



THE FOUNDATIONS
OF MECHANICS
AND THERMODYNAMICS

SELECTED PAPERS

BY

W. NOLL

WITH A PREFACE BY C. TRUESDELL

SPRINGER-VERLAG
BERLIN HEIDELBERG NEW YORK 1974

WALTER NOLL

**Professor of Mathematics
Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA**

CLIFFORD AMBROSE TRUESDELL, III

**Professor of Rational Mechanics
The Johns Hopkins University, Baltimore, Maryland, USA**

ISBN-13: 978-3-642-65819-8

e-ISBN-13: 978-3-642-65817-4

DOI: 10.1007/978-3-642-65817-4

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law where copies are made for other than private use, a fee is payable to the publisher, the amount of the fee to be determined by agreement with the publisher. © by Springer-Verlag Berlin Heidelberg 1974. Library of Congress Catalog Card Number 74-1651. Softcover reprint of the hardcover 1st edition 1974

Preface

WALTER NOLL's influence upon research into the foundations of mechanics and thermodynamics in the past twenty years is plain, everywhere acknowledged. Less obvious is the wide effect his writings have exerted upon those who apply mechanics to special cases, but it is witnessed by the now common use of terms, concepts, and styles of argument he introduced, use sometimes by young engineers who have learnt them in some recent textbook and hence take them for granted, often with no idea whence they come.

The purpose of this volume of reprints is to put into general hands those of NOLL's works which presently promise broadest and most fertile service to students of the thermomechanics of deformable bodies. This purpose explains the selection of the contents, about half of the pages he has published in periodicals or proceedings. The complete list of his works, which follows this preface, serves also as Table, the articles reprinted here being indicated by their respective page numbers at the right.

Had influence already manifest been the basis of the choice, it would have dictated inclusion of No. 2, which was NOLL's thesis, but by now that work is in part obsolete, in part available through the intermediacy of a dozen books and a hundred papers. As a single example, I may adduce the fact that in Volume 1 of *A Supplement to the Oxford English Dictionary*, 1972, that thesis is the earliest source quoted in the article "constitutive equation", although, with a lapse of accuracy scarcely to be expected, NOLL's name is misspelled there. Equally certainly, on the basis of influence I should have had to exclude his latest development of the foundations, presented in No. 35, since its impact is not yet widely felt. In the past NOLL's major works have often seemed forbiddingly abstract and remote at first sight, and five to ten years have gone by before they came to be widely understood and applied. Had quality been the criterion, it would have forbidden me to exclude No. 3, his paper on statistical mechanics, or No. 33, his study of GIBBS' phase rule, but their subjects are peripheral to the field of this volume. Instead, the reader will find below two largely expository papers, Nos. 24 and 25, which present in typical simplicity and elegance propositions in algebra and geometry now frequently applied in continuum mechanics.

This preface is no place for an obituary of a man not yet fifty and in full power of work, but I will record the circumstances in which NOLL began to study the foundations of mechanics.

Students well trained in mathematics, as mathematicians understand that discipline, rarely take any interest in mechanics. Students of mechanics, all too often aping their teachers, neglect the prime lesson the history of that discipline affords, namely, that "the paradise of the mathematical sciences" flowers and bears its best fruit when mathematicians cultivate it with their most powerful tools – indeed, design those tools for that purpose. After the war the relict

Preface

German scholars, against odds now not only forgotten but also hard to imagine, were striving to revivify the life of the mind which the mental and physical barbarity preached and practised by the -isms and -acies of 1933–1946 had all but eradicated. Thinking that among the disciples of these elders, restorers rather than progressives, I might find a student or two who would wish to master new mathematics but grasp it and use it with the wholeness of earlier times, in 1952 I wrote to Mr. HAMEL, one of the few then remaining mathematicians from the classical mould, to ask him to name some young men fit to study for the doctorate in The Graduate Institute for Applied Mathematics at Indiana University, flourishing at that time though soon to be destroyed by the jealous ambition of the local, stereotyped pure. Having just retired from the Technische Universität in Charlottenburg, he passed my inquiry on to Mr. SZABŐ, in whose institute there NOLL was then an assistant. Although Mr. NOLL informs me that he had attended only one course by HAMEL, and that “there were very few students, and none of us understood what he was talking about”, nevertheless I like to think of WALTER NOLL as our link with the great Göttingen school, with the HILBERT who wrote:

What a vital nerve would be cut off from mathematics by rooting out geometry and mathematical physics! On the contrary, I think that wherever mathematical ideas come up, whether from the theory of knowledge or in geometry, or out of the theories of natural science, mathematics ought to investigate the principles underlying these ideas and by means of a simple and complete system of axioms establish them in such a way that in deduction the new ideas shall be no whit inferior to the old arithmetic concepts.

In Berlin NOLL had followed courses by E. SCHMIDT and HASSE, the former, like HAMEL, once a doctorand under HILBERT and the latter once a colleague. The volume of the *Journal of Rational Mechanics and Analysis* in which NOLL's thesis is published is dedicated *in memoriam* to HAMEL. The present volume of reprints will mark the twentieth anniversary of the defense of that thesis on August 9, 1954, in my office. The room itself, like all that we valued of our circumstances then, has since been destroyed.

Nowadays, when the common student seeks a secure berth by grafting himself upon some modest little professor whom he regards as prone to foster painlessly his limaceous glide toward a dissertation not too strenuous or, even better, to draught it for him, tradition is moribund, and we lightly disregard it. Mathematics once was transmitted almost like a priesthood, through novitiate, trial, and the laying on of hands. The burden lay upon the candidate, not upon his professor. Before NOLL came to Indiana, a year at the Sorbonne had given him some of the abstract, direct mathematics of BOURBAKI, which then seemed to him unrelated to natural science, but from which he afterward created what is now become the common dialect of continuum mechanics. That year brought him into contact with living exponents of another tradition, even more powerful than HILBERT's: the didactic, systematic rationalism of DESCARTES.

C. TRUESDELL

Publications of Walter Noll

The numbers prefixed to the papers reprinted in this volume are set in **bold** type in the list below. The pages on which they appear here are indicated at the right.

In the texts reprinted, both such errata as have been published before and also some further ones are included. Whenever it was feasible to do so, the text itself has been corrected, but in some few cases errata are listed at the ends of the works to which they pertain.

BOOKS

- I. *The Non-Linear Field Theories of Mechanics, Encyclopedia of Physics*, Volume III/3, 602 pages, Berlin-Heidelberg-New York, Springer-Verlag, 1965 (with C. TRUESDELL)
- II. *Viscometric Flows of Non-Newtonian Fluids, Theory and Experiment*, Springer Tracts in Natural Philosophy, Volume 5, 130 pages, Berlin-Heidelberg-New York, Springer-Verlag, 1966 (with B. D. COLEMAN and H. MARKOVITZ)

PAPERS AND NOTES

1. Eine Bemerkung zur Schwarzschen Ungleichheit, *Mathematische Nachrichten*, Volume 7, pp. 55–59 (1952) (with E. MOHR)
2. On the Continuity of the Solid and Fluid States, *Journal of Rational Mechanics and Analysis*, Volume 4, pp. 3–81 (1955)
3. Die Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik, *Journal of Rational Mechanics and Analysis*, Volume 4, pp. 627–646 (1955)
4. Verschiebungsfunktionen für elastische Schwingungsprobleme, *Zeitschrift für angewandte Mathematik und Mechanik*, Volume 37, pp. 81–87 (1957)
5. On the Uniqueness and Non-existence of Stokes Flow, *Archive for Rational Mechanics and Analysis*, Volume 1, pp. 97–106 (1957) (with R. FINN)
6. On the Rotation of an Incompressible Continuous Medium in Plane Motion, *Quarterly of Applied Mathematics*, Volume 15, pp. 317–319 (1957)
7. On Exterior Boundary Value Problems in Linear Elasticity, *Archive for Rational Mechanics and Analysis*, Volume 2, pp. 191–196 (1958) (with R. J. DUFFIN)
8. A Mathematical Theory of the Mechanical Behavior of Continuous Media, *Archive for Rational Mechanics and Analysis*, Volume 2, pp. 197–226 (1958)

9. The Foundations of Classical Mechanics in the Light of Recent Advances in Continuum Mechanics, pp. 266–281 of *The Axiomatic Method, with Special Reference to Geometry and Physics* (Symposium at Berkeley, 1957), Amsterdam, North-Holland Publishing Co., 1959 32–47
10. On Certain Steady Flows of General Fluids, *Archive for Rational Mechanics and Analysis*, Volume 3, pp. 289–303 (1959) (with B. D. COLEMAN) 49–63
11. Helical Flow of General Fluids, *Journal of Applied Physics*, Volume 30, pp. 1508–1512 (1959) (with B. D. COLEMAN)
12. Conditions for Equilibrium at Negative Absolute Temperatures, *The Physical Review*, Volume 115, pp. 262–265 (1959) (with B. D. COLEMAN)
13. On the ThermoStatics of Continuous Media, *Archive for Rational Mechanics and Analysis*, Volume 4, pp. 97–128 (1959) (with B. D. COLEMAN) 65–96
14. An Approximation Theorem for Functionals with Applications in Continuum Mechanics, *Archive for Rational Mechanics and Analysis*, Volume 6, pp. 355–370 (1960) (with B. D. COLEMAN) 97–112
15. Recent Results in the Continuum Theory of Viscoelastic Fluids, *Annals of the New York Academy of Science*, Volume 89, pp. 672–714 (1961) (with B. D. COLEMAN)
16. Foundations of Linear Viscoelasticity, *Reviews of Modern Physics*, Volume 33, pp. 239–249 (1961) (with B. D. COLEMAN) 113–123
17. Normal Stresses in Second-order Viscoelasticity, *Transactions of the Society of Rheology*, Volume 5, pp. 41–46 (1961) (with B. D. COLEMAN)
18. Steady Extension of Incompressible Simple Fluids, *The Physics of Fluids*, Volume 5, pp. 840–843 (1962) (with B. D. COLEMAN)
19. Simple Fluids with Fading Memory, pp. 530–552 of *Second-Order Effects in Elasticity, Plasticity, and Fluid Dynamics* (Symposium at Haifa, 1962), Oxford etc., Pergamon Press, 1964 (with B. D. COLEMAN)
20. Motions with Constant Stretch History, *Archive for Rational Mechanics and Analysis*, Volume 11, pp. 97–105 (1962) 125–133
21. La Mécanique Classique, Basée sur un Axiome d'Objectivité, pp. 47–56 of *La Méthode Axiomatique dans les Mécaniques Classiques et Nouvelles* (Colloque International, Paris, 1959), Paris, Gauthier-Villars, 1963 135–144
22. The Thermodynamics of Elastic Materials with Heat Conduction and Viscosity, *Archive for Rational Mechanics and Analysis*, Volume 13, pp. 167–178 (1963) (with B. D. COLEMAN) 145–156

23. Material Symmetry and Thermoelastic Inequalities in Finite Elastic Deformations, *Archive for Rational Mechanics and Analysis*, Volume 15, pp. 87–111 (1964) (with B. D. COLEMAN) 157–181
24. Euclidean Geometry and Minkowskian Chronometry, *American Mathematical Monthly*, Volume 71, pp. 129–144 (1964) 183–198
25. Proof of the Maximality of the Orthogonal Group in the Unimodular Group, *Archive for Rational Mechanics and Analysis*, Volume 18, pp. 100–102 (1965) 200–202
26. The Equations of Finite Elasticity, pp. 93–101 of *Symposium on Applications of Nonlinear Partial Differential Equations in Mathematical Physics* (1964), Providence, American Mathematical Society, 1965
27. The Foundations of Mechanics, pp. 159–200 of *Non-Linear Continuum Theories* (C. I. M. E. Lectures, 1965), Rome, Cremonesi, 1966
28. Space-Time Structures in Classical Mechanics, pp. 28–34 of *Delaware Seminar in the Foundations of Physics*, Berlin-Heidelberg-New York, Springer-Verlag, 1967 204–210
29. Materially Uniform Simple Bodies with Inhomogeneities, *Archive for Rational Mechanics and Analysis*, Volume 27, pp. 1–32 (1967) 211–242
30. Inhomogeneities in Materially Uniform Simple Bodies, pp. 239–246 of *IUTAM Symposium on Mechanics of Generalized Continua* (1967), Berlin-Heidelberg-New York, Springer-Verlag, 1968
31. Quasi-invertibility in a Staircase Diagram, *Proceedings of the American Mathematical Society*, Volume 23, pp. 1–4 (1969)
32. Representations of Certain Isotropic Tensor Functions, *Archiv der Mathematik*, Volume 21, pp. 87–90 (1970)
33. On certain Convex Sets of Measures and on Phases of Reacting Mixtures, *Archive for Rational Mechanics and Analysis*, Volume 38, pp. 1–12 (1970)
34. Annihilators of Linear Differential Operators, *Journal d'Analyse*, Volume 24, pp. 205–284 (1971) (with H. D. DOMBROWSKI)
35. A New Mathematical Theory of Simple Materials, *Archive for Rational Mechanics and Analysis*, Volume 48, pp. 1–50 (1972) 243–292
36. Lectures on the Foundations of Continuum Mechanics and Thermodynamics, *Archive for Rational Mechanics and Analysis*, Volume 52, pp. 62–92 (1973) 294–324

Several of the above-listed papers have been reprinted already in collective volumes, as follows:

Nos. 2, 8, 10, 11, and 14 in

Continuum Mechanics II, The Rational Mechanics of Materials, International Science Review Series, Volume VIII, Part 2, New York etc., Gordon & Breach, 1965.

Nos. 13 and 16 in

Continuum Mechanics III, Foundations of Elasticity Theory, International Science Review Series, Volume VIII, Part 3, New York etc., Gordon & Breach, 1965.

Nos. 8, 29, and 30 in

Continuum Theory of Inhomogeneities in Simple Bodies, New York, Springer-Verlag, 1968.

A Mathematical Theory of the Mechanical Behavior of Continuous Media

WALTER NOLL

Communicated by C. TRUESDELL

Contents

1. Introduction	198
I. Local kinematics	
2. Basic concepts	200
3. Deformations and linear transformations	200
4. Local configurations	202
5. Gradients	203
6. Rotation and strain tensors	204
7. Histories	205
8. Rate of strain and spin	205
9. Rational expressions for the rates	206
II. The general constitutive equation	
10. Basic concepts	207
11. The principle of objectivity of material properties	208
12. The principle of determinism for the stress	209
13. The general constitutive equation	210
14. Material isomorphisms	211
15. Constitutive functionals	212
16. The local isotropy group	213
III. Simple materials	
17. Simple constitutive functionals	214
18. Simple materials	216
19. The isotropy group	217
20. Isotropic and anisotropic solids	218
21. Fluids	219
22. Constitutive equations involving the Cauchy-Green tensors	220
IV. Special classes of materials	
23. Materials of the differential type	221
24. Materials of the rate type	223
References	226

1. Introduction

Until not long ago continuum mechanics meant to most people the theories of inviscid and linearly viscous fluids and of linearly elastic solids. However, the behavior of only few real materials can be described adequately by these classical theories. Experimental scientists, who had to deal with real materials, developed a science of non-classical materials called rheology. But they did not succeed in fitting their experimental results into a general mathematical framework. Most of the rheological theories are either one-dimensional, and hence appropriate at best for particular experimental situations, or are confined to infinitesimal deformations, in which case they are only of limited use, because large deformations occur easily in the materials these theories are intended to describe.

In the last few years a number of mathematicians have succeeded in devising general three-dimensional theories which are valid for large deformations*. Guided directly by physical experience or by the one-dimensional theories of the rheologists, these authors have proposed various special hypotheses intended to characterize the mechanical behavior of large classes of materials in arbitrary motions. Using the principle of objectivity of material properties (*cf.* Section 11 of this paper) and usually also the assumption of material isotropy** they then derived the general constitutive equation (also called rheological equation of state or stress-strain relation) compatible with the original hypothesis. In many cases they also solved particular problems in order to make possible a comparison with experimental results***.

This paper does not start with a special hypothesis. It starts (Chapter II) with a general principle, called *the principle of determinism for the stress*, which is implicit in all physical experience and applies to any material whatsoever as long as only its mechanical behavior is considered and as long as there are no constraints. From this principle and the principle of objectivity the most general constitutive equation for all materials is derived.

Constitutive equations characterize material properties. But material properties are of a local nature and may change from particle to particle in a body. It turned out to be necessary, therefore, to develop concepts which describe the local behavior of a motion at a particular particle. This is done in Chapter I. Some of the concepts introduced there are similar to those of the theory of "jets" introduced by EHRESMANN [2] into modern differential geometry. The treatment of strain and related notions given in Chapter I is, I believe, more concise and direct than previous treatments. It is based on the well known polar decomposition theorem for linear transformations.

With only one exception**** the ideal materials previously considered in the literature are special cases of what is called a *simple material* in Chapter III of

* A review of the work done up to 1953 was given by TRUESDELL [12]. For more recent researches we refer to [1], [4], [5], [8], [9], [13].

** These two things were not clearly distinguished for some time. A clarification was given by the author in [5], where the term "principle of isotropy of space" was used for the principle of objectivity.

*** *Cf.* [3], [5], [10], [12], [13], [14].

**** TRUESDELL's theory of Maxwellian fluids (*cf.* [12], Chapter V D).

the present paper*. There is reason to believe that most real materials are adequately covered by this category. A detailed discussion of the theory of these simple materials is given. It includes new and very general definitions of isotropic and anisotropic solids and of fluids in terms of the special invariances of the corresponding constitutive equations. Fluids are always isotropic in the sense in which the term is used in this paper. I believe that physically observed anisotropies in fluids are not really intrinsic in these fluids but depend on the past history of the substance.

In Chapter IV it is shown how two large classes of ideal materials fit into the general framework. In the isotropic case, these two classes were introduced by RIVLIN & ERICKSEN [8] and COTTER & RIVLIN [1]. The general non-isotropic case is discussed here, too.

The theory of simple materials presented in Chapter III generalizes the memory theory of linear visco-elasticity. The theory of materials of the rate type (Section 24) generalizes the differential operator theory of linear visco-elasticity. The materials of the differential type (Section 23) correspond to the special case when only the stress and not the stress rates occur in the basic equation.

Notation and terminology. Physicists customarily use symbols to denote physical quantities. They do not use symbols for the functions that relate these quantities, except sometimes the symbol f , which then stands for any function. Pure mathematicians, on the other hand, distinguish carefully between functions and their values, and they use different symbols for different functions. Both methods have advantages and disadvantages. I have found it necessary here to follow the pure mathematicians. Thus, in this paper, F means something entirely different in nature from $F(t)$. The first is a function, and the second a value of that function and hence a constant.

The terms "function" and "functional" are used here as synonyms for "mapping". A mapping is an operation that assigns to each element of one set, called the *domain* of the mapping, a value in another set. A mapping that has the same values as a given mapping α on a subset of the domain of α but is not defined outside this subset is called the *restriction* of α to the subset in question. Sometimes, if no confusion can arise, we use the same symbol α for such a restriction.

If α and β are two suitable mappings, then $\alpha \circ \beta$ denotes their *composition* defined by $(\alpha \circ \beta)(X) = \alpha(\beta(X))$. The inverse of a one-to-one mapping α is denoted by α^{-1} . If \mathcal{F} and \mathcal{G} are two sets, then $\mathcal{F} \times \mathcal{G}$ means the set of all ordered pairs of elements of \mathcal{F} and \mathcal{G} .

In the case of linear transformations F, G of a vector-space into itself we omit the symbol \circ and write simply FG for their composition. The inverse of a linear transformation G is denoted by G^{-1} , its transpose by G^T . The identity transformation is denoted by I . A linear transformation Q is orthogonal if and only if $QQ^T = I$, i.e., $Q^T = Q^{-1}$. The term *tensor* is used as a synonym for linear transformation. Tensors of order higher than two do not occur in this paper.

* More general materials are those "of order n ", discussed by the author in [7]. They include TRUESDELL's theory of Maxwellian fluids. Simple materials are those of order one.

Points and vectors of a Euclidean space are denoted by boldface letters, except when they are values of mappings. If \mathbf{x} and \mathbf{y} are two points, then $\mathbf{y} - \mathbf{x}$ denotes the vector leading from \mathbf{x} to \mathbf{y} . If \mathbf{x} is a point and \mathbf{u} a vector, then $\mathbf{x} + \mathbf{u}$ denotes the point \mathbf{y} uniquely determined by $\mathbf{u} = \mathbf{y} - \mathbf{x}$.

I. Local kinematics

2. Basic concepts

Precise axiomatic definitions of the basic concepts used in this paper are given in [6]. Here is a brief summary:

A *body* \mathcal{B} is a smooth manifold of elements X, Z, \dots , called *particles*. The *configurations* $\varphi, \vartheta, \dots$ of \mathcal{B} are the elements of a set of one-to-one mappings of \mathcal{B} into a three-dimensional Euclidean point space \mathcal{E} . The vector space associated with \mathcal{E} will be denoted by \mathcal{V} . The *mass-distribution* m of \mathcal{B} is a measure defined on all Borel subsets of \mathcal{B} . If φ is a configuration of \mathcal{B} , then there is a corresponding *mass density* ρ_φ such that

$$(2.1) \quad m(\mathcal{C}) = \int_{\varphi(\mathcal{C})} \rho_\varphi(\mathbf{X}) dV$$

for all Borel subsets \mathcal{C} of \mathcal{B} . A *motion* is a one-parameter family $\{\vartheta_\tau\}$, $-\infty < \tau < \infty$, of configurations. The parameter τ is called the *time*. A motion will often simply be denoted by ϑ , which is then regarded as a point-valued function of two variables, a particle and a time. The mass density corresponding to the configuration ϑ_τ is denoted by $\rho(\tau)$.

A body, its configurations, its mass-distribution, and its possible motions are subject to the axioms given in [6]. These axioms insure that the customary notions and operations are meaningful. However, in some of the considerations of this paper, continuity and differentiability conditions stronger than those implied by the axioms of [6] have to be assumed. It will be clear from the context when this is the case, and we shall not mention these conditions explicitly.

3. Deformations and linear transformations

A smooth homeomorphism δ which maps a neighborhood of the null-vector \mathbf{O} of the vector space \mathcal{V} onto another such neighborhood and which maps \mathbf{O} into itself,

$$(3.1) \quad \delta(\mathbf{O}) = \mathbf{O},$$

shall be called a *local homeomorphism* of \mathcal{V} . We define an equivalence relation among all local homeomorphisms by local identity: $\delta \sim \tilde{\delta}$ if and only if $\delta(\mathbf{P}) = \tilde{\delta}(\mathbf{P})$ for all \mathbf{P} in some neighborhood of \mathbf{O} , however small. The resulting equivalence classes Δ will be called *deformations*. The equivalence class of $\delta \circ \hat{\delta}$ depends only on the equivalence classes Δ and $\hat{\Delta}$ of δ and $\hat{\delta}$, and it will therefore be denoted by $\Delta \circ \hat{\Delta}$. With the law of composition thus defined, the deformations form a group denoted by \mathcal{D} .

A *linear transformation* G , i.e., a mapping of \mathcal{V} onto itself such that

$$(3.2) \quad G(\mathbf{P}_1 + \mathbf{P}_2) = G(\mathbf{P}_1) + G(\mathbf{P}_2), \quad G(a\mathbf{P}) = aG(\mathbf{P}),$$

is a local homeomorphism provided that it is regular, *i.e.*, provided that $G\mathbf{P} = \mathbf{O}$ is possible only for $\mathbf{P} = \mathbf{O}$. Two different linear transformations are never equivalent in the sense defined above. Hence a regular linear transformation defines a unique deformation and may thus be regarded as a special deformation. The regular linear transformations thus form a subgroup of the group of deformations \mathcal{D} , called the *linear group* and denoted by \mathcal{L} .

The gradient at $\mathbf{P} = \mathbf{O}$ of a local homeomorphism δ is a local property of δ and hence depends only on the equivalence class Δ to which δ belongs. We can, therefore, define the *gradient* $\nabla\Delta$ of a deformation Δ by

$$(3.3) \quad \nabla\Delta = \nabla\delta(\mathbf{O}) \quad \text{if } \delta \in \Delta.$$

The gradient $\nabla\Delta$ is a regular linear transformation and hence itself a deformation. Since, by the chain rule,

$$(3.4) \quad \nabla(\Delta \circ \hat{\Delta}) = (\nabla\Delta)(\nabla\hat{\Delta}),$$

it follows that the gradient operation ∇ is an endomorphism of the group \mathcal{D} . Its kernel is the group \mathcal{N} of all *null-deformations*, *i.e.*, all deformations whose gradient is the identity I . The quotient group \mathcal{D}/\mathcal{N} consists of the equivalence classes of deformations defined by the equivalence relation

$$(3.5) \quad \Delta \sim \tilde{\Delta} \quad \text{if and only if } \nabla\Delta = \nabla\tilde{\Delta}.$$

The deformations in each class have the same gradient and any representing local homeomorphisms differ only by small terms of order one. The relation (3.5) may be read: Δ and $\tilde{\Delta}$ are equal up to a small term of order one. Each class has a unique representative which is a linear transformation. We have

$$(3.6) \quad \Delta \sim \nabla\Delta,$$

i.e. any deformation differs from its gradient only by a small term of order one. In the case when $\Delta = G$ is a linear transformation, we have

$$(3.7) \quad \nabla G = G.$$

The quotient group \mathcal{D}/\mathcal{N} is isomorphic to the linear group \mathcal{L} .

A deformation Δ will be called *isochoric* if there is a volume preserving local homeomorphism δ in the class Δ . The isochoric deformations form a subgroup \mathcal{I} of \mathcal{D} . The null-deformations are isochoric, and \mathcal{N} is a subgroup of \mathcal{I} . A linear transformation G is isochoric if and only if

$$(3.8) \quad |\det G| = 1.$$

The isochoric linear transformations form a subgroup \mathcal{U} of \mathcal{L} , called the *unimodular group*. They are also called *unimodular transformations*. The group of *orthogonal transformations* is a subgroup of \mathcal{U} ; it will be denoted by \mathcal{O} . The space of all *symmetric linear transformations* will be denoted by \mathcal{S} . The set of all *positive definite* and symmetric linear transformations will be denoted by \mathcal{S}_+ . It is a subset of the linear group \mathcal{L} .

4. Local configurations

Consider a neighborhood $\mathcal{N}(X)$ of a particle X in a body \mathcal{B} , *i.e.*, a part of \mathcal{B} containing X . Let ψ be a smooth homeomorphism of $\mathcal{N}(X)$ into the vector space \mathcal{V} mapping X itself into the null-vector,

$$(4.1) \quad \psi(X) = \mathbf{0}.$$

We define an equivalence relation among all such homeomorphisms by the condition that $\psi \sim \tilde{\psi}$ if and only if $\psi(Z) = \tilde{\psi}(Z)$ for all Z in some neighborhood of X . The resulting equivalence classes Φ will be called the *local configurations* of X . The set of all local configurations of X will be denoted by \mathcal{C}_X .

Let δ be a local homeomorphism of \mathcal{V} . The equivalence class of the mapping $\delta \circ \psi$ depends only on the equivalence class Φ of ψ and on the deformation Δ of δ . Therefore, it is meaningful to speak about the local configuration $\Delta \circ \Phi$ of X , which is the composite of the local configuration Φ of X and the deformation Δ . If Φ and $\hat{\Phi}$ are two local configurations of X , we define

$$(4.2) \quad \Delta = \hat{\Phi} \circ \Phi^{-1} \quad \text{if} \quad \hat{\Phi} = \Delta \circ \Phi.$$

This Δ will be called the *deformation from the configuration Φ to the configuration $\hat{\Phi}$* .

Let φ be a configuration of the body \mathcal{B} . Then $\psi = \varphi - \varphi(X)$ maps \mathcal{B} into the vector space \mathcal{V} and has the property (4.1). Its equivalence class will be denoted by $\Phi(\varphi, X)$, and it will be called the *localization* at X of φ . Assume that a motion ϑ of \mathcal{B} is given. We then use the notation

$$(4.3) \quad \Theta(\tau) = \Phi(\vartheta_\tau, X),$$

and we call Θ the *localization* at X of the motion ϑ . We note that Θ is a function of a real variable whose values are in \mathcal{C}_X . Any sufficiently smooth function of this type will be called a *local motion* of X .

Let Θ be a local motion and Φ a local configuration of X . We then write

$$(4.4) \quad \Omega(\tau) = \Theta(\tau) \circ \Phi^{-1}$$

and call Ω the *deformation function* of X in the local motion Θ relative to the *local reference configuration* Φ . Ω is a function of a real variable with values in \mathcal{D} . Any sufficiently smooth function of this type will be called a *deformation function*.

In the special case when $\Phi = \Theta(t)$ in (4.4), where t is a particular time, we use the notation

$$(4.5) \quad \Omega_t(\tau) = \Theta(\tau) \circ \Theta(t)^{-1}$$

and call Ω_t the *deformation function* of X in Θ relative to the time t .

Let φ be a configuration of the body \mathcal{B} and ρ_φ the corresponding mass density. The value $\rho_\varphi(\varphi(X))$ depends only on the localization $\Phi(\varphi, X)$ of φ at X . We can, therefore, define

$$(4.6) \quad \rho_\Phi = \rho_\varphi(\varphi(X)), \quad \Phi = \Phi(\varphi, X),$$

and call ρ_Φ the *mass density* of X in the local configuration Φ .

5. Gradients

Assume that two local configurations Φ and $\hat{\Phi}$ of a particle X are given. The regular linear transformation

$$(5.1) \quad G = G(\hat{\Phi}, \Phi) = \nabla \Delta, \quad \Delta = \hat{\Phi} \circ \Phi^{-1}$$

will then be called the *deformation gradient* from Φ to $\hat{\Phi}$.

In the case when a local motion Θ and a local reference configuration Φ are given, we use the notation

$$(5.2) \quad F(\tau) = \nabla \Omega(\tau) = G(\Theta(\tau), \Phi),$$

$$(5.3) \quad F_i(\tau) = \nabla \Omega_i(\tau) = G(\Theta(\tau), \vartheta(t)).$$

Assume that $\hat{\Phi}$ is another local reference configuration, and let

$$(5.4) \quad \hat{F}(\tau) = \nabla \hat{\Omega}(\tau) = G(\Theta(\tau), \hat{\Phi}).$$

Then, by the chain rule for gradients, we have

$$(5.5) \quad F(\tau) = \hat{F}(\tau) G,$$

where $G = G(\hat{\Phi}, \Phi)$ is defined by (5.1). We note the following formulas:

$$(5.6) \quad G(\hat{\Phi}, \Phi) = G(\Phi, \hat{\Phi})^{-1},$$

$$(5.7) \quad F(\tau) = F_i(\tau) F(t), \quad F_i(t) = I.$$

We define an equivalence relation among all local configurations $\Phi \in \mathcal{C}_X$ by

$$(5.8) \quad \Phi \sim \tilde{\Phi} \quad \text{if and only if} \quad G(\tilde{\Phi}, \Phi) = I.$$

The corresponding equivalence classes M will be called the *configuration gradients* at X . The class of local configurations equivalent to a given $\Phi \in \mathcal{C}_X$ is denoted by $\nabla \Phi$, so that

$$(5.9) \quad M = \nabla \Phi \quad \text{if and only if} \quad \Phi \in M.$$

The equivalence class of the local configuration $\Delta \circ \Phi$, where Δ is a deformation, depends only on the equivalence class of Φ and on the gradient $\nabla \Delta$ of Δ . Therefore, it is meaningful to speak about the configuration gradient GM , which is the product of the configuration gradient M and the regular linear transformation G . This product is characterized by the property

$$(5.10) \quad \nabla(\Delta \circ \Phi) = (\nabla \Delta) (\nabla \Phi).$$

If M and \hat{M} are two configuration gradients, we define

$$(5.11) \quad G = \hat{M} M^{-1} \quad \text{if} \quad \hat{M} = G M.$$

We have

$$(5.12) \quad G(\hat{\Phi}, \Phi) = \hat{M} M^{-1}, \quad M = \nabla \Phi, \quad \hat{M} = \nabla \hat{\Phi},$$

which shows that $G(\hat{\Phi}, \Phi)$ depends only on the gradients of Φ and $\hat{\Phi}$. In particular, the deformation gradient function F in (5.2) depends only on $\nabla \Phi$, and we may call

$$(5.13) \quad F(\tau) = (\nabla \Theta(\tau)) M^{-1}$$

the deformation gradient at time τ of the local motion Θ relative to the configuration gradient M as a reference.

