defining the symmetry properties of the material are then given by

$$\underset{\sim}{A}_1,\underset{\sim}{A}_2,\underset{\sim}{A}_3,\underset{\sim}{A}_4 = \begin{Vmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{Vmatrix}, \begin{Vmatrix} 1 & . & . \\ . & -1 & . \\ . & . & -1 \end{Vmatrix}, \begin{Vmatrix} -1 & . & . \\ . & 1 & . \\ . & . & -1 \end{Vmatrix}, \begin{Vmatrix} -1 & . & . \\ . & -1 & . \\ . & . & 1 \end{Vmatrix}. \quad (2.1)$$

Let y_i and $(A_k y)_i$ denote the components of the absolute vector $\underset{\sim}{y}$ when referred to the reference frames x and $\underset{\sim}{A}_k\underset{\sim}{x}$ respectively. These components are related by the equations

$$(A_k y)_i = T_{ij}^{(k)} y_j \quad (2.2)$$

where the matrices $\underset{\sim}{T}_k = \underset{\sim}{T}(\underset{\sim}{A}_k) = \left\| T_{ij}^{(k)} \right\|$ are given by

$$\underset{\sim}{T}_1, \underset{\sim}{T}_2, \underset{\sim}{T}_3, \underset{\sim}{T}_4 = \begin{Vmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{Vmatrix}, \begin{Vmatrix} 1 & . & . \\ . & -1 & . \\ . & . & -1 \end{Vmatrix}, \begin{Vmatrix} -1 & . & . \\ . & 1 & . \\ . & . & -1 \end{Vmatrix}, \begin{Vmatrix} -1 & . & . \\ . & -1 & . \\ . & . & 1 \end{Vmatrix}. \tag{2.3}$$

The restrictions of the form (1.7) given by

$$P(\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_N) = P(\underset{\sim}{T}_k \underset{\sim}{y}_1, \ldots, \underset{\sim}{T}_k \underset{\sim}{y}_N) \quad (k=1, \ldots, 4) \tag{2.4}$$

may be written more explicitly as

$$\begin{aligned} & P(y_1^{(1)}, y_2^{(1)}, y_3^{(1)}, \ldots, y_1^{(N)}, y_2^{(N)}, y_3^{(N)}) \\ & = P(y_1^{(1)}, -y_2^{(1)}, -y_3^{(1)}, \ldots, y_1^{(N)}, -y_2^{(N)}, -y_3^{(N)}) \\ & = P(-y_1^{(1)}, y_2^{(1)}, -y_3^{(1)}, \ldots, -y_1^{(N)}, y_2^{(N)}, -y_3^{(N)}) \\ & = P(-y_1^{(1)}, -y_2^{(1)}, y_3^{(1)}, \ldots, -y_1^{(N)}, -y_2^{(N)}, y_3^{(N)}). \end{aligned} \tag{2.5}$$

In order to determine the integrity basis, we employ the following obvious theorem.

Theorem 1. Let P be a polynomial function of the real quantities $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m$ which satisfies the

relation

$$P(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m) = P(\alpha_1, \ldots, \alpha_n, -\beta_1, \ldots, -\beta_m). \quad (2.6)$$

Then P is expressible as a polynomial in the quantities

$$\alpha_i (i=1,\ldots,n) \text{ and } \beta_j\beta_k (j,k=1,\ldots,m). \quad (2.7)$$

With (2.5) and theorem 1, we then readily see that an integrity basis for polynomial functions of $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_N$ which are invariant under the group $\{\underset{\sim}{A}\}$ defined by (2.1) is formed by the quantities

$$y_1^{(i)}y_1^{(j)},\ y_2^{(i)}y_2^{(j)},\ y_3^{(i)}y_3^{(j)},\ y_1^{(i)}y_2^{(j)}y_3^{(k)} (i,j,k=1,2,\ldots,N). \quad (2.8)$$

There are a number of theorems such as theorem 1 above which enable us in principle to determine the integrity basis for a wide variety of problems. However, the integrity bases which are obtained upon application of such theorems will generally contain a number of redundant terms which must be eliminated. This may prove to be a matter of some difficulty. In order to avoid such difficulties, we may employ an iterative procedure. We outline below the application of this procedure to the generation of the multilinear elements

of the integrity basis for functions of N vectors $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_N$ which are invariant under the group (2.1).

We first determine the number P_n of linearly independent invariants which are multilinear in n vectors $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_n$. The transformation properties of the 3^n quantities $y_i^{(1)} y_j^{(2)} \ldots y_k^{(n)}$ (i,j,k=1,2,3) under change of reference frames are described by the Kronecker nth power $\underset{\sim}{T}_k^{[n]}$ of the matrices $\underset{\sim}{T}_k$. The number P_n is obtained by taking the mean value over the group $\{\underset{\sim}{A}\}$ of the trace of the matrices $\underset{\sim}{T}_k^{[n]}$. Since the trace of the Kronecker nth power of a matrix $\underset{\sim}{T}_k$ is equal to the nth power of the tr $\underset{\sim}{T}_k$, we have

$$P_n = \frac{1}{4} \sum_{k=1}^{4} (\text{tr } \underset{\sim}{T}_k)^n. \tag{2.9}$$

From (2,3), we see that

$$(\text{tr } \underset{\sim}{T}_1, \text{ tr } \underset{\sim}{T}_2, \text{ tr } \underset{\sim}{T}_3, \text{ tr } \underset{\sim}{T}_4) = (3, -1, -1, -1). \tag{2.10}$$

With (2.9) and (2.10), we then have

$$P_1 = 0, \; P_2 = 3, \; P_3 = 6, \; P_4 = 21, \; \ldots \; . \tag{2.11}$$

We now proceed to generate the multilinear elements of the integrity basis. From (2.11), we see that there are no invariants of degree 1 in $\underset{\sim}{y}_1$ and three linearly independent

invariants of degree 1, 1 in $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$. With (2.5), we readily see that

$$y_1^{(1)}y_1^{(2)},\ y_2^{(1)}y_2^{(2)},\ y_3^{(1)}y_3^{(2)} \tag{2.12}$$

are invariant under $\{\underset{\sim}{A}\}$. They are obviously linearly independent. The multilinear elements of the integrity basis of degree two are then comprised of the $\binom{N}{2}$ sets of invariants obtained by replacing $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$ in the set of invariants (2.12) by all possible sets of two different vectors chosen from $\underset{\sim}{y}_1$, ..., $\underset{\sim}{y}_N$. From (2.11), we see there are six linearly independent invariants of degree 1, 1, 1 in $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, $\underset{\sim}{y}_3$. With (2.5), we readily see that these are given by

$$y_i^{(1)}y_j^{(2)}y_k^{(3)} (ijk = 123, 132, 213, 231, 312, 321). \tag{2.13}$$

The multilinear elements of the integrity basis of degree three are then comprised of the $\binom{N}{3}$ sets of invariants obtained by replacing $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, $\underset{\sim}{y}_3$ in (2.13) by all possible sets of three different vectors chosen from $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, ..., $\underset{\sim}{y}_N$. From (2.11), we see that there are 21 linearly independent invariants of degree 1, 1, 1, 1 in $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, $\underset{\sim}{y}_3$, $\underset{\sim}{y}_4$. However, we readily verify that there are also 21 linearly independent invariants of this degree which may be obtained as products of invariants of the form (2.12). Hence there are no invariants of degree 1, 1, 1, 1 in $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$, $\underset{\sim}{y}_3$, $\underset{\sim}{y}_4$ which are required as elements

of the integrity basis.

This iteration process must be terminated at some stage and it is necessary to determine by one means or another an upper bound on the degree of the elements of the integrity basis. For the case under consideration, we may employ a result which says that since the group $\{\underset{\sim}{A}\}$ is comprised of four transformations, the elements of the integrity basis must be of degree four or less. This enables us to state that the typical multilinear elements of the integrity basis for polynomial functions of N vectors invariant under the group (2.1) are given by (2.12) and (2.13). The determination of the non-linear elements of the integrity basis may be carried out in a similar fashion.

In the iterative procedure described above, we know the number P_n of linearly independent invariants of degree n. We determine by inspection the invariants of degree n which may be obtained as products of invariants of degree less than n. Suppose there are R_n such invariants. They are not necessarily all linearly independent. Suppose then that Q_n of these R_n invariants are linearly independent. We must determine these Q_n invariants and then determine $P_n - Q_n$ additional invariants $I_1, I_2, \ldots$ such that the $I_1, I_2, \ldots$ together with the Q_n invariants which are products of lower

order invariants form a set of P_n linearly independent invariants. The I_1, I_2, ... are then elements of the integrity basis. This can be a very formidable problem when P_n and Q_n are large. In sections 3 and 5, we discuss methods which essentially reduce this problem to a number of smaller problems which may usually be solved much more readily.

3. Invariants of symmetry type $(n_1 n_2 \ldots n_r)$.

Let $I_1, \ldots, I_n$ be a set of linearly independent invariants which are multilinear in the quantities $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_M$. We choose as an example the invariants (2.13). Let

$$I_1 = y_1^{(1)} y_2^{(2)} y_3^{(3)}, \quad I_2 = y_1^{(1)} y_2^{(3)} y_3^{(2)}, \quad I_3 = y_1^{(2)} y_2^{(3)} y_3^{(1)},$$
$$I_4 = y_1^{(2)} y_2^{(1)} y_3^{(3)}, \quad I_5 = y_1^{(3)} y_2^{(1)} y_3^{(2)}, \quad I_6 = y_1^{(3)} y_2^{(2)} y_3^{(1)}. \tag{3.1}$$

Let s be that permutation of the numbers 1, ..., M which carries 1 into i_1, ..., M into i_M. The permutation s applied to the subscripts on the tensors $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_M$ transforms the invariant $I_j(\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_M)$ into the invariant

$$sI_j(\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_M) = I_j(\underset{\sim}{A}_{i_1}, \ldots, \underset{\sim}{A}_{i_M}). \tag{3.2}$$

We assume that the space spanned by $I_1, \ldots, I_n$ is invariant under the group $\underset{\sim}{S}_M$ of all M! permutations of 1, ..., M, i.e., the invariants $sI_j(\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_M)$ are expressible as linear combinations of $I_1, \ldots, I_n$. Thus,

$$sI_j = I_k D_{kj}(s). \tag{3.3}$$

For example, consider the transformation properties of the quantities (3.1) under permutation of the vectors $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \underset{\sim}{y}_3$ among themselves. The symmetric group $\underset{\sim}{S}_3$ is comprised of

six permutations

$$e,\ (12),\ (13),\ (23),\ (123),\ (132). \tag{3.4}$$

In table 1 below, we list the quantities $sI_j(\underset{\sim}{y}_1, \underset{\sim}{y}_2, \underset{\sim}{y}_3)$ for $j = 1, \ldots, 6$ and for all permutations s belonging to the set (3.4).

Table 1

	I_1	I_2	I_3	I_4	I_5	I_6
e	I_1	I_2	I_3	I_4	I_5	I_6
(12)	I_4	I_3	I_2	I_1	I_6	I_5
(13)	I_6	I_5	I_4	I_3	I_2	I_1
(23)	I_2	I_1	I_6	I_5	I_4	I_3
(123)	I_3	I_4	I_5	I_6	I_1	I_2
(132)	I_5	I_6	I_1	I_2	I_3	I_4

The matrices $\underset{\sim}{D}(s)$ appearing in (3.3) are then seen to be given by

$$\underset{\sim}{D}(e) = \begin{Vmatrix} 1 & . & . & . & . & . \\ . & 1 & . & . & . & . \\ . & . & 1 & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & 1 & . \\ . & . & . & . & . & 1 \end{Vmatrix}, \quad \underset{\sim}{D}(12) = \begin{Vmatrix} . & . & . & 1 & . & . \\ . & . & 1 & . & . & . \\ . & 1 & . & . & . & . \\ 1 & . & . & . & . & . \\ . & . & . & . & . & 1 \\ . & . & . & . & 1 & . \end{Vmatrix}$$

$$\underset{\sim}{D}(13) = \begin{Vmatrix} . & . & . & . & . & 1 \\ . & . & . & . & 1 & . \\ . & . & . & 1 & . & . \\ . & . & 1 & . & . & . \\ . & 1 & . & . & . & . \\ 1 & . & . & . & . & . \end{Vmatrix}, \quad \underset{\sim}{D}(23) = \begin{Vmatrix} . & 1 & . & . & . & . \\ 1 & . & . & . & . & . \\ . & . & . & . & . & 1 \\ . & . & . & . & 1 & . \\ . & . & . & 1 & . & . \\ . & . & 1 & . & . & . \end{Vmatrix}, \quad (3.$$

$$\underset{\sim}{D}(123) = \begin{Vmatrix} . & . & . & . & 1 & . \\ . & . & . & . & . & 1 \\ 1 & . & . & . & . & . \\ . & 1 & . & . & . & . \\ . & . & 1 & . & . & . \\ . & . & . & 1 & . & . \end{Vmatrix}, \quad \underset{\sim}{D}(132) = \begin{Vmatrix} . & . & 1 & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & 1 & . \\ . & . & . & . & . & 1 \\ 1 & . & . & . & . & . \\ . & 1 & . & . & . & . \end{Vmatrix}$$