If Φ and $\hat{\Phi}$ are two local configurations at X , then the corresponding densities ϱ_Φ and $\varrho_{\hat{\Phi}}$ are related by

$$(5.14) \quad \varrho_{\hat{\Phi}} = |\det G(\hat{\Phi}, \Phi)| \varrho_\Phi.$$

It follows from (5.8) that $\varrho_{\hat{\Phi}} = \varrho_\Phi$ if $\hat{\Phi}$ and Φ have the same configuration gradient. Hence we can define

$$(5.15) \quad \varrho_M = \varrho_\Phi \quad \text{if} \quad M = \nabla \Phi,$$

and we call ϱ_M the mass density for the configuration gradient M at X . Equation (5.12) then implies

$$(5.16) \quad \varrho_{GM} = |\det G| \varrho_M.$$

It follows that

$$(5.17) \quad \varrho_M = \varrho_{\hat{M}} \quad \text{if and only if} \quad G = \hat{M}M^{-1} \quad \text{is unimodular.}$$

6. Rotation and strain tensors

Of fundamental importance for the local kinematics of continuous media is the well known *polar decomposition theorem*:

Let F be any regular linear transformation. Then there are unique decompositions

$$(6.1) \quad F = RU, \quad F = VR$$

in which R is orthogonal and U and V are symmetric and positive definite, *i.e.*, $R \in \mathcal{O}$, $U, V \in \mathcal{S}_+$. The following relations are valid:

$$(6.2) \quad U^2 = F^T F, \quad V^2 = F F^T,$$

$$(6.3) \quad V = R U R^T, \quad V^2 = R U^2 R^T.$$

This theorem, when applied to a deformation gradient $F = G(\hat{\Phi}, \Phi)$, gives rise to the following terminology: We call R the *rotation tensor*, U the *right strain tensor*, and V the *left strain tensor* of the deformation from Φ to $\hat{\Phi}$. The squares

$$(6.4) \quad C = U^2 = F^T F, \quad B = V^2 = F F^T = R C R^T$$

will be called the *right* and the *left Cauchy-Green tensors* of the deformation from Φ to $\hat{\Phi}$.

The strain tensors U and V describe adequately what is meant physically by strain because their eigenvalues are the principal extension ratios of the deformation from Φ to $\hat{\Phi}$. However, it is often of advantage to use the Cauchy-Green tensors rather than the strain tensors as measures of strain because their components are rational functions of the components of the deformation gradient F , while the components of U and V are complicated irrational functions of the components of F . Both C and B are, of course, symmetric and positive definite. Their eigenvalues are the squares of the principal extension ratios of the deformation.

In the special case when F is the tensor function defined by (5.2) R, U, V, C , and B will also be tensor functions of a real variable. If F is replaced by the tensor function F_t defined in (5.3), we use the analogous notation R_t, U_t, V_t, C_t , and B_t . These functions have the same smoothness properties as F or F_t , respectively. We note that

$$(6.5) \quad F_t(t) = R_t(t) = U_t(t) = V_t(t) = C_t(t) = B_t(t) = I.$$

7. Histories

Let \mathcal{F} be an arbitrary set. The class of all functions with values in \mathcal{F} whose domain is the negative real axis will then be denoted by \mathcal{F}^* .

Let α be any function of a real variable with values in \mathcal{F} . We then define $\alpha^t \in \mathcal{F}^*$ by

$$(7.1) \quad \alpha^t(s) = \alpha(t + s) \quad \text{for } s \leq 0,$$

and we call it the *history up to time t* of the function α .

For functions $\alpha^* \in \mathcal{F}^*$ we generally use the notation

$$(7.2) \quad \alpha_0 = \alpha^*(0).$$

Then

$$(7.3) \quad \alpha_0^t = \alpha^t(0) = \alpha(t).$$

Let Θ be a local motion of a particle X , so that $\Theta(\tau) \in \mathcal{C}_X$. The corresponding history $\Theta^t \in \mathcal{C}_X^*$, defined by

$$(7.4) \quad \Theta^t(s) = \Theta(t + s) \quad \text{for } s \leq 0,$$

will then be called the *kinematical history up to t* of X . If, in addition, a local reference configuration Φ is given and if $\Omega = \Theta \circ \Phi^{-1}$ is the corresponding deformation function, we call $\Omega^t \in \mathcal{D}^*$, defined by

$$(7.5) \quad \Omega^t(s) = \Omega(t + s) \quad \text{for } s \leq 0,$$

the *deformation history up to t* of X .

8. Rate of strain and spin

Consider a local motion Θ of a particle X . Let F_t be the corresponding deformation gradient function defined by (5.3) and F_t^t the history of F_t up to time t defined according to (7.1).

We then call \star

$$(8.1) \quad E(t) = \dot{F}_t(t) = \dot{F}_t^t(0)$$

the *velocity gradient* at time t of the local motion Θ . Similarly \star

$$(8.2) \quad E_n(t) = \overset{(n)}{F}_t(t) = \overset{(n)}{F}_t^t(0), \quad n = 0, 1, 2, \dots$$

\star A superposed dot denotes the operation of differentiation and a superposed (n) its n^{th} iteration.

is called the n^{th} *acceleration gradient* at time t of Θ . We have, by (6.5) and (8.1),

$$(8.3) \quad E_0 = I, \quad E_1 = E.$$

If Θ is the localization at X of a global motion ϑ , then $E(t)$ ($E_n(t)$) actually coincides with the gradient with respect to position of the velocity (n^{th} acceleration) at $\mathbf{x} = \vartheta_t(X)$ and at time t .

The polar decomposition

$$(8.4) \quad F_t(\tau) = R_t(\tau) U_t(\tau)$$

defines the rotation tensor $R_t(\tau)$ and the right strain tensor $U_t(\tau)$ of the deformation from the configuration at time t to the configuration at time τ . The histories R_t^t and U_t^t up to time t of R_t and U_t are defined according to (7.1). The tensor

$$(8.5) \quad W(t) = \dot{R}_t(t) = \dot{R}_t^t(0)$$

is called the *spin* at time t , and the tensor

$$(8.6) \quad D(t) = \dot{U}_t(t) = \dot{U}_t^t(0)$$

the *rate of strain* at time t . Similarly, we define the n^{th} *spin* W_n and the n^{th} *rate of strain* D_n by

$$(8.7) \quad W_n(t) = \dot{R}_t^n(t) = \dot{R}_t^{n,t}(0),$$

$$(8.8) \quad D_n(t) = \dot{U}_t^n(t) = \dot{U}_t^{n,t}(0).$$

Replacing U_t in (8.8) by the right Cauchy-Green tensor $C_t = U_t^2$, we get a tensor

$$(8.9) \quad A_n(t) = \dot{C}_t^n(t) = \dot{C}_t^{n,t}(0),$$

which we shall call the n^{th} *Rivlin-Ericksen tensor**. By (6.5), (8.5), and (8.6) we have

$$(8.10) \quad R_0 = D_0 = A_0 = I, \quad R_1 = R, \quad D_1 = D.$$

One could give definitions similar to (8.8) and (8.9) by replacing the right strain and Cauchy-Green tensors by their left counterparts. But the tensors defined in this way are of little interest. Moreover, for $n=1$ one would get nothing new, because

$$(8.11) \quad D(t) = \dot{U}_t(t) = \dot{V}_t(t),$$

as is not hard to prove.

9. Rational expressions for the rates

It is possible to find explicit expressions for W_n , D_n , and A_n as polynomials in the acceleration gradients E_k and their transpositions E_k^T , $k=1, 2, \dots, n$. Differentiating $C_t = F_t^T F_t$ [cf. (6.4)] n times, by the product rule we obtain

$$C_t^n(\tau) = \sum_{k=0}^n \binom{n}{k} F_t^{k,T}(\tau) F_t^{(n-k)}(\tau).$$

* RIVLIN & ERICKSEN [8] recognized the importance of these tensors and used them extensively.

Substituting $\tau = t$ yields, according to (8.2) and (8.9),

$$(9.1) \quad A_n = \sum_{k=0}^n \binom{n}{k} E_k^T E_{n-k}.$$

Since $E_0 = I$ by (8.3), this may be written in the form

$$(9.2) \quad A_n = E_n + E_n^T + \sum_{k=1}^{n-1} \binom{n}{k} E_k^T E_{n-k}.$$

Differentiating $C_i^{(\tau)} = U_i^2(\tau)$ and then substituting $\tau = t$ we get

$$(9.3) \quad A_n = 2D_n + \sum_{k=1}^{n-1} \binom{n}{k} D_k D_{n-k},$$

and hence

$$(9.4) \quad D_n = \frac{1}{2} \left\{ A_n - \sum_{k=1}^{n-1} \binom{n}{k} D_k D_{n-k} \right\}.$$

This is a recursion formula which can be used to find explicit expressions for the D_n as polynomials in the A_k , $k = 1, 2, \dots, n$. Hence, after substitution of (9.2), one would also obtain explicit expressions for the D_n as polynomials in the E_k and E_k^T , $k = 1, 2, \dots, n$.

Differentiating $F_i(\tau) = R_i(\tau) U_i(\tau)$ and then putting $\tau = t$, we find

$$(9.5) \quad E_n = W_n + D_n + \sum_{k=1}^{n-1} \binom{n}{k} W_k D_{n-k},$$

and hence

$$(9.6) \quad W_n = E_n - D_n - \sum_{k=1}^{n-1} \binom{n}{k} W_k D_{n-k}.$$

This is again a recursive formula which permits us to express W_n as a polynomial in the E_k and D_k . Since the D_k are polynomials in the E_k and E_k^T , we can also find expressions for the W_n as polynomials in the E_k and E_k^T , $k = 1, 2, \dots, n$.

In the special case $n = 1$ we find

$$(9.7) \quad D = \frac{1}{2} A_1 = \frac{1}{2} (E + E^T), \quad W = \frac{1}{2} (E - E^T).$$

It follows that the spin W is skew ($W^T = -W$). The higher spins W_n , $n > 1$ are not necessarily skew. Of course, the D_n and A_n are all symmetric. For $n = 2$ we find

$$(9.8) \quad \begin{aligned} A_2 &= E_2 + E_2^T + 2E^T E, \\ D_2 &= \frac{1}{4} (A_2 - \frac{1}{2} A_1^2) = \frac{1}{2} [E_2 + E_2^T - (E + E^T)^2] + E^T E, \\ W_2 &= \frac{1}{2} [E_2 - E_2^T - E^T (E - E^T)]. \end{aligned}$$

II. The general constitutive equation

10. Basic concepts

A *dynamical process* for a body \mathcal{B} is defined as a motion ϑ of \mathcal{B} coupled with a system of forces for \mathcal{B} at each time τ , subject to the principle of linear momentum and the principle of angular momentum (*cf.* [6], Section 5). A system of forces

can be split into body forces and contact forces. If sufficient continuity assumptions are made, the contact forces are determined, for each time τ , by a field of symmetric *stress tensors* $S(\tau)$. If a motion ϑ and smooth fields of symmetric tensors $S(\tau)$ are arbitrarily prescribed, one can always find a dynamical process such that $S(\tau)$ is the corresponding stress tensor field at time τ . Appropriate body forces can be chosen, for example, by putting the mutual body forces equal to zero and by letting the density \mathbf{b} of the external body forces be given by

$$(10.1) \quad \mathbf{b} = \dot{\mathbf{v}} - \frac{1}{\rho} \operatorname{div} S,$$

where $\dot{\mathbf{v}}$ is the acceleration of the motion ϑ . Then CAUCHY'S law of motion ([6], (5.25)) holds, and the principles of dynamics are satisfied.

If we subsequently talk about a *process* for a body \mathcal{B} we mean a pair $\{\vartheta, S\}$, where ϑ is an arbitrary motion of \mathcal{B} and S an arbitrary time family of symmetric stress fields. We can always adjoin suitable body forces to a given process so as to make it a dynamical process.

A *local process* for a particle $X \in \mathcal{B}$ is defined as a pair $\{\Theta, S\}$, where Θ is a local motion of X and S a symmetric tensor function of a real variable. The *localization* at X of a process $\{\vartheta, S\}$ is defined by replacing ϑ by its localization Θ at X and S by its values for X . We shall use the same symbol S for these values and write $S(\tau) = S(\tau, X)$. No confusion can arise because, in all local considerations, the particle X will be fixed.

A *constitutive assumption* is a restrictive condition on the possible dynamical processes a body can undergo, and it characterizes the material properties of the body. Restrictive conditions on the possible motions alone are particular constitutive assumptions; they are called *constraints*. Examples of constraints are rigidity (every possible motion is a rigid motion) and incompressibility (every possible motion is isochoric). In this paper, we shall assume that there are no constraints, *i.e.* that all motions ϑ are possible. We shall investigate constitutive assumptions in the form of functional relations between the stress S and the motion ϑ of a process $\{\vartheta, S\}$. Such relations will be called *constitutive equations*.

11. The principle of objectivity of material properties

A *change of frame* is a transformation of space and time specified by a point-valued function \mathbf{c} of a real variable, a function Q of a real variable whose values are orthogonal transformations, and a real constant a . It transforms a pair $\{\mathbf{x}, \tau\}$ consisting in a point \mathbf{x} and a time τ into the pair $\{\mathbf{x}', \tau'\}$ given by*

$$(11.1) \quad \mathbf{x}' = \mathbf{c}(\tau) + Q(\tau)(\mathbf{x} - \mathbf{O}),$$

$$(11.2) \quad \tau' = \tau - a,$$

where \mathbf{O} is an arbitrary point, the same for all transformations. Vectors $\mathbf{u} \in \mathcal{V}$ transform according to

$$(11.3) \quad \mathbf{u}' = Q(\tau)\mathbf{u}.$$

* TOUPIN has analyzed these transformations recently [11]; he calls them "Euclidean transformations of space-time".

If G is a linear transformation and $\mathbf{u} \in \mathcal{V}$, then $G\mathbf{u} \in \mathcal{V}$. Hence, by (11.3), $(G\mathbf{u})' = Q(\tau)G\mathbf{u}$. Defining the transform G' of G by $(G\mathbf{u})' = G'\mathbf{u}'$, we get $Q(\tau)G\mathbf{u} = G'Q(\tau)\mathbf{u}$, and hence the following law of transformation for second order tensors:

$$(11.4) \quad G' = Q(\tau)GQ(\tau)^T.$$

Two dynamical processes are called *equivalent* if they are related by a change of frame as made precise in Definition 7 of [6]. Under such a change of frame the contact forces transform in an objective manner, *i.e.*, according to the law (11.3). Hence the stress tensor must transform according to the law (11.4). The motion ϑ determines the position of the particles and hence transforms according to (11.1). Disregarding the body forces, we can say that *two processes* $\{\vartheta, S\}$ and $\{\vartheta', S'\}$ are equivalent if they are related by a change of frame in the form

$$(11.5) \quad \begin{aligned} \vartheta'(Z, \tau') &= \mathbf{c}(\tau) + Q(\tau)[\vartheta(Z, \tau) - \mathbf{O}], \\ S'(Z, \tau') &= Q(\tau)S(Z, \tau)Q^T(\tau), \\ \tau' &= \tau - a. \end{aligned}$$

Constitutive equations are subject to the following invariance requirement:

Principle of objectivity of material properties. *If a process $\{\vartheta, S\}$ is compatible with a constitutive equation, then also all processes $\{\vartheta', S'\}$ equivalent to it must be compatible with the same constitutive equation.*

This is a special case of the general principle of objectivity stated in [6]. Its physical meaning is simply that the material properties of a body should not depend on the observer, no matter how he moves.

12. The principle of determinism for the stress

We ask for guiding principles which will enable us to find the most general form of a constitutive equation. The following two are implied by physical experience:

(i) The stress at a particle X should depend only on the physical state of an arbitrarily small neighborhood of X . The state of parts distant from X should have no direct influence on the stress at X . This condition is implicit in the concept of a contact force.

(ii) The physical state of a body at a time t should depend only on its past history, *i.e.*, on what happened to it at times $\tau \leq t$, and not on its future, *i.e.*, on what will happen to it at times $\tau > t$. This condition expresses the causality of natural processes.

The physical history of a body consists of several components: its kinematical history, its thermodynamic history, its electromagnetic history, its chemical history, etc. In reality, each of these components may influence the stress. But, in this investigation, we shall disregard all non-mechanical influences.

Principle of determinism. *The stress $S(t)$ at a particle X and at time t is determined by the past history of the motion of an arbitrarily small neighborhood of X .*

In precise terms, this principle states that

(a) the stress $S(t)$ at X is a *functional* \mathfrak{F}_t of the motion ϑ ,

$$(12.1) \quad S(t) = \mathfrak{F}_t(\vartheta)$$

the domain of \mathfrak{F}_t being the class of all possible motions and its values being symmetric tensors,

(b) for any two motions ϑ and $\bar{\vartheta}$ which coincide in some neighborhood of X for times $\tau \leq t$ the value of the functional \mathfrak{F}_t is the same, *i.e.*,

$$(12.2) \quad \mathfrak{F}_t(\vartheta) = \mathfrak{F}_t(\bar{\vartheta})$$

whenever

$$(12.3) \quad \vartheta(Z, \tau) = \bar{\vartheta}(Z, \tau)$$

for $\tau \leq t$ and Z in some neighborhood $\mathcal{N}(X)$ of X , however small.

13. The general constitutive equation

Not any arbitrary functional \mathfrak{F}_t with the Property (b) is permissible in (12.1), because (12.1), as a constitutive equation, is subject to the principle of objectivity stated in Section 11. This principle states that (12.1) must hold equally if the process $\{\vartheta, S\}$ is replaced by any equivalent process $\{\vartheta', S'\}$, *i.e.*, that

$$(13.1) \quad S'(t') = \mathfrak{F}_{t'}(\vartheta')$$

must be valid for any process $\{\vartheta', S'\}$ related to $\{\vartheta, S\}$ by a transformation of the form (11.5). Choosing $Q = I$, $a = 0$, and $\mathbf{c}(\tau) = \mathbf{O} - [\vartheta_\tau(X) - \mathbf{O}]$ in (11.5), we get $S' = S$, $\tau' = \tau$, and $\vartheta'_\tau = \mathbf{O} + [\vartheta_\tau - \vartheta_\tau(X)]$. It follows that $S(t) = S'(t) = \mathfrak{F}_t(\vartheta')$ holds and hence that $S(t)$ depends only on the vector-valued function ψ defined by $\psi(Z, \tau) = \vartheta(Z, \tau) - \vartheta(X, \tau)$. Moreover, the Property (b) of \mathfrak{F}_t implies that $S(t)$ can depend only on the local behavior of ψ in an arbitrarily small neighborhood of X . This means that $S(t)$ is determined by the localization Θ at X of the motion ϑ as defined by (4.3). Thus, (12.1) reduces to the form

$$(13.2) \quad S(t) = \mathfrak{F}_t(\Theta)$$

where \mathfrak{F}_t is a functional with the property that

$$(13.3) \quad \mathfrak{F}_t(\Theta) = \mathfrak{F}_t(\bar{\Theta})$$

whenever

$$(13.4) \quad \Theta(\tau) = \bar{\Theta}(\tau) \quad \text{for} \quad \tau \leq t.$$

Now we consider another equivalent process by choosing $\mathbf{c}(\tau) = \mathbf{O}$, $Q = I$, and $a = t$ in (11.5). We then have

$$\tau' = \tau - t, \quad t' = t - t = 0, \quad S'(t') = S'(0) = S(t),$$

and, for the corresponding local motion Θ ,

$$\Theta'(\tau') = \Theta(\tau) = \Theta(t + \tau').$$

It follows from (13.4) that

$$(13.5) \quad S(t) = S'(0) = \mathfrak{F}_0(\Theta').$$

By (13.4) we have $\mathfrak{F}_0(\Theta') = \mathfrak{F}_0(\bar{\Theta}')$ whenever $\Theta'(\tau') = \bar{\Theta}'(\tau')$ for $\tau' \leq 0$, which shows that the value $\mathfrak{F}_0(\Theta')$ depends only on the restriction of the function $\Theta'(\tau')$ to $\tau' \leq 0$. But this restriction is nothing but the kinematical history $\Theta^t \in \mathcal{C}_X^*$ up to time t as defined by (7.4). Omitting the index 0 in (13.5), we see that the constitutive equation (13.2) reduces to the form

$$(13.6) \quad S(t) = \mathfrak{F}(\Theta^t),$$

where Θ^t is the kinematical history of X for the local motion Θ . The form of the functional \mathfrak{F} is independent of t and depends only on the particle X . Its domain is \mathcal{C}_X^* .

We finally consider an equivalent process by choosing $\mathbf{c}(\tau) = \mathbf{O}$ and $\mathbf{a} = 0$ in (11.5), leaving the orthogonal tensor function Q arbitrary. The principle of objectivity then implies that the functional \mathfrak{F} of (13.6) must satisfy the relation

$$(13.7) \quad Q_0 \mathfrak{F}(\Theta^*) Q_0^T = \mathfrak{F}(Q^* \circ \Theta^*), \quad Q_0 = Q^*(0),$$

for all kinematical histories $\Theta^* \in \mathcal{C}_X^*$ and all orthogonal tensor functions $Q^* \in \mathcal{O}^*$.

It is not hard to see that, conversely, the principle of determinism and the principle of objectivity are automatically satisfied for any equation of the form (13.6) provided the functional \mathfrak{F} has the property (13.7). Hence (13.6) with (13.7) is the most general constitutive equation. An equation of this form restricts the class of all possible local processes $\{\Theta, S\}$ for the particle X and characterizes the local material properties of X .

If we subsequently speak about a particle X , we always assume that a functional of the type discussed above is associated with it. We call it the *functional of the particle X* .

14. Material isomorphisms

The nature of the domain \mathcal{C}_X^* of the functional \mathfrak{F} of the particle X varies with the particle X . Hence there is no direct way to compare the functionals of different particles. It is desirable to render such a comparison possible, because only then can a precise meaning be given to the statement that two different particles consist of the same material.

A body was defined in [6] as a certain mathematical structure. As in the case with any such structure, it is meaningful to talk about isomorphisms between bodies. An *isomorphism* of a body \mathcal{B} onto a body $\bar{\mathcal{B}}$ is a one-to-one mapping γ of \mathcal{B} onto $\bar{\mathcal{B}}$ such that

(a) the configurations $\bar{\varphi}$ of $\bar{\mathcal{B}}$ are of the form

$$(14.1) \quad \bar{\varphi} = \varphi \circ \gamma^{-1},$$

(b) the mass distributions m and \bar{m} of \mathcal{B} and $\bar{\mathcal{B}}$ are related so that

$$(14.2) \quad m(\mathcal{C}) = \bar{m}(\gamma(\mathcal{C}))$$

for all Borel subsets \mathcal{C} in \mathcal{B} .

Assume that X and \bar{X} are two particles and consider the isomorphisms γ , if any, of neighborhoods $\mathcal{N}(X)$ onto neighborhoods $\mathcal{N}(\bar{X})$ which map X into $\bar{X} = \gamma(X)$. We define an equivalence relation among all such isomorphisms by the condition that $\gamma \sim \tilde{\gamma}$ if and only if $\gamma(Z) = \tilde{\gamma}(Z)$ for Z in some neighborhood of X , however small. The resulting equivalence classes Γ will be called the *local isomorphisms* of X onto \bar{X} . If Γ is such a local isomorphism then the local configurations $\bar{\Phi}$ of \bar{X} are of the form

$$(14.3) \quad \bar{\Phi} = \Phi \circ \Gamma^{-1},$$

where Φ is a local configuration of X .

Definition 1. A local isomorphism Γ of a particle X onto a particle \bar{X} will be called a *material isomorphism* of X onto \bar{X} provided the functionals \mathfrak{F} and $\bar{\mathfrak{F}}$ of X and \bar{X} are related by

$$(14.4) \quad \bar{\mathfrak{F}}(\Theta^*) = \mathfrak{F}(\Theta^* \circ \Gamma^{-1})$$

for all kinematical histories $\Theta^* \in \mathcal{C}_X^*$.

We shall say that two particles consist of the same material if they are materially isomorphic to each other.

15. Constitutive functionals

Assume that a local motion Θ and a local reference configuration Φ of a particle X are given. Let Θ^t and Ω^t be the corresponding kinematical history and deformation history, respectively, as defined in Section 7. By (4.4) they are related by

$$(15.1) \quad \Theta^t = \Omega^t \circ \Phi.$$

Hence the general constitutive equation (13.6) may be rewritten in the form

$$(15.2) \quad S(t) = \mathfrak{F}(\Omega^t \circ \Phi).$$

Definition 2. A functional \mathfrak{G} whose domain is the set \mathcal{D}^* of all deformation histories and whose values are symmetric tensors is called a *constitutive functional* provided it has the following property: For all deformation histories $\Omega^* \in \mathcal{D}^*$ and all orthogonal tensor functions $Q^* \in \mathcal{O}^*$ the relation

$$(15.3) \quad Q_0 \mathfrak{G}(\Omega^*) Q_0^T = \mathfrak{G}(Q^* \circ \Omega^*), \quad Q_0 = Q^*(0),$$

holds.

It is clear that the functional \mathfrak{G}_Φ defined by

$$(15.4) \quad \mathfrak{G}_\Phi(\Omega^*) = \mathfrak{G}(\Omega^*; \Phi) = \mathfrak{F}(\Omega^* \circ \Phi)$$

is a constitutive functional, because the property (15.3) for $\mathfrak{G} = \mathfrak{G}_\Phi$ is equivalent to the property (13.7) for \mathfrak{F} . Hence the most general constitutive equation may be formulated in this way:

Theorem 1. *Given a particle X and a local reference configuration Φ of X , there is a constitutive functional \mathfrak{G}_Φ such that, for any local process $\{\Theta, S\}$ of X , the stress S is related to the local motion Θ by*

$$(15.5) \quad S(t) = \mathfrak{G}_\Phi(\Omega^t) = \mathfrak{G}(\Omega^t; \Phi),$$

where Ω^t is the deformation history of X defined in Section 7. The constitutive functionals \mathfrak{G}_Φ and $\mathfrak{G}_{\hat{\Phi}}$ corresponding to the local reference configurations Φ and $\hat{\Phi}$ are related, for all deformation histories $\Omega^* \in \mathcal{D}^*$, by

$$(15.6) \quad \mathfrak{G}(\Omega^* \circ \Delta; \Phi) = \mathfrak{G}(\Omega^*; \hat{\Phi}),$$

where $\Delta = \hat{\Phi} \circ \Phi^{-1}$ is the deformation from Φ to $\hat{\Phi}$.

The relation (15.6) is an immediate consequence of (15.4). We call \mathfrak{G}_Φ the constitutive functional of the particle X relative to the local reference configuration Φ .

Since the nature of the domain \mathcal{D}^* of constitutive functionals is not related to particles, as was the case with the functionals \mathfrak{F} in (13.6), they characterize materials in a manner that is independent of the particular particle. The following theorem is a consequence of Definition 1 and (15.4):

Theorem 2. *Two particles X and \bar{X} are materially isomorphic, i.e., consist of the same material, if and only if there is a local configuration Φ of X and a local isomorphism Γ of X onto \bar{X} such that*

$$(15.7) \quad \mathfrak{G}_\Phi(\Omega^*) = \mathfrak{G}_{\bar{\Phi}}(\Omega^*), \quad \bar{\Phi} = \Phi \circ \Gamma^{-1},$$

for all $\Omega^* \in \mathcal{D}^*$.

16. The local isotropy group

Let Γ be a *material automorphism* of X , i.e., a material isomorphism of X onto itself. By Theorem 2, there is a local configuration Φ such that

$$(16.1) \quad \mathfrak{G}(\Omega^*; \Phi) = \mathfrak{G}(\Omega^*; \hat{\Phi}), \quad \hat{\Phi} = \Phi \circ \Gamma^{-1}.$$

The material automorphism Γ and the deformation $\Delta = \hat{\Phi} \circ \Phi^{-1}$ from Φ to $\hat{\Phi}$ are related by

$$(16.2) \quad \Delta = \Phi \circ \Gamma^{-1} \circ \Phi^{-1}, \quad \Gamma = \hat{\Phi}^{-1} \circ \Delta^{-1} \circ \Phi.$$

Since Γ , an isomorphism, preserves the mass distribution, it follows that Δ must be an isochoric deformation.

In addition to (16.1) we have the relation (15.6) between the two functionals \mathfrak{G}_Φ and $\mathfrak{G}_{\hat{\Phi}}$. Combining these two relations we obtain

$$(16.3) \quad \mathfrak{G}_\Phi(\Omega^* \circ \Delta) = \mathfrak{G}_\Phi(\Omega^*).$$

Conversely, if (16.3) holds for an isochoric deformation $\Delta \in \mathcal{I}$, then $\Gamma = \hat{\Phi}^{-1} \circ \Delta^{-1} \circ \Phi$ can easily be seen to be a material automorphism of X .

Definition 3. Let \mathfrak{G} be a constitutive functional. The group \mathcal{G}^l of all isochoric deformations $\Delta \in \mathcal{I}$ with the property that

$$(16.4) \quad \mathfrak{G}(\Omega^* \circ \Delta) = \mathfrak{G}(\Omega^*)$$

holds for all $\Omega^* \in \mathcal{D}^*$ is called the local isotropy group of \mathfrak{G} .

\mathcal{G}^l is a subgroup of the group \mathcal{I} of all isochoric deformations. The material automorphisms of a particle also form a group. The relations (16.2) establish an isomorphism of this group with the local isotropy group \mathcal{G}_Φ^l of \mathfrak{G}_Φ . The local isotropy groups \mathcal{G}_Φ^l and $\mathcal{G}_{\hat{\Phi}}^l$ corresponding to different local configurations Φ and $\hat{\Phi}$ will in general be different. But they are all isomorphic to each other because they are isomorphic to the group of material automorphisms of X . In fact, it is not hard to see that $\mathcal{G}_{\hat{\Phi}}^l$ is conjugate to \mathcal{G}_Φ^l :

$$(16.5) \quad \mathcal{G}_{\hat{\Phi}}^l = \Delta \circ \mathcal{G}_\Phi^l \circ \Delta^{-1}, \quad \Delta = \hat{\Phi} \circ \Phi^{-1}$$

III. Simple materials

17. Simple constitutive functionals

It may happen that the values $\mathfrak{G}(\Omega^*)$ of a constitutive functional \mathfrak{G} are not affected if $\Omega^*(s)$, for each $s \leq 0$, is changed by a small term of order one. By (3.6), $\Omega^*(s)$ differs from $\nabla \Omega^*(s)$ only by such a small term.

Definition 4. A constitutive functional \mathfrak{G} is said to be simple if

$$(17.1) \quad \mathfrak{G}(\Omega^*) = \mathfrak{G}(\nabla \Omega^*)$$

for all deformation histories $\Omega^* \in \mathcal{D}^*$.

The condition (17.1) is equivalent to the following: If $\nabla \Omega^*(s) = \nabla \tilde{\Omega}^*(s)$ for all $s \leq 0$, then

$$(17.2) \quad \mathfrak{G}(\Omega^*) = \mathfrak{G}(\tilde{\Omega}^*).$$

The values of a simple constitutive functional \mathfrak{G} are determined for all $\Omega^* \in \mathcal{D}^*$ if they are known for the histories of linear transformations $F^* \in \mathcal{L}^*$. In other words, \mathfrak{G} is determined by its restriction to \mathcal{L}^* .

From now on we assume that \mathfrak{G} is simple, and we use the same symbol \mathfrak{G} for its restriction to \mathcal{L}^* . By (15.3) \mathfrak{G} satisfies the relation

$$(17.3) \quad Q_0 \mathfrak{G}(F^*) Q_0^T = \mathfrak{G}(Q^* F^*), \quad Q_0 = Q^*(0),$$

for all $F^* \in \mathcal{L}^*$ and all $Q^* \in \mathcal{O}^*$.

Let $G \in \mathcal{L}$ be an arbitrary regular linear transformation with the polar decomposition

$$(17.4) \quad G = P T, \quad P \in \mathcal{O}, \quad T \in \mathcal{S}_+.$$

For any $F^* \in \mathcal{L}^*$ we define

$$(17.5) \quad \hat{F}^* = F^* G^{-1} = \hat{R}^* \hat{U}^*, \quad \hat{R}^* \in \mathcal{O}^*, \quad \hat{U}^* \in \mathcal{S}_+^*,$$

where \hat{R}^* and \hat{U}^* are determined by the polar decomposition of \hat{F}^* . We then have

$$F^* = \hat{F}^* G = \hat{R}^* \hat{U}^* P T = (\hat{R}^* P) (\hat{P}^T U^* P) T.$$

Substituting $\hat{R}^* P$ for Q^* and $(P^T \hat{U}^* P) T$ for F^* in (17.3), we see that

$$(17.6) \quad \mathfrak{G}(F^*) = \hat{R}_0 P \mathfrak{G}((P^T \hat{U}^* P) T) P^T \hat{R}_0^T, \quad \hat{R}_0 = \hat{R}^*(0).$$

We now define a functional \mathfrak{R} with domain $\mathcal{S}_+^* \times \mathcal{S}_+$ by

$$(17.7) \quad \mathfrak{G}(U^* T) = \mathfrak{R}(U^*; T)$$

for all $U^* \in \mathcal{S}_+^*$ and $T \in \mathcal{S}_+$. Equation (17.6) then shows that

$$(17.8) \quad \mathfrak{G}(F^*) = \hat{R}_0 P \mathfrak{R}(P^T \hat{U}^* P; T) P^T \hat{R}_0^T.$$

The functional \mathfrak{R} has the property that

$$(17.9) \quad \mathfrak{R}(U^*; T) = \mathfrak{R}(\tilde{U}^*; \tilde{T}) \quad \text{if} \quad U^* T = \tilde{U}^* \tilde{T}.$$

Conversely, if \mathfrak{R} is any functional with domain $\mathcal{S}_+^* \times \mathcal{S}_+$ and with the property (17.9), then (17.8) defines a simple constitutive functional \mathfrak{G} . The variables on the right side of (17.8) are defined in terms of F^* and an arbitrary $G \in \mathcal{L}$. For the special choice $G = I$ we get

$$(17.10) \quad \mathfrak{G}(F^*) = R_0 \mathfrak{R}(U^*; I) R_0^T, \quad R_0 = R^*(0),$$

where U^* and R^* are defined by the polar decomposition

$$(17.11) \quad F^* = R^* U^*, \quad R^* \in \mathcal{O}^*, \quad U^* \in \mathcal{S}_+^*.$$

In the special case when $G = F_0 = F^*(0)$ (17.8) becomes

$$(17.12) \quad \mathfrak{G}(F^*) = R_0 \mathfrak{R}(R_0^T U_*^* R_0; U_0) R_0^T,$$

where $R_0 = R^*(0)$, $U_0 = U^*(0)$, and where $U_*^* \in \mathcal{S}_+^*$ is determined by the polar decomposition

$$(17.13) \quad F^* F_0^{-1} = R_*^* U_*^*, \quad R_*^* \in \mathcal{O}^*, \quad U_*^* \in \mathcal{S}_+^*.$$

We define a functional \mathfrak{R}_1 with domain \mathcal{S}_+^* by

$$(17.14) \quad \mathfrak{R}_1(U^*) = \mathfrak{R}(U^*; I) \quad \text{for} \quad U^* \in \mathcal{S}_+^*.$$

Then (17.10) takes the form

$$(17.15) \quad \mathfrak{G}(F^*) = R_0 \mathfrak{R}_1(U^*) R_0^T.$$

Conversely, if \mathfrak{R}_1 is an arbitrary functional with domain \mathcal{S}_+^* , then (17.15) defines a simple constitutive functional.