The matrices $\underset{\sim}{D}(s)$ given by (3.5) which describe the transformation properties of the invariants (3.1) under the permutations of $\underset{\sim}{S}_3$ are said to form a matrix representation of degree 6 of the symmetric group $\underset{\sim}{S}_3$. Thus, to every element s of $\underset{\sim}{S}_3$ there corresponds a matrix $\underset{\sim}{D}(s)$ such that to the product u = ts of two permutations corresponds the matrix

$$\underset{\sim}{D}(u) = \underset{\sim}{D}(t)\ \underset{\sim}{D}(s). \tag{3.6}$$

For example,

$$(13)\ (23) = (\ 132) \tag{3.7}$$

and we see from (3.5) that

$$\underset{\sim}{D}(13)\ \underset{\sim}{D}(23) = \underset{\sim}{D}(132). \tag{3.8}$$

The invariants (3.1) are said to form the carrier space of the representation

$$\underset{\sim}{\Gamma} = \{\underset{\sim}{D}(e), \underset{\sim}{D}(12), \ldots, \underset{\sim}{D}(132)\}. \tag{3.9}$$

Consider now the invariants defined by

$$J_1 = I_1 + I_2 + I_3 + I_4 + I_5 + I_6,$$

$$J_2 = I_1 + I_4 - I_5 - I_6 \ , \ J_3 = I_3 + I_6 - I_1 - I_2, \tag{3.10}$$

$$J_4 = I_1 + I_6 - I_3 - I_4 \ , \ J_5 = I_4 + I_5 - I_1 - I_2,$$

$$J_6 = I_1 + I_3 + I_5 - I_2 - I_4 - I_6.$$

From table 1, we readily obtain the transformation properties of the invariants $J_1, \ldots, J_6$ under the permutations of $\underset{\sim}{y}_1$, $\underset{\sim}{y}_2$ and $\underset{\sim}{y}_3$. We list the quantities sJ_i in table 2.

Table 2

	J_1	J_2	J_3	J_4	J_5	J_6
e	J_1	J_2	J_3	J_4	J_5	J_6
(12)	J_1	J_2	$-J_2-J_3$	J_5	J_4	$-J_6$
(13)	J_1	J_3	J_2	J_4	$-J_4-J_5$	$-J_6$
(23)	J_1	$-J_2-J_3$	J_3	$-J_4-J_5$	J_5	$-J_6$
(123)	J_1	J_3	$-J_2-J_3$	$-J_4-J_5$	J_4	J_6
(132)	J_1	$-J_2-J_3$	J_2	J_5	$-J_4-J_5$	J_6

We thus have

$$s\, J_p = J_q\, H_{qp}(s) \tag{3.11}$$

where the matrices $\underset{\sim}{H}(s)$ form a matrix representation of $\underset{\sim}{S}_3$ which is said to be equivalent to the representation $\underset{\sim}{D}(s)$. From table 2, we see that the matrices $\underset{\sim}{H}(s)$ are all of the form

$$\underset{\sim}{H}(s) = \begin{array}{|ccc|} \hline \underset{\sim}{K}(s) & & \\ & \underset{\sim}{L}(s) & \\ & & \underset{\sim}{M}(s) \\ & & \underset{\sim}{N}(s) \\ \hline \end{array} \tag{3.12}$$

where $\underset{\sim}{K}$, $\underset{\sim}{L}$, $\underset{\sim}{M}$ and $\underset{\sim}{N}$ are 1 x 1, 2 x 2, 2 x 2 and 1 x 1 matrices respectively and where all of the non-zero components of $\underset{\sim}{H}$ appear in the matrices $\underset{\sim}{K}$, $\underset{\sim}{L}$, $\underset{\sim}{M}$ and $\underset{\sim}{N}$. The sets of matrices $\underset{\sim}{K}(s)$, ..., $\underset{\sim}{N}(s)$ are listed in table 3.

Table 3

s	e	(12)	(13)	(23)	(123)	(132)
$\underset{\sim}{K}(s)$	1	1	1	1	1	1
$\underset{\sim}{L}(s)$	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}1 & -1\\ 0 & -1\end{matrix}$	$\begin{matrix}0 & 1\\ 1 & 0\end{matrix}$	$\begin{matrix}-1 & 0\\ -1 & 1\end{matrix}$	$\begin{matrix}0 & -1\\ 1 & -1\end{matrix}$	$\begin{matrix}-1 & 1\\ -1 & 0\end{matrix}$
$\underset{\sim}{M}(s)$	$\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$	$\begin{matrix}0 & 1\\ 1 & 0\end{matrix}$	$\begin{matrix}1 & -1\\ 0 & -1\end{matrix}$	$\begin{matrix}-1 & 0\\ -1 & 1\end{matrix}$	$\begin{matrix}-1 & 1\\ -1 & 0\end{matrix}$	$\begin{matrix}0 & -1\\ 1 & -1\end{matrix}$
$\underset{\sim}{N}(s)$	1	-1	-1	-1	1	1

The sets of matrices $\underset{\sim}{K}(s), \ldots, \underset{\sim}{N}(s)$ also form representations of the symmetric group $\underset{\sim}{S}_3$. The invariants J_1, (J_2, J_3), (J_4, J_5) and J_6 form the carrier spaces for the representations

$$\underset{\sim}{\Gamma}_1 = \{\underset{\sim}{K}(s)\}, \ \underset{\sim}{\Gamma}_2 = \{\underset{\sim}{L}(s)\}, \ \underset{\sim}{\Gamma}_3 = \{\underset{\sim}{M}(s)\}, \ \underset{\sim}{\Gamma}_4 = \{\underset{\sim}{N}(s)\} \tag{3.13}$$

respectively. We say that the representation $\underset{\sim}{\Gamma} = \{\underset{\sim}{D}(s)\}$ has been decomposed into the direct sum of the representations $\underset{\sim}{\Gamma}_1, \ldots, \underset{\sim}{\Gamma}_4$. If a matrix representation can be decomposed in this fashion, it is said to be reducible. If not, it is said to be irreducible. Each of the representations (3.13) are irreducible representations of the symmetric group $\underset{\sim}{S}_3$.
The quantity

$$\text{char}\ \{\underset{\sim}{K}(s)\} = \Big[\text{tr}\ \underset{\sim}{K}(e), \ \text{tr}\ \underset{\sim}{K}(12), \ \text{tr}\ \underset{\sim}{K}(13), \ \text{tr}\ \underset{\sim}{K}(23), \ \text{tr}\ \underset{\sim}{K}(123), \ \text{tr}\ \underset{\sim}{K}(132)\Big] \tag{3.14}$$

is referred to as the character of the representation $\underset{\sim}{\Gamma}_1 = \{\underset{\sim}{K}(s)\}$. There are only three inequivalent irreducible representations associated with the symmetric group $\underset{\sim}{S}_3$. These are denoted by

$$(3), \ (21), \ (111). \tag{3.15}$$

The characters of these irreducible representations are given [1] by

$$\text{char } (3) = \left[1,1,1,1,1,1\right] , \quad \text{char } (21) = \left[2,0,0,0,-1,-1\right] ,$$
$$\text{char } (111) = \left[1,-1,-1,-1,1,1\right] . \tag{3.16}$$

We see from table 3 that

$$\text{char } \{\underset{\sim}{K}(s)\} = \text{char } (3), \quad \text{char } \{\underset{\sim}{L}(s)\} = \text{char } \{\underset{\sim}{M}(s)\} = \text{char } (21),$$
$$\text{char } \{\underset{\sim}{N}(s)\} = \text{char } (111). \tag{3.17}$$

This reflects the fact that the character of any irreducible representation of $\underset{\sim}{S}_3$ must equal either char (3), char (21) or char (111).

The invariant J_1 defined by (3.10) forms the carrier space for the irreducible representation $\{\underset{\sim}{K}(s)\}$ for which char $\{\underset{\sim}{K}(s)\}$ is equal to char (3). We then refer to J_1 as a set of invariants of symmetry type (3). The invariants J_2 and J_3 defined by (3.10) form the carrier space for the representation $\{\underset{\sim}{L}(s)\}$. Since char $\{\underset{\sim}{L}(s)\}$ = char (21), we refer to J_2 and J_3 as a set of invariants of symmetry type (21). Similarly the invariants J_4, J_4 and J_6 defined by (3.10) are referred to as sets of invariants of symmetry types (21) and (111) respectively.

The number of inequivalent irreducible representations of the symmetric group $\underset{\sim}{S}_M$ is equal to the number of partitions of M, i.e., the number of solutions in positive integers of the equation

$$n_1 + n_2 + \ldots + n_r = M, \; n_1 \geq n_2 \geq \ldots \geq n_r. \tag{3.18}$$

For example, the partitions of 4 are given by 4,31,22,211,1111 and the inequivalent irreducible representations of $\underset{\sim}{S}_4$ are denoted by

$$(4), \; (31), \; (22), (211), \; (1111). \tag{3.19}$$

Similarly, the partitions of 5 are given by 5,41,32,311,221, 2111,11111 and the inequivalent irreducible representations of $\underset{\sim}{S}_5$ are denoted by

$$(5), \; (41), \; (32), \; (311), \; (221), \; (2111), \; (11111). \tag{3.20}$$

The characters of the various irreducible representations of $\underset{\sim}{S}_M$ may be found in the literature for $M \leq 15$ (see [1] for M=1, ..., 8). If a set of invariants $J_1, \ldots, J_p$ forms the carrier

space for an irreducible representation $\underset{\sim}{\Gamma}$ of $\underset{\sim}{S}_M$ for which char $\underset{\sim}{\Gamma}$ = char $(n_1 n_2 \ldots n_r)$, we say that the invariants J_1, ..., J_p form a set of invariants of symmetry type $(n_1 n_2 \ldots n_r)$.

In the next section, we give an example to show how we may employ the notion of a set of invariants of symmetry type $(n_1 n_2 \ldots n_r)$ to ease the burden of computation involved in determining the multilinear elements of an integrity basis.

4. Integrity basis for N symmetric second-order traceless tensors $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_N$ - the proper orthogonal group.

We outline the computation yielding the multilinear elements of this integrity basis. We borrow from the discussion of Spencer and Rivlin [2] the following results.

(i) Every multilinear element of the integrity basis involves at most six tensors and is of the form

$$\operatorname{tr} \underset{\sim}{A}_i \underset{\sim}{A}_j \ldots \underset{\sim}{A}_k . \tag{4.1}$$

(ii) The trace of a matrix product formed from symmetric 3 x 3 matrices is unaltered by cyclic permutation of the factors in the product and is also unaltered if the order of the factors in the product is reversed.

For example, we have

$$\begin{aligned} \operatorname{tr} \underset{\sim}{A}_1 \underset{\sim}{A}_2 \underset{\sim}{A}_3 &= \operatorname{tr} \underset{\sim}{A}_2 \underset{\sim}{A}_3 \underset{\sim}{A}_1 = \operatorname{tr} \underset{\sim}{A}_3 \underset{\sim}{A}_1 \underset{\sim}{A}_2 \\ &= \operatorname{tr} \underset{\sim}{A}_3 \underset{\sim}{A}_2 \underset{\sim}{A}_1 = \operatorname{tr} \underset{\sim}{A}_1 \underset{\sim}{A}_3 \underset{\sim}{A}_2 = \operatorname{tr} \underset{\sim}{A}_2 \underset{\sim}{A}_1 \underset{\sim}{A}_3 . \end{aligned} \tag{4.2}$$

We may readily compute (see [3]) the number $p_{(n_1 n_2 \ldots)}$ of sets of linearly independent invariants of symmetry type $(n_1 n_2 \ldots n_r)$. These quantities are listed in table 4.