We also define a functional \mathfrak{R}_2 as the restriction of \mathfrak{R} obtained by allowing for its first variable only those functions $U_*^* \in \mathcal{S}_+^*$ whose value for $s=0$ is the identity,

$$(17.16) \quad U_*^*(0) = I.$$

We denote the set of all U_*^* with the property (17.16) by \mathcal{S}_{+*}^* . The domain of \mathfrak{R}_2 is then $\mathcal{S}_{+*}^* \times \mathcal{S}_+$, and it is identical with \mathfrak{R} in this domain. The function

U_*^* defined by (17.13) has the property (17.16), and the same is true for $R_0^T U_*^* R_0$. Hence (17.12) has the form

$$(17.17) \quad \mathfrak{G}(F^*) = R_0 \mathfrak{R}_2(R_0^T U_*^* R_0; U_0) R_0^T.$$

Conversely, if \mathfrak{R}_2 is an arbitrary functional with domain $\mathcal{S}_{+*}^* \times \mathcal{S}_+$, then (17.17) defines a simple constitutive functional \mathfrak{G} .

18. Simple materials

We say that the material at a particle X is *simple* or, briefly, that X is simple if the constitutive functional \mathfrak{G}_Φ of X , for some local configuration Φ of X , is simple. We assume from now on that this is the case. Let $\hat{\Phi}$ be another local configuration of X and let $\Delta = \hat{\Phi} \circ \Phi^{-1}$ be the deformation from Φ to $\hat{\Phi}$. For every deformation history $\Omega^* \in \mathcal{D}^*$, by (3.7) and the chain rule (3.4), we have

$$(18.1) \quad \nabla(\Omega^* \circ \Delta) = \nabla[(\nabla\Omega^*) \circ \Delta].$$

It follows from (17.2) that

$$\mathfrak{G}(\Omega^* \circ \Delta; \hat{\Phi}) = \mathfrak{G}(\nabla\Omega^* \circ \Delta; \hat{\Phi}),$$

and hence from (15.6) that

$$\mathfrak{G}(\Omega^*; \hat{\Phi}) = \mathfrak{G}(\nabla\Omega^*; \hat{\Phi}),$$

which shows that $\mathfrak{G}_{\hat{\Phi}}$ is also simple. Therefore, if the material at a particle is simple then the constitutive functional \mathfrak{G}_Φ is simple for all local reference configurations Φ .

It follows from (15.6) that the functionals \mathfrak{G}_Φ and $\mathfrak{G}_{\hat{\Phi}}$ corresponding to two local configurations Φ and $\hat{\Phi}$ are related by

$$(18.2) \quad \mathfrak{G}(F^*G; \Phi) = \mathfrak{G}(F^*; \hat{\Phi}),$$

where $G = G(\hat{\Phi}, \Phi)$ is the deformation gradient from Φ to $\hat{\Phi}$ defined by (5.1). It follows from (18.2) that

$$\mathfrak{G}(F^*; \Phi) = \mathfrak{G}(F^*; \hat{\Phi}) \quad \text{if} \quad G(\hat{\Phi}, \Phi) = I.$$

This means, according to (5.8), that $\mathfrak{G}(F^*; \Phi)$ depends only on the gradient $M = \nabla\Phi$ of the local configuration Φ . We can therefore define

$$(18.3) \quad \mathfrak{G}_M(F^*) = \mathfrak{G}(F^*; M) = \mathfrak{G}(F^*; \Phi) \quad \text{if} \quad M = \nabla\Phi.$$

The following theorem is a corollary of Theorem 1.

Theorem 3. *Let X be a simple particle and let M be a configuration gradient of X . Then there is a simple constitutive functional \mathfrak{G}_M such that, for any local process $\{\Theta, S\}$, the stress is related to the local motion by the constitutive equation*

$$(18.4) \quad S(t) = \mathfrak{G}_M(F^t) = \mathfrak{G}(F^t; M),$$

where $F^t \in \mathcal{L}^*$ is the history of the deformation gradient of the local motion Θ relative to M as a reference. The functionals \mathfrak{G}_M and $\mathfrak{G}_{\hat{M}}$ corresponding to two configuration

gradients are related by

$$(18.5) \quad \mathfrak{G}(F^*G; M) = \mathfrak{G}(F^*; \hat{M}), \quad G = \hat{M}M^{-1},$$

for all $F^* \in \mathcal{L}^*$.

The results of the previous section enable us to put the constitutive equation into various other forms:

Theorem 4. *The constitutive equation for a simple particle X may be written, with reference to a fixed configuration gradient M of X , in one of the following four forms:*

(a) *There is a functional \mathfrak{R}_1 with domain \mathcal{S}_+^* such that the stress is given by*

$$(18.6) \quad S(t) = R(t) \mathfrak{R}_1(U^t) R(t)^T,$$

where $R(t)$ is the rotation tensor and U^t the history of the right strain tensor relative to M as a reference.

(b) *There is a functional \mathfrak{R} with domain $\mathcal{S}_+^* \times \mathcal{S}$ and with the property (17.9) such that, for any configuration gradient \hat{M} as a reference, the stress is given by*

$$(18.7) \quad S(t) = \hat{R}(t) P \mathfrak{R}(P^T \hat{U}^t P; T_r) P^T \hat{R}(t)^T,$$

where $\hat{R}(t)$ is the rotation tensor and \hat{U}^t the history of the right strain tensor relative to \hat{M} as a reference, and where P is the rotation tensor and T_r the right strain tensor of the deformation from M to \hat{M} , $\hat{M}M^{-1} = PT$.

(c) *In the special case when \hat{M} is the gradient of the configuration at some time t_0 the equation (18.7) takes the form*

$$(18.8) \quad S(t) = R_{t_0}(t) R(t_0) \mathfrak{R}(R(t_0)^T U_{t_0}^t R(t_0); U(t_0)) R(t_0)^T R_{t_0}(t)^T,$$

where the rotation tensors R and R_{t_0} and the right strain tensors U and U_{t_0} are taken relative to M and the configuration at time t_0 as a reference, respectively.

(d) *There is a functional \mathfrak{R}_2 with domain $\mathcal{S}_{+*}^* \times \mathcal{S}_+$ such that the stress is given by*

$$(18.9) \quad S(t) = R(t) \mathfrak{R}_2(R^T(t) U_t^t R(t); U(t)) R(t)^T,$$

where R , U , and U_t are defined as before.

The forms (18.6) and (18.9) have the advantage that the functionals \mathfrak{R}_1 and \mathfrak{R}_2 are not subject to any restrictive condition.

19. The isotropy group

If \mathfrak{G} is a simple constitutive functional, then it follows from (17.2) that $\mathfrak{G}(\Omega^*) = \mathfrak{G}(\Omega^* \circ \Delta)$ holds for all null-deformations $\Delta \in \mathcal{N}$, i.e., whenever $\nabla \Delta = I$. Hence, by definition 3, the local isotropy group \mathcal{G}^l of \mathfrak{G} contains the group \mathcal{N} of all null-deformations as a normal subgroup. The quotient group $\mathcal{G}^l / \mathcal{N}$ is isomorphic to a group of unimodular transformations.

Definition 5. *The group \mathcal{G} of all unimodular transformations H with the property that*

$$(19.1) \quad \mathfrak{G}(F^*H) = \mathfrak{G}(F^*)$$

holds for all $F^* \in \mathcal{L}^*$ is called the isotropy group of the simple constitutive functional \mathfrak{G} .

\mathcal{G} is a subgroup of the group \mathcal{U} of all unimodular transformations and it is isomorphic to $\mathcal{G}^l | \mathcal{N}$.

Theorem 5. An orthogonal transformation Q is an element of the isotropy group \mathcal{G} of a constitutive functional \mathfrak{G} if and only if one of the following conditions is satisfied:

(a) For all $F^* \in \mathcal{L}^*$

$$(19.2) \quad Q \mathfrak{G}(F^*) Q^T = \mathfrak{G}(Q F^* Q^T).$$

(b) For all $U^* \in \mathcal{S}_+^*$ and all $T \in \mathcal{S}_+$

$$(19.3) \quad Q \mathfrak{R}(U^*; T) Q^T = \mathfrak{R}(Q U^* Q^T; Q T Q^T),$$

where \mathfrak{R} is defined by (17.7).

(c) For all $U^* \in \mathcal{S}_+^*$

$$(19.4) \quad Q \mathfrak{R}_1(U^*) Q^T = \mathfrak{R}_1(Q U^* Q^T),$$

where \mathfrak{R}_1 is defined by (17.14).

(d) For all $U_*^* \in \mathcal{S}_{+*}^*$ and all $T \in \mathcal{S}_+$

$$(19.5) \quad Q \mathfrak{R}_2(U_*^*; T) Q^T = \mathfrak{R}_2(Q U_*^* Q^T; Q T Q^T)$$

where \mathfrak{R}_2 is defined in Section 17.

Proof. If we substitute the constant tensor Q for the tensor function Q^* in (17.3), and $Q F^*$ for F^* and Q for H in (19.1), we see that (19.2) holds if and only if $Q \in \mathcal{G}$. The equivalence of the condition (a) with (b), (c), and (d) follows directly from the definitions of \mathfrak{R} , \mathfrak{R}_1 , and \mathfrak{R}_2 .

If \mathcal{G} contains the full orthogonal group \mathcal{O} , then (19.2), (19.3), (19.4), and (19.5) are valid for all orthogonal Q .

The isotropy group \mathcal{G}_M of the constitutive functional \mathfrak{G}_M of a particle will, in general, depend on the choice of the configuration gradient M . But, as in the case of local isotropy, the groups \mathcal{G}_M and $\mathcal{G}_{\hat{M}}$ corresponding to two configuration gradients M and \hat{M} are conjugate:

$$(19.6) \quad \mathcal{G}_{\hat{M}} = G \mathcal{G}_M G^{-1}, \quad G = \hat{M} M^{-1}.$$

It follows from (18.5) that $H \in \mathcal{G}_M$ if and only if

$$(19.7) \quad \mathfrak{G}(F^*; M) = \mathfrak{G}(F^*; H M)$$

for all $F^* \in \mathcal{L}^*$.

20. Isotropic and anisotropic solids

We say that a constitutive functional \mathfrak{G} defines a *solid* if its isotropy group is a subgroup of the orthogonal group, i.e., if $\mathcal{G} \subset \mathcal{O}$. A particle X is said to be a solid particle if there is a configuration gradient M of X such that \mathfrak{G}_M defines a solid. A solid is called an *isotropic solid* if the isotropy group of its defining

functional is the full orthogonal group, *i.e.*, if $\mathcal{G} = \mathcal{O}$. Let X be an isotropic solid particle. Any configuration gradient M such that $\mathcal{G}_M = \mathcal{O}$ is then called an *undistorted state* of X .

The following theorem follows immediately from Theorems 4 and 5.

Theorem 6. *Let X be an isotropic solid particle. Its constitutive equation may be written, with reference to an undistorted state M of X , in one of the following forms:*

$$(20.1) \quad S(t) = \mathfrak{R}_1(R(t) U^t R^T(t)),$$

$$(20.2) \quad S(t) = \hat{R}(t) \mathfrak{R}(\hat{U}^t; T_1) \hat{R}(t)^T,$$

$$(20.3) \quad S(t) = R_{t_0}(t) \mathfrak{R}(U_{t_0}^t; V(t_0)) R_{t_0}(t),$$

$$(20.4) \quad S(t) = \mathfrak{R}_2(U_i^t; V(t)).$$

The notation of Theorem 4 applies here. In addition, $V = R U R^T$ is the left strain tensor relative to M as a reference and T_1 is the left strain tensor of the deformation from M to \hat{M} . The functionals \mathfrak{R} , \mathfrak{R}_1 and \mathfrak{R}_2 satisfy the conditions (b), (c), and (d) of Theorem 5 for all orthogonal transformations Q .

A solid is called *anisotropic* if the isotropy group of its defining functional is a proper subgroup of the orthogonal group, *i.e.*, if $\mathcal{G} \subset \mathcal{O}$ and $\mathcal{G} \neq \mathcal{O}$. Material symmetries in anisotropic solids, such as *orthotropy*, *transverse isotropy*, and the various types of *crystal symmetry* are defined according to the special nature of the isotropy group \mathcal{G} .

21. Fluids

We say that a constitutive functional defines a *fluid* if its isotropy group \mathcal{G} is the full unimodular group \mathcal{U} , *i.e.*, if $\mathcal{G} = \mathcal{U}$. A particle X is said to be a fluid particle if, for some configuration gradient M of X , the corresponding constitutive functional defines a fluid, *i.e.*, if $\mathcal{G}_M = \mathcal{U}$. Let \hat{M} be any other configuration gradient. By (19.6) $\mathcal{G}_{\hat{M}}$ is conjugate to $\mathcal{G}_M = \mathcal{U}$. But \mathcal{U} is a normal subgroup of \mathcal{L} and hence coincides with all its conjugates. It follows that, if X is a fluid particle, then $\mathcal{G}_M = \mathcal{U}$ for all configuration gradients M of X .

For a fluid it follows from (19.7) that

$$(21.1) \quad \mathfrak{G}(F^*; M) = \mathfrak{G}(F^*; \hat{M})$$

whenever $H = \hat{M} \hat{M}^{-1}$ is unimodular. But, by (5.17), this is the case if and only if the densities ϱ_M and $\varrho_{\hat{M}}$ coincide. It follows that the value $\mathfrak{G}(F^*; M)$ can depend only on the density ϱ_M . Therefore, we can define a functional \mathfrak{H} with domain $\mathcal{L}^* \times \mathcal{R}_+$ (\mathcal{R}_+ = set of positive real numbers) such that

$$(21.2) \quad \mathfrak{G}(F^*; M) = \mathfrak{H}(F^*; \varrho_M)$$

for all $F^* \in \mathcal{L}^*$ and all configuration gradients M of X . Let $F^* = R^* U^*$ be the polar decomposition of F^* . Substituting R^* for Q^* and U^* for F^* in (17.3) and using (21.2) we see that

$$(21.3) \quad \mathfrak{G}(F^*; M) = R_0 \mathfrak{H}(U^*; \varrho_M) R_0^T, \quad R_0 = R^*(0),$$

which shows that \mathfrak{S} is determined by its restriction to $\mathcal{S}_+^* \times \mathcal{R}_+$. We use the same symbol \mathfrak{S} for this restriction. Since $\mathcal{G}_M = \mathcal{U}$ contains the orthogonal group it follows from (19.2) that \mathfrak{S} satisfies the relation

$$(21.4) \quad Q \mathfrak{S}(U^*; d) Q^T = \mathfrak{S}(Q U^* Q^T; d)$$

for all $U^* \in \mathcal{S}_+^*$, all $Q \in \mathcal{O}$, and all $d > 0$.

Theorem 7. *The constitutive equation for a fluid particle X may be written in one of the following forms:*

(a) *There is a functional \mathfrak{S} with domain $\mathcal{S}_+^* \times \mathcal{R}_+$ and with the property (21.4) such that the stress is given by*

$$(21.5) \quad S(t) = R(t) \mathfrak{S}(U^t; \varrho_M) R(t)^T,$$

where $R(t)$ is the rotation tensor and U^t the history of the right strain tensor relative to an arbitrary configuration gradient M with mass density ϱ_M .

(b) *In the special case when M is the configuration gradient at some time t_0 , the equation (21.5) takes the form*

$$(21.6) \quad S(t) = R_{t_0}(t) \mathfrak{S}(U_{t_0}^t; \varrho(t_0)) R_{t_0}(t)^T.$$

(c) *There is a functional \mathfrak{S}_1 with domain $\mathcal{S}_{+*}^* \times \mathcal{R}_+$ and with the property that*

$$(21.7) \quad Q \mathfrak{S}_1(U_*^*; d) Q^T = \mathfrak{S}_1(Q U_*^* Q^T; d)$$

for all $U_*^* \in \mathcal{S}_{+*}^*$, all $Q \in \mathcal{O}$, and all $d > 0$, such that the stress is given by

$$(21.8) \quad S(t) = \mathfrak{S}_1(U_*^t; \varrho(t)).$$

Proof. The part (a) and its special case (b) follow from (21.3) and Theorem 3. The part (c) follows from (b) by choosing $t_0 = t$ and by defining \mathfrak{S}_1 to be the restriction of \mathfrak{S} to $\mathcal{S}_{+*}^* \times \mathcal{R}_+$.

We note that an arbitrary functional \mathfrak{S}_1 with domain $\mathcal{S}_{+*}^* \times \mathcal{R}_+$ and with the property (21.7) may define a fluid. The constitutive equation (21.8) is intrinsic in the sense that it does not depend on the choice of a reference configuration. The constitutive equation of a solid cannot be put into such an intrinsic form.

22. Constitutive equations involving the Cauchy-Green tensors

As we pointed out in Section 6, it is often better to use the Cauchy-Green tensors C and B instead of the strain tensors U and V . In order to do so, we define new functionals $\overline{\mathfrak{R}}$, $\overline{\mathfrak{R}}_1$, $\overline{\mathfrak{R}}_2$, $\overline{\mathfrak{S}}$ and $\overline{\mathfrak{S}}_1$, of the same nature as the corresponding functionals without the superposed bars defined in Sections 17 and 20, by the following formulas:

$$(22.1) \quad \overline{\mathfrak{R}}(U^{*2}; T^2) = U_0 \mathfrak{R}(U^*; T) U_0,$$

$$(22.2) \quad \overline{\mathfrak{R}}_1(U^{*2}) = U_0 \mathfrak{R}_1(U^*) U_0,$$

$$(22.3) \quad \overline{\mathfrak{R}}_2(U_*^{*2}; T^2) = \mathfrak{R}_2(U_*^*; T),$$

$$(22.4) \quad \overline{\mathfrak{S}}(U^{*2}; d) = U_0 \mathfrak{S}(U^*; d) U_0,$$

$$(22.5) \quad \overline{\mathfrak{S}}_1(U_*^{*2}; d) = \mathfrak{S}_1(U_*^*; d),$$

valid for all $U^* \in \mathcal{S}_+^*$ ($U_0 = U^*(0)$), all $U_*^* \in \mathcal{S}_{+*}^*$ ($U_*^*(0) = I$), all $T \in \mathcal{S}_+$, and all $d > 0$.

It is not hard to see that the general constitutive equations (18.6), (18.8), and (18.9) of Theorem 4 for simple materials then take the form

$$(22.6) \quad F(t)^T S(t) F(t) = \bar{\mathfrak{R}}_1(C^t), \star$$

$$(22.7) \quad F_{i_0}(t)^T S(t) F_{i_0}(t) = R(t_0) \bar{\mathfrak{R}}(R(t_0)^T C_{i_0}^t R(t_0); C(t_0)) R(t_0)^T,$$

$$(22.8) \quad S(t) = R(t) \bar{\mathfrak{R}}_2(R(t)^T C_i^t R(t); C(t)) R(t)^T.$$

In the case of isotropic solids, the functionals $\bar{\mathfrak{R}}$, $\bar{\mathfrak{R}}_1$, and $\bar{\mathfrak{R}}_2$ satisfy, for all orthogonal Q , the functional relations obtained from (19.3), (19.4), and (19.5) by superposing bars. Moreover, the simplified constitutive equations (20.3) and (20.4) for isotropic solids take the form

$$(22.9) \quad F_{i_0}(t)^T S(t) F_{i_0}(t) = \bar{\mathfrak{R}}(C_{i_0}^t; B(t_0)),$$

$$(22.10) \quad S(t) = \bar{\mathfrak{R}}_2(C_i^t; B(t)).$$

The functionals $\bar{\mathfrak{G}}$ and $\bar{\mathfrak{G}}_1$ have the same properties (21.4) and (21.7) as the corresponding functionals without the superposed bars. The constitutive equations (21.6) and (21.8) for fluids take the form

$$(22.11) \quad F_{i_0}(t)^T S(t) F_{i_0}(t) = \bar{\mathfrak{G}}(C_{i_0}^t; \varrho(t_0)),$$

$$(22.12) \quad S(t) = \bar{\mathfrak{G}}_1(C_i^t; \varrho(t)).$$

IV. Special classes of materials

23. Materials of the differential type

The value $\mathfrak{G}(F^*)$ of a simple constitutive functional is determined by the values $F^*(s)$ of the tensor function F^* for $s \leq 0$. It may happen that $\mathfrak{G}(F^*)$ depends only on the values $F^*(s)$ for s very near to zero. If F^* has sufficiently many continuous derivatives then $F^*(s)$ may be approximated, for small values of s , by its Taylor expansion up to some order n . This Taylor expansion is determined by the value of F^* and its derivatives up to the order n at $s = 0$, *i.e.*, by

$$(23.1) \quad F_0 = F^*(0), \quad \dot{F}_0 = \dot{F}^*(0), \dots, \overset{(n)}{F}_0 = \overset{(n)}{F}^*(0).$$

Definition 6. A simple constitutive functional \mathfrak{G} is said to be of the differential type if

$$(23.2) \quad \mathfrak{G}(F^*) = \mathfrak{G}(\tilde{F}^*)$$

whenever

$$(23.3) \quad \overset{(k)}{F}^*(0) = \overset{(k)}{\tilde{F}}^*(0), \quad k = 0, 1, \dots, n.$$

We have seen that every simple constitutive functional \mathfrak{G} has a representation of the form (17.17) in terms of a functional \mathfrak{R}_2 . Since \mathfrak{R}_2 is just a restriction of \mathfrak{R}

* This form, in other notation, has been proposed independently by GREEN & RIVLIN [4].

which is defined by (17.7) in terms of \mathfrak{G} , we have

$$(23.4) \quad \mathfrak{R}_2(U_*^*; T) = \mathfrak{G}(U_*^* T)$$

for all $U_*^* \in \mathcal{S}_{+*}^*$ and all $T \in \mathcal{S}_+$. If \mathfrak{G} is of the differential type then its value depends only on the values at $s=0$ of its argument and its first n derivatives. Hence, by (23.4), the value $\mathfrak{R}_2(U_*^*; T)$ depends only on the first n derivatives of U_*^* at $s=0$, since $U_*^*(0) = I$. It follows that there is a function \mathfrak{f} of $n+1$ symmetric tensor variables such that

$$(23.5) \quad \mathfrak{R}_2(U_*^*; T) = \mathfrak{f}(\dot{U}_*^*(0), \ddot{U}_*^*(0), \dots, \overset{(n)}{U}_*^*(0); T)$$

for all $U_*^* \in \mathcal{S}_{+*}^*$ and all $T \in \mathcal{S}_+$. In the case when $U_*^* = U_t^*$ is the history up to time t of the right strain tensor relative to the configuration at time t of a local motion, the derivative $\overset{(k)}{U}_t^*(0)$ coincides by (8.8) with the k^{th} rate of strain $D_k(t)$. For simplicity we use the notation

$$(23.6) \quad \mathfrak{f}(D_1, D_2, \dots, D_n; T) = \mathfrak{f}(D_k; T).$$

We say that a particle X is of differential type if its constitutive functional G_M , for some configuration gradient M , is of the differential type. It is not hard to see that $G_{\hat{M}}$, for any other configuration gradient \hat{M} , is then also of the differential type. The following theorem is a consequence of (18.9) and the remarks made above:

Theorem 8. *The constitutive equation of a particle X of differential type may be written, with reference to a configuration gradient M of X , in the following form:*

There is a symmetric-tensor-valued function \mathfrak{f} of $n+1$ symmetric tensor variables such that the stress is given by

$$(23.7) \quad S(t) = R(t) \mathfrak{f}(R^T(t) D_k(t) R(t); U(t)) R(t)^T,$$

where $D_k(t)$ is the k^{th} rate of strain and where the rotation tensor $R(t)$ and the right strain tensor $U(t)$ are taken relative to M as a reference.

In a material of the differential type the stress depends only on the immediate past of the motion and not on its course at times long ago.

Since the rates of strain D_k can be expressed as polynomials in the acceleration gradients E_k and E_k^T as shown in Section 9, it follows that the constitutive equation (23.7) is a relation involving the displacement gradient $G = RU$, the acceleration gradients, and the stress.

In the case of *isotropic solids*, it follows from Theorem 6 and (23.5) that \mathfrak{f} must satisfy the relation

$$(23.8) \quad Q \mathfrak{f}(D_k; T) Q^T = \mathfrak{f}(Q D_k Q^T; Q T Q^T)$$

for all $D_k \in \mathcal{S}$, $k=1, 2, \dots, n$, all $Q \in \mathcal{O}$, and all $T \in \mathcal{S}_+$. A function with this property is called an *isotropic tensor function*. By (20.4) the constitutive equation (23.7) reduces for isotropic solids to

$$(23.9) \quad S(t) = \mathfrak{f}(D_k(t); V(t)).$$

The constitutive equation for a *fluid* of the differential type has the form

$$(23.10) \quad S(t) = \mathfrak{h}(D_k(t); \varrho(t))$$

where \mathfrak{h} is an isotropic function of n symmetric tensor variables and one positive scalar variable, *i.e.*, it satisfies the relation

$$(23.11) \quad Q \mathfrak{h}(D_k; d) Q^T = \mathfrak{h}(Q D_k Q^T; d)$$

for all $D_k \in \mathcal{S}$, $k = 1, 2, \dots, n$, all $d > 0$, and all $Q \in \mathcal{O}$. This is an immediate consequence of Theorem 7, (21.8).

If we use the alternate forms (22.8), (22.10), and (22.12) of the general constitutive equations we arrive at the following forms for materials of the differential type: In the general case,

$$(23.12) \quad S(t) = R(t) \bar{\mathfrak{f}}(R(t)^T A_k(t) R(t); C(t)) R(t)^T.$$

For isotropic solids*,

$$(23.13) \quad S(t) = \bar{\mathfrak{f}}(A_k(t); B(t)).$$

For fluids,

$$(23.14) \quad S(t) = \bar{\mathfrak{h}}(A_k(t); \varrho(t)).$$

In these equations $\bar{\mathfrak{f}}$ and $\bar{\mathfrak{h}}$ are of the same type as \mathfrak{f} and \mathfrak{h} ; in (23.12) $\bar{\mathfrak{f}}$ may be arbitrary but in (23.13) it must be isotropic. Of course, $\bar{\mathfrak{h}}$ is also isotropic. $A_k(t)$ is the k^{th} Rivlin-Ericksen tensor, defined by (8.9) and related to the acceleration gradients by (9.2).

24. Materials of the rate type

The general constitutive equation of a simple material in the form (18.8) may be rewritten as

$$(24.1) \quad \bar{S}_{t_0}(t) = \mathfrak{R}(\bar{U}_{t_0}^t; U(t_0))$$

where

$$(24.2) \quad \bar{S}_{t_0}(t) = R(t_0)^T R_{t_0}(t)^T S(t) R_{t_0}(t) R(t_0)$$

and

$$(24.3) \quad \bar{U}_{t_0}^t(s) = \bar{U}_{t_0}(t+s) = R(t_0)^T U_{t_0}(t+s) R(t_0).$$

Keeping t_0 and $U(t_0)$ fixed, we may interpret (24.1) in the following manner: Assuming that the function \bar{U}_{t_0} defined by

$$(24.4) \quad \bar{U}_{t_0}(t) = R(t_0) U_{t_0}(t) R(t_0)^T$$

is given, the function \bar{S}_{t_0} defined by (24.2) is completely determined. In other words, (24.1) defines an operation on functions \bar{U}_{t_0} with values in \mathcal{S}_+ which gives functions \bar{S}_{t_0} with values in \mathcal{S} . It may happen that this operation is defined by the process of solution of a differential equation for \bar{S}_{t_0} in the form

$$(24.5) \quad \mathfrak{f}(\bar{S}_{t_0}(t), \dot{\bar{S}}_{t_0}(t), \dots, \overset{(m)}{\bar{S}}_{t_0}(t); \bar{U}_{t_0}(t), \dot{\bar{U}}_{t_0}(t), \dots, \overset{(n)}{\bar{U}}_{t_0}(t); U(t_0)) = 0,$$

* This form was first derived by RIVLIN & ERICKSEN [8].

where \mathfrak{f} is a symmetric-tensor-valued function of $m+n+3$ symmetric tensor variables. For simplicity we use a notation similar to (23.6), so that (24.5) becomes

$$(24.6) \quad \mathfrak{f}(\overset{(l)}{S}_{i_0}(t); \overset{(k)}{U}_{i_0}(t); U(t_0)) = 0.$$

If the function \bar{U}_{i_0} is given, then (24.6) is a differential equation of order m for the function \bar{S}_{i_0} . We assume that the form of \mathfrak{f} is such that there is a unique solution \bar{S}_{i_0} which assumes given initial values $\overset{(l)}{S}_{i_0}(t_1)$, $l=0, 1, \dots, m-1$, no matter how we choose $U(t_0)$, $\overset{(k)}{U}_{i_0}$, t_1 , and $\overset{(l)}{S}_{i_0}(t_1)$.

Since t_0 is arbitrary, in (24.6) we may make the special choice $t_0=t$ obtaining

$$(24.7) \quad \mathfrak{f}(\overset{(l)}{S}_i(t); \overset{(k)}{U}_i(t); U(t)) = 0.$$

By (24.4) and (8.8) we have

$$(24.8) \quad \overset{(k)}{U}_i(t) = R^T(t) D_k(t) R(t).$$

The tensor function \hat{S}_i defined by

$$(24.9) \quad \hat{S}_i(t) = \overline{R_{i_0}(t)^T \overset{(l)}{S}(t) R_{i_0}(t)} \Big|_{t_0=t}$$

will be called the l^{th} *invariant stress rate*. If we carry out the differentiation in (24.9) according to the product rule, and if we observe (8.7), we see that

$$(24.10) \quad \hat{S}_i = \sum_{\substack{p, q, r=0, \dots, l \\ p+q+r=l}} \frac{l!}{p! q! r!} W_p^T \overset{(q)}{S} W_r;$$

thus \hat{S}_i can be expressed explicitly in terms of the stress S , its time derivatives $\overset{(q)}{S}$ up to the order l , and the spins W_p up to the order l . We have $\hat{S}_0 = S$. For $l=1$, we get the *invariant stress rate**

$$(24.11) \quad \hat{S}_1 = \dot{S} = \dot{S} - W S + S W.$$

Observing (24.2), (24.9), and (24.8), we see that (24.7) has the form

$$(24.12) \quad \mathfrak{f}(R(t)^T \hat{S}_i(t) R(t); R(t)^T D_k(t) R(t); U(t)) = 0.$$

A material with a constitutive equation of this form will be called a *material of the rate type*.

It must be noted that (24.12) is not really a complete constitutive equation. The stress is not determined by the local motion alone but only when, in addition, initial values $\hat{S}_i(t_1)$, $l=0, 1, \dots, m-1$, for some initial time t_1 , are given. These initial values, on the other hand, should be determined by the history of the local motion up to the time t_1 . A constitutive equation of the type (24.12) characterizes not a single material but a family of materials depending on m symmetric tensor parameters.

* It was introduced by ZAREMBA [15].

In the case of *isotropic solids* it follows from Theorem 6 that the tensor function \bar{f} in (24.12) must be isotropic and that the constitutive equation reduces to

$$(24.13) \quad \bar{f}(\hat{S}_i(t); D_k(t); V(t)) = 0.$$

Fluids of the rate type are described by an equation of the form

$$(24.14) \quad g(\hat{S}_i(t); D_k(t); \varrho(t)) = 0,$$

where g is an isotropic tensor function of $m+n+1$ tensor variables and one positive scalar variable. However, not every equation of the form (24.14) defines a class of fluids of the rate type. It may also define a class of isotropic solids, because the stress may depend on the right strain tensor relative to some reference configuration through the initial values.

Starting from the general constitutive equations (22.7), (22.9), and (22.11), one can easily derive alternate forms for the constitutive equations of the rate type:

In the general case we obtain

$$(24.15) \quad \bar{f}(R(t)^T \tilde{S}_i(t) R(t); R(t)^T A_k(t) R(t); C(t)) = 0$$

where $A_k(t)$ is the k^{th} Rivlin-Ericksen tensor (9.2) and where $\tilde{S}_i(t)$ is the i^{th} stress flux defined by

$$(24.16) \quad \tilde{S}_i(t) = \overline{F_{i_0}(t)^T S(t) F_{i_0}(t)} \Big|_{t_0=t}.$$

We find that \tilde{S}_i may be expressed explicitly in terms of the stress S , its derivatives up to the order m , and the acceleration gradients E_i up to the order m by the formula

$$(24.17) \quad \tilde{S}_i = \sum_{\substack{a, p, r=0, \dots, l \\ p+q+r=l}} \frac{l!}{p! q! r!} E_p^T \overset{(q)}{S} E_r.$$

For $l=0$ we have $\tilde{S}_0 = S$, and for $l=1$ we get the *stress flux**

$$(24.18) \quad \tilde{S}_1 = \tilde{S} = \dot{S} + E^T S + S E.$$

In the case of isotropic solids \bar{f} is isotropic and (24.15) reduces to

$$(24.19) \quad \bar{f}(\tilde{S}_i(t); A_k(t); B(t)) = 0^{**}.$$

For fluids we get

$$(24.20) \quad \bar{h}(\tilde{S}_i(t); A_k(t); \varrho(t)) = 0,$$

where \bar{h} is an isotropic tensor function.

Acknowledgement. The research leading to this paper was sponsored in part by the National Science Foundation under Grant NSF-G 5250 and in part by the Air Force Office of Scientific Research under Contract AF 18(600)-1138 with Carnegie Institute of Technology.

* It differs from the flux introduced by CAUCHY (cf. the discussion given by TRUESDELL in [12], Section 55 bis), in which there are minus signs on the right.

** This form was first derived by COTTER & RIVLIN [1].

References

- [1] COTTER, B., & R. S. RIVLIN: Tensors associated with time-dependent stress. *Quart. Appl. Math.* **13**, 177—182 (1955).
- [2] EHRESMANN, C.: Introduction à la théorie des structures infinitésimales et des pseudo-groupes de Lie. *Colloques internationaux du centre national de la recherche scientifique. Géométrie différentielle*, 97—110. Paris 1953.
- [3] GREEN, A. E.: Simple extension of a hypo-elastic body of grade zero. *J. Rational Mech. Anal.* **5**, 637—642 (1956).
- [4] GREEN, A. E., & R. S. RIVLIN: The mechanics of non-linear materials with memory. *Arch. Rational Mech. Anal.* **1**, 1—34 (1957).
- [5] NOLL, W.: On the continuity of the solid and fluid states. *J. Rational Mech. Anal.* **4**, 3—81 (1955).
- [6] NOLL, W.: The foundations of classical mechanics in the light of recent advances in continuum mechanics. *Proceedings of the Berkeley Symposium on the Axiomatic Method*, 1958.
- [7] NOLL, W.: On the foundations of the mechanics of continuous media. *Carnegie Institute of Technology, Technical Report no. 17, Air Force Office of Scientific Research*, 1957.
- [8] RIVLIN, R. S., & J. L. ERICKSEN: Stress-deformation relations for isotropic materials. *J. Rational Mech. Anal.* **4**, 323—425 (1955).
- [9] RIVLIN, R. S.: Further remarks on stress-deformation relations for isotropic materials. *J. Rational Mech. Anal.* **4**, 681—702 (1955).
- [10] RIVLIN, R. S.: Solution of some problems in the exact theory of viscoelasticity. *J. Rational Mech. Anal.* **5**, 179—188 (1956).
- [11] TOUPIN, R. A.: World invariant kinematics. *Arch. Rational Mech. Anal.* **1**, 181—211 (1958).
- [12] TRUESDELL, C.: The mechanical foundations of elasticity and fluid dynamics. *J. Rational Mech. Anal.* **1**, 125—300 (1952) and **2**, 593—616 (1953).
- [13] TRUESDELL, C.: Hypo-elasticity. *J. Rational Mech. Anal.* **4**, 83—133 (1955).
- [14] TRUESDELL, C.: Hypo-elastic shear. *J. Appl. Physics* **27**, 441—447 (1956).
- [15] ZAREMBA, S.: Sur une forme perfectionnée de la théorie de la relaxation. *Bull. Int. Acad. Sci. Cracovie* **1903**, 594—614.

Carnegie Institute of Technology
Pittsburgh, Pennsylvania

(Received July 31, 1958)

**THE FOUNDATIONS OF CLASSICAL MECHANICS
IN THE LIGHT OF RECENT ADVANCES
IN CONTINUUM MECHANICS**

WALTER NOLL

**THE FOUNDATIONS OF CLASSICAL MECHANICS
IN THE LIGHT OF RECENT ADVANCES
IN CONTINUUM MECHANICS ¹**

WALTER NOLL

1. Introduction. It is a widespread belief even today that classical mechanics is a dead subject, that its foundations were made clear long ago, and that all that remains to be done is to solve special problems. This is not so. It is true that the mechanics of systems of a finite number of mass points has been on a sufficiently rigorous basis since Newton. Many textbooks on theoretical mechanics dismiss continuous bodies with the remark that they can be regarded as the limiting case of a particle system with an increasing number of particles. They cannot. The erroneous belief that they can had the unfortunate effect that no serious attempt was made for a long period to put classical continuum mechanics on a rigorous axiomatic basis. Only the recent advances in the theory of materials other than perfect fluids and linearly elastic solids have revived the interest in the foundations of classical mechanics. A clarification of these foundations is of importance also for the following reason. It is known that continuous matter is really made up of elementary particles. The basic laws governing the elementary particles are those of quantum mechanics. The science that provides the link between these basic laws and the laws describing the behavior of gross matter is statistical mechanics. At the present time this link is quite weak, partly because the mathematical difficulties are formidable, and partly because the basic laws themselves are not yet completely clear. A rigorous theory of continuum mechanics would give at least some precise information on what kind of gross behavior the basic laws ought to predict.

I want to give here a brief outline of an axiomatic scheme for continuum mechanics, and I shall attempt to introduce the same level of rigor and clarity as is now customary in pure mathematics. The mathematical

¹ The results presented in this paper were obtained in the course of research sponsored by the U.S. Air Force Office of Scientific Research under contract no. AF 18 (600)-1138 with Carnegie Institute of Technology.

structures involved are quite complex, and some fine details have to be omitted in order not to overburden the paper with technicalities.

Notation: Points and vectors in Euclidean space will be indicated by bold face letters. If \mathbf{x} and \mathbf{y} are two points, then $\mathbf{x} - \mathbf{y}$ denotes the vector determined by the ordered pair (\mathbf{y}, \mathbf{x}) . If \mathbf{x} is a point and \mathbf{v} a vector, then $\mathbf{x} + \mathbf{v}$ denotes the point uniquely determined by $(\mathbf{x} + \mathbf{v}) - \mathbf{x} = \mathbf{v}$. The word “smooth” will be used instead of “continuously differentiable”. Some equations will be valid only up to a set of measure zero. It will be clear from the context when this is the case.

2. Bodies.

DEFINITION 1: A BODY is a set \mathfrak{B} endowed with a structure defined by
 (a) a set Φ of mappings of \mathfrak{B} into a three-dimensional Euclidean point space E , and
 (b) a real valued set function m defined for a set of subsets of \mathfrak{B}
 subject to seven axioms as follows:

- (S.1) Every mapping $\varphi \in \Phi$ is one-to-one.
- (S.2) For each $\varphi \in \Phi$, the image $B = \varphi(\mathfrak{B})$ is a region in the space E , a region being defined as a compact set with piecewise smooth boundaries.
- (S.3) If $\varphi \in \Phi$ and $\psi \in \Phi$ then the mapping $\chi = \psi \circ \varphi^{-1}$ of $\varphi(\mathfrak{B})$ onto $\psi(\mathfrak{B})$ can be extended to a smooth homeomorphism of E onto itself.
- (S.4) If χ is a smooth homeomorphism of E onto itself and if $\varphi \in \Phi$, then also $\chi \circ \varphi \in \Phi$.

These four axioms give \mathfrak{B} the structure of a piece of a differentiable manifold that is isomorphic to a region in Euclidean three-space. The following three axioms give \mathfrak{B} the structure of a measure space.

- (M.1) m is a non-negative measure, defined for all Borel subsets \mathfrak{C} of \mathfrak{B} .
- (M.2) For each $\varphi \in \Phi$, the measure $\mu_\varphi = m \circ \varphi^{-1}$ induced by m on the region $B = \varphi(\mathfrak{B})$ in space is absolutely continuous relative to the Lebesgue measure in B . Hence it has a density ρ_φ so that

$$(2.1) \quad m(\mathfrak{C}) = \int_{\varphi(\mathfrak{C})} \rho_\varphi(\mathbf{x}) dV.$$

- (M.3) For each $\varphi \in \Phi$ the density ρ_φ is positive and bounded.

² The symbol \circ denotes the composition of mappings and a superposed -1 denotes the inverse of a mapping.

We use the following terminology: The elements X, Y, \dots of \mathfrak{B} are the **PARTICLES** of the body. The mappings $\varphi \in \Phi$ are the **CONFIGURATIONS** of the body. The point $\boldsymbol{x} = \varphi(X)$ is the **POSITION** of the particle X in the configuration φ . The set function m is the **MASS DISTRIBUTION** of the body. The number $m(\mathfrak{C})$ is the **MASS** of the set \mathfrak{C} . Here and subsequently we refer to Borel sets simply as sets. The density ρ_φ is the **MASS DENSITY** of \mathfrak{B} in the configuration φ . Note that it would have been sufficient to require the existence of ρ_φ only for one particular configuration φ . It then follows that the mass density exists also for all other configurations.

A compact subset \mathfrak{P} of \mathfrak{B} with piecewise smooth boundaries will be called a **PART** of the body \mathfrak{B} . It may again be regarded as a body whose configurations are the restrictions to \mathfrak{P} of the configurations of \mathfrak{B} and whose mass distribution is the restriction of the mass distribution of \mathfrak{B} to the subsets of \mathfrak{P} . Two parts \mathfrak{P} and \mathfrak{Q} will be called **SEPARATE** if

$$\mathfrak{P} \cap \mathfrak{Q} \subset \overline{\mathfrak{P}} \cap \overline{\mathfrak{Q}},$$

where $\overline{\mathfrak{P}}$ denotes the boundary of \mathfrak{P} .

3. Kinematics

DEFINITION 2: A **MOTION** of a body \mathfrak{B} is a one-parameter family $\{\theta_t\}$, $-\infty < t < \infty$, of configurations $\theta_t \in \Phi$ of \mathfrak{B} such that

(K.1) *The derivative*

$$(3.1) \quad \boldsymbol{v}(X, t) = \frac{d}{dt} \theta_t(X)$$

exists for all $X \in \mathfrak{B}$ and all t , it is a continuous function of X and t jointly, and it is a smooth function of X .

(K.2) *The derivative*

$$(3.2) \quad \dot{\boldsymbol{v}}(X, t) = \frac{d}{dt} \boldsymbol{v}(X, t) = \frac{d^2}{dt^2} \theta_t(X)$$

exists piecewise and is piecewise continuous in X and t jointly.

The parameter t is called the **TIME**. Derivatives with respect to t will be denoted by superposed dots. $\boldsymbol{v}(X, t)$ is called the **VELOCITY** of the particle X at time t . $\dot{\boldsymbol{v}}(X, t)$ is called the **ACCELERATION** of X at t .

Let h be any real, vector, or tensor valued function of X and t , and assume that $h(X, t)$ is smooth in X and t jointly. We may then associate

with h the function \hat{h} defined by

$$(3.3) \quad \hat{h}(\mathbf{x}, t) = h(\theta_t^{-1}(\mathbf{x}), t)$$

for $-\infty < t < \infty$ and $\mathbf{x} \in \theta_t(\mathfrak{B})$. By the chain rule of differentiation we have

$$(3.4) \quad \dot{h}(X, t) = \dot{\hat{h}}(\theta_t(X), t) + \nabla \hat{h}(\theta_t(X), t) \cdot \mathbf{v}(X, t),$$

where $\nabla \hat{h}$ denotes the gradient of \hat{h} with respect to \mathbf{x} . It is customary in the literature to use the same symbol for h and \hat{h} , to omit the independent variables, and to distinguish $\dot{\hat{h}}$ from \dot{h} by writing $\dot{\hat{h}} = \frac{\partial \hat{h}}{\partial t}$. Equation (3.4) then takes the familiar form

$$(3.5) \quad \dot{h} = \frac{\partial h}{\partial t} + \mathbf{v} \cdot \text{grad } h.$$

The LINEAR MOMENTUM at time t of a set $\mathfrak{C} \subset \mathfrak{B}$ is defined by

$$(3.6) \quad \mathbf{g}(\mathfrak{C}; t) = \int_{\mathfrak{C}} \mathbf{v}(X, t) dm.$$

It follows from (K.1) and (K.2) that $\mathbf{g}(\mathfrak{C}, t)$ is piecewise smooth in t . As a function of \mathfrak{C} it is a vector valued measure, absolutely continuous relative to m with density \mathbf{v} .

The ANGULAR MOMENTUM at time t of a set $\mathfrak{C} \subset \mathfrak{B}$, relative to a point $\mathbf{O} \in E$, is defined by

$$(3.7) \quad \mathbf{h}(\mathfrak{C}; t; \mathbf{O}) = \int_{\mathfrak{C}} [\theta_t(X) - \mathbf{O}] \times \mathbf{v}(X, t) dm.$$

It is piecewise smooth in t , and, as a function of \mathfrak{C} , it is a vector valued measure.

4. Forces

DEFINITION 3: A SYSTEM OF BODY FORCES for a body \mathfrak{B} is a family $\{\mathbf{B}_{\mathfrak{P}}\}$ of vector valued set functions subject to the following axioms:

- (B.1) For each part \mathfrak{P} of \mathfrak{B} , $\mathbf{B}_{\mathfrak{P}}$ is a vector valued measure defined on the Borel subsets of \mathfrak{P} .
- (B.2) For each \mathfrak{P} , $\mathbf{B}_{\mathfrak{P}}$ is absolutely continuous relative to the mass distribution m of \mathfrak{P} . Hence it has a density $\mathbf{b}_{\mathfrak{P}}$ so that

$$(4.1) \quad \mathbf{B}_{\mathfrak{P}}(\mathfrak{C}) = \int_{\mathfrak{C}} \mathbf{b}_{\mathfrak{P}}(X) dm.$$

(B.3) The density $\mathbf{b}_{\mathfrak{B}}$ is bounded, i.e.

$$|\mathbf{b}_{\mathfrak{B}}(X)| < k < \infty,$$

where k is independent of \mathfrak{B} and $X \in \mathfrak{B}$.

DEFINITION 4: A SYSTEM OF CONTACT FORCES for a body \mathfrak{B} is a family $\{\mathbf{C}_{\mathfrak{B}}\}$ of vector valued set functions subject to the following axioms:

(C.1) For each part \mathfrak{B} of \mathfrak{B} , $\mathbf{C}_{\mathfrak{B}}$ is a vector valued measure defined on the Borel subsets of \mathfrak{B} .

(C.2) $\mathbf{C}_{\mathfrak{B}}(\mathfrak{C}) = \mathbf{C}_{\mathfrak{B}}(\mathfrak{C} \cap \overline{\mathfrak{B}})$.

(C.3) If $\mathfrak{c} \subset \overline{\mathfrak{D}}$, $\mathfrak{c} \subset \overline{\mathfrak{B}}$, and $\mathfrak{B} \subset \mathfrak{D}$, then

$$\mathbf{C}_{\mathfrak{B}}(\mathfrak{c}) = \mathbf{C}_{\mathfrak{D}}(\mathfrak{c}).$$

(C.4) If $\varphi \in \Phi$ is any configuration of \mathfrak{B} and if $\overline{P} = \varphi(\overline{\mathfrak{B}})$, then the induced measure $\mathbf{C}_{\mathfrak{B}} \circ \varphi^{-1}$, when restricted to the Borel subsets of the boundary surface \overline{P} of $P = \varphi(\mathfrak{B})$, is absolutely continuous relative to the Lebesgue surface measure on \overline{P} . Hence it has a density $\mathbf{s}(\mathfrak{B}, \varphi)$ so that

$$(4.2) \quad \mathbf{C}_{\mathfrak{B}}(\mathfrak{c}) = \int_{\varphi(\mathfrak{c})} \mathbf{s}(\mathfrak{B}, \varphi; \mathbf{x}) dA$$

for all Borel subsets $\mathfrak{c} \subset \overline{\mathfrak{B}}$.

(C.5) The density $\mathbf{s}(\mathfrak{B}, \varphi)$ is bounded, i.e.

$$|\mathbf{s}(\mathfrak{B}, \varphi; \mathbf{x})| < l < \infty,$$

where l does not depend on \mathfrak{B} or $\mathbf{x} \in \varphi(\overline{\mathfrak{B}})$.

As in the case of a mass distribution, it would, have been sufficient in (C.4) to require the existence of $\mathbf{s}(\mathfrak{B}, \varphi)$ only for a particular $\varphi \in \Phi$. The existence of \mathbf{s} for all other configurations is then an automatic consequence. The axiom (C.2) means that $\mathbf{C}_{\mathfrak{B}}$ is essentially a vector measure on the boundary $\overline{\mathfrak{B}}$.

It is useful to consider surfaces in \mathfrak{B} as being *oriented*, and to employ the operation of addition of oriented surfaces in the sense of algebraic topology. The boundary $\overline{\mathfrak{B}}$ of a part \mathfrak{B} of \mathfrak{B} will be regarded as oriented in such a way that the positive side of $\overline{\mathfrak{B}}$ is exterior to \mathfrak{B} . If \mathfrak{B} and \mathfrak{D} are two separate parts of \mathfrak{B} , then

$$(4.3) \quad \overline{\mathfrak{B} \cup \mathfrak{D}} = \overline{\mathfrak{B}} + \overline{\mathfrak{D}}.$$

This is true because the common boundary of \mathfrak{B} and \mathfrak{D} , if any, appears

twice with opposite orientation on the right side of (4.3) and hence cancels. We shall say that the surface c is a *PIECE* of the surface \mathfrak{d} if c is a subset of \mathfrak{d} and if the orientation of c is induced by \mathfrak{d} . The significance of the axiom (C.3) is brought out by the following theorem:

THEOREM I: *There is a vector valued function S , defined for all oriented surfaces c in \mathfrak{B} , such that*

$$(4.4) \quad C_{\mathfrak{P}}(c) = S(c)$$

whenever c is a piece of the boundary $\overline{\mathfrak{P}}$ of \mathfrak{P} . We say that $S(c)$ is the CONTACT FORCE ACTING ACROSS THE ORIENTED SURFACE c .

Proof: For each c which is not a piece of $-\overline{\mathfrak{B}}$ we can find a part $\mathfrak{D}(c)$ of \mathfrak{B} such that c is a piece of $\overline{\mathfrak{D}(c)}$. We then define $S(c) = C_{\mathfrak{D}(c)}(c)$. Now let \mathfrak{P} be an arbitrary part of \mathfrak{B} and let c be a piece of $\overline{\mathfrak{P}}$. We then have

$$\begin{aligned} c \subset \overline{\mathfrak{P}}, \quad c \subset \overline{\mathfrak{D}(c)}, \quad c \subset \overline{\mathfrak{D}(c) \cap \mathfrak{P}}, \\ \mathfrak{P} \cap \mathfrak{D}(c) \subset \mathfrak{P}, \quad \mathfrak{P} \cap \mathfrak{D}(c) \subset \mathfrak{D}(c). \end{aligned}$$

Applying axiom (C.3) twice, we get

$$C_{\mathfrak{P}}(c) = C_{\mathfrak{P} \cap \mathfrak{D}(c)}(c), \quad C_{\mathfrak{D}(c)}(c) = C_{\mathfrak{P} \cap \mathfrak{D}(c)}(c).$$

Hence

$$C_{\mathfrak{P}}(c) = C_{\mathfrak{D}(c)}(c) = S(c).$$

If c is a part of $-\overline{\mathfrak{B}}$ we define

$$(4.5) \quad S(c) = -S(-c).$$

It follows from theorem I and axiom (C.4) that there is a vector valued function $s(c, \varphi; \mathbf{x})$ such that

$$(4.6) \quad S(c) = \int_{\varphi(c)} s(c, \varphi; \mathbf{x}) dA.$$

Also, if $\mathbf{x} \in \varphi(\mathfrak{d}) \subset \varphi(c)$ and if \mathfrak{d} is a piece of c , then

$$(4.7) \quad s(c, \varphi; \mathbf{x}) = s(\mathfrak{d}, \varphi; \mathbf{x}).$$

If c_1 and c_2 are two pieces of a surface c and if $c = c_1 + c_2$, then

$$(4.8) \quad S(c) = S(c_1) + S(c_2).$$

This is true because $C_{\mathfrak{P}}$, as a measure, is additive and because, by axiom (C.4) the value of $C_{\mathfrak{P}}$ for the common boundary curve of c_1 and c_2 is zero.

DEFINITION 5: A SYSTEM OF FORCES for a body \mathfrak{B} is a family of vector valued measures $\{F_{\mathfrak{P}}\}$ such that, for each part \mathfrak{P} of \mathfrak{B} , $F_{\mathfrak{P}}$ is defined on the subsets of \mathfrak{P} and such that the $F_{\mathfrak{P}}$ have decompositions

$$(4.9) \quad F_{\mathfrak{P}} = B_{\mathfrak{P}} + C_{\mathfrak{P}},$$

where $\{B_{\mathfrak{P}}\}$ is a system of body forces and $\{C_{\mathfrak{P}}\}$ is a system of contact forces.

It is not hard to see that the decomposition (4.9), if it exists, is automatically unique.

We use the following terminology: The measure $F_{\mathfrak{P}}$ is the FORCE acting on the part \mathfrak{P} of \mathfrak{B} . The vector $F_{\mathfrak{P}}(\mathfrak{P})$ is the RESULTANT FORCE acting on \mathfrak{P} . Let \mathfrak{P} and \mathfrak{Q} be two separate parts of \mathfrak{B} . The vector measure

$$(4.10) \quad F_{\mathfrak{P},\mathfrak{Q}} = F_{\mathfrak{P}} - F_{\mathfrak{P} \cup \mathfrak{Q}}$$

defined on the subsets of \mathfrak{P} , is the MUTUAL FORCE exerted on \mathfrak{P} by \mathfrak{Q} . The mutual force exerted on a part \mathfrak{P} of \mathfrak{B} by the closure of its complement is denoted by $F_{(\mathfrak{P})}$ and it is called the INTERNAL FORCE acting on \mathfrak{P} . The restriction of $F_{\mathfrak{P}}$ to a part \mathfrak{P} of \mathfrak{B} is the EXTERNAL FORCE acting on \mathfrak{P} . A similar terminology and notation will be used when "force" is replaced by "body force" or by "contact force".

Let $\{F_{\mathfrak{P}}\}$ be a system of forces for a body \mathfrak{B} , $\varphi \in \mathcal{D}$ a configuration of \mathfrak{B} , and $\mathbf{O} \in E$ a point in space. The MOMENT about \mathbf{O} of the force $F_{\mathfrak{P}}$ acting on the part \mathfrak{P} of \mathfrak{B} in the configuration φ is the vector valued measure $M(F_{\mathfrak{P}}, \varphi, \mathbf{O})$ defined by

$$(4.11) \quad M(F_{\mathfrak{P}}, \varphi, \mathbf{O}; \mathfrak{C}) = \int_{\mathfrak{C}} [\varphi(X) - \mathbf{O}] \times dF_{\mathfrak{P}}$$

for the subsets \mathfrak{C} of \mathfrak{P} . The vector $M(F_{\mathfrak{P}}, \varphi, \mathbf{O}; \mathfrak{P})$ is the RESULTANT MOMENT about \mathbf{O} acting on \mathfrak{P} .

5. Dynamical processes

DEFINITION 6: A DYNAMICAL PROCESS is a triple $\{\mathfrak{B}, \theta_t, F_{\mathfrak{P},t}\}$, where \mathfrak{B} is a body, θ_t is a motion of \mathfrak{B} , and $F_{\mathfrak{P},t}$ is a one-parameter family of systems of forces for \mathfrak{B} , subject to the following two axioms:

(D.1) Principle of linear momentum: For all parts \mathfrak{P} of \mathfrak{B} and all times t ,

$$(5.1) \quad F_{\mathfrak{P},t}(\mathfrak{P}) = \dot{\mathbf{g}}(\mathfrak{P}; t),$$

where \mathbf{g} is defined by (3.6). In words: The resultant force acting on the part \mathfrak{P} is equal to the rate of change of the linear momentum of \mathfrak{P} .

(D.2) Principle of angular momentum: Let $\mathbf{O} \in E$ be any point in space.

Then for all parts \mathfrak{P} of \mathfrak{B} and all times t ,

$$(5.2) \quad \mathbf{M}(F_{\mathfrak{P},t}, \theta_t, \mathbf{O}; \mathfrak{P}) = \dot{\mathbf{h}}(\mathfrak{P}; t; \mathbf{O}),$$

where \mathbf{h} and \mathbf{M} are defined by (3.7) and (4.11), respectively. In words: The resultant moment about \mathbf{O} acting on a part \mathfrak{P} is equal to the rate of change of the angular momentum of \mathfrak{P} relative to \mathbf{O} .

It would have been sufficient to require that (5.2) holds for a particular $\mathbf{O} \in E$. It is then automatically valid for all points in space. Also, (5.2) remains valid if the fixed point \mathbf{O} is replaced by the variable mass center

$$(5.3) \quad \mathbf{c}(\mathfrak{P}, t) = \mathbf{O} + \frac{1}{m(\mathfrak{P})} \int_{\mathfrak{P}} (\theta_t(X) - \mathbf{O}) dm$$

of the part \mathfrak{P} . These statements can be proved in the classical manner.

We now prove a number of important theorems. For simplicity we omit the variable t ; we write

$$(5.4) \quad \mathbf{s}(\mathbf{c}; \mathbf{x}) = \mathbf{s}(\mathbf{c}, \theta_t; \mathbf{x})$$

for the density of the contact force as defined by (4.6).

THEOREM II: For any two separate parts \mathfrak{P} and \mathfrak{Q} of \mathfrak{B} we have

$$(5.5) \quad F_{\mathfrak{P},\mathfrak{Q}}(\mathfrak{P}) = -F_{\mathfrak{Q},\mathfrak{P}}(\mathfrak{Q})$$

i.e. the resultant mutual force exerted on \mathfrak{P} by \mathfrak{Q} is equal and opposite to the resultant mutual force exerted on \mathfrak{Q} by \mathfrak{P} .

Proof: We apply axiom (D.1) to \mathfrak{P} , \mathfrak{Q} , and $\mathfrak{P} \cup \mathfrak{Q}$:

$$(5.6) \quad F_{\mathfrak{P}}(\mathfrak{P}) = \dot{\mathbf{g}}(\mathfrak{P}), \quad F_{\mathfrak{Q}}(\mathfrak{Q}) = \dot{\mathbf{g}}(\mathfrak{Q}), \quad F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P} \cup \mathfrak{Q}) = \dot{\mathbf{g}}(\mathfrak{P} \cup \mathfrak{Q}).$$

Since $\mathfrak{P} \cap \mathfrak{Q}$ has no mass by (M.2), it follows from (3.6) that

$$\mathbf{g}(\mathfrak{P} \cup \mathfrak{Q}) = \mathbf{g}(\mathfrak{P}) + \mathbf{g}(\mathfrak{Q});$$

hence, by (5.6),

$$F_{\mathfrak{P}}(\mathfrak{P}) + F_{\mathfrak{Q}}(\mathfrak{Q}) = F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P} \cup \mathfrak{Q}).$$

It is not hard to see that $F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P} \cap \mathfrak{Q}) = 0$. Hence

$$F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P} \cup \mathfrak{Q}) = F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P}) + F_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{Q}).$$

The assertion follows now from the definition (4.10).

THEOREM III (reaction principle)³: *The contact force $S(c)$ acting across c is opposite to the contact force acting across $-c$, i.e.*

$$(5.7) \quad S(c) = -S(-c)$$

Proof: If c is a piece of $\overline{\mathfrak{B}}$, then (5.7) is true by the definition (4.5). If not, it is possible to find two separate parts \mathfrak{P} and \mathfrak{Q} such that $\mathfrak{P} \cap \mathfrak{Q} = c$ (see Fig. 1). We orient c such that it is a piece of $\overline{\mathfrak{P}}$. Then $-c$

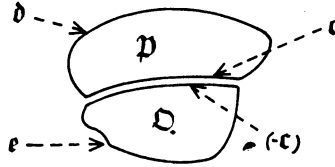


Fig. 1

will be a piece of $\overline{\mathfrak{Q}}$. The surfaces $\overline{\mathfrak{P}}$, $\overline{\mathfrak{Q}}$, and $\overline{\mathfrak{P} \cup \mathfrak{Q}}$ have decompositions

$$\overline{\mathfrak{P}} = c + b, \quad \overline{\mathfrak{Q}} = (-c) + e, \quad \overline{\mathfrak{P} \cup \mathfrak{Q}} = b + e.$$

It follows from theorem I and (4.8) that

$$C_{\mathfrak{P}}(\mathfrak{P}) = S(c) + S(b), \quad C_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P}) = S(b)$$

and hence that

$$C_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{P}) = C_{\mathfrak{P}}(\mathfrak{P}) - C_{\mathfrak{P} \cup \mathfrak{Q}}(\mathfrak{P}) = S(c).$$

Similarly, we obtain

$$C_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{Q}) = S(-c).$$

For the total resultant mutual forces, we get

$$(5.8) \quad \begin{aligned} F_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{P}) &= B_{\mathfrak{P}, \mathfrak{Q}}(\mathfrak{P}) + S(c) \\ F_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{Q}) &= B_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{Q}) + S(-c). \end{aligned}$$

Application of theorem II gives

$$(5.9) \quad S(c) + S(-c) = -[B_{\mathfrak{P}, \mathfrak{Q}}(\mathfrak{P}) + B_{\mathfrak{Q}, \mathfrak{P}}(\mathfrak{Q})].$$

Using axiom (M.3) one can show that the parts \mathfrak{P} and \mathfrak{Q} can be chosen

³ Various statements, mostly quite vague, pass under the title "principle of action and reaction" in the literature. All of these statements, when made precise, are provable theorems in the theory presented here.

such that their masses $m(\mathfrak{B})$ and $m(\mathfrak{Q})$ are arbitrarily small. Axiom (B.3) then implies that the right side of (5.9) can be made arbitrarily small in absolute value. It follows that the left side of (5.9) must vanish. Q.e.d.

As a corollary, it follows that

$$(5.10) \quad \mathcal{S}(c_1 + c_2) = \mathcal{S}(c_1) + \mathcal{S}(c_2),$$

no matter whether c_1 and c_2 are pieces of $c = c_1 + c_2$, as in (4.8), or not. Hence \mathcal{S} may be regarded as an additive vector valued function of oriented surfaces in \mathfrak{B} . Another corollary is that the statement of theorem II remains true if “mutual force” there is replaced by “mutual contact force” or by “mutual body force”.

THEOREM IV (stress principle)⁴: *There is a vector valued function $\mathbf{s}(\mathbf{x}, \mathbf{n})$, where $\mathbf{x} \in \theta_t(\mathfrak{B})$ and where \mathbf{n} is a unit vector, such that*

$$(5.11) \quad \mathbf{s}(c; \mathbf{x}) = \mathbf{s}(\mathbf{x}, \mathbf{n})$$

whenever $\theta_t(c)$ has the unit normal \mathbf{n} at $\mathbf{x} \in \theta_t(c)$, directed towards the positive side of the oriented surface $\theta_t(c)$, the orientation of $\theta_t(c)$ being induced by the orientation of c .