Table 4

$(n_1 n_2 \ldots)$	(2)	(3)	(4)	(22)	(5)	(41)	(32)	(221)	(11111)
$p_{(n_1 n_2 \ldots)}$	1	1	1	2	1	1	1	1	1
$q_{(n_1 n_2 \ldots)}$	0	0	1	1	1	1	1	0	0
$N_{(n_1 n_2 \ldots)}$	1	1	1	2	1	4	5	5	1

$(n_1 n_2 \ldots)$	(6)	(42)	(321)	(3111)	(222)
$p_{(n_1 n_2 \ldots)}$	2	3	1	1	2
$q_{(n_1 n_2 \ldots)}$	2	3	1	0	2
$N_{(n_1 n_2 \ldots)}$	1	9	16	10	5

In table 4, $N_{(n_1 n_2 \ldots)}$ denotes the number of invariants comprising a set of invariants of symmetry type $(n_1 n_2 \ldots)$ and $q_{(n_1 n_2 \ldots)}$ denotes the number of sets of invariants which may be obtained as products of lower order invariants. The computation then

proceeds as follows.

(i) Invariants linear in $\underset{\sim}{A}_1$.

Since tr $\underset{\sim}{A}_1 = 0$, there are no linearly independent invariants of degree one in $\underset{\sim}{A}_1$.

(ii) Invariants multilinear in $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$.

There is only a single linearly independent invariant of this degree which is given by

$$\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2. \tag{4.3}$$

We see that

$$e \text{ tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 = \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2, \quad (12) \text{ tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 = \text{tr } \underset{\sim}{A}_2\underset{\sim}{A}_1 = \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2. \tag{4.4}$$

Thus, tr $\underset{\sim}{A}_1\underset{\sim}{A}_2$ forms the carrier space for a matrix representation $\underset{\sim}{D}(s)$ of degree one which is given by

$$\underset{\sim}{D}(e) = \left\| 1 \right\| , \quad \underset{\sim}{D}(12) = \left\| 1 \right\| . \tag{4.5}$$

From (4.5) and the character tables for S_2 given in [1], we see that char $\{\underset{\sim}{D}(s)\}$ = [1,1] = char (2). Hence the invariant tr $\underset{\sim}{A}_1\underset{\sim}{A}_2$ forms a set of invariants of symmetry type (2).

(iii) Invariants multilinear in $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$, $\underset{\sim}{A}_3$.

There is only a single linearly independent invariant of this degree which is given by

$$\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3. \tag{4.6}$$

With (4.2), we see that

$$\begin{aligned} \text{e} \quad & \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3 = \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3, \\ (12) \quad & \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3 = \text{tr } \underset{\sim}{A}_2\underset{\sim}{A}_1\underset{\sim}{A}_3 = \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3, \ \ldots, \\ (132) \quad & \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3 = \text{tr } \underset{\sim}{A}_3\underset{\sim}{A}_1\underset{\sim}{A}_2 = \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3. \end{aligned} \tag{4.7}$$

Thus, $\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3$ is readily seen to form the carrier space for a representation $\{\underset{\sim}{E}(s)\}$ of $\underset{\sim}{S}_3$ such that char $\{\underset{\sim}{E}(s)\}$ = char (3). Hence, $\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3$ forms a set of invariants of symmetry type (3).

(iv) Invariants multilinear in $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$, $\underset{\sim}{A}_3$, $\underset{\sim}{A}_4$.

From table 4, we see that there are five linearly independent invariants of this degree which form one set of invariants of symmetry type (4) and two sets of invariants of symmetry type (22). We may obtain three invariants as products of invariants of the form (4.3). These are given by

$$\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 \text{ tr } \underset{\sim}{A}_3\underset{\sim}{A}_4, \ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_3 \text{ tr } \underset{\sim}{A}_2\underset{\sim}{A}_4, \ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_4 \text{ tr } \underset{\sim}{A}_2\underset{\sim}{A}_3. \tag{4.8}$$

We may upon investigating the manner in which the invariants (4.8) behave under permutations of the tensors $A_1, \ldots, A_4$ establish that the invariants (4.8) may be split into a set of invariants of symmetry type (4) and a set of invariants of symmetry type (22). Thus, $q_{(4)} = q_{(22)} = 1$. We then see that we require one set of invariants of symmetry type (22) as elements of the integrity basis (since $p_{(22)} = 2$ and $q_{(22)} = 1$). This set is given by

$$I_1(\underset{\sim}{A}_1, \underset{\sim}{A}_2, \underset{\sim}{A}_3, \underset{\sim}{A}_4) = Y \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \mathrm{tr}\, \underset{\sim}{A}_1\underset{\sim}{A}_3\underset{\sim}{A}_2\underset{\sim}{A}_4,$$

$$I_2(\underset{\sim}{A}_1, \underset{\sim}{A}_2, \underset{\sim}{A}_3, \underset{\sim}{A}_4) = Y \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad \mathrm{tr}\, \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3\underset{\sim}{A}_4, \quad (22) \tag{4.9}$$

where Y (...) denotes a Young symmetry operator (see [3] for a discussion of the properties of Y(...) and further references). For example, we have

$$Y \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \{e + (12)\}\{e + (34)\}\{e - (13)\}\{e - (24)\}. \tag{4.10}$$

(v) Invariants multilinear in $A_1, \ldots, A_5$.

There are 10 invariants of this degree which may be obtained as products of the invariants (4.3) and (4.6). These

are given by

$$\text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 \text{ tr } \underset{\sim}{A}_3\underset{\sim}{A}_4\underset{\sim}{A}_5,\ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_3 \text{ tr } \underset{\sim}{A}_2\underset{\sim}{A}_4\underset{\sim}{A}_5,\ \ldots,\ \text{tr } \underset{\sim}{A}_4\underset{\sim}{A}_5 \text{ tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3. \quad (4.11)$$

The ten invariants (4.11) form the carrier space for a representation which is denoted by (2).(3) and is referred to as the direct product of the representations (2) and (3). The decomposition of representations $(n_1n_2\ldots).(m_1m_2\ldots)$ has been considered by Murnaghan [4]. We see from tables given in [4] that

$$(2).(3) = (5) + (41) + (32) \quad (4.12)$$

and hence

$$q_{(5)} = q_{(41)} = q_{(32)} = 1. \quad (4.13)$$

Then from table 4 we see that one set of invariants of symmetry types (11111) and (221) are required as elements of the integrity basis. These will be given by

$$Y \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \text{ tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3\underset{\sim}{A}_4\underset{\sim}{A}_5 \quad (4.14)$$

and

$$Y \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & \end{pmatrix} \quad \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3\underset{\sim}{A}_4\underset{\sim}{A}_5, \ \ldots \ . \tag{4.15}$$

(vi) Invariants multilinear in $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_6$.