Proof: Let c_1 and c_2 be two surfaces in \mathfrak{B} tangent to each other at $X = \theta_t^{-1}(\mathbf{x})$. The surfaces $c_1 = \theta_t(c_1)$ and $c_2 = \theta_t(c_2)$ in space E are then tangent to each other at the point \mathbf{x} . We assume that \mathbf{n} is their unit normal at \mathbf{x} and that c_1 and c_2 are oriented in such a way that \mathbf{n} is directed

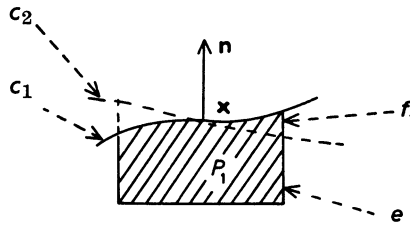


Fig. 2

toward the positive side of c_1 and c_2 . Consider the region P_1 bounded by a piece d_1 of c_1 , a piece of a circular cylinder f of radius r whose axis is \mathbf{n}

⁴ The assertion of this theorem appears in all of the past literature as an assumption. It has been proposed occasionally that one should weaken this assumption and allow the stress to depend not only on the tangent plane at \mathbf{x} , but also on the curvature of the surface c at \mathbf{x} . The theorem given here shows that such dependence on the curvature, or on any other local property of the surface at \mathbf{x} , is impossible.

and by a plane perpendicular to \mathbf{n} at a distance r from \mathbf{x} as shown in Fig. 2. The region P_2 is defined in a similar manner. Denote the common boundary of P_1 and P_2 on the cylinder and the plane by e . The boundaries \bar{P}_1 and \bar{P}_2 then have decompositions into separate pieces of the form

$$(5.12) \quad \bar{P}_1 = d_1 + e + f_1, \quad \bar{P}_2 = d_2 + e + f_2,$$

where f_1 and f_2 are pieces of the cylinder f . We denote the surface area of a surface c by $A(c)$ and the volume of a region P by $V(P)$. It is not hard to see that then

$$(5.13) \quad A(d_i) = \pi r^2 + o(r^2),$$

$$(5.14) \quad A(f_i) = o(r^2),$$

$$(5.15) \quad V(P_i) = o(r^2)$$

for $i = 1, 2$. $\mathfrak{P}_1 = \theta_t^{-1}(P_1)$ and $\mathfrak{P}_2 = \theta_t^{-1}(P_2)$ will be parts of \mathfrak{B} for small enough r , except when $\mathbf{x} \in \theta_t(\mathfrak{B})$, and \mathbf{n} is directed toward the interior of $\theta_t(\mathfrak{B})$. Applying axiom (D.1) to \mathfrak{P}_1 and \mathfrak{P}_2 gives

$$(5.16) \quad F_{\mathfrak{P}_i}(\mathfrak{P}_i) = B_{\mathfrak{P}_i}(\mathfrak{P}_i) + C_{\mathfrak{P}_i}(\mathfrak{P}_i) = \int_{\mathfrak{P}_i} \dot{\mathbf{v}} dm, \quad i = 1, 2.$$

By (4.1) and (4.4) this may be written in the form

$$(5.17) \quad S(\bar{\mathfrak{P}}_i) = \int_{\mathfrak{P}_i} (\dot{\mathbf{v}} - \mathbf{b}_{\mathfrak{P}_i}) dm.$$

By (4.6), (4.7), (4.8), and (5.12) we have

$$(5.18) \quad S(\bar{\mathfrak{P}}_i) = \int_{d_i} \mathbf{s}(c_i) dA + \int_{f_i} \mathbf{s}(f_i) dA + \int_e \mathbf{s}(e) dA; \quad i = 1, 2,$$

where $\bar{f}_i = \theta_t^{-1}(f_i)$, $e = \theta_t^{-1}(e)$. Subtracting the two equations (5.18) and using (5.17), we get

$$(5.19) \quad \int_{d_1} \mathbf{s}(c_1) dA - \int_{d_2} \mathbf{s}(c_2) dA = \\ = \int_{\mathfrak{P}_1} (\dot{\mathbf{v}} - \mathbf{b}_{\mathfrak{P}_1}) dm - \int_{\mathfrak{P}_2} (\dot{\mathbf{v}} - \mathbf{b}_{\mathfrak{P}_2}) dm - \int_{f_1} \mathbf{s}(f_1) dA + \int_{f_2} \mathbf{s}(f_2) dA.$$

Since $\dot{\mathbf{v}}$, $\mathbf{b}_{\mathfrak{P}_i}$ and the mass density are bounded by constants independent of \mathfrak{B} , according to the axioms (K.2), (B.3), and (M.3), it follows from (5.15) that

$$\int_{\mathfrak{P}_i} (\dot{\mathbf{v}} - \mathbf{b}_{\mathfrak{P}_i}) dm = o(r^2), \quad i = 1, 2.$$

Similarly, it follows from axiom (C.5) and from (5.14) that

$$\int_{f_i} \mathbf{s}(f_i) dA = o(r^2), \quad i = 1, 2.$$

Hence, by (5.19),

$$\int_{d_1} \mathbf{s}(c) dA = \int_{d_2} \mathbf{s}(c_2) dA + o(r^2).$$

Dividing by πr^2 and using (5.13), we get

$$(5.20) \quad \frac{\int_{d_1} \mathbf{s}(c_1) dA}{A(d_1)} = \frac{\int_{d_2} \mathbf{s}(c_2) dA}{A(d_2)} + \frac{o(r^2)}{\pi r^2}.$$

By a theorem on measures with density, we have

$$\lim_{r \rightarrow 0} \frac{\int_{d_i} \mathbf{s}(c_i) dA}{A(d_i)} = \mathbf{s}(c_i; \mathbf{x}), \quad i = 1, 2.$$

Thus, letting r go to zero in (5.20), we finally obtain

$$\mathbf{s}(c_1; \mathbf{x}) = \mathbf{s}(c_2; \mathbf{x}),$$

which shows that $\mathbf{s}(c; \mathbf{x})$ has the same value for all surfaces c with the same unit normal \mathbf{n} . The exceptional case when $\mathbf{x} \in \theta_i^{-1}(\overline{\mathfrak{B}})$ and \mathbf{n} is directed toward the interior of $\theta_i(\mathfrak{B})$ is taken care of by the definition (4.5).

The vector $\mathbf{s}(\mathbf{x}, \mathbf{n})$ is called the **STRESS** acting at \mathbf{x} across the surface element with unit normal \mathbf{n} . By (4.6) the contact force $\mathbf{S}(c)$ acting across c is given by

$$(5.21) \quad \mathbf{S}(c) = \int_{\theta_i(c)} \mathbf{s}(\mathbf{x}, \mathbf{n}) dA,$$

where \mathbf{n} is the unit normal at \mathbf{x} to the oriented surface $\theta_i(c)$. By theorem II we have

$$(5.22) \quad \mathbf{s}(\mathbf{x}, \mathbf{n}) = -\mathbf{s}(\mathbf{x}, -\mathbf{n}).$$

The following two additional assumptions suffice to ensure the validity of the classical theorems of continuum mechanics:

- (a) The stress $\mathbf{s}(\mathbf{x}, \mathbf{n})$, for each \mathbf{n} , is a smooth function of $\mathbf{x} \in \theta_i(\mathfrak{B})$.

(b) For almost all $X \in \mathfrak{B}$, the limit

$$(5.23) \quad \mathbf{b}(X) = \lim_{\mathfrak{B} \rightarrow X} \frac{1}{m(\mathfrak{B})} \mathbf{B}_{\mathfrak{B}}(\mathfrak{B}),$$

where \mathfrak{B} is a neighborhood of X contracting to X , exists.

Under these assumptions, one can prove the following theorems in the classical manner:

(1) There is a field of linear transformations $S(\mathbf{x})$, $\mathbf{x} \in \theta_t(\mathfrak{B})$, such that

$$(5.24) \quad \mathbf{s}(\mathbf{x}, \mathbf{n}) = S(\mathbf{x})\mathbf{n}.$$

$S(\mathbf{x})$ is called the STRESS TENSOR at \mathbf{x} .

(2) The stress tensor $S(\mathbf{x})$ is symmetric.

(3) Cauchy's equation of motion

$$(5.25) \quad \operatorname{div} S + \rho \mathbf{b} = \rho \dot{\mathbf{v}}$$

holds, where S is the stress tensor, ρ is the mass density, $\dot{\mathbf{v}}$ is the acceleration, and \mathbf{b} is defined by (5.23).

6. Equivalence of dynamical processes. The position of a particle can be specified physically not in an absolute sense but only relative to a given *frame of reference*. Such a frame is a set of objects whose mutual distances change very little in time, like the walls of a laboratory, the fixed stars, or the wooden horses on a merry-go-round. In classical physics, a change of frame corresponds to a transformation of space and time which preserves distances and time intervals. It is well known that the most general such transformation is of the form

$$(6.1) \quad \begin{aligned} \mathbf{x}^* &= \mathbf{c}(t) + Q(t)(\mathbf{x} - \mathbf{O}), \\ t^* &= t + a, \end{aligned}$$

where $\mathbf{c}(t)$ is a point valued function of t , $Q(t)$ is a function of t whose values are orthogonal transformations, a is a real constant, and \mathbf{O} is a point, which may be fixed once and for all. We assume that $\mathbf{c}(t)$ and $Q(t)$ are twice continuously differentiable. A change of frame (6.1) also induces a transformation on vectors and tensors. A vector \mathbf{u} , for example, is transformed into

$$(6.2) \quad \mathbf{u}^* = Q(t)\mathbf{u}.$$

Let $\{\mathfrak{B}, \theta_t, \mathbf{F}_{\mathfrak{B}, t}\}$ be a dynamical process. A change of frame $\{\mathbf{c}, Q, a\}$

will transform the motion θ_t into a new motion $\theta_{t'}$ defined by

$$(6.3) \quad \theta_{t'}(X) = \mathbf{c}(t - a) + Q(t - a)[\theta_{t-a}(X) - \mathbf{O}].$$

The velocities and the accelerations of the two motions θ_t and $\theta_{t'}$ are, in general, not related by the transformation formula (6.2) for vectors. They depend on the choice of the frame of reference. We say that they are not *objective*. However, there are objective kinematical quantities, for example the rate of deformation tensor.

If we wish to assume that forces have an objective meaning we would have to require that $\mathbf{F}_{\mathfrak{B},t}(\mathfrak{C})$ transforms according to the law (6.2) under a change of frame. However, when this assumption is made, a dynamical process does not transform into a dynamical process because the axioms (D.1) and (D.2) are not preserved, except when \mathbf{c} is linear in t and Q is constant. It is this difficulty which has led to the concept of absolute space and which has caused much controversy in the history of mechanics. A clarification was finally given by Einstein in his general theory of relativity, in which gravitational forces and inertial forces cannot be separated from each other in an objective manner. If we wish to stay in the realm of classical mechanics we may resolve the paradox by sacrificing the objectivity of the external body forces while retaining the objectivity of the essential types of forces, the contact forces and the mutual body forces. This can be done by assuming that the forces transform according to a law of the form

$$(6.4) \quad \mathbf{F}'_{\mathfrak{B},t}(\mathfrak{C}) = Q(t - a)\mathbf{F}_{\mathfrak{B},t-a}(\mathfrak{C}) + \mathbf{I}(\mathfrak{C}, t).$$

Here $\mathbf{I}(\mathfrak{C}, t)$ will be called the INERTIAL FORCE acting on \mathfrak{C} due to the change of frame $\{\mathbf{c}, Q, a\}$.

DEFINITION 7: Two dynamical processes $\{\mathfrak{B}, \theta_t, \mathbf{F}_{\mathfrak{B},t}\}$ and $\{\mathfrak{B}, \theta_{t'}, \mathbf{F}'_{\mathfrak{B},t}\}$ are called EQUIVALENT if there is a change of frame $\{\mathbf{c}, Q, a\}$ such that $\theta_{t'}$ and $\mathbf{F}'_{\mathfrak{B},t}$ are related to θ_t and $\mathbf{F}_{\mathfrak{B},t}$ by (6.3) and (6.4).

The classical analysis of relative motion shows that the inertial force $\mathbf{I}(\mathfrak{C}, t)$ is necessarily of the form

$$(6.5) \quad \mathbf{I}(\mathfrak{C}, t) = \int_{\mathfrak{C}} \mathbf{i}(X, t) dm$$

with

$$(6.6) \quad \mathbf{i}(X, t) = \ddot{\mathbf{c}}(t - a) + 2V(t - a)[\mathbf{v}'(X, t) - \dot{\mathbf{c}}(t - a)] + [V^2(t - a) - \dot{V}(t - a)][\theta_{t'}(X) - \mathbf{c}(t - a)]$$

where ν' is the velocity of the motion θ_t' , and where $V(t)$ is defined by

$$(6.7) \quad V(t) = \dot{Q}(t)Q(t)^{-1}.$$

It is not hard to see that the inertial force I gives a contribution only to the external body forces and that the contact forces and the mutual forces transform according to (6.2) and hence are objective. The external body forces and the inertial forces cannot be separated from each other in an objective manner. Experience shows that, for the body consisting of the entire solar system, there are frames relative to which the external body forces nearly vanish. These are the classical Galilean frames. Two equivalent dynamical processes really correspond to the same physical process, viewed only from two different frames of reference.

7. Constitutive assumptions. An axiom that characterizes the particular material properties of a body is called a CONSTITUTIVE ASSUMPTION. It restricts the class of dynamical processes the body can undergo. A familiar example is the assumption that the body is rigid. It restricts the possible motions to those in which the distance between any two particles remains unchanged in time. More important for modern continuum mechanics are constitutive assumptions in the form of functional relations between the stress tensor S and the motion θ_t . Such relations are called CONSTITUTIVE EQUATIONS (sometimes also rheological equations of state or stress-strain relations). A classical example is the constitutive equation for linear viscous fluids

$$(7.1) \quad S = (-p + \lambda \operatorname{tr} D)I + 2\mu D,$$

where D is the rate of deformation tensor, I is the unit tensor, p is the pressure, and λ and μ are viscosity constants. A wide variety of constitutive equations have been investigated in recent years⁵, and a general theory of such equations has been developed [2].

Constitutive assumptions are subject to a general restriction:

PRINCIPLE OF OBJECTIVITY: *If a dynamical process is compatible with a constitutive assumption then all processes equivalent to it must also be compatible with this constitutive assumption. In other words, constitutive assumptions must be invariant under changes of frame.*

This principle, although implicitly used by many scientists in the history of mechanics, was stated explicitly first by Oldroyd [3] and was

⁵ A review of the literature and a bibliography is given in [1].

clarified further by the author [4]. It is of great importance in the theory of constitutive equations.

8. Unsolved problems. The axiomatic treatment given here is still too special. It does not cover concentrated forces, contact couples and body couples, sliding, impact, rupture, and other discontinuities, singularities, and degeneracies. It would be desirable to have a universal scheme which covers any conceivable situation.

A more fundamental physical problem is to find a rigorous unified theory of continuum mechanics and thermodynamics. Classical thermodynamics deals only with equilibrium states and hence is not adequate for processes with fast changes of state in time. Such a unified theory should lead to further restrictive conditions on the form of constitutive equations and hence to more definite and realistic theories for special materials. Also, a satisfactory connection with statistical mechanics can be expected only after such a theory has been developed.

References

- [1] NOLL, W., ERICKSEN, J. L. and TRUESDELL, C., *The Non-linear Field Theories of Mechanics*. Article to appear in the Encyclopedia of Physics.
- [2] ———, *A general theory of constitutive equations*. To appear in Archive for Rational Mechanics and Analysis.
- [3] OLDROYD, J. G., *On the formulation of rheological equations of state*. Proceedings of the Royal Society of London (A) 200 (1950), 523–541.
- [4] NOLL, W., *On the Continuity of the Solid and Fluid States*. Journal of Rational Mechanics and Analysis 4 (1955), 3–81.

On Certain Steady Flows of General Fluids

BERNARD D. COLEMAN & WALTER NOLL

Contents

Introduction	289
1. Basic concepts	290
2. A class of steady flows	292
3. Reduction of the general constitutive equation	293
4. Rectilinear flow	295
5. Simple shearing flow	296
6. Flow through a channel	297
7. Poiseuille flow	299
8. Couette flow	300

Introduction

In a recent article on the mechanical behavior of continuous media [1], the physical concept of a *fluid* was given a mathematical definition. We believe that this definition covers almost all real fluids (whenever thermal and other non-mechanical effects can be disregarded) and is more general than any proposed previously because it accounts for all hereditary effects including stress relaxation. The general fluids covered by this definition may even exhibit, in some situations, physical phenomena which are usually attributed to solids.

Here we show that certain steady flow problems can be solved assuming only the definition of a general fluid and incompressibility. The class of problems considered includes most of the classical flows: simple shearing flow, channel flow (*i.e.* flow between fixed plates), Poiseuille flow, and Couette flow. It also includes torsional and other flows for which we do not work out the details, because, under reasonable body forces, such flows are not compatible with the dynamical equations unless inertia is neglected.

The solutions obtained here are in terms of three unspecified real functions of one variable which depend on the particular material; we call them *material functions*. Our solutions are simply related. If experimental measurements are combined with any one of the solutions to determine the three material functions, then complete stress and velocity profiles can be predicted for the other flows. Thus, our work really establishes correspondence principles for the flows considered. It is hoped that our results may help experimental rheologists to rationalize the treatment of their data by giving precise meaning to such concepts as "shear-dependent viscosity".

If one uses a special theory, either molecular or phenomenological, to solve a particular flow problem of our class (say, simple shearing flow), then our three

material functions can be readily identified*. The solutions to the other flow problems can then be obtained from the formulae given here without further reference to the original special theory**.

It should be pointed out that our three material functions do not determine the behavior of the fluid in all experimental situations, since the functional that predicts the response of a material to all situations may contain vastly more information than the three material functions introduced here.

Some of the solutions and relations obtained here have appeared in the literature, but the derivations given have rested upon special assumptions about the nature of the fluid considered. Of these previous investigations, the most general is RIVLIN's study of flows in fluids of the differential type [4]. Such fluids do not have long range memory and hence constitute only a special class of the general fluids considered here. It turns out, however, that RIVLIN's solutions (including those for helical flow, which will be discussed in a future note) remain valid for general fluids.

This paper can be read independently of reference [1] if the reader is willing to accept equations (1.1)–(1.3) below.

Notation. We denote vectors and points by bold face Latin or Greek minuscules: \mathbf{v} , \mathbf{x} , $\boldsymbol{\xi}$.

Second order tensors are always denoted by light face Latin majuscules: A , B , I , Q . The transpose of a tensor Q is written Q^T . The unit or "metric" tensor is denoted by I . The symbol $t_{i,j}$ denotes the *physical* components of T relative to an orthogonal, but not necessarily Cartesian, coordinate system, except when a statement is made to the contrary. We use the summation convention throughout. The matrix of the physical components of T is denoted by (T) or $\|t_{i,j}\|$, and the trace of T is written $\text{tr } T$.

Partial differentiation is indicated as follows: $\partial_y = \partial/\partial y$ or $\partial_i = \partial/\partial x_i$.

1. Basic concepts

In a general fluid ([1], §§ 21 and 22) the stress tensor S at time t depends on the history of the motion up to t and is given by a constitutive equation which may be stated in the form (22.12) of [1]. In this paper we discuss only isochoric motions; we do not make the dependence of S on the density $\rho(t)$ explicit, and we write

$$S(t) = \mathfrak{F}(C_t^i). \quad (1.1)$$

Here \mathfrak{F} is a functional of the tensor-valued function C_t^i and satisfies the relation

$$Q \mathfrak{F}(C_t^i) Q^T = \mathfrak{F}(Q C_t^i Q^T) \quad (1.2)$$

for any constant orthogonal tensor Q (cf. [1], (21.7)). The function C_t^i is the history of the right Cauchy-Green tensor taken with respect to the configuration

* A summary of results pertaining to the normal stress effect for special theories has been presented by MARKOVITZ [2]. The present investigation gives a proof of his equations (12)–(16), which he offered as conjectures.

** For example, much of the work in Chapter IV of [3] could have been avoided if the present results had been known to the author at the time.

at time t and is defined by

$$C_i^t(s) = C_i(t+s) = F_i^T(t+s) F_i(t+s), \quad s \leq 0, \tag{1.3}$$

where $F_i(t+s)$ is the displacement gradient at time $t+s$ relative to the configuration at time t (cf. [I], §§ 6 and 7).

In this paper we consider homogeneous *incompressible* fluids. The stress is then determined by the history of the motion only up to a hydrostatic pressure p , and (1.1) must be replaced by

$$T(t) = \mathfrak{F}(C_i^t), \tag{1.4}$$

where

$$T = S + pI \tag{1.5}$$

is called the *extra-stress*. Moreover, the tensor-valued functional \mathfrak{F} is determined only up to an arbitrary scalar-valued functional of C_i^t . We remove this indeterminacy by the normalization

$$\text{tr } T = \text{tr}[\mathfrak{F}(C_i^t)] = 0. \tag{1.6}$$

The pressure p then reduces to the mean pressure

$$p = -\frac{1}{3} \text{tr } S, \tag{1.7}$$

and T is the deviatoric part of S . We note that the extra stress T must vanish if the motion reduces to a state of rest, in which case we have $C_i^t(s) = I = \text{const}$. Hence

$$0 = \mathfrak{F}(I). \tag{1.8}$$

Consider a particle which at time t finds itself at the position \mathbf{x} , with general coordinates x^i , and which was at time $t+s$, $s \leq 0$, at the point \mathfrak{X} , with coordinates ξ^i . For the dependence of \mathfrak{X} on \mathbf{x} , t , and $t+s$, we write

$$\mathfrak{X} = \mathfrak{D}_t(\mathbf{x}, t+s). \tag{1.9}$$

The function \mathfrak{D}_t , with coordinates $\vartheta_{(t)}^i$, is the displacement function relative to the configuration at time t .

It follows from (1.3) that the covariant components of $C_i^t(s)$, relative to the natural basis at \mathbf{x} of the coordinate system, are

$$c_{(t)ij}^{(t)}(s) = f_{(t)i}^k(t+s) f_{(t)j}^l(t+s) g_{kl}(t+s), \tag{1.10}$$

where

$$f_{(t)j}^i(t+s) = \partial_j[\vartheta_{(t)}^i(\mathbf{x}, t+s)], \tag{1.11}$$

and $g_{kl}(t+s)$ denotes the covariant components of the unit tensor I at $\mathfrak{X} = \mathfrak{D}_t(\mathbf{x}, t+s)$. If the velocity $\mathbf{v}(\mathfrak{X}, t+s)$ is given as a function of the position \mathfrak{X} and the time $t+s$, the displacement function

$$\mathfrak{X}(s) = \mathfrak{D}_t(\mathbf{x}, t+s) \tag{1.12}$$

is that solution of the differential equation

$$\frac{d\mathfrak{X}}{ds} = \mathbf{v}(\mathfrak{X}, t+s) \tag{1.13}$$

which satisfies the initial condition

$$\mathfrak{X}(0) = \mathfrak{D}_t(\mathbf{x}, t) = \mathbf{x}. \tag{1.14}$$

If the motion is a steady flow, the velocity $\mathbf{v}(\boldsymbol{\xi})$ depends on only the position $\boldsymbol{\xi}$ and not on the time $t+s$. In this case $\mathfrak{D}_t(\mathbf{x}, t+s)$, and hence also $F_t(t+s)$ and $C_t^i(s)$, are independent of the reference time t . It follows from (1.4) that the extra-stress T must be independent of time.

We assume that the body forces \mathbf{g} (per unit mass) have a single-valued potential ψ :

$$\mathbf{g} = -\text{grad } \psi. \tag{1.15}$$

We introduce the *modified pressure*:

$$\varphi = \dot{p} + \rho \psi. \tag{1.16}$$

It coincides with the ordinary pressure p if there are no body forces. The dynamical equations take the form

$$\text{div } T - \text{grad } \varphi = \rho \dot{\mathbf{v}}. \tag{1.17}$$

2. A class of steady flows

We consider here steady flows for which the velocity field $\mathbf{v}(\mathbf{x})$ has the contravariant components

$$\{v^1, v^2, v^3\} = \{0, v(x^1), 0\} \tag{2.1}$$

in some appropriate *orthogonal* coordinate system x^1, x^2, x^3 . The matrix $\|g_{ij}\|$ of the covariant components of the unit tensor I is then diagonal. We further assume that the g_{ij} are independent of x^2 , so that $\|g_{ij}\|$ has the form

$$\|g_{ij}(\mathbf{x})\| = \left\| \begin{array}{ccc} [g_1(x^1, x^3)]^2 & 0 & 0 \\ 0 & [g_2(x^1, x^3)]^2 & 0 \\ 0 & 0 & [g_3(x^1, x^3)]^2 \end{array} \right\|, \tag{2.2}$$

where g_1, g_2 , and g_3 are three positive functions of x^1 and x^3 . The flows defined in this manner are isochoric, because the equation

$$\text{div } \mathbf{v} = 0, \quad \text{i.e.,} \quad (\sqrt{g})^{-1} \partial_i (\sqrt{g} v^i) = 0, \tag{2.3}$$

where $\sqrt{g} = g_1 g_2 g_3$, is obviously satisfied.

In the special case of the velocity field (2.1), integration of the differential equation (1.13) with the initial condition (1.14) yields the simple result

$$\xi^1 = x^1, \quad \xi^2 = x^2 + s v(x^1), \quad \xi^3 = x^3. \tag{2.4}$$

From (2.4), (1.12) and (1.11) we obtain

$$\|f_{(t)i}^i(t+s)\| = \left\| \begin{array}{ccc} 1 & 0 & 0 \\ s v'(x^1) & 1 & 0 \\ 0 & 0 & 1 \end{array} \right\|, \tag{2.5}$$

where v' denotes the derivative of v . Equations (2.4) and (2.2) imply that

$$g_i(\xi^1, \xi^3) = g_i(x^1, x^3), \quad i = 1, 2, 3; \tag{2.6}$$

i. e., that the covariant components of the unit tensor remain constant along the paths of the particles. Using equations (1.10), (2.5), (2.2), and (2.6) we obtain

the following expression for the covariant components of the tensor $C_i^t(s)$:

$$\|c_{ij}^{(t)}(s)\| = \left\| \begin{array}{ccc} [g_1(x^1, x^3)]^2 + [s g_2(x^1, x^3) v'(x^1)]^2 & s [g_2(x^1, x^3)]^2 v'(x^1) & 0 \\ s [g_2(x^1, x^3)]^2 v'(x^1) & [g_2(x^1, x^3)]^2 & 0 \\ 0 & 0 & [g_3(x^1, x^3)]^2 \end{array} \right\|. \quad (2.7)$$

The matrix $(C_i^t(s))$ of the physical components of $C_i^t(s)$ is

$$(C_i^t(s)) = \left\| \frac{c_{ij}^{(t)}(s)}{g_i(x^1, x^3) g_j(x^1, x^3)} \right\| = \left\| \begin{array}{ccc} 1 + \kappa^2 s^2 & \kappa s & 0 \\ \kappa s & 1 & 0 \\ 0 & 0 & 1 \end{array} \right\|, \quad (2.8)$$

where

$$\kappa = \frac{g_2(x^1, x^3)}{g_1(x^1, x^3)} v'(x^1). \quad (2.9)$$

In the remainder of this paper we shall always use the physical components of second order tensors.

3. Reduction of the general constitutive equation

The results of the previous section may be summarized in this way: If the velocity field has the form (2.1) in an orthogonal coordinate system obeying (2.2), the function C_i^t has the form

$$C_i^t(s) = I + s A + s^2 B, \quad (3.1)$$

where the matrices (A) , (B) of the physical components of A and B are given by

$$(A) = \kappa \left\| \begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right\|, \quad (B) = \kappa^2 \left\| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right\|. \quad (3.2)$$

When (3.1) holds, the function C_i^t is completely determined by A and B , and the functional \mathfrak{F} of (1.4) reduces to a function \mathfrak{h} of the two tensors A and B , so that

$$T = \mathfrak{h}(A, B). \quad (3.3) \star$$

It follows from (1.2) that the tensor function \mathfrak{h} must be isotropic; *i.e.*,

$$\mathfrak{h}(Q A Q^T, Q B Q^T) = Q \mathfrak{h}(A, B) Q^T \quad (3.4)$$

must hold for all orthogonal tensors Q . It is a consequence of this fact that the component functions \mathfrak{h}_{ij} of \mathfrak{h} are the same for all *orthonormal* bases. Since physical components are components relative to orthonormal bases, the functions \mathfrak{h}_{ij} do not depend on the coordinate system or the point under consideration, if physical components are used.

The normalization (1.6) of the functional \mathfrak{F} implies that the function \mathfrak{h} is normalized by

$$\text{tr } T = \text{tr} [\mathfrak{h}(A, B)] = 0. \quad (3.5)$$

* We note that, if (3.1) holds, A and $2B$ coincide with the first two Rivlin-Ericksen tensors A_1 and A_2 as defined in [I], (8.9). Equation (3.3) has the form of a constitutive equation of the differential type (*cf.* [I] (23.14)).

Furthermore, it follows from (1.8) that

$$0 = \mathfrak{h}(0, 0). \quad (3.6)$$

We now consider the orthogonal tensor Q whose physical components are given by the matrix

$$(Q) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix}. \quad (3.7)$$

Then, by (3.2), we have

$$Q A Q^T = A, \quad Q B Q^T = B. \quad (3.8)$$

Hence, from (3.4) we get

$$Q T Q^T = T \quad (3.9)$$

for $T = \mathfrak{h}(A, B)$. Denoting the *physical components* of T by t_{ij} , we find that the matrix of the physical components of $Q T Q^T$ is given by

$$(Q T Q^T) = \begin{vmatrix} t_{11} & t_{12} & -t_{13} \\ t_{21} & t_{22} & -t_{23} \\ -t_{31} & -t_{32} & t_{33} \end{vmatrix}. \quad (3.10)$$

According to (3.9), this matrix must coincide with the matrix $(T) = \|t_{ij}\|$, which is possible only if $t_{13} = t_{31} = t_{23} = t_{32} = 0$. Hence (T) must have the form

$$(T) = \begin{vmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ 0 & 0 & t_{33} \end{vmatrix}. \quad (3.11)$$

By (3.5), we have

$$t_{11} + t_{22} + t_{33} = 0. \quad (3.12)$$

From the fact that the component functions \mathfrak{h}_{ij} are independent of the coordinate system, and from (3.2) and (3.3), it follows that the physical components t_{ij} of T are functions of κ alone. It is clear from (3.11) and (3.12) that these components t_{ij} may be expressed in terms of three independent functions of κ . We choose

$$t_{12} = \tau(\kappa), \quad t_{11} - t_{33} = \sigma_1(\kappa), \quad t_{22} - t_{33} = \sigma_2(\kappa). \quad (3.13)$$

For the physical components $s_{ij} = t_{ij} - p \delta_{ij}$ of the stress tensor we also have

$$s_{12} = s_{21} = \tau(\kappa), \quad (3.14)$$

$$s_{11} - s_{33} = \sigma_1(\kappa), \quad s_{22} - s_{33} = \sigma_2(\kappa). \quad (3.15)$$

The three functions τ , σ_1 , and σ_2 depend only on the material, not on the particular flow; they are *material functions*.

The functions τ , σ_1 , and σ_2 are not completely arbitrary, as may be seen from the following argument: For a second choice of Q in (3.4) we take

$$(Q) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}. \quad (3.16)$$

Equation (3.2) then yields

$$QAQ^T = -A, \quad QBQ^T = B, \tag{3.17}$$

and, from (3.14), we get

$$(QTQ^T) = \begin{vmatrix} t_{11} & -t_{12} & 0 \\ -t_{21} & t_{22} & 0 \\ 0 & 0 & t_{33} \end{vmatrix} \tag{3.18}$$

It is clear from (3.2) that the conjugation operation (3.17) is equivalent to replacing κ by $-\kappa$. Thus, by (3.13), equation (3.18) states that

$$\tau(-\kappa) = -\tau(\kappa), \quad \sigma_i(\kappa) = \sigma_i(-\kappa); \tag{3.19}$$

i.e., that τ must be an *odd* function, while σ_1 and σ_2 must be *even* functions. No further restrictions are imposed on τ , σ_1 , and σ_2 by the isotropy condition (3.4).

The condition (3.6) implies that the material functions τ , σ_1 , and σ_2 must vanish for $\kappa = 0$:

$$\tau(0) = \sigma_1(0) = \sigma_2(0) = 0. \tag{3.20}$$

A further restrictive condition on τ follows from the principle that the dissipation of energy must be positive. This condition is

$$\kappa \tau(\kappa) > 0, \quad \kappa \neq 0; \tag{3.21}$$

i.e. $\tau(\kappa)$ must have the same sign as κ . If it is assumed that τ is twice continuously differentiable, then (3.20) and (3.21) imply that $\tau(\kappa)$ must be a strictly increasing function of κ in some interval $-\kappa_0 \leq \kappa \leq +\kappa_0$ around $\kappa = 0$. In this interval τ will have a strictly increasing and odd inverse, τ^{-1} . In the remainder of this paper, whenever τ^{-1} occurs, we assume that we are in the range in which τ^{-1} is defined. If τ^{-1} is not defined over the entire real axis then the flows which we discuss below will not occur at high speeds.

We can give a precise meaning to the "shear-dependent viscosity" used in the rheological literature by identifying it with the material function η :

$$\eta(\kappa) = \frac{\tau(\kappa)}{\kappa}. \tag{3.22}$$

Clearly η must be an *even* function. Since $\tau(0) = 0$ and τ is twice differentiable, η is differentiable, even at $\kappa = 0$, and

$$\eta'(0) = 0. \tag{3.23}$$

If we assume that σ_1 and σ_2 are differentiable, we also have

$$\sigma'_1(0) = \sigma'_2(0) = 0. \tag{3.24}$$

If a polynomial approximation for η , σ_1 , and σ_2 , in a neighborhood of $\kappa = 0$, is used to fit experimental data, only even powers of κ can occur.

4. Rectilinear flow

If the velocity field has the form (2.1) in a Cartesian coordinate system

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \tag{4.1}$$

then we call the motion *rectilinear flow*. In this case

$$g_1 = g_2 = g_3 = 1, \quad (4.2)$$

and hence, by (2.9),

$$\kappa = v'(x). \quad (4.3)$$

By (3.13), the extra stresses t_{ij} are then functions of x only, and the dynamical equations (1.17) become

$$\partial_x t_{xx} - \partial_x \varphi = 0, \quad \partial_x t_{xy} - \partial_y \varphi = 0, \quad \partial_z \varphi = 0. \quad (4.4)$$

A simple analysis shows that these equations are satisfied only if

$$t_{xy}(x) = -ax + b, \quad (4.5)$$

$$\varphi = \varphi(x, y) = -ay + t_{xx}(x) - c, \quad (4.6)$$

where a , b , and c are constants.

There are two important special cases of rectilinear flow which we shall discuss in the next two sections.

5. Simple shearing flow

This is a rectilinear flow between an infinite plate I at rest and an infinite plate II moving with a constant speed V parallel to plate I. We denote the distance between the plates by d . We choose the y and z axes to be in the plane of plate I and such that plate II moves in the y direction. The x axis is then perpendicular to the plates with $x=0$ at plate I. Assuming that the fluid adheres to the walls, we have the boundary conditions

$$v(0) = 0, \quad v(d) = V. \quad (5.1)$$

We assume now that the modified pressure φ , defined by (1.16), has no gradient in direction of the flow; *i.e.*, that $\partial_y \varphi = 0$. We then have $a=0$ in (4.6). Hence, by (3.13), (4.3) and (4.5), the shearing stress

$$t_{xy}(x) = \tau(v'(x)) = b \quad (5.2)$$

is constant. It follows that $v'(x) = \tau^{-1}(b)$ must be constant and hence that $v(x)$ is linear. The conditions (5.1) then imply that

$$v(x) = \frac{V}{d} x, \quad (5.3)$$

and (4.3) becomes

$$\kappa = \frac{V}{d} = \text{const.} \quad (5.4)$$

According to (3.14), the shearing stress is

$$s_{xy} = \tau\left(\frac{V}{d}\right). \quad (5.5)$$

Since s_{xy} is equal to the tangential force per unit area that must be applied to the moving plate II in order to produce the flow, equation (5.5) can be used to determine the material function τ experimentally.

It follows from (5.4) and (3.13) that the t_{ij} must actually be constant. Hence, by (4.6), since $a = 0$,

$$\varphi = t_{xx} - c = \text{const.} \tag{5.6}$$

Using (1.16) and $s_{xx} = t_{xx} - p$, we find

$$s_{xx} = c - \varrho \psi, \tag{5.7}$$

and (3.15) becomes

$$s_{xx} - s_{zz} = \sigma_1 \left(\frac{V}{d}\right), \quad s_{yy} - s_{zz} = \sigma_2 \left(\frac{V}{d}\right). \tag{5.8}$$

Hence, the matrix of the stress tensor is

$$(S) = \left\| \begin{array}{ccc} c - \varrho \psi & \tau \left(\frac{V}{d}\right) & 0 \\ \tau \left(\frac{V}{d}\right) & c - \varrho \psi + \sigma_2 \left(\frac{V}{d}\right) - \sigma_1 \left(\frac{V}{d}\right) & 0 \\ 0 & 0 & c - \varrho \psi - \sigma_1 \left(\frac{V}{d}\right) \end{array} \right\|. \tag{5.9}$$

Since the normal stresses can be measured, at least in principle, equations (5.8) can be used to determine the material functions σ_1 and σ_2 .

6. Flow through a channel

This is a rectilinear flow between two parallel infinite plates which are both at rest. We now place the y and z axes in a plane half-way between the plates; the x axis is then perpendicular to the plates. Let the distance between the plates be d ; it follows that the plates are in the planes $x = \pm \frac{1}{2} d$.

The fluid adheres to the plates. For rectilinear flow, to get any motion at all we must have $a \neq 0$ in (4.6). When a is positive, there is a flow in the positive y -direction. From (3.13), (4.3) and (4.5) we get $v'(x) = \frac{-1}{\tau}(-ax + b)$. The boundary conditions are $v(\pm \frac{1}{2} d) = 0$. The fact that $\frac{-1}{\tau}$ is an odd function implies $b = 0$; *i.e.*,

$$v'(x) = -\frac{1}{\tau}(ax). \tag{6.1}$$

After integration we obtain the velocity profile

$$v(x) = -\int_{-\frac{1}{2}d}^x \frac{-1}{\tau}(a\xi) d\xi. \tag{6.2}$$

The total applied force f in the direction of the flow, exerted on a column of fluid with cross-section \mathcal{A} and lying between two planes $y = y_I$ and $y = y_{II}$, is given by

$$f = \left(\int_{\mathcal{A}} s_{yy} dA \right)_{y=y_{II}} - \left(\int_{\mathcal{A}} s_{yy} dA \right)_{y=y_I} - \int_{y_I}^{y_{II}} \int_{\mathcal{A}} (\partial_y \psi) \varrho dA dy. \tag{6.3}$$

This reduces to

$$f = \int_{\mathcal{A}} \{ (s_{yy} - \varrho \psi)_{y=y_{II}} - (s_{yy} - \varrho \psi)_{y=y_I} \} dA. \tag{6.4}$$

By (1.16) and (4.6) we have

$$s_{yy} - \varrho \psi = t_{yy} - p - \varrho \psi = t_{yy} - \varphi = ay + t_{yy}(x) - t_{xx}(x) + c,$$

hence

$$s_{yy} - \rho\psi = ay + h(x). \quad (6.5)$$

Substitution into (6.4) yields

$$f = a(y_{II} - y_I)A,$$

where A is the area of the cross-section of the column. We thus have

$$a = \frac{f}{A(y_{II} - y_I)}, \quad (6.6)$$

which gives a physical meaning to the parameter a . It is the *applied force per unit volume in the direction of the flow*. It is the driving force that produces the flow. If there are no body forces, a reduces to what is usually called the *pressure head* per unit length. It must be produced by a pump or similar device. If the only body force is gravity, acting in the direction of flow, and if there is no pressure head, then a is simply the *specific weight* of the fluid. In any case, a is a measurable quantity. We call it the *driving force*.

The volume discharge per unit time through a cross-section of the channel one unit deep is

$$Q = \int_{-\frac{1}{2}d}^{\frac{1}{2}d} v(x) dx. \quad (6.7)$$

Integrating by parts, remembering the boundary conditions $v(\pm\frac{1}{2}d) = 0$, and using (6.1), we find

$$Q = - \int_{-\frac{1}{2}d}^{\frac{1}{2}d} x v'(x) dx = \int_{-\frac{1}{2}d}^{\frac{1}{2}d} x \tau^{-1}(ax) dx. \quad (6.8)$$

We finally obtain

$$Q = \frac{2}{a^2} \int_0^{\frac{1}{2}ad} \xi \tau^{-1}(\xi) d\xi. \quad (6.9)$$

Thus, the experimentally measurable rate of discharge Q is related to the measurable driving force a through the material function τ which can be determined by independent measurements.

Equation (6.9) is easily solved for the function τ^{-1} in terms of the function $Q = Q(a)$:

$$\tau^{-1}\left(\frac{ad}{2}\right) = \frac{2}{ad^2} \frac{d}{da} [a^2 Q(a)]. \quad (6.10)$$

This formula may be used to calculate the material function τ when Q has been measured as a function of a .

The normal stresses are obtained from (4.6), (3.15), (4.3) and (6.1):

$$\begin{aligned} s_{xx} &= ay + \rho\psi - c, \\ s_{xx} - s_{zz} &= \sigma_1(\tau^{-1}(ax)), \\ s_{yy} - s_{zz} &= \sigma_2(\tau^{-1}(ax)). \end{aligned} \quad (6.11)$$

If the material functions τ , σ_1 , and σ_2 are known from independent measurements, these formulae may be used to predict the measurable normal stresses. It will be noticed that the normal stresses are determined only up to a constant hydrostatic pressure of magnitude c .

If only τ is known, then (6.11) may be used to determine σ_1 and σ_2 by measuring $s_{xx} - s_{zz}$ and $s_{yy} - s_{zz}$. If the potential ψ of the body forces is known, then a measurement of s_{xx} just determines the value of the constant c .

7. Poiseuille flow

We consider a steady flow through an infinite circular pipe of radius R and use cylindrical coordinates with the z -axis coincident with the axis of the pipe. We label the coordinates as follows:

$$r = x^1, \quad z = x^2, \quad \vartheta = x^3. \tag{7.1}$$

We say that we have Poiseuille flow if in this coordinate system the velocity field has the form (2.1) and if the fluid adheres to the walls of the pipe; *i.e.*

$$v(R) = 0. \tag{7.2}$$

The coordinate system r, z, ϑ is orthogonal and obeys (2.2) with

$$g_1 = g_2 = 1, \quad g_3 = r. \tag{7.3}$$

Therefore, we find that (2.9) becomes

$$\kappa = v'(r). \tag{7.4}$$

By (3.13), the extra stresses t_{ij} are then functions of r only, and the dynamical equations (1.17) become

$$\begin{aligned} \partial_r t_{rr} + r^{-1}(t_{rr} - t_{\vartheta\vartheta}) - \partial_r \varphi &= 0, \\ \partial_r t_{rz} + r^{-1}t_{rz} - \partial_z \varphi &= 0, \\ -\partial_\vartheta \varphi &= 0. \end{aligned} \tag{7.5}$$

An elementary analysis shows that these equations are satisfied only if

$$t_{rz} = -\frac{1}{2}ar + br^{-1}, \tag{7.6}$$

$$\varphi = \varphi(r, z) = -az + h(r), \tag{7.7}$$

$$h'(r) = \partial_r t_{rr} + r^{-1}(t_{rr} - t_{\vartheta\vartheta}), \tag{7.8}$$

where a and b are constants and h is an arbitrary function of r . Since the shearing stress $s_{rz} = t_{rz}$ must be continuous at the center of the pipe (*i.e.* for $r=0$), we must put $b=0$ in (7.6). According to (3.14) and (7.4), we thus have

$$s_{rz} = -\frac{1}{2}ar = \tau(v'(r));$$

hence, since τ is an odd function,

$$v'(r) = -\tau^{-1}\left(\frac{1}{2}ar\right). \tag{7.9}$$

This equation, together with the boundary condition (7.2), determines the velocity profile:

$$v(r) = \int_r^R \tau^{-1}\left(\frac{1}{2}a\xi\right) d\xi. \tag{7.10}$$

An investigation analogous to that given in § 6 for the case of a flow through a channel shows that the parameter a is again the applied force per unit volume in the direction of the flow. We again call a the *driving force*. It reduces to the pressure head per unit length if there are no body forces. If the flow takes place in a vertical pipe under the action of gravity alone, a is just the specific weight of the fluid.

The volume discharge per unit time through a cross-section of the pipe is

$$Q = 2\pi \int_0^R v(r) r dr. \quad (7.11)$$

After integration by parts and use of (7.2) and (7.9), we find*

$$Q = \frac{8\pi}{a^3} \int_0^{\frac{1}{2}aR} \xi^2 \bar{\tau}^{-1}(\xi) d\xi. \quad (7.12)$$

Thus, if the material function τ is known, one can predict, for pipe flow, the dependence of the rate of discharge Q on the driving force a .

As in the case of channel flow, (7.12) is easily inverted:

$$\bar{\tau}^{-1}(\frac{1}{2}aR) = \frac{1}{\pi a^2 R^3} \frac{d}{da} [a^3 Q(a)]. \quad (7.13)$$

If $Q = Q(a)$ has been measured, this formula may be used to calculate the material function τ .

To calculate the normal stresses, we observe that (3.13), (7.4), (7.9) and the fact that σ_1 is an even function imply

$$t_{rr} - t_{\theta\theta} = \sigma_1(\bar{\tau}^{-1}(\frac{1}{2}ar)). \quad (7.14)$$

Integration of (7.8) yields

$$h(r) = t_{rr} - \int_r^R \xi^{-1} \sigma_1(\bar{\tau}^{-1}(\frac{1}{2}a\xi)) d\xi - c, \quad (7.15)$$

where c is a constant. Substituting (7.15) into (7.7) and using $\varphi = \rho + \varrho\psi$ and $s_{rr} = t_{rr} - \rho$, we obtain

$$s_{rr} = az + \varrho\psi + \int_r^R \xi^{-1} \sigma_1(\bar{\tau}^{-1}(\frac{1}{2}a\xi)) d\xi + c. \quad (7.16)$$

By (3.15), we have

$$s_{rr} - s_{\theta\theta} = \sigma_1(\bar{\tau}^{-1}(\frac{1}{2}ar)), \quad s_{zz} - s_{\theta\theta} = \sigma_2(\bar{\tau}^{-1}(\frac{1}{2}ar)). \quad (7.17)$$

The remarks made in § 6 about the prediction of measurable normal stresses in channel flow also apply here.

8. Couette flow

Here we consider a steady flow between two infinite coaxial cylinders; the inner cylinder has radius R_1 , and the outer cylinder has radius R_2 . Using cylindrical coordinates with the z axis coincident with the axis of the cylinders, we

* Equivalent relations have appeared previously in the literature (*cf.* [5], pp. 249 and 250). The derivation given here, we believe, is the first that is not based on special assumptions.

now label the coordinates

$$r = x^1, \quad \vartheta = x^2, \quad z = x^3. \tag{8.1}^*$$

If, in these coordinates, the velocity field has the form (2.1), we speak about Couette flow. We then have

$$g_1 = g_3 = 1, \quad g_2 = r, \tag{8.2}$$

so that (2.9) becomes

$$\kappa = r v'(r). \tag{8.3}$$

Note that $v(r)$ has now the physical dimension of an angular velocity. By (3.13) and (8.3), the extra stresses t_{ij} depend only on r . The dynamical equations (1.17) become

$$\begin{aligned} \partial_r t_{rr} + r^{-1}(t_{rr} - t_{\vartheta\vartheta}) - \partial_r \varphi &= -\rho r v^2, \\ \partial_r t_{r\vartheta} + 2r^{-1}t_{r\vartheta} - r^{-1}\partial_\vartheta \varphi &= 0, \\ -\partial_z \varphi &= 0. \end{aligned} \tag{8.4}$$

An analysis of these equations, using the fact that φ must be single valued and hence periodic in ϑ , shows that φ can depend only on r and that $t_{r\vartheta} = M/(2\pi r^2)$, where M is a constant. From (3.13) and (8.3), it follows that

$$t_{r\vartheta} = \tau(r v'(r)) = \frac{M}{2\pi r^2}. \tag{8.5}$$

The moment per unit height exerted on the fluid inside the cylindrical surface $r = \text{const.}$ is, by (8.5), equal to

$$r(2\pi r) t_{r\vartheta} = M \tag{8.6}$$

and is independent of r . From (8.5) we get

$$v'(r) = \frac{1}{r} \tau^{-1}\left(\frac{M}{2\pi r^2}\right). \tag{8.7}$$

We consider further the special case in which the inner cylinder remains at rest while the outer cylinder rotates with constant angular velocity Ω . Assuming that the fluid adheres to the walls, we then have

$$v(R_1) = 0, \quad v(R_2) = \Omega. \tag{8.8}$$

M is the moment per unit height that must be applied to the outer cylinder in order to produce the flow.

Integration of (8.7) with the boundary condition (8.8) yields

$$\Omega = \frac{1}{2} \int_{M/(2\pi R_2^2)}^{M/(2\pi R_1^2)} \frac{1}{\xi} \tau^{-1}(\xi) d\xi. \tag{8.9}^{**}$$

Thus, a knowledge of the material function τ makes it possible to predict the angular velocity Ω of the outer cylinder as a function of the moment per unit height M . If the gap between the two cylinders is very small $[(R_2 - R_1)/R_1 \ll 1]$,

* Note that this differs from the labeling (7.1) used for Poiseuille flow.

** A derivation, based on special assumptions, of a formula equivalent to (8.9) is given in [6], § 3.

the formula (8.9) may be approximated by

$$\Omega \approx \frac{R_2 - R_1}{R_1} \tau^{-1} \left(\frac{M}{2\pi R_2^2} \right). \tag{8.10}$$

We wish to invert (8.9) so as to be able to calculate the material function τ from an experimental determination of $\Omega = \Omega(M)$. Differentiation of (8.9) with respect to M gives

$$2M \frac{d\Omega(M)}{dM} = \tau^{-1} \left(\frac{M}{2\pi R_1^2} \right) - \tau^{-1} \left(\frac{M}{2\pi R_2^2} \right). \tag{8.11}$$

With the abbreviations

$$F(M) = 2M \frac{d\Omega(M)}{dM}, \quad \alpha = \frac{R_1^2}{R_2^2}, \tag{8.12}$$

equation (8.11) becomes

$$F(M) = \tau^{-1} \left(\frac{M}{2\pi R_1^2} \right) - \tau^{-1} \left(\alpha \frac{M}{2\pi R_1^2} \right). \tag{8.13}$$

If we replace here M by $\alpha^n M$ and sum from $n = 0$ to $n = N$, we get

$$\sum_{n=0}^N F(\alpha^n M) = \tau^{-1} \left(\frac{M}{2\pi R_1^2} \right) - \tau^{-1} \left(\alpha^{N+1} \frac{M}{2\pi R_1^2} \right). \tag{8.14}$$

Since $R_1 < R_2$, we have $\alpha < 1$ and hence $\lim_{n \rightarrow \infty} \alpha^{n+1} = 0$. The continuity of τ^{-1} at zero and the fact that $\tau^{-1}(0) = 0$ (cf. (3.20)) insure that

$$\lim_{n \rightarrow \infty} \tau^{-1} \left(\alpha^{n+1} \frac{M}{2\pi R_1^2} \right) = 0.$$

Hence, (8.14) yields*

$$\tau^{-1} \left(\frac{M}{2\pi R_1^2} \right) = \sum_{n=0}^{\infty} F(\alpha^n M). \tag{8.15}$$

Equations (8.12) and (8.15) make it possible to calculate τ if $\Omega = \Omega(M)$ is known. When the gap between the cylinders is small, τ can be calculated directly from (8.10).

We now determine the normal stresses. From (3.13), (8.3), and (8.7), we get

$$t_{rr} - t_{\theta\theta} = \sigma_1 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right) - \sigma_2 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right). \tag{8.16}$$

Substitution of (8.16) into the first equation of (8.4) and integration yields, after using $s_{rr} = t_{rr} - p$ and $\varphi = p + \varrho\psi$,

$$s_{rr} = \varrho\psi - \int_{R_1}^r \left\{ \varrho \xi [v(\xi)]^2 + \frac{1}{\xi} \sigma_1 \left(\tau^{-1} \left(\frac{M}{2\pi \xi^2} \right) \right) - \frac{1}{\xi} \sigma_2 \left(\tau^{-1} \left(\frac{M}{2\pi \xi^2} \right) \right) \right\} d\xi + c, \tag{8.17}$$

where $v(\xi)$, by (8.7) and (8.8), is

$$v(\xi) = \int_{R_1}^{\xi} \frac{1}{\zeta} \tau^{-1} \left(\frac{M}{2\pi \zeta^2} \right) d\zeta. \tag{8.18}$$

* Note added in proof: It has been brought to our attention that an equivalent inversion of (8.9) was published by KRIEGER & ELROD [7].

By (3.15), we have

$$s_{rr} - s_{zz} = \sigma_1 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right), \quad s_{\theta\theta} - s_{zz} = \sigma_2 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right), \quad (8.19)$$

The remarks made at the end of § 6 apply here also.

When the potential ψ of the body forces is independent of r , (8.17) yields the following expression for the difference of the normal stresses at the outer and inner cylinders:

$$s_{rr}(R_2) - s_{rr}(R_1) = - \int_{R_1}^{R_2} \left\{ \rho r [v(r)]^2 + \frac{1}{r} \sigma_1 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right) - \frac{1}{r} \sigma_2 \left(\tau^{-1} \left(\frac{M}{2\pi r^2} \right) \right) \right\} dr. \quad (8.20)$$

When $(R_2 - R_1)/R_1 \ll 1$, this formula may be approximated by

$$s_{rr}(R_2) - s_{rr}(R_1) \approx \frac{R_2 - R_1}{R_1} \left[\sigma_2 \left(\tau^{-1} \left(\frac{M}{2\pi R_2^2} \right) \right) - \sigma_1 \left(\tau^{-1} \left(\frac{M}{2\pi R_2^2} \right) \right) \right]. \quad (8.21)$$

The error in both the approximate formulae (8.10) and (8.21) is of the order of $[(R_2 - R_1)/R_1]^2$. The inertial term $\rho r [v(r)]^2$ in the integral on the right side of (8.20) makes a contribution to $s_{rr}(R_2) - s_{rr}(R_1)$ of the order of $[(R_2 - R_1)/R_1]^3$ and hence does not appear in (8.21).

Acknowledgement. The research leading to this paper was partly supported by the National Science Foundation under Grant NSF-G 5250 to Carnegie Institute of Technology.

References

- [1] NOLL, W.: A mathematical theory of the mechanical behavior of continuous media. *Arch. Rational Mech. Anal.* **2**, 197–226 (1958).
- [2] MARKOVITZ, H.: Normal stress effect in polyisobutylene solutions. II. Classification and application of rheological theories. *Trans. Society of Rheology* **1**, 37–52 (1957).
- [3] NOLL, W.: On the continuity of the solid and fluid states. *J. Rational Mech. Anal.* **4**, 3–81 (1955).
- [4] RIVLIN, R. S.: Solution of some problems in the exact theory of viscoelasticity. *J. Rational Mech. Anal.* **5**, 179–188 (1956).
- [5] HERMANS, J. J.: Flow properties of disperse systems, Chapter 5. Amsterdam: North-Holland Publishing Co. 1953.
- [6] MOONEY, M.: Explicit formulas for slip and fluidity. *J. Rheology* **2**, 210–222 (1931).
- [7] KRIEGER, I. M., & H. ELROD: Direct determination of the flow curves of non-Newtonian fluids. II. Shearing rate in the concentric cylinder viscometer. *J. Appl. Phys.* **27**, 134–136 (1953).

Mellon Institute
Pittsburgh, Pennsylvania
and
Carnegie Institute of Technology
Pittsburgh, Pennsylvania

(Received March 13, 1959)

On the Thermostatistics of Continuous Media

BERNARD D. COLEMAN & WALTER NOLL

Contents

1. Mechanical preliminaries	101
2. Thermomechanic states	102
3. The caloric equation of state	103
4. The isotropy group	103
5. Forces, stresses, and work	104
6. Definition of thermal equilibrium	107
7. Conditions for thermal equilibrium	107
8. The fundamental postulates	109
9. An alternative axiomatization	110
10. Infinitesimal deformations from an arbitrary state	112
11. Simple fluids	113
12. Isotropic materials	115
13. The free energy	117
14. Thermal stability	119
15. Mechanical stability	121
16. Gibbs' thermostatistics of fluids	125
References	128

Introduction

In this article we regard thermostatistics as being that branch of thermodynamics which deals with bodies which are at rest at the present time and which, for all practical purposes, may be regarded as having been at rest at all times in the past.

We attempt to develop here a rigorous theory of thermostatistics for continuous bodies in arbitrary states of strain. The thermodynamics of chemical reactions, phase transitions, and capillarity is not discussed. Our aim is to derive some of the fundamental laws of hydrostatics and elastostatics from thermodynamic principles. Among these laws are the existence of elastic potentials for stress-strain relations, the known inequalities of hydrostatics, and some new inequalities for hydrostatics and elastostatics.

In his classic work, "On the Equilibrium of Heterogeneous Substances", J. W. GIBBS [1] laid down criteria for determining whether a given (global) state of a body is thermodynamically stable. He used these criteria to derive particular equations and inequalities which represent conditions (in some cases necessary

and other cases sufficient) for various special states to be stable. The equations GIBBS obtained as necessary conditions for thermodynamic stability are now recognized as fundamental laws in physical chemistry. GIBBS also derived inequalities which, apparently because they are in obvious accord with everyday experience and thus might be mistakenly called trivial, have attracted relatively little attention and are sometimes not even mentioned in modern thermodynamics courses. For example, in his treatment of homogeneous systems at rest under uniform hydrostatic pressure, GIBBS showed that a necessary condition for such a system to be in a stable state is that both its heat capacity at constant volume and its adiabatic modulus of compression be non-negative. It is inequalities of this type which are emphasized in the present paper. We take, however, a point of view different from that of GIBBS.

In the classical treatments of thermostatics (e.g., [1], [2]) the adjective *stable* is used in two senses. It is sometimes used as a modifier for the word *equilibrium*; i.e. one refers to "states of stable equilibrium"; or it is used as a modifier for the word *state*; i.e. one refers to "stable states". In this paper we never use the word *stable* in the former sense. The theory which we develop here makes a careful distinction between *local* states, referring to a material point in a body, and *global* states, referring to the body as a whole. A local thermomechanic state is specified by giving the entropy density and the local configuration at a material point. A global thermomechanic state, on the other hand, is specified only when the entropy field and the complete configuration are specified for the entire body. We regard *thermal equilibrium* to be a property of *local states*. We consider just one type of thermal equilibrium. We define a state of thermal equilibrium as a local thermomechanic state which minimizes an appropriate potential rather than as a state at which a first variation vanishes. We regard *stability* as a property of only *global states*. We consider several types of stable states, defined as global thermomechanic states which minimize certain energy integrals subject to different constraints.

Our theory is based on two physical postulates. The first asserts that, at a material point, any local thermomechanic state can be an equilibrium state provided the local temperature and local forces have appropriate values. The second postulate is essentially the assumption that, at least in continuum mechanics, absolute temperatures are never negative. We believe that these physical postulates, which are stated in terms of our definition of equilibrium, contain the physical content for the statics of continuous media of the First and Second Laws of Thermodynamics. From our postulates we prove relationships between the stress-strain equation and the caloric equation of state, and we derive various inequalities restricting the form of the caloric equation of state. We should like to propose that the inequalities which we obtain for the finite theory of elasticity answer some of the questions raised by C. TRUESDELL [3] in his recent article, "Das ungelöste Hauptproblem der endlichen Elastizitätstheorie".

Although our definition of thermal equilibrium is new, some of the definitions of the stability of global states which we propose for study are similar to stability definitions considered by GIBBS [1] and J. HADAMARD [4]. In particular, our concepts of isothermal and adiabatic stability at *fixed boundary* are closely related

to, but not identical to, *Hadamard stability**. We briefly discuss GIBBS' theory of the stability of fluid phases in § 16. In a future article we hope to give a discussion of GIBBS' theory of the stability of fluid mixtures.

We regard the main tasks of the science of thermostatics to be, first, the exploration of the consequences for the caloric equation of state of the existence of local states of thermal equilibrium and, second, the derivation of useful necessary and sufficient criteria for global states to be stable. In the present paper, § 6—§ 13 are devoted to the first task and § 14—§ 16 deal briefly with the second. From our present point of view, we should say that the great classical thermodynamicists, GIBBS and DUHEM, devoted their main efforts to the second task.

It will be noticed that in this paper we never mention such notions as "reversible processes" and "quasi-static processes"; in fact, our theory of thermostatics, being a truly statical theory, has no need of "processes" at all.

In writing the present paper we have striven for a level of mathematical rigor comparable to that of works in pure mathematics rather than to that customary in physics.

Notation and basic mathematical concepts. We often find it convenient to distinguish between functions and their values. The basic local thermodynamic variables are denoted by light face Greek minuscules: $\varepsilon, \psi, \eta, \vartheta, \dots$. Symbols such as $\hat{\varepsilon}, \bar{\varepsilon}, \bar{\bar{\varepsilon}} \dots$ and $\hat{\psi}, \check{\psi}, \bar{\psi} \dots$ represent real valued *functions* whose values are the thermodynamic variables ε and ψ .

We denote vectors and points of the three-dimensional Euclidean space \mathcal{E} by bold face Latin minuscules: $\mathbf{v}, \mathbf{x}, \mathbf{y} \dots$.

Second order tensors are denoted by light face Latin majuscules: F, U, Q, R, I . However, we reserve the symbols X and Z to represent material points of a physical body. The term tensor is used as a synonym for linear transformation. Tensors of order higher than two do not occur in this paper. For the trace of the tensor F we write $\text{tr } F$ and for the determinant of F we write $\det F$. We say that F is invertible if F has an inverse F^{-1} ; *i.e.* if $\det F \neq 0$. The transpose of F is denoted by F^T . The identity transformation is written I . For the composition, or product, of two linear transformations A and B we write simply AB .

* Hadamard stability requires (roughly) that the first variation of the integral of the elastic potential vanish, and that the second variation be non-negative, for all smooth variations in the state of strain which are compatible with a fixed boundary. This sort of stability is necessary but not sufficient for stability at fixed boundary as we define it here. In the theory of the propagation of waves in a perfectly elastic solid, Hadamard stability of a particular rest state implies the reality of all roots of the wave velocity equation for acceleration waves of arbitrary direction which might impinge on an object in that state. J. L. ERICKSEN & R. A. TOUPIN [5] have recently considered a modification of Hadamard stability in which they require that the second variation of the integral of the elastic potential be strictly positive. They use their definition of stability to prove uniqueness theorems in the theory of small deformations superimposed on large. R. HILL [15] also has recently discussed relationships between uniqueness and stability. In the third article of his "Recherches sur l'élasticité" P. DUHEM [6] formulated several definitions of stability which are applicable to bodies with fixed and partially free surfaces; he also derived several necessary conditions on the equation of state for particular states of strain to be stable.

Let $\mathbf{h}(\mathbf{x})$ be a function for which both the range and the domain consist of either vectors or points in Euclidean space \mathcal{E} . Assume that for \mathbf{x} in some region the derivative

$$\left. \frac{d}{ds} \mathbf{h}(\mathbf{x} + s \mathbf{v}) \right|_{s=0} = \nabla \mathbf{h}(\mathbf{x}; \mathbf{v}) \quad (1)$$

exists for all \mathbf{v} and is continuous in \mathbf{x} . It is the content of a fundamental theorem of analysis that $\nabla \mathbf{h}(\mathbf{x}; \mathbf{v})$ is then a linear function of \mathbf{v} , and hence we can write

$$\nabla \mathbf{h}(\mathbf{x}; \mathbf{v}) = [\nabla \mathbf{h}(\mathbf{x})] \mathbf{v}, \quad (2)$$

where $\nabla \mathbf{h}(\mathbf{x})$ is a linear transformation (tensor), called the *gradient* of \mathbf{h} at \mathbf{x} .

Similarly, the gradient of a real valued function $\zeta(F)$ of a tensor variable F is a tensor valued function $\zeta_F(F)$ defined by the relation

$$\left. \frac{d}{ds} \zeta(F + sA) \right|_{s=0} = \text{tr} [\zeta_F(F) A], \quad (3)$$

where A is an arbitrary tensor. If Cartesian coordinates are used, and if $\|f_{ij}\|$ is the matrix of F , then the matrix of ζ_F is given by

$$\|\zeta_F(F)\| = \left\| \frac{\partial \zeta}{\partial f_{ji}} \right\|,$$

where i is the row and j the column index.

We make frequent use of the following theorem, called the *polar decomposition theorem*: *Any invertible tensor F has unique decompositions*

$$F = RU = VR \quad (4)$$

where R is orthogonal (i.e., $RR^T = I$) and U, V are positive definite symmetric tensors (i.e., $U = U^T, V = V^T$, and the proper numbers of U and V are all real and greater than zero). In addition, we have

$$U = R^T V R, \quad U^2 = F^T F, \quad V^2 = F F^T. \quad (5)$$

Consider a smooth (i.e., continuously differentiable) real valued function $\zeta(\mathbf{w})$ whose domain \mathcal{W} is a region in a finite dimensional vector space. The function ζ is called *strictly convex* if either of the following two equivalent conditions are satisfied:

(a) For all \mathbf{w}_1 and $\mathbf{w}_2 \neq \mathbf{w}_1$ in \mathcal{W} and all positive α, β with $\alpha + \beta = 1$, the inequality

$$\zeta(\alpha \mathbf{w}_1 + \beta \mathbf{w}_2) < \alpha \zeta(\mathbf{w}_1) + \beta \zeta(\mathbf{w}_2) \quad (6)$$

holds.

(b) For all \mathbf{w} and $\mathbf{w}^* \neq \mathbf{w}$ in \mathcal{W} the inequality

$$\zeta(\mathbf{w}^*) - \zeta(\mathbf{w}) - (\mathbf{w}^* - \mathbf{w}) \cdot \nabla \zeta(\mathbf{w}) > 0 \quad (7)$$

is satisfied.

When \mathcal{W} is a region in the space of all tensors, we use the notation of (3) and the convexity inequality (7) becomes

$$\zeta(F^*) - \zeta(F) - \text{tr} [(F^* - F) \zeta_F(F)] > 0. \quad (8)$$

For a twice continuously differentiable function $\zeta(\mathbf{w})$ to be strictly convex in \mathcal{W} , it is sufficient that the second gradient $\nabla\nabla\zeta(\mathbf{w})$ be positive definite for \mathbf{w} in \mathcal{W} . This condition is not necessary, however: if $\zeta(\mathbf{w})$ is convex, it follows only that $\nabla\nabla\zeta(\mathbf{w})$ is positive *semidefinite*.

1. Mechanical preliminaries

We give a brief summary of those concepts from the mechanics of continuous media that are relevant to the present investigation. For a detailed discussion we refer to [7] and [8].

A *body* \mathcal{B} is a smooth manifold of elements X, Z, \dots , called *material points**. A *configuration* f of \mathcal{B} is a smooth one-to-one mapping of \mathcal{B} onto a region in a three-dimensional Euclidean point space \mathcal{E} . The point $\mathbf{x} = f(X)$ is the *position* of the material point X in the configuration f . The *mass distribution* m of \mathcal{B} is a measure defined on all Borel subsets of \mathcal{B} . For the total mass of \mathcal{B} we write $m(\mathcal{B})$. To each configuration f of \mathcal{B} corresponds a mass density ρ .