There are 15 + 30 + 10 = 55 invariants of this degree which may be obtained as products of lower order invariants. These are given by

$$\begin{aligned}
&1.\ \ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 \ \text{tr } \underset{\sim}{A}_3\underset{\sim}{A}_4 \ \text{tr } \underset{\sim}{A}_5\underset{\sim}{A}_6, \ \ldots, \text{ 15 invariants,} && \\
&\qquad\qquad\qquad\qquad\qquad\qquad (6) + (42) + (222); && \\
&2.\ \ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2 \ I_j(\underset{\sim}{A}_3, \underset{\sim}{A}_4, \underset{\sim}{A}_5, \underset{\sim}{A}_6) \ (j=1,2), \ \ldots, \text{ 30 invariants,} && (4.16) \\
&\qquad\qquad\qquad\qquad\qquad\qquad (42) + (321) + (222); && \\
&3.\ \ \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3 \ \text{tr } \underset{\sim}{A}_4\underset{\sim}{A}_5\underset{\sim}{A}_6, \ \ldots, \text{ 10 invariants,} && \\
&\qquad\qquad\qquad\qquad\qquad\qquad (6) + (42). &&
\end{aligned}$$

In (4.16), we have listed on the right the irreducible representations into which the representations for which the invariants $(4.16)_1, \ldots, (4.16)_3$ form the carrier spaces may be decomposed. We then see from (4.16) that

$$q_{(6)} = q_{(222)} = 2, \ q_{(42)} = 3, \ q_{(321)} = 1. \tag{4.17}$$

From table 4, we see that

$$p_{(6)} = p_{(222)} = 2,\ p_{(42)} = 3,\ p_{(321)} = 1 \text{ and } p_{(3111)} = 1. \quad (4.18)$$

With (4.17) and (4.18), we see that there must be a single set of 10 invariants of symmetry type (3111) present in the integrity basis. This is given by

$$Y \begin{pmatrix} 1 & 2 & 3 \\ 4 & & \\ 5 & & \\ 6 & & \end{pmatrix} \text{tr } \underset{\sim}{A}_1\underset{\sim}{A}_2\underset{\sim}{A}_3\underset{\sim}{A}_4\underset{\sim}{A}_5\underset{\sim}{A}_6, \ \ldots \ . \quad (4.19)$$

We thus see that the typical multilinear elements of the integrity basis are given by the invariants (4.3), (4.6), (4.9), (4.14), (4.15) and (4.19). The non-linear elements of the integrity basis are readily obtained once the multilinear elements have been determined. Details of the procedure involved in obtaining the non-linear elements from the multilinear are given by Smith [3].

5. Reduction of invariant - theoretic problems to standard form - the crystallographic groups.

We consider the problem of determining the general form of the polynomial $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ which is subject to the restrictions

$$P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma}) = P\left[\underset{\sim}{T}_k\underset{\sim}{\beta}, \ldots, \underset{\sim}{S}_k\underset{\sim}{\gamma}\right] \quad (k=1, \ldots, N). \tag{5.1}$$

The set of matrices

$$\underset{\sim}{T}_k = \underset{\sim}{T}(\underset{\sim}{A}_k) = \left\| T_{ij}^{(k)} \right\| \tag{5.2}$$

forms a matrix representation $\underset{\sim}{\Gamma}$ of the crystallographic group $\{\underset{\sim}{A}\} = \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_N$ and describes the transformation properties of the components $\beta_1, \ldots, \beta_p$ of $\underset{\sim}{\beta}$. The $\beta_1, \ldots, \beta_p$ form the carrier space for the representation $\underset{\sim}{\Gamma}$. We may in some cases be able to determine p quantities

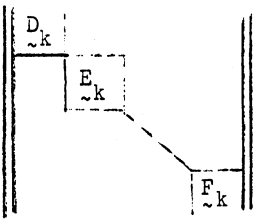
$$\phi_i = Q_{ij}\beta_j \quad (i,j=1, \ldots, p) \tag{5.3}$$

where $\det Q_{ij} \neq 0$ such that the matrices

$$\underset{\sim}{Q}\underset{\sim}{T}_k\underset{\sim}{Q}^{-1} \quad (k=1, \ldots, N) \tag{5.4}$$

which describe the transformation properties of the components $\phi_1, \ldots, \phi_p$ of $\underset{\sim}{\phi}$ are all expressible in the form

$$\underset{\sim}{Q}\underset{\sim}{T}_k\underset{\sim}{Q}^{-1} = \quad (k=1, \ldots, N). \tag{5.5}$$



In (5.5), the only non-zero elements of the matrices $\underset{\sim}{Q}\underset{\sim}{T}_k\underset{\sim}{Q}^{-1}$ appear in the matrices $\underset{\sim}{D}_k, \ldots, \underset{\sim}{F}_k$. The representation $\underset{\sim}{\Gamma} = \{\underset{\sim}{T}_k\}$ is then said to be reducible. If no such decomposition is possible, the representation $\underset{\sim}{\Gamma}$ is said to be irreducible. If the decomposition process (5.5) has been carried out as far as possible, then the representation $\underset{\sim}{\Gamma}$ is said to have been decomposed into the direct sum of irreducible representations $\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_r$. This is denoted by

$$\underset{\sim}{\Gamma} = \underset{\sim}{\Gamma}_1 + \underset{\sim}{\Gamma}_2 + \ldots + \underset{\sim}{\Gamma}_r \tag{5.6}$$

where

$$\underset{\sim}{\Gamma}_1 = \{\underset{\sim}{D}_k\}, \underset{\sim}{\Gamma}_2 = \{\underset{\sim}{E}_k\}, \ldots, \underset{\sim}{\Gamma}_r = \{\underset{\sim}{F}_k\}. \tag{5.7}$$

The quantities

$$\underset{\sim}{\lambda} = (\phi_1, \ldots, \phi_q), \ \underset{\sim}{\psi} = (\phi_{q+1}, \ldots, \phi_m), \ \ldots,$$

$$\underset{\sim}{\chi} = (\phi_{s+1}, \ldots, \phi_r) \tag{5.8}$$

whose components form the carrier spaces for the irreducible representations $\underset{\sim}{\Gamma}_1 = \{\underset{\sim}{D}_k\}$, $\underset{\sim}{\Gamma}_2 = \{\underset{\sim}{E}_k\}$, ..., $\underset{\sim}{\Gamma}_r = \{\underset{\sim}{F}_k\}$ are referred to as the basic quantities associated with the irreducible representations $\underset{\sim}{\Gamma}_1$, $\underset{\sim}{\Gamma}_2$, ..., $\underset{\sim}{\Gamma}_r$ respectively.