Consider a neighborhood $\mathcal{N}(X)$ of a material point in a body; *i.e.*, a part of the body containing X in its interior. Let \mathbf{g} be a smooth homeomorphism of $\mathcal{N}(X)$ into the three-dimensional vector space \mathcal{V} such that X itself is mapped into the zero vector $\mathbf{0}$. The inverse mapping of \mathbf{g} is denoted by $\bar{\mathbf{g}}^1$. Let \mathbf{g}_1 and \mathbf{g}_2 be two such homeomorphisms. The composition $\mathbf{g}_2 \circ \bar{\mathbf{g}}_1^1$ of \mathbf{g}_2 and $\bar{\mathbf{g}}_1^1$ is defined by

$$(\mathbf{g}_2 \circ \bar{\mathbf{g}}_1^1)(\mathbf{x}) = \mathbf{g}_2(\bar{\mathbf{g}}_1^1(\mathbf{x})).$$

It is a mapping of a neighborhood of $\mathbf{0}$ onto another neighborhood of $\mathbf{0}$. We define an equivalence relation " \sim " among all these homeomorphisms by the condition that $\mathbf{g}_1 \sim \mathbf{g}_2$ if the gradient of the mapping $\mathbf{g}_2 \circ \bar{\mathbf{g}}_1^1$ at $\mathbf{0}$ is the identity I . The resulting equivalence classes will be called the *local configurations*** M of X . If M_1 is the equivalence class of \mathbf{g}_1 and M_2 the equivalence class of \mathbf{g}_2 then the gradient at $\mathbf{0}$ of $\mathbf{g}_2 \circ \bar{\mathbf{g}}_1^1$, *i.e.*

$$G = \nabla(\mathbf{g}_2 \circ \bar{\mathbf{g}}_1^1)(\mathbf{0}), \quad (1.1)$$

depends only on M_1 and M_2 . We write

$$G = M_2 M_1^{-1}, \quad M_2 = G M_1, \quad (1.2)$$

and call G the *deformation gradient* from M_1 to M_2 ; G is an invertible linear transformation.

It is often convenient to employ a *local reference configuration* M_r and to characterize the other local configurations

$$M = F M_r, \quad (1.3)$$

by their deformation gradients F from the local reference configuration M_r . If, in this way, two local configurations M_1 and M_2 correspond, respectively, to F_1 and F_2 then the deformation gradient G from M_1 to M_2 is given by

$$G = F_2 F_1^{-1}, \quad F_2 = G F_1. \quad (1.4)$$

* The term "particle" is often used. We prefer "material point" to avoid confusion with molecules and other physical particles.

** The term "configuration gradient" was used in [7].

The *rotation tensor* R , the *right stretch tensor* U , and the *left stretch tensor* V of a deformation gradient F are defined by the unique polar decompositions

$$F = RU = VR, \quad (1.5)$$

where R is orthogonal, while U and $V = RUR^T$ are symmetric and positive definite. We note that U and V have the same proper numbers; these proper numbers are called the *principal stretches* v_1, v_2, v_3 . A deformation gradient G is called a *pure stretch* if its rotation tensor reduces to the identity I ; *i.e.*, if G is symmetric and positive definite and hence coincides with its own right and left stretch tensors.

The mass densities at X corresponding to the local configurations M_1 and M_2 are denoted, respectively, by ϱ_1 and ϱ_2 . We have

$$\varrho_2 = \frac{1}{|\det G|} \varrho_1, \quad (1.6)$$

where G is related to M_1 and M_2 by (1.2).

2. Thermomechanic states

A global thermomechanic state, or simply a *state*, of a body \mathcal{B} is a pair $\{\mathbf{f}, \eta\}$ consisting of a configuration \mathbf{f} of \mathcal{B} and a scalar field η defined on \mathcal{B} ; η is called the *entropy distribution* of the state.

A local thermomechanic state, or simply a *local state*, of a material point X is defined as a pair (M, η) consisting of a local configuration M of X and a real number η , called the *entropy density* (per unit mass) of the local state**.

In the following we often use a local reference configuration M_* and, according to (1.3), characterize the other local configurations M by the deformation gradients F from M_* . We then use the pair (F, η) to characterize the local states.

Two local states (F, η) and (F', η') will be called *equivalent* if they differ only by a change of frame of reference. The local configuration transforms under a change of frame according to the law $F' = QF$ where Q is orthogonal. We assume that the entropy density η is *objective*; *i.e.*, it remains invariant under a change of frame. Thus, the local states (F, η) and (F', η') are equivalent if and only if

$$F' = QF, \quad \eta' = \eta \quad (2.1)$$

for some orthogonal Q .

We say that two global states $\{\mathbf{f}, \eta\}$ and $\{\mathbf{f}', \eta'\}$ are equivalent if they differ by only a change of frame. This is the case if and only if

$$\eta'(X) = \eta(X), \quad F'(X) = QF(X) \quad (2.2)$$

for all X in the body and some orthogonal tensor Q independent of X . Here, $F(X)$ and $F'(X)$ are the deformation gradients at X corresponding to \mathbf{f} and \mathbf{f}' respectively.

* The term "strain tensor" was used in [7].

** In this article, pairs in braces, $\{\}$, always refer to *global* properties; the elements of such pairs are fields over \mathcal{B} . On the other hand, pairs in brackets, $(\)$, always refer to *local* properties and have elements which are either real numbers or tensors. Note that the symbol η in $\{\mathbf{f}, \eta\}$ and (M, η) denotes different entities; in the first case η denotes a field while in the second case it denotes a number. No confusion should arise, however.

3. The caloric equation of state

A *material* is characterized by a real valued function of local states, whose values ε are called the *energy densities* (per unit mass) of the local states. We pick a fixed local reference configuration M_r and characterize the state (M, η) by the pair (F, η) where $F = M M_r^{-1}$. We write

$$\varepsilon = \hat{\varepsilon}(F, \eta). \quad (3.1)$$

It is assumed here that the function $\hat{\varepsilon}$ has continuous derivatives with respect to F and η^* .

We assume that the energy density is objective; *i.e.* invariant under a change of frame. It follows from (2.1) that the function $\hat{\varepsilon}$ must satisfy the relation

$$\hat{\varepsilon}(QF, \eta) = \hat{\varepsilon}(F, \eta) \quad (3.2)$$

for all orthogonal Q . Using the polar decomposition (1.5) and putting $Q = R^T$ in (3.2) we see that

$$\varepsilon = \hat{\varepsilon}(F, \eta) = \hat{\varepsilon}(U, \eta); \quad (3.3)$$

i.e., that the energy density is determined by the right stretch tensor U and the entropy η .

The function $\hat{\varepsilon}$ in (3.3) depends on the choice of the local reference configuration M_r . The function $\hat{\varepsilon}'$ corresponding to some other local reference configuration M'_r is related to $\hat{\varepsilon}$ by

$$\hat{\varepsilon}'(F, \eta) = \hat{\varepsilon}(FG, \eta), \quad (3.4)$$

where $G = M'_r M_r^{-1}$ is the deformation gradient from M_r to M'_r .

The equation (3.3) characterizes the thermal and mechanical properties of a material in statics. It is called the *caloric equation of state* of the material.

4. The isotropy group

It may happen that the energy function $\hat{\varepsilon}$ remains the same function if the local reference configuration M_r is changed to another local reference configuration $M'_r = H M_r$ with the same density. It follows from (3.4) that $\hat{\varepsilon}$ then satisfies the relation

$$\hat{\varepsilon}(F, \eta) = \hat{\varepsilon}(FH, \eta). \quad (4.1)$$

Since M'_r and M_r have the same density, it is clear from (1.6) that $|\det H| = 1$; *i.e.*, H is a unimodular transformation. The unimodular transformations H for which (4.1) holds form a group, called the *isotropy group* \mathcal{G} of $\hat{\varepsilon}$ or of the material defined by $\hat{\varepsilon}$. This group depends, in general, on the choice of the local reference configuration, but it can be shown that the groups corresponding to two different local configurations are always conjugate and hence isomorphic.

We say that the energy function $\hat{\varepsilon}$ defines a *simple fluid* if its isotropy group \mathcal{G} is the full unimodular group \mathcal{U} . If $\mathcal{G} = \mathcal{U}$ for one reference configuration, then $\mathcal{G} = \mathcal{U}$ for all reference configurations. A material point is called a *fluid material point* if its energy function defines a simple fluid. The caloric equation

* For the application to physical situations it is necessary to limit the domain of $\hat{\varepsilon}$ to a region in the space of local configurations and an interval on the η -axis. We do not supply the mathematical details which may arise in the consideration of limitations of this kind.

of state (3.3) then reduces to the form

$$\hat{\varepsilon} = \hat{\varepsilon}(F, \eta) = \bar{\varepsilon}(v, \eta), \quad (4.2)$$

where

$$v = \frac{1}{\varrho} = |\det F| \frac{1}{\varrho_r} \quad (4.3)$$

is the *specific volume* of the local configuration $M = FM_r$; ϱ and ϱ_r are the mass densities corresponding to M and M_r . The function $\bar{\varepsilon}$ in (4.2) does not depend on the choice of the reference configuration.

We say that a material point is an *isotropic material point* if the isotropy group of its energy function $\hat{\varepsilon}$, relative to some local reference configuration, contains the orthogonal group \mathcal{O} . Those local reference configurations of the material point for which \mathcal{G} contains \mathcal{O} are said to be *undistorted*. A simple fluid is isotropic, and all of its local configurations are undistorted. For any isotropic material, it follows from (3.2) and (4.1) that $\hat{\varepsilon}$ satisfies the relation

$$\hat{\varepsilon}(QUQ^T, \eta) = \hat{\varepsilon}(U, \eta) \quad (4.4)$$

for all symmetric and positive definite U and all orthogonal Q , provided the local reference configuration for $\hat{\varepsilon}$ is undistorted. Taking $Q = R$, so that $V = RUR^T$ is the left stretch tensor, we see that for isotropic material points the caloric equation of state (3.3) may be written in the form

$$\varepsilon = \hat{\varepsilon}(F, \eta) = \hat{\varepsilon}(V, \eta). \quad (4.5)$$

It is a further consequence of (4.4) that for each fixed value of η , ε may be expressed as a symmetric function of the three principal stretches v_1, v_2, v_3 :

$$\varepsilon = \hat{\varepsilon}(F, \eta) = \hat{\varepsilon}(V, \eta) = \bar{\varepsilon}(v_1, v_2, v_3; \eta) = \bar{\varepsilon}(v_j, \eta). \quad (4.6)$$

It may also be expressed as a function of the three principal invariants I_V, II_V, III_V of V and U :

$$\varepsilon = \hat{\varepsilon}(V, \eta) = \bar{\varepsilon}(I_V, II_V, III_V; \eta). \quad (4.7)$$

We say that the energy function $\hat{\varepsilon}$ defines a *simple solid* if its isotropy group \mathcal{G} is contained as a subgroup in the orthogonal group \mathcal{O} . A material point is called a *solid material point* if its energy function $\hat{\varepsilon}$, relative to some local configuration as a reference, defines a simple solid. The local reference configurations with this property are again called the *undistorted* states of the solid material point. For an *isotropic simple solid*, the isotropy group \mathcal{G} is identical to the orthogonal group \mathcal{O} .

Throughout the rest of this paper, whenever we discuss isotropic materials it is to be understood that the local reference configuration for the energy density function is undistorted, unless the reference configuration is explicitly specified.

5. Forces, stresses, and work

A *system of forces* is a system of vector valued measures, one for each part of the body \mathcal{B} under consideration*. One must distinguish between contact and body forces. The contact force acting across an oriented surface element in \mathcal{B} will be denoted by $d\mathbf{c}$.

* For a detailed axiomatic treatment *cf.* [8].

Definition of mechanical equilibrium. *In order that a body \mathcal{B} be in **mechanical equilibrium** under a given system of forces, two conditions must be fulfilled for each part \mathcal{P} of \mathcal{B} : (a) the sum of the forces acting on \mathcal{P} must vanish, and (b) the sum of the moments, about any point, of the forces acting on \mathcal{P} must vanish.*

The condition (a), called the *force condition*, depends only on the body and the force system, not on the configuration of the body. The condition (b), called the *moment condition*, does depend on the configuration of the body; *i.e.*, for a given force system, the moment condition may be satisfied for one configuration but not for another.

The force condition alone implies that, for each configuration, the contact forces $d\mathbf{c}$ arise from a stress-tensor S , so that

$$d\mathbf{c} = S\mathbf{n}dA, \quad (5.1)$$

where \mathbf{n} is the unit normal vector of the oriented surface element and dA its area in the configuration under consideration. For fixed contact forces $d\mathbf{c}$, the stress tensor S will be different for different configurations.

We consider now a neighborhood $\mathcal{N}(X)$ of a material point X and assume that a system of contact forces $d\mathbf{c}$ is given for $\mathcal{N}(X)$. Let \mathbf{f}_r be a fixed reference configuration and \mathbf{f} some other configuration of $\mathcal{N}(X)$. If $d\mathbf{c}$ is such that the force condition is satisfied, then (5.1) is valid for all configurations; we can write for the reference configuration \mathbf{f}_r , in particular,

$$d\mathbf{c} = S_r\mathbf{n}_r dA_r, \quad (5.2)$$

where \mathbf{n}_r is the unit normal of the oriented surface element in the reference configuration \mathbf{f}_r , and dA_r is the area of the surface element in \mathbf{f}_r . We denote the position vector, in the configuration \mathbf{f} , of a typical material point Z in $\mathcal{N}(X)$, relative to the position of X as origin, by \mathbf{p} , and we consider the tensor K defined by

$$K = \frac{1}{v(\mathcal{N}(X))} \int_{\overline{\mathcal{N}(X)}} \mathbf{p} \otimes d\mathbf{c}, \quad (5.3)$$

where $\overline{\mathcal{N}(X)}$ denotes the boundary surface of $\mathcal{N}(X)$ and $v(\mathcal{N}(X))$ the volume of $\mathcal{N}(X)$ in the configuration \mathbf{f} , and where \otimes denotes a tensor product. If the force condition is satisfied, the relation (5.1) is valid, and we have

$$K^T = \frac{1}{v(\mathcal{N}(X))} \int_{\overline{\mathcal{N}(X)}} S\mathbf{n} \otimes \mathbf{p} dA.$$

In the limit as $\mathcal{N}(X)$ shrinks to X , we obtain, after using Green's theorem,

$$S^T = \lim_{\mathcal{N}(X) \rightarrow X} K. \quad (5.4)$$

The same argument, with the configuration \mathbf{f} replaced by the reference configuration \mathbf{f}_r , gives

$$S_r^T = \lim_{\mathcal{N}(X) \rightarrow X} \frac{1}{v_r(\mathcal{N}(X))} \int_{\overline{\mathcal{N}(X)}} \mathbf{p}_r \otimes d\mathbf{c}, \quad (5.5)$$

where $v_r(\mathcal{N}(X))$ is the volume of $\mathcal{N}(X)$ in the reference configuration and \mathbf{p}_r the position vector, in the reference configuration, of a typical material point Z

in $\mathcal{N}(X)$, relative to the position of X as origin. The position vector \mathbf{p} of Z in the configuration \mathbf{f} is related to \mathbf{p}_r by the relation

$$\mathbf{p} = F \mathbf{p}_r + o(|\mathbf{p}_r|) \quad (5.6)$$

where F is the gradient at X of the deformation from \mathbf{f}_r to \mathbf{f} and where

$$\lim_{d \rightarrow 0} \frac{o(d)}{d} = 0.$$

Substitution of (5.6) into (5.3) and use of (5.4) and (5.5) yields

$$S = \frac{\rho}{\rho_r} F S_r, \quad (5.7)$$

where ρ and ρ_r are, respectively, the mass densities at X in the configurations \mathbf{f} and \mathbf{f}_r .

The skew part of K , defined by (5.3), is the moment about X , per unit volume, of the contact forces $d\mathbf{c}$ acting on $\mathcal{N}(X)$ in the configuration \mathbf{f} . If the moment condition is satisfied for the configuration \mathbf{f} , then the total moment (*i.e.* the moment of the contact forces *and* the body forces) about X in \mathbf{f} must vanish. Since the moment per unit volume about X of the body forces on $\mathcal{N}(X)$ goes to zero as $\mathcal{N}(X)$ shrinks to X , it follows from (5.4) that S must be symmetric if the moment condition is satisfied in \mathbf{f} .

We say that a material point X is in *local mechanical equilibrium*, when the body is in a given configuration and under a given force system, if the stress tensor S exists at X and is symmetric.

The local behavior at X of a system of contact forces is completely determined by the tensor S_r defined by (5.2). It is called the *Kirchhoff tensor** of the system. For a given force system, the Kirchhoff tensor depends only on the choice of the reference configuration and remains the same if the actual configuration is changed. From (5.7) we see that the existence of the Kirchhoff tensor S_r and the symmetry of $F S_r$ are necessary and sufficient conditions for local mechanical equilibrium at a material point in the local configuration determined by F .

In order that a body \mathcal{B} in a configuration \mathbf{f} be in mechanical equilibrium, it is not sufficient that all its material points be in local mechanical equilibrium; *i.e.*, that the stress tensor exist and be symmetric at each material point. Global mechanical equilibrium will prevail only if, in addition, Cauchy's law

$$\operatorname{div} S + \rho \mathbf{b} = \mathbf{0} \quad (5.8)$$

is satisfied. In this equation, S , ρ , and the density \mathbf{b} of the body forces are to be regarded as fields with domain $\mathbf{f}(\mathcal{B})$.

We consider now a smooth one-parameter family of configurations $\mathbf{f}(s)$ with deformation gradients $F(s)$ at X . The *work per unit mass* done on $\mathcal{N}(X)$ by the contact forces $d\mathbf{c}$ along the path of configurations $\mathbf{f}(s)$ from $s=s_1$ to $s=s_2$ is defined by

$$w = \frac{1}{m(\mathcal{N}(X))} \int_{s_1}^{s_2} \left[\int_{\mathcal{N}(X)} \frac{d\mathbf{p}}{ds} \cdot d\mathbf{c} \right] ds, \quad (5.9)$$

* Cf. TRUESDELL [9], (26.5).

where $m(\mathcal{N}(X))$ is the mass of $\mathcal{N}(X)$ and $\mathbf{p}(s)$ denotes the position vector, in the configuration $\mathbf{f}(s)$, of a typical material point in $\mathcal{N}(X)$. Assuming that the contact forces $d\mathbf{c}$ are independent of s , we obtain

$$w = \frac{1}{m(\mathcal{N}(X))} \left[\int_{\mathcal{N}(X)} \mathbf{p}(s_2) \cdot d\mathbf{c} - \int_{\mathcal{N}(X)} \mathbf{p}(s_1) \cdot d\mathbf{c} \right]. \quad (5.10)$$

Observing (5.3), (5.4) and (5.7), and taking the limit $\mathcal{N}(X) \rightarrow X$, we get

$$\rho_r w = \text{tr} [F(s_2) S_r] - \text{tr} [F(s_1) S_r]. \quad (5.11)$$

This relation shows that $-\frac{1}{\rho_r} \text{tr}(F S_r)$ has the physical meaning of the potential energy, per unit mass, of the local contact forces.

6. Definition of thermal equilibrium

A *force temperature pair* for a material point X is a pair (S_r, ϑ) consisting of a tensor S_r , to be interpreted as the Kirchhoff tensor of a system of contact forces at X , and a real number ϑ , to be interpreted as the temperature at X .

Let a force temperature pair (S_r, ϑ) be given and consider the function

$$\hat{\lambda}(F, \eta) = \hat{\varepsilon}(F, \eta) - \frac{1}{\rho_r} \text{tr}(F S_r) - \eta \vartheta. \quad (6.1)$$

To help motivate the definition of thermal equilibrium given below, we make the following remarks. According to (5.11) the term $-\frac{1}{\rho_r} \text{tr}(F S_r)$ is the potential energy, per unit mass, of the local contact forces. The term $-\eta \vartheta$ may be interpreted as a thermal potential energy. Thus, the value $\lambda = \hat{\lambda}(F, \eta)$ gives a kind of free energy per unit mass of the local state (F, η) when under the action of the force temperature pair (S_r, ϑ) .

Definition of thermal equilibrium. *The local state (F, η) is called a **state of thermal equilibrium under a given force temperature pair** (S_r, ϑ) if*

- (a) *the stress tensor $S = (\rho/\rho_r) F S_r$ is symmetric,*
- (b) *the inequality*

$$\hat{\lambda}(F^*, \eta^*) > \hat{\lambda}(F, \eta) \quad (6.2)$$

holds for all states $(F^, \eta^*) \neq (F, \eta)$ such that*

$$F^* = G F, \quad (6.3)$$

where G is symmetric and positive definite.

The condition (a) means that F corresponds to a local configuration in local mechanical equilibrium (cf. § 5). The condition (b) means that a change of state increases the free energy λ provided that the configuration of the changed state is related to the original configuration by a pure stretch G (cf. § 1).

7. Conditions for thermal equilibrium

In this section we show that, for a local state (F, η) to be a state of thermal equilibrium under the force temperature pair (S_r, ϑ) , the following three conditions

are necessary and sufficient:

(α) The stress tensor $S = \frac{\rho}{\rho_r} F S_r$, is given by the *stress relation**

$$S = \rho F \hat{\varepsilon}_F(F, \eta). \quad (7.1)$$

(β) The temperature ϑ is given by the *temperature relation*

$$\vartheta = \hat{\varepsilon}_\eta(F, \eta). \quad (7.2)$$

(γ) The inequality

$$\hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F, \eta) - \text{tr}[(F^* - F) \hat{\varepsilon}_F(F, \eta)] - (\eta^* - \eta) \hat{\varepsilon}_\eta(F, \eta) > 0 \quad (7.3)$$

holds if $(F^*, \eta^*) \neq (F, \eta)$ and F^* is related to F by $F^* = GF$, where G is positive definite and symmetric.

We assume first that (F, η) is a state of thermal equilibrium and prove the validity of (α), (β), and (γ). By (6.2) and (6.3), the function $\hat{\lambda}(GF, \eta^*)$ of the symmetric tensor variable G and the scalar variable η^* has a minimum for $G = I$ and $\eta^* = \eta$. By a theorem of calculus, it follows that the derivatives of $\hat{\lambda}(GF, \eta^*)$ with respect to G and η^* must vanish for $G = I$ and $\eta^* = \eta$. If we set the derivative of $\hat{\lambda}(GF, \eta^*)$ with respect to η^* equal to zero at $\eta^* = \eta$, we obtain the temperature relation (7.2). The gradient of $\hat{\lambda}(GF, \eta^*)$ with respect to G may be computed using the formula (3) of the mathematical preliminaries and (6.1); we obtain the equation

$$\text{tr} \left\{ \left[F \hat{\varepsilon}_F(F, \eta) - \frac{1}{\rho_r} F S_r \right] A \right\} = 0, \quad (7.4)$$

which is valid for arbitrary symmetric tensors A . Using (5.7) the equation (7.4) may be rewritten in the form

$$\text{tr} \{ [\rho F \hat{\varepsilon}_F(F, \eta) - S] A \} = 0. \quad (7.5)$$

By the condition (a) of the definition of thermal equilibrium, S is symmetric. It follows from (3.2) and Theorem I of reference [10], p. 42, that $\rho F \hat{\varepsilon}_F(F, \eta)$ is also symmetric. Thus, the tensor $\rho F \hat{\varepsilon}_F(F, \eta) - S$ is symmetric. On the other hand, (7.5) can be valid for arbitrary symmetric A only if $\rho F \hat{\varepsilon}_F(F, \eta) - S$ is skew; whence it follows that $\rho F \hat{\varepsilon}_F(F, \eta) - S$ must vanish, which proves (7.1). The inequality (7.3) is obtained simply by substitution of (7.1) and (7.2) into the inequality (6.2), after $\hat{\lambda}$ is replaced by its definition (6.1).

We assume now that the conditions (α), (β), and (γ) are satisfied. It then follows from (7.1), (3.2) and the theorem of reference [10] mentioned above that the stress tensor S must be symmetric, so that condition (a) of the definition of thermal equilibrium is satisfied. Furthermore, the Kirchhoff tensor is given by

$$S_r = \rho_r \hat{\varepsilon}_F(F, \eta). \quad (7.6)$$

Substitution of (7.6) and (7.2) into the inequality (7.3) gives the inequality (6.2); hence condition (b) of the definition of equilibrium is also satisfied.

* This is the familiar stress-strain relation of finite elasticity theory (cf. [10], (16.4)).

8. The fundamental postulates

We are now able to lay down our two fundamental postulates:

Postulate I. For every local state (F, η) for which $\hat{\varepsilon}(F, \eta)$ is defined there exists a force temperature pair (S_r, ϑ) such that (F, η) is a state of thermal equilibrium under (S_r, ϑ) .

Postulate II. The energy function $\hat{\varepsilon}(F, \eta)$ is strictly increasing in η for each fixed F .

Postulate I and the results of the previous section yield the following theorems:

Theorem 1. The force temperature pair (S_r, ϑ) which makes the local state (F, η) a state of thermal equilibrium is given by

$$S_r = \rho_r \hat{\varepsilon}_F(F, \eta), \tag{8.1}$$

$$\vartheta = \hat{\varepsilon}_\eta(F, \eta). \tag{8.2}$$

Theorem 2. The energy function $\hat{\varepsilon}$ obeys the inequality

$$\hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F, \eta) - \text{tr}[(F^* - F) \hat{\varepsilon}_F(F, \eta)] - (\eta^* - \eta) \hat{\varepsilon}_\eta(F, \eta) > 0 \tag{8.3}$$

for any two local states (F, η) and (F^*, η^*) , in the domain of definition of $\hat{\varepsilon}$, which are related by

$$F^* = GF, \tag{8.4}$$

where G is symmetric and positive definite.

The discussion of the previous section shows that Theorem 2 is equivalent to Postulate I. In fact, if we are given a state (F, η) , we can define a force temperature pair (S_r, ϑ) according to (8.1) and (8.2) and then use the inequality (8.3) to prove that (F, η) is in equilibrium under (S_r, ϑ) .

The inequality (8.3) of Theorem 2 is a restricted convexity condition on the function $\hat{\varepsilon}$. If we take, in particular, $F^* = F$, then (8.3) reduces to

$$\hat{\varepsilon}(F, \eta^*) - \hat{\varepsilon}(F, \eta) - (\eta^* - \eta) \hat{\varepsilon}_\eta(F, \eta) > 0 \tag{8.5}$$

for $\eta^* \neq \eta$. This inequality is the content of the following corollary to Theorem 2:

Theorem 3. For each fixed local configuration, the energy density is given by a strictly convex function of the entropy density.

This theorem is equivalent to the statement that $\hat{\varepsilon}_\eta(F, \eta)$ must be a strictly increasing function of η for each fixed F . It follows that the equation (8.2) can be solved for η in a unique manner:

$$\eta = \tilde{\eta}(F, \vartheta). \tag{8.6}$$

Here, $\tilde{\eta}$ is a strictly increasing function* of ϑ for each F . The fact that (8.6) is obtained by solving (8.2) for η is expressed by the identity

$$\hat{\varepsilon}_\eta[F, \tilde{\eta}(F, \vartheta)] = \vartheta. \tag{8.7}$$

* The specific heat c at fixed strain is given by $c = \vartheta \tilde{\eta}_\vartheta(F, \vartheta)$. Hence, it is a consequence of Theorem 3 that c/ϑ is never negative and, for each F , is strictly positive except possibly for a nowhere dense set of values of ϑ .

If we take $\eta^* = \eta$ in (8.3), we obtain

$$\hat{\varepsilon}(F^*, \eta) - \hat{\varepsilon}(F, \eta) - \text{tr}[(F^* - F) \hat{\varepsilon}_F(F, \eta)] > 0; \quad (8.8)$$

this inequality holds whenever $F^* = GF$, where $G \neq I$ is symmetric and positive definite.

A local state (F, η) is called a *natural state* if the corresponding stress (8.4) vanishes. Keeping the entropy fixed, we may use the local configuration of the natural state as the reference configuration, so that $F = I$ and $\hat{\varepsilon}_F(I, \eta) = \frac{1}{\rho} S = 0$. In this case, the inequality (8.8), by (8.4), reduces to

$$\hat{\varepsilon}(G, \eta) > \hat{\varepsilon}(I, \eta), \quad (8.9)$$

which is valid for arbitrary symmetric and positive definite $G \neq I$. Replacing G by the right stretch tensor U of an arbitrary deformation gradient F and using (3.3), we see that

$$\hat{\varepsilon}(F, \eta) \geq \hat{\varepsilon}(I, \eta); \quad (8.10)$$

this expression becomes an equality only when F is orthogonal; *i.e.*, when (F, η) is equivalent to (I, η) . Hence, the *energy density is smallest in a natural state*. It should be pointed out that this observation, though important for the theory of simple solids, is vacuous for fluids. For, we shall prove in § 11 that the stress on a fluid material point in thermal equilibrium is always a strictly positive pressure; thus, for a fluid there is no natural state.

We note that the restriction (8.4) on the inequality (8.3) of Theorem 2 is essential for application of the present theory to physical situations. This restriction means that the local configurations corresponding to F^* and F must be related by a pure stretch. If, for example, these local configurations were related by a rotation so that $F^* = QF$, with Q an orthogonal transformation, then the left side of (8.8) would reduce to $\text{tr}[(Q - I)F \hat{\varepsilon}_F(F, \eta)]$, since $\hat{\varepsilon}(F^*)$ would equal $\hat{\varepsilon}(F)$ by (3.2). The stress relation (7.4) shows that the left side of (8.8) would then become $\frac{1}{\rho} \text{tr}[(Q - I)S]$. One can show that this expression can be made negative by an appropriate choice of Q if S has at least one negative proper number. Thus, the inequality (8.8), were it to hold for arbitrary pairs F, F^* , would exclude the possibility of thermal equilibrium under compression stresses, which is certainly not in accord with experience*.

9. An alternative axiomatization

In this section we hope to make clear our reasons for assuming Postulate II and to motivate further our definition of equilibrium.

It follows from Postulate II that, for each fixed F , the caloric equation of state has a unique solution for η :

$$\eta = \hat{\eta}(F, \varepsilon), \quad (9.1)$$

and that the function $\hat{\eta}$ is strictly increasing in ε for each F . This one-to-one correspondence between ε and η at each F makes it possible to give an alternative

* It has also been pointed out by HILL [15] that an assumption of unrestricted convexity of $\hat{\varepsilon}$ in the deformation gradient would lead to unacceptable physical behavior.

axiomatization of our present theory of thermostatistics by taking ε and F as independent variables and defining thermal equilibrium in terms of the function $\hat{\eta}$. In such a formulation a local thermomechanic state is characterized by a pair (F, ε) , and thermal equilibrium is defined as follows:

Alternative definition of thermal equilibrium. *The local state (F, ε) is called a state of thermal equilibrium under the force temperature pair (S_r, ϑ) , with $\vartheta \neq 0$, if*

- (a) *the stress tensor $S = (\rho|\rho_r)F S_r$ is symmetric,*
- (b) *the inequality*

$$\hat{\eta}(F, \varepsilon) > \hat{\eta}(F^*, \varepsilon^*) + \frac{1}{\vartheta \rho_r} \text{tr}[(F^* - F) S_r] - \frac{\varepsilon^* - \varepsilon}{\vartheta} \tag{9.2}$$

holds for all states $(F^, \varepsilon^*) \neq (F, \varepsilon)$ such that $F^* = GF$, where G is symmetric and positive definite.*

Theorem 4. *The definition of thermal equilibrium given in § 6 and the alternative definition of thermal equilibrium are equivalent (for $\vartheta \neq 0$) if Postulate II is assumed.*

Proof. In § 7 we showed that, under the original definition of § 6, in order for a state (F, η) to be a state of thermal equilibrium for the force temperature pair (S_r, ϑ) it is necessary that

$$\vartheta = \hat{\varepsilon}_\eta(F, \eta). \tag{9.3}$$

By a very similar argument it can be shown, using the alternative definition of thermal equilibrium, that in order for the state (F, ε) to be a state of thermal equilibrium it is necessary that

$$\frac{1}{\vartheta} = \hat{\eta}_\varepsilon(F, \varepsilon). \tag{9.4}$$

Now, by Postulate II, the functions $\hat{\varepsilon}$ and $\hat{\eta}$ are strictly increasing in η and ε , respectively, for fixed F . Hence, ϑ cannot be negative if (S_r, ϑ) is to be a force temperature pair for some state of thermal equilibrium, regardless of which of the two definitions is used. Since we here assume $\vartheta \neq 0$, we have $\vartheta > 0$, and (9.2) can be multiplied by ϑ and then rearranged to give

$$-\vartheta \hat{\eta}(F^*, \varepsilon^*) - \frac{1}{\rho_r} \text{tr}(F^* S_r) + \varepsilon^* > -\vartheta \hat{\eta}(F, \varepsilon) - \frac{1}{\rho_r} \text{tr}(F S_r) + \varepsilon. \tag{9.5}$$

Noting the relations

$$\begin{aligned} \eta^* &= \hat{\eta}(F^*, \varepsilon^*), & \eta &= \hat{\eta}(F, \varepsilon), \\ \varepsilon^* &= \hat{\varepsilon}(F^*, \eta^*), & \varepsilon &= \hat{\varepsilon}(F, \eta), \end{aligned} \tag{9.6}$$

and (6.1), we see that (9.5) is equivalent to (6.2). The requirement that $(F, \eta) \neq (F^*, \eta^*)$ and the requirements on $G = F^* F^{-1}$ are the same for (6.2) and (9.5). The condition (a) is obviously the same in both definitions; hence the definitions are equivalent, q.e.d.

From a certain point of view the alternative definition of thermal equilibrium given in this section is more fundamental than the original definition of § 6. The alternative definition is more closely related to the physical notion that, since entropy tends to increase, equilibrium states should be, in some sense, states of maximum entropy. The definition of § 6 is closely related to the idea, which is often used in mechanics, that equilibrium states should be, in some

sense, states of minimum potential. It should be emphasized that the two definitions are equivalent only if Postulate II is assumed; *i.e.*, only if states of negative temperature are excluded. Of course, negative temperatures never occur in continuum mechanics, but there are subjects in which they do occur (*cf.* [11], [12]). Statistical mechanical considerations suggest that for systems capable of negative temperatures a practical definition of thermal equilibrium should be based on the idea of maximum entropy.

10. Infinitesimal deformations from an arbitrary state

Here we consider the classical theory of *infinitesimal deformations* from an arbitrary initial configuration. We make no attempt to justify the use of the theory of infinitesimal deformations as an approximation to the theory of finite deformations.

In the theory of infinitesimal deformations one considers cases in which $F^* = GF$ is obtained from F by superimposing an infinitesimal deformation. The *infinitesimal strain tensor* E is defined as the *symmetric part* of $G - I$.

In the special case in which G is positive definite and symmetric (*i.e.*, when F^* is related to F by a pure stretch) we have

$$E = G - I, \quad (10.1)$$

and the excess energy $\hat{\varepsilon}(GF, \eta) - \hat{\varepsilon}(F, \eta)$ is a function of E alone:

$$\sigma(E) = \hat{\varepsilon}(GF, \eta) - \hat{\varepsilon}(F, \eta). \quad (10.2)$$

Equation (10.2) is valid approximately even when G is not symmetric.

In the infinitesimal theory it is assumed (i) that (10.2) is valid exactly for all G and (ii) that the excess energy is exactly given by the sum,

$$\sigma(E) = \sigma_1(E) + \sigma_2(E), \quad (10.3)$$

of a term $\sigma_1(E)$ linear in E and a term $\sigma_2(E)$ quadratic in E .

By taking the gradient of (10.2) with respect to E and then putting $E = 0$, it is easily shown that the linear term $\sigma_1(E)$ must be given by

$$\sigma_1(E) = \text{tr}[EF \hat{\varepsilon}_F(F, \eta)]. \quad (10.4)$$

Hence, using the stress relation (7.1), we have

$$\sigma_1(E) = \frac{1}{\rho} \text{tr}(ES), \quad (10.5)$$

where S is the stress of the original state (F, η) .

Now, the fundamental inequality (8.8) may be written

$$\hat{\varepsilon}(GF, \eta) - \hat{\varepsilon}(F, \eta) - \text{tr}[(G - I)F \hat{\varepsilon}_F(F, \eta)] > 0. \quad (10.6)$$

From (10.1), (10.2) we get

$$\sigma(E) - \text{tr}[EF \hat{\varepsilon}_F(F, \eta)] > 0, \quad (10.7)$$

and it follows from (10.3), (10.4) and (10.5) that

$$\sigma_2(E) = \sigma(E) - \frac{1}{\rho} \text{tr}(ES) > 0. \quad (10.8)$$

This inequality is the content of the following theorem:

Theorem 5. *For an infinitesimal deformation superimposed, at fixed entropy, on an arbitrary state, the excess energy is the sum of a positive definite quadratic form in the infinitesimal strain tensor E of the superimposed strain and a linear term $\frac{1}{\rho} \text{tr}(ES)$, where ρ is the density and S the stress corresponding to the original state.*^e

If the original state is a natural state in which the stress vanishes, the above theorem reduces to the familiar statement that the strain energy is a positive definite quadratic form in the infinitesimal strain tensor. For isotropic materials, this statement is equivalent to following well known inequalities for the Lamé constants:

$$\mu > 0, \quad 3\lambda + 2\mu > 0, \quad (10.9)$$

which state that the shear modulus and the compression modulus must be positive.

11. Simple fluids

For simple fluids we have, by (4.2) and (4.3),

$$\hat{\varepsilon}(F, \eta) = \bar{\varepsilon}(|\det F| v_r, \eta), \quad (11.1)$$

where $v_r = 1/\rho_r$ is the specific volume in the reference configuration. Taking the gradient of (11.1) with respect to F , we obtain

$$\hat{\varepsilon}_F(F, \eta) = \bar{\varepsilon}_v(v, \eta) v F^{-1}, \quad (11.2)$$

where $v = |\det F| v_r$. On substituting (11.2) into the fundamental inequality (8.3) and using (8.4), we obtain

$$\bar{\varepsilon}(v^*, \eta^*) - \bar{\varepsilon}(v, \eta) - v \bar{\varepsilon}_v(v, \eta) \text{tr}(G - I) - (\eta^* - \eta) \bar{\varepsilon}_\eta(v, \eta) > 0, \quad (11.3)$$

which must hold for all positive definite symmetric $G = FF^{*-1}$ whenever either $G \neq I$ or $\eta \neq \eta^*$.

We assume now that $v^* = v$; *i.e.*, that $|\det F^*| = |\det F|$, which means that G is unimodular. We also choose $\eta^* = \eta$. Then (11.3) reduces to

$$-v \bar{\varepsilon}_v(v, \eta) \text{tr}(G - I) > 0, \quad (11.4)$$

which must be valid for all symmetric positive definite unimodular tensors $G \neq I$. Let g_1, g_2, g_3 be the proper numbers of G . We then have

$$g_i > 0, \quad g_1 g_2 g_3 = 1 \quad (11.5)$$

and

$$\text{tr}(G - I) = g_1 + g_2 + g_3 - 3. \quad (11.6)$$

Using the fact that the arithmetic mean is greater than the geometric mean,

$$\frac{g_1 + g_2 + g_3}{3} > \sqrt[3]{g_1 g_2 g_3}, \quad (11.7)$$

we see that (11.5) and (11.6) imply

$$\text{tr}(G - I) > 0. \quad (11.8)$$

Hence, it follows from (11.4) that

$$\bar{\epsilon}_v(v, \eta) < 0 \quad (11.9)$$

for all v and η for which $\bar{\epsilon}$ is defined. Thus, $\bar{\epsilon}(v, \eta)$ must be a strictly decreasing function of v for each fixed η .

Substitution of (11.2) into (7.1) shows that the stress relation reduces to

$$S = -\bar{p}(v, \eta) I, \quad (11.10)$$

where

$$\bar{p}(v, \eta) = -\bar{\epsilon}_v(v, \eta) \quad (11.11)$$

is the hydrostatic pressure. By (11.9) it is positive.

For further exploitation of (11.3) we choose $G = \alpha I$, $\alpha > 0$. Since

$$\alpha^3 = |\det G| = \left| \frac{\det F^*}{\det F} \right| = \frac{v^*}{v},$$

we have

$$G = \left(\sqrt[3]{\frac{v^*}{v}} \right) I. \quad (11.12)$$

Substitution of (11.12) into (11.3) yields the inequality

$$\bar{\epsilon}(v^*, \eta^*) - \bar{\epsilon}(v, \eta) - 3v\bar{\epsilon}_v(v, \eta) \left(\sqrt[3]{\frac{v^*}{v}} - 1 \right) - (\eta^* - \eta)\bar{\epsilon}_\eta(v, \eta) > 0, \quad (11.13)$$

which must be valid for all v, v^*, η, η^* except, of course, when both $v = v^*$ and $\eta = \eta^*$. In order to understand the significance of this inequality we introduce the new variable

$$\nu = \sqrt[3]{v} \quad (11.14)$$

and define the function $\bar{\bar{\epsilon}}$ by

$$\bar{\bar{\epsilon}}(v, \eta) = \bar{\epsilon}(\nu, \eta) = \bar{\epsilon}(\sqrt[3]{v}, \eta). \quad (11.15)$$

A straightforward calculation shows that (11.13) is equivalent to the inequality

$$\bar{\bar{\epsilon}}(v^*, \eta^*) - \bar{\bar{\epsilon}}(\nu, \eta) - (v^* - \nu)\bar{\bar{\epsilon}}_\nu(\nu, \eta) - (\eta^* - \eta)\bar{\bar{\epsilon}}_\eta(\nu, \eta) > 0, \quad (11.16)$$

which states that $\bar{\bar{\epsilon}}(\nu, \eta)$ is strictly convex in ν and η jointly. If $\bar{\bar{\epsilon}}(\nu, \eta)$ happens to possess continuous second derivatives, it follows that the matrix

$$\begin{vmatrix} \bar{\bar{\epsilon}}_{\nu\nu} & \bar{\bar{\epsilon}}_{\nu\eta} \\ \bar{\bar{\epsilon}}_{\eta\nu} & \bar{\bar{\epsilon}}_{\eta\eta} \end{vmatrix} \quad (11.17)$$

must be positive semi-definite.

We summarize in the following theorem:

Theorem 6. *For a simple fluid in thermal equilibrium, the stress S reduces to a hydrostatic pressure $S = -\bar{p}(v, \eta)I$. The pressure $\bar{p}(v, \eta) = -\bar{\epsilon}_v(v, \eta)$ is always positive. The energy density $\bar{\bar{\epsilon}}(\nu, \eta)$ is a strictly convex function of the cube root ν of the specific volume and the entropy η jointly.*

It is not hard to show that, for simple fluids, the positivity of $\bar{p}(v, \eta)$ and the convexity of $\bar{\bar{\epsilon}}(\nu, \eta)$ are not only necessary but also sufficient conditions for the validity of the fundamental inequality (11.3). Hence, these conditions are also sufficient conditions for the validity of Postulate I for simple fluids.

12. Isotropic materials

For isotropic materials in general, if we pick an undistorted state as reference, we have, by (4.5),

$$\hat{\varepsilon}(F, \eta) = \hat{\varepsilon}(VR, \eta) = \hat{\varepsilon}(V, \eta), \tag{12.1}$$

where V is the left stretch tensor, defined by the polar decomposition $F = VR$. On computing the gradient of (12.1) with respect to V , we find

$$\hat{\varepsilon}_F(F, \eta) = R^T \hat{\varepsilon}_V(V, \eta). \tag{12.2}$$

If we substitute (12.2) into (7.1) and again use $F = VR$, we see that the stress relation may be written in the form

$$S = \rho V \hat{\varepsilon}_V(V, \eta) = \rho \hat{\varepsilon}_V(V, \eta) V. \tag{12.3}$$

On substituting (8.4), (12.1) and (12.2) into the fundamental inequality (8.3) and observing that

$$\hat{\varepsilon}(F^*) = \hat{\varepsilon}(GF) = \hat{\varepsilon}(GVR) = \hat{\varepsilon}(GV),$$

we obtain

$$\hat{\varepsilon}(GV, \eta^*) - \hat{\varepsilon}(V, \eta) - \text{tr}[(G - I) V \hat{\varepsilon}_V(V, \eta)] - (\eta^* - \eta) \hat{\varepsilon}_\eta(V, \eta) > 0. \tag{12.4}$$

This inequality must be valid for all η, η^* and all symmetric and positive definite G and V , except, of course, when both $\eta = \eta^*$ and $G = I$.

We consider now the special case when G and V commute; *i.e.*, when

$$V^* = GV \tag{12.5}$$

is symmetric. In this case the tensors V and V^* have an orthonormal basis of proper vectors in common. The matrices of V and V^* , relative to this basis, are

$$\|V\| = \begin{vmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ 0 & 0 & v_3 \end{vmatrix}, \quad \|V^*\| = \begin{vmatrix} v_1^* & 0 & 0 \\ 0 & v_2^* & 0 \\ 0 & 0 & v_3^* \end{vmatrix}, \tag{12.6}$$

where the v_i and the v_i^* are the proper numbers of V and V^* , respectively. The matrix of $\hat{\varepsilon}_V(V, \eta)$ is

$$\|\hat{\varepsilon}_V(V, \eta)\| = \begin{vmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{vmatrix}, \tag{12.7}$$

where

$$\varepsilon_i = \bar{\varepsilon}_i(v_j, \eta) = \frac{\partial}{\partial v_i} \bar{\varepsilon}(v_j, \eta) \tag{12.8}$$

are the partial derivatives of the function (4.6). Substitution of (12.5), (4.6), (12.6), and (12.7) into (12.4) gives the inequality

$$\bar{\varepsilon}(v_j^*, \eta^*) - \bar{\varepsilon}(v_j, \eta) - \sum_{i=1}^3 (v_i^* - v_i) \bar{\varepsilon}_i(v_j, \eta) - (\eta^* - \eta) \bar{\varepsilon}_\eta(v_j, \eta) > 0, \tag{12.9}$$

which is valid except when $\eta^* = \eta$ and $v_i = v_i^*$ for all i . We have thus proved *

* We have shown that for isotropic materials the inequality (12.9) is a *necessary* condition for validity of the fundamental inequality (8.3). At the present time, it is an open matter as to whether (12.9) is *sufficient* for the validity of (8.3) in the isotropic case, or whether further inequalities which are independent of (12.9) can be deduced from Postulate I for isotropic materials.

Theorem 7. *For an isotropic material, the energy density $\bar{e}(v_j, \eta)$ is a strictly convex function of the principal stretches v_j and the entropy density η jointly.*

If \bar{e} happens to be twice continuously differentiable, it follows that the matrix

$$\begin{vmatrix} \bar{e}_{11} & \bar{e}_{12} & \bar{e}_{13} & \bar{e}_{1\eta} \\ \bar{e}_{21} & \bar{e}_{22} & \bar{e}_{23} & \bar{e}_{2\eta} \\ \bar{e}_{31} & \bar{e}_{32} & \bar{e}_{33} & \bar{e}_{3\eta} \\ \bar{e}_{\eta 1} & \bar{e}_{\eta 2} & \bar{e}_{\eta 3} & \bar{e}_{\eta\eta} \end{vmatrix} \quad (12.10)$$

is positive semidefinite. Here the indices 1, 2, 3, and η denote the derivatives of \bar{e} with respect to v_1, v_2, v_3 , and η , respectively.

A corollary of the convexity inequality (12.9) is

Theorem 8. *For an isotropic material, the functions $\bar{e}_i(v_j, \eta)$, defined by (12.8), have the property that $v_i > v_k$ implies $\bar{e}_i(v_j, \eta) > \bar{e}_k(v_j, \eta)$.*

Proof. Without loss of generality, we take $i = 1$ and $k = 2$. We then choose $v_1^* = v_2, v_2^* = v_1, v_3^* = v_3$, and $\eta^* = \eta$. Since $\bar{e}(v_j, \eta)$ is a symmetric function of the principal stretches v_j (cf. § 4), and since the v_j^* differ from the v_j only by their order, we have

$$\bar{e}(v_j^*, \eta) = \bar{e}(v_j, \eta).$$

Hence (12.9) reduces to

$$-(v_2 - v_1) \bar{e}_1(v_j, \eta) - (v_1 - v_2) \bar{e}_2(v_j, \eta) > 0;$$

i.e.,

$$(v_1 - v_2) [\bar{e}_1(v_j, \eta) - \bar{e}_2(v_j, \eta)] > 0.$$

Thus, if $v_1 > v_2$, then $\bar{e}_1(v_j, \eta) > \bar{e}_2(v_j, \eta)$, *q.e.d.*

In an isotropic material, the left stretch tensor V and the stress tensor S have an orthonormal basis e_i of proper vectors in common. The e_i determine the principal axes of stress. It follows from (12.3), (12.6), and (12.7) that the principal stresses are given by

$$s_i = \rho v_i \bar{e}_i(v_j, \eta). \quad (12.11)$$

When measured per unit area in the undistorted reference state, these principal stresses must be replaced by

$$s'_i = \rho_r \bar{e}_i(v_j, \eta). \quad (12.12)$$

Hence Theorem 8 has the following simple physical interpretation:

Theorem 8a. *If, at a given value η , the principal stretch v_i is greater than the principal stretch v_k , then the principal stress, measured per unit area in the undistorted reference state, in the direction of v_i is greater than that in the direction of v_k .*

It should be noted that the statement of this theorem does not necessarily remain valid if the principal stresses are measured per unit area of the deformed state \star , except when these stresses are all positive; *i.e.*, except in a state of pure tension.

* Such a statement was proposed as a postulate by M. BAKER & J. L. ERICKSEN [13]. In our theory, only the modification given by Theorem 8a is valid. Related inequalities have been studied by J. BARTA [14].

13. The free energy

It is often useful to employ the deformation gradient F and the temperature ϑ , rather than F and the entropy η , as the independent variables. This is possible because, by (8.2) and (8.6), there is a one-to-one correspondence between η and ϑ for each fixed F .

The *free energy function* $\tilde{\psi}$ is defined by

$$\tilde{\psi}(F, \vartheta) = \hat{\varepsilon}[F, \tilde{\eta}(F, \vartheta)] - \vartheta \tilde{\eta}(F, \vartheta), \quad (13.1)$$

where the entropy function $\tilde{\eta}$ is defined in (8.6) as the unique solution of the equation (8.2). The values ψ of the free energy function $\tilde{\psi}$ are called *free energy densities* \star .

Differentiation of (13.1) with respect to F , using the chain rule, gives

$$\tilde{\psi}_F(F, \vartheta) = \hat{\varepsilon}_F[F, \tilde{\eta}(F, \vartheta)] + \hat{\varepsilon}_\eta[F, \tilde{\eta}(F, \vartheta)] \tilde{\eta}_F(F, \vartheta) - \vartheta \tilde{\eta}_F(F, \vartheta).$$

It follows from (8.7) that the last two terms cancel, so that

$$\tilde{\psi}_F(F, \vartheta) = \hat{\varepsilon}_F[F, \tilde{\eta}(F, \vartheta)]. \quad (13.2)$$

Differentiation of (13.1) with respect to ϑ gives

$$\tilde{\psi}_\vartheta(F, \vartheta) = -\tilde{\eta}(F, \vartheta). \quad (13.3)$$

From Theorems 1 and 2, (13.2), and (13.3) we get

Theorem 9. *For the force temperature pair (S_r, ϑ) to make the local state (F, η) a state of thermal equilibrium, it is necessary and sufficient that S_r and η be given by*

$$S_r = \varrho_r \tilde{\psi}_F(F, \vartheta), \quad (13.4)$$

$$\eta = -\tilde{\psi}_\vartheta(F, \vartheta). \quad (13.5)$$

On multiplying (13.4) on the left by $(\varrho/\varrho_r)F$ and noting that $S = (\varrho/\varrho_r)FS_r$, we get the following form for the stress relation:

$$S = \varrho F \tilde{\psi}_F(F, \vartheta). \quad (13.6)$$

Assuming that two temperatures ϑ, ϑ^* and two deformation gradients F, F^* are given, we now put

$$\eta = \tilde{\eta}(F, \vartheta) = -\tilde{\psi}_\vartheta(F, \vartheta)$$

$$\eta^* = \tilde{\eta}(F^*, \vartheta^*) = -\tilde{\psi}_\vartheta(F^*, \vartheta^*). \quad (13.7)$$

By substituting (13.4)–(13.7) into the fundamental inequality (8.3) of Theorem 2, we obtain

Theorem 10. *The free energy function $\tilde{\psi}$ obeys the inequality*

$$\tilde{\psi}(F^*, \vartheta^*) - \tilde{\psi}(F, \vartheta) - \text{tr}[(F^* - F)\tilde{\psi}_F(F, \vartheta)] - (\vartheta^* - \vartheta)\tilde{\psi}_\vartheta(F^*, \vartheta^*) > 0 \quad (13.8)$$

for any two pairs (F, ϑ) and $(F^*, \vartheta^*) \neq (F, \vartheta)$ in the domain of definition of $\tilde{\psi}$ which are related by

$$F^* = GF, \quad (13.9)$$

where G is symmetric and positive definite.

\star The term ‘‘Helmholtz free energy per unit mass’’ would also be in accord with common usage.

As Theorem 2, so also Theorem 10 is equivalent to Postulate I.

If in (13.8) we take the special case $F^* = F$ and interchange ϑ and ϑ^* , we obtain

$$\tilde{\psi}(F, \vartheta^*) - \tilde{\psi}(F, \vartheta) - (\vartheta^* - \vartheta) \tilde{\psi}_\vartheta(F, \vartheta) < 0. \quad (13.10)$$

This inequality, which is valid for all F and all $\vartheta^* \neq \vartheta$, states that the free energy function $\tilde{\psi}(F, \vartheta)$ is *strictly concave* in ϑ for each F .

Putting $\vartheta^* = \vartheta$ in (13.8) gives the following restricted convexity of $\tilde{\psi}$ in F :

$$\tilde{\psi}(F^*, \vartheta) - \tilde{\psi}(F, \vartheta) - \text{tr}[(F^* - F) \tilde{\psi}_F(F, \vartheta)] > 0, \quad (13.11)$$

the restriction being the condition (13.9).

The considerations and results of § 11 and § 12 on simple fluids and isotropic materials remain valid if the energy function $\hat{\varepsilon}$ is replaced by the free energy function $\tilde{\psi}$, except that the convexity of $\hat{\varepsilon}(F, \eta)$ in η corresponds to the concavity of $\tilde{\psi}(F, \vartheta)$ in ϑ . We summarize the relevant results.

For a simple fluid, the free energy density reduces to a function of the specific volume v and the temperature ϑ only:

$$\psi = \tilde{\psi}(F, \vartheta) = \bar{\psi}(v, \vartheta). \quad (13.12)$$

The stress reduces to a hydrostatic pressure given by

$$S = -\bar{p}(v, \vartheta) I, \quad \bar{p}(v, \vartheta) = -\bar{\psi}_v(v, \vartheta). \quad (13.13)$$

The pressure is always positive. The function $\bar{\psi}$, giving the free energy as a function of the cube root ν of the specific volume and the temperature,

$$\bar{\psi}(\nu, \vartheta) = \bar{\psi}(\nu^3, \vartheta), \quad (13.14)$$

satisfies the inequality

$$\bar{\psi}(\nu^*, \vartheta^*) - \bar{\psi}(\nu, \vartheta) - (\nu^* - \nu) \bar{\psi}_\nu(\nu, \vartheta) - (\vartheta^* - \vartheta) \bar{\psi}_\vartheta(\nu^*, \vartheta^*) > 0. \quad (13.15)$$

This inequality implies that $\bar{\psi}(\nu, \vartheta)$ is strictly convex in ν for each ϑ and strictly concave in ϑ for each ν .

For isotropic materials in general, the free energy reduces to a function of the temperature ϑ and the three principal stretches v_1, v_2, v_3 , computed relative to an undistorted state:

$$\tilde{\psi}(F, \vartheta) = \bar{\psi}(v_1, v_2, v_3; \vartheta). \quad (13.16)$$

The function $\bar{\psi}$ is symmetric and strictly convex in the variables v_1, v_2, v_3 ; $\bar{\psi}$ is strictly concave in ϑ . Theorems 8 and 8a remain valid if $\bar{\varepsilon}$ is replaced by $\bar{\psi}$; *i.e.*, if the temperature, rather than the entropy is fixed at a given value. The stress relation may be written in the form

$$S = \rho V \tilde{\psi}_V(V, \vartheta), \quad (13.17)$$

where V is the left stretch tensor.

The forms, (13.6), (13.13), and (13.17), of the stress relation are useful in discussing experiments involving equilibrium states for which the temperature is controlled, while the forms, (7.1), (11.10), and (12.3), are appropriate for discussing experiments involving equilibrium states for which the entropy is controlled.

14. Thermal stability

Consider a body \mathcal{B} and a global thermomechanic state $\{\mathbf{f}, \eta\}$ of \mathcal{B} , defined by a configuration \mathbf{f} of \mathcal{B} and an entropy distribution η of \mathcal{B} ; (cf. § 2). Let the caloric equation of state of the material point X of \mathcal{B} be given by

$$\varepsilon(X) = \hat{\varepsilon}[F(X), \eta(X); X]. \quad (14.1)$$

Here $F(X)$ is the deformation gradient at X of the configuration \mathbf{f} relative to some reference configuration \mathbf{f}_r . We do not assume that the body is homogeneous, and hence the function $\hat{\varepsilon}$ may depend explicitly on X as indicated in (14.1). The *total entropy* of \mathcal{B} in the given state is defined by

$$H = \int_{\mathcal{B}} \eta(X) \, dm \quad (14.2)$$

and the *total internal energy* of \mathcal{B} by

$$E = \int_{\mathcal{B}} \hat{\varepsilon}[F(X), \eta(X); X] \, dm. \quad (14.3)$$

In this section we shall deal with situations in which the deformation gradient $F(X)$ is kept fixed at each X while the entropy field $\eta = \eta(X)$ is varied. It will not be necessary to make the dependence of ε on F explicit, and the following abbreviated notation will be convenient:

$$\hat{\varepsilon}[F(X), \eta(X); X] = \varepsilon(X, \eta(X)). \quad (14.4)$$

Definition of thermal stability. Let $\{\mathbf{f}, \eta\}$ be a state of \mathcal{B} and let E and H be, respectively, the total internal energy and total entropy corresponding to the state $\{\mathbf{f}, \eta\}$. We say that $\{\mathbf{f}, \eta\}$ is a **thermally stable** state of \mathcal{B} if every other state $\{\mathbf{f}, \eta^*\}$, with the same configuration as $\{\mathbf{f}, \eta\}$ and the same total entropy as $\{\mathbf{f}, \eta\}$,

$$H^* = \int_{\mathcal{B}} \eta^*(X) \, dm = H = \int_{\mathcal{B}} \eta(X) \, dm, \quad (14.5)$$

has a greater total internal energy than the state $\{\mathbf{f}, \eta\}$; i.e.,

$$E^* = \int_{\mathcal{B}} \varepsilon(X, \eta^*(X)) \, dm > E = \int_{\mathcal{B}} \varepsilon(X, \eta(X)) \, dm. \quad (14.6)$$

We give another condition, equivalent to the one given above, which could also be used to define thermal stability.

Theorem 11. A state $\{\mathbf{f}, \eta\}$ of \mathcal{B} is thermally stable if and only if every other state $\{\mathbf{f}, \eta^*\}$ with the same configuration as $\{\mathbf{f}, \eta\}$ and the same total energy as $\{\mathbf{f}, \eta\}$,

$$E^* = \int_{\mathcal{B}} \varepsilon(X, \eta^*(X)) \, dm = E = \int_{\mathcal{B}} \varepsilon(X, \eta(X)) \, dm, \quad (14.7)$$

has a lower total entropy than the state $\{\mathbf{f}, \eta\}$; i.e.,

$$H = \int_{\mathcal{B}} \eta(X) \, dm > H^* = \int_{\mathcal{B}} \eta^*(X) \, dm. \quad (14.8)$$

Proof. We show that the hypothesis of Theorem 11 is necessary for thermal stability by showing that if there exists a state $\{\mathbf{f}, \eta_1\}$ (with η_1 not identical to η) which obeys the equation (14.7) of Theorem 11 but violates (14.8), then there must exist a state $\{\mathbf{f}, \eta_2\}$ (with η_2 not identical to η) which obeys the equation

(14.5) of the definition of thermal stability but which does not obey (14.6). Let η_1 be the entropy density distribution which obeys (14.7) but not (14.8); we construct η_2 as follows:

$$\eta_2(X) = \eta_1(X) + \frac{H - H_1}{m(\mathcal{B})}, \quad (14.9)$$

where H_1 is the total entropy corresponding to η_1 . The total entropy corresponding to η_2 is

$$H_2 = \int_{\mathcal{B}} \eta_2(X) dm = H. \quad (14.10)$$

Hence, the state $\{f, \eta_2\}$ obeys the equation (14.5) of the definition. We have assumed that η_2 is not identical to η and that $H_1 \geq H$. If $H_1 = H$, then η_2 is the same as η_1 and hence different from η . In this trivial case of $\eta_2 = \eta_1$, it follows from the fact that η_1 obeys (14.7) that

$$\int_{\mathcal{B}} \varepsilon(X, \eta_2(X)) dm = \int_{\mathcal{B}} \varepsilon(X, \eta_1(X)) dm = \int_{\mathcal{B}} \varepsilon(X, \eta(X)) dm. \quad (14.11)$$

If $H_1 > H$, then $\eta_2(X) < \eta_1(X)$ for all X in \mathcal{B} . It then follows from Postulate II of § 8 and the assumption that η_1 obeys (14.7) that

$$\int_{\mathcal{B}} \varepsilon(X, \eta_2(X)) dm < \int_{\mathcal{B}} \varepsilon(X, \eta_1(X)) dm = \int_{\mathcal{B}} \varepsilon(X, \eta(X)) dm. \quad (14.12)$$

It is clear from (14.12) that η_2 is not identical to η . Hence, whenever $H_1 \geq H$, we have, by the construction (14.9), a state $\{f, \eta_2\}$ with η_2 different from η but with $H_2 = H$ and

$$E_2 = \int_{\mathcal{B}} \varepsilon(X, \eta_2(X)) dm \leq \int_{\mathcal{B}} \varepsilon(X, \eta(X)) dm = E. \quad (14.13)$$

Thus, a violation of the hypothesis of Theorem 11 implies the existence of a state different from $\{f, \eta\}$ which obeys (14.5) yet violates (14.6).

The sufficiency of the hypothesis of Theorem 11 is proved analogously by starting with a state which obeys (14.5) of the definition, but not (14.6), and then using Postulate II to construct a state which obeys (14.7) of the theorem, but which violates (14.8).

The main result of the present section is the following theorem:

Theorem 12. *A state $\{f, \eta\}$ of a body is thermally stable if and only if it is of uniform temperature; i.e., if and only if*

$$\vartheta = \varepsilon_{\eta}(X, \eta(X)) \quad (14.14)$$

is a constant, independent of the material point X .

Proof. To show the necessity of $\vartheta = \text{constant}$, we observe that, by (14.6), the function $\eta(X)$ is the solution of the variational problem

$$\int_{\mathcal{B}} \varepsilon(X, \eta^*(X)) dm = \text{Minimum} \quad (14.15)$$

subject to the constraint (14.5). It follows that the first variation of

$$\int_{\mathcal{B}} [\varepsilon(X, \eta^*(X)) - \alpha \eta^*(X)] dm \quad (14.16)$$

must vanish for $\eta^* = \eta$. Here α is a constant Lagrange parameter. We obtain

$$\alpha = \varepsilon_{\eta}(X, \eta(X)) = \vartheta = \text{constant}. \quad (14.17)$$

To prove the sufficiency of $\vartheta = \text{constant}$, we substitute the function values $F(X)$, $\eta(X)$ and $\eta^*(X)$ for F , η and η^* in the convexity inequality (8.5). Using the abbreviation (14.4) and the equation (14.14), we get

$$\varepsilon(X, \eta^*(X)) - \varepsilon(X, \eta(X)) - [\eta^*(X) - \eta(X)] \vartheta \geq 0. \quad (14.18)$$

This inequality must be strict for some X if η^* and η are different continuous functions. If ϑ is a constant and if (14.5) holds, then integration of (14.18) over the body \mathcal{B} gives the inequality (14.6), which proves that $\{\mathbf{f}, \eta\}$ is thermally stable, *q. e. d.*

15. Mechanical stability

Consider a state $\{\mathbf{f}, \eta\}$ of a body \mathcal{B} . According to Postulate I of §8 it is possible to find a temperature field ϑ and a stress field S such that every material point of \mathcal{B} is in *thermal equilibrium* for the force temperature field defined by S and ϑ . In fact, S and ϑ are given by the stress relation (7.1) and the temperature relation (7.2), respectively. If a field of body forces \mathbf{b} is given, then the state $\{\mathbf{f}, \eta\}$ will be a state of *mechanical equilibrium* if Cauchy's condition

$$\text{Div } S + \rho \mathbf{b} = 0 \quad (15.1)$$

holds. If $\{\mathbf{f}, \eta\}$ is such that every material point is in thermal equilibrium, it is always possible to choose \mathbf{b} such that the state $\{\mathbf{f}, \eta\}$ is a state of mechanical equilibrium. We need only to define \mathbf{b} by (15.1). We say that the fields S , ϑ , and \mathbf{b} , given by (7.1), (7.2) and (15.1), make $\{\mathbf{f}, \eta\}$ a state of equilibrium. We call S , ϑ and \mathbf{b} , respectively, the stress, temperature, and body force fields of $\{\mathbf{f}, \eta\}$.

We investigate the possible meaning that can be given to the statement that an equilibrium state $\{\mathbf{f}, \eta\}$ is *stable*. First, we require that it be *thermally stable* which, according to Theorem 12, means that the temperature ϑ must be uniform. In addition, we require that some condition of *mechanical stability* be satisfied. One must distinguish between various types of *isothermal mechanical stability* and *adiabatic mechanical stability*.

In the case of isothermal mechanical stability, one compares the given equilibrium state $\{\mathbf{f}, \eta\}$ with a class of states $\{\mathbf{f}^*, \eta^*\}$ corresponding to the same uniform temperature $\vartheta = \hat{\vartheta}(F, \eta)$ as the given state. Each of these states is characterized by its configuration \mathbf{f}^* alone, because the corresponding entropy distribution is then determined by

$$\eta = \tilde{\eta}(F^*, \vartheta). \quad (15.2)$$

External forces or boundary conditions must be prescribed for each of the comparison configurations \mathbf{f}^* . The configuration \mathbf{f} is called *stable* if the increase in the total free energy would always be greater than the work done on the body by the external forces if the configuration were to be deformed into any of the comparison configurations \mathbf{f}^* . We give more precise definitions in two special cases.

Definition of isothermal stability at fixed boundary (IFB stability). *An equilibrium state $\{\mathbf{f}, \eta\}$ is called IFB stable if $\{\mathbf{f}, \eta\}$ has a uniform temperature ϑ*

and if for every state $\{\mathbf{f}^*, \eta\}$ which satisfies the following conditions:

- (a) \mathbf{f}^* lies in a prescribed neighborhood* of \mathbf{f} ,
- (b) $\mathbf{f}^*(X) = \mathbf{f}(X)$, when X belongs to $\bar{\mathcal{B}}$,
- (c) the temperature corresponding to $\{\mathbf{f}^*, \eta\}$ is equal to ϑ for all X in \mathcal{B} ,