If the matrices $\underset{\sim}{T}_k = \underset{\sim}{T}(\underset{\sim}{A}_k)$ form a representation $\underset{\sim}{\Gamma}$ of the group $\{\underset{\sim}{A}\}$, then the matrices $\underset{\sim}{Q}\underset{\sim}{T}_k\underset{\sim}{Q}^{-1}$ also form a representation of $\{\underset{\sim}{A}\}$ which is said to be <u>equivalent</u> to the representation $\underset{\sim}{\Gamma}$. Associated with a <u>crystallographic group</u> which is a finite group, there is only a finite number of inequivalent irreducible representations $\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_r$. Any representation $\underset{\sim}{\Gamma}$ of $\{\underset{\sim}{A}\}$ can be decomposed into the direct sum of the irreducible representations $\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_r$. This says that we may always split an arbitrary quantity $\underset{\sim}{\beta} = (\beta_1, \ldots, \beta_p)$ which forms the carrier space for $\underset{\sim}{\Gamma}$ into sets of the basic quantities $\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \ldots, \underset{\sim}{\psi}_1, \underset{\sim}{\psi}_2, \ldots$, $\underset{\sim}{\chi}_1, \underset{\sim}{\chi}_2, \ldots$ associated with the various irreducible representations $\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_r$ of $\{\underset{\sim}{A}\}$. Thus, the problem of determining the integrity basis for polynomial functions $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ of any

number of quantities $\underset{\sim}{\beta}$, ..., $\underset{\sim}{\gamma}$ of arbitrary type which are invariant under $\{\underset{\sim}{A}\}$ is always reducible to the problem of determining the integrity basis for polynomial functions $P^*(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \ldots, \underset{\sim}{\psi}_1, \underset{\sim}{\psi}_2, \ldots, \underset{\sim}{\chi}_1, \underset{\sim}{\chi}_2, \ldots)$ of the basic quantities associated with the irreducible representations $\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_r$ of $\{\underset{\sim}{A}\}$.

Example. The crystal class D_2.

The set of matrices $\{\underset{\sim}{A}\} = \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_4$ defining the symmetry properties of the material are given by (2.1). The irreducible representations associated with the group $\{\underset{\sim}{A}\} = \underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_4$ are all of degree one and are listed in table 5. The quantities which form the carrier spaces for representations equivalent to $\underset{\sim}{\Gamma}_1, \ldots, \underset{\sim}{\Gamma}_4$ will be denoted by $a_1, a_2, \ldots, d_1, d_2, \ldots$ respectively.

Table 5

$\{\underset{\sim}{A}\}$	$\underset{\sim}{A}_1$	$\underset{\sim}{A}_2$	$\underset{\sim}{A}_3$	$\underset{\sim}{A}_4$	Basic Quantities
$\underset{\sim}{\Gamma}_1$	1	1	1	1	$a_1, a_2, \ldots$
$\underset{\sim}{\Gamma}_2$	1	1	-1	-1	$b_1, b_2, \ldots$
$\underset{\sim}{\Gamma}_3$	1	-1	1	-1	$c_1, c_2, \ldots$
$\underset{\sim}{\Gamma}_4$	1	-1	-1	1	$d_1, d_2, \ldots$

Let y_i and $(A_k y)_i$, G_{ij} and $(A_k G)_{ij}$ denote the components of an absolute vector $\underset{\sim}{y}$ and an absolute second-order symmetric tensor $\underset{\sim}{G}$ when referred to the reference frames $\underset{\sim}{x}$ and $\underset{\sim k}{x}$ respectively. We may readily compute the $(A_k y)_i$ and $(A_k G)_{ij}$ from tensor transformation rules. The results are listed in table 6.

Table 6

	$\underset{\sim}{A}_1$	$\underset{\sim}{A}_2$	$\underset{\sim}{A}_3$	$\underset{\sim}{A}_4$
$(A_k y)_1$	y_1	y_1	$-y_1$	$-y_1$
$(A_k y)_2$	y_2	$-y_2$	y_2	$-y_2$
$(A_k y)_3$	y_3	$-y_3$	$-y_3$	y_3
$(A_k G)_{11}$	G_{11}	G_{11}	G_{11}	G_{11}
$(A_k G)_{22}$	G_{22}	G_{22}	G_{22}	G_{22}
$(A_k G)_{33}$	G_{33}	G_{33}	G_{33}	G_{33}
$(A_k G)_{12}$	G_{12}	$-G_{12}$	$-G_{12}$	G_{12}
$(A_k G)_{13}$	G_{13}	$-G_{13}$	G_{13}	$-G_{13}$
$(A_k G)_{23}$	G_{23}	G_{23}	$-G_{23}$	$-G_{23}$

With table 5 and 6, we see that y_1 and G_{23}, y_2 and G_{31}, y_3 and G_{12} are basic quantities associated with the irreducible representations $\underset{\sim}{\Gamma}_2$, $\underset{\sim}{\Gamma}_3$ and $\underset{\sim}{\Gamma}_4$ of $\{\underset{\sim}{A}\}$ respectively. Also G_{11}, G_{22} and G_{33} are basic quantities associated with the

irreducible representation $\underset{\sim}{\Gamma}_1$ of $\{\underset{\sim}{A}\}$. Let us now employ the notation

$$\begin{aligned} a_1 &= G_{11},\ a_2 = G_{22},\ a_3 = G_{33}, \\ b_1 &= y_1,\ b_2 = G_{23}, \\ c_1 &= y_2,\ c_2 = G_{31}, \\ d_1 &= y_3,\ d_2 = G_{12}. \end{aligned} \tag{5.9}$$

Then the problem of determining the general form of the function $P(y_i, G_{ij})$ which is invariant under $\{\underset{\sim}{A}\}$ may be reduced to the problem of determining the general form of the polynomial function

$$P(a_i, b_j, c_k, d_m) \quad (i=1,2,3;\ j=1,2;\ k=1,2;\ m=1,2) \tag{5.10}$$

which is subject to the restrictions

$$\begin{aligned} &P(a_i, b_j, c_k, d_m) = P(a_i, b_j, -c_k, -d_m) \\ &= P(a_i, -b_j, c_k, -d_m) = P(a_i, -b_j, -c_k, d_m). \end{aligned} \tag{5.11}$$

Also the problem of determining the general form of the polynomial $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ which is invariant under the group $\{\underset{\sim}{A}\}$ defined by (2.1) and where the quantities appearing as arguments of P may be arbitrary in number and in type is then

readily seen to be reducible to the form (5.10) and (5.11). The only difference would be that the range of the subscripts i, j, k, m would differ from that of (5.10).

The above example is a particularly simple case. For the less tractable of the crystal classes, we proceed as follows. Let $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_N$ denote the crystallographic group and let

$$\underset{\sim}{\Gamma} \; : \; \underset{\sim}{T}(\underset{\sim}{A}_1), \ldots, \underset{\sim}{T}(\underset{\sim}{A}_N) \tag{5.12}$$

denote the representation which gives the relation between the components β_i and $(A_k\beta)_i$ of $\underset{\sim}{\beta}$ in the various equivalent reference frames $\underset{\sim}{x}, \ldots, \underset{\sim}{A}_N\underset{\sim}{x}$ associated with the crystal class considered. Then, $\underset{\sim}{\Gamma}$ may be decomposed into the direct sum

$$\underset{\sim}{\Gamma} = n_1\underset{\sim}{\Gamma}_1 + n_2\underset{\sim}{\Gamma}_2 + \ldots + n_r\underset{\sim}{\Gamma}_r \tag{5.13}$$

where $\underset{\sim}{\Gamma}_1, \ldots, \underset{\sim}{\Gamma}_r$ are the irreducible representations associated with the crystallographic group $\{\underset{\sim}{A}\}$ and where n_i denotes the number of times $\underset{\sim}{\Gamma}_i$ appears in the decomposition of $\underset{\sim}{\Gamma}$. The number n_i appearing in (5.13) may be computed [5] from the formula

$$n_i = \frac{1}{N} \sum_{k=1}^{N} \operatorname{tr} \underset{\sim}{T}(\underset{\sim}{A}_k) \operatorname{tr} \overline{\underset{\sim}{F}}(\underset{\sim}{A}_k) \tag{5.14}$$

where the matrices $\underset{\sim}{F}(\underset{\sim}{A}_i)$ are those defining the irreducible representation $\underset{\sim}{\Gamma}_i$ and where tr $\overline{\underset{\sim}{F}}$ denotes the complex conjugate of the trace of the matrix $\underset{\sim}{F}$. The components $\phi_1, \phi_2, \ldots$ of the quantity $\underset{\sim}{\phi}$ which form the carrier space for the irreducible representation $\underset{\sim}{\Gamma}_i$ are linear combinations of the components $\beta_1, \ldots, \beta_p$ of $\underset{\sim}{\beta}$ and may be obtained [5] upon application of the formula

$$\phi_j = \sum_{k=1}^{N} \overset{\circ}{F}_{j1}(\underset{\sim}{A}_k^{-1}) T_{pq}(\underset{\sim}{A}_k) \beta_q . \tag{5.15}$$

The application of (5.15) may become very tedious. A computer program is being written which should routinely produce the ϕ_j. Thus, for any crystallographic group we may always reduce the problem of determining the form of a function $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ of any number of quantities $\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma}$ of any arbitrary type to that of determining the general form of a function $P(\underset{\sim}{\phi}_1, \underset{\sim}{\phi}_2, \ldots, \underset{\sim}{\phi}_{n_1}, \underset{\sim}{\psi}_1, \underset{\sim}{\psi}_2, \ldots, \underset{\sim}{\psi}_{n_2}, \ldots)$ where the $\underset{\sim}{\phi}_i, \underset{\sim}{\psi}_i, \ldots$ are basic quantities associated with the various irreducible representations of the group $\{\underset{\sim}{A}\}$.

6. Results available in the literature

If we apply the procedure of section 5, we find that application of theorem 1 (see section 2) and a generalization of this theorem suffice to give complete results for all of the crystal classes except the cubic crystal classes. These results are given in [5]. Complete results for the cubic crystal classes may be obtained upon application of a method discussed in [3] which is a generalization of the procedure outlined in sections 3 and 4. Results for the full orthogonal group and the proper orthogonal group are confined mainly to cases involving only vectors and second-order tensors. We list some references below where the integrity bases for many cases of interest may be found. In this list, tensor means a symmetric second-order tensor.

(see next page)

Symmetry	Variables	Reference
Crystal classes	1 vector, 1 tensor	[6]
Crystal classes	M vectors	[7]
Crystal classes (except cubic crystals)	Arbitrary	[5]
Transverse isotropy	M vectors, N tensors	[8]
Hemihedral isotropy	M vectors, N tensors	[2],[9]
Holohedral isotropy	M polar vectors, N axial vectors P tensors	[10]

We note that it has been shown by Wineman and Pipkin [11] that the assumption that the functions $P(\underset{\sim}{\beta}, \ldots, \underset{\sim}{\gamma})$ are polynomials may be removed.

References

[1] Littlewood, D. E.: The Theory of Group Characters, 2nd ed. Oxford (1950).

[2] Spencer, A.J.M. and R. S. Rivlin: Arch. Rational Mech. Anal. 9, 45 (1962).

[3] Smith, G. F.: Accad. Nazionale Lincei, Serie 8, Volume 9, 51 (1968).

[4] Murnaghan, F. D.: Amer. J. Math. 59, 437 (1937).

[5] Kiral, E. and G. F. Smith: Forthcoming.

[6] Smith, G. F. and M. M. Smith and R. S. Rivlin: Arch. Rational Mech. Anal. 12, 93 (1963).

[7] Smith, G. F. and R. S. Rivlin: Arch. Rational Mech. Anal. 15, 169 (1964).

[8] Adkins, J. E.: Arch. Rational Mech. Anal. 5, 263 (1960).

[9] Spencer, A.J.M.: Arch. Rational Mech. Anal. 18, 51 (1965).

[10] Smith, G. F.: Arch. Rational Mech. Anal. 18, 282 (1965).

[11] Wineman, A. S. and A. C. Pipkin: Arch. Rational Mech. Anal. 17, 184 (1964).

CENTRO INTERNAZIONALE MATEMATICO ESTIVO
(C.I.M.E.)

E. VARLEY

NON LINEAR CONTINUUM THEORIES IN MECHANICS AND PHYSICS AND THEIR APPLICATIONS

Testo non pervenuto

Corso tenuto a Bressanone dal 3 all'11 settembre 1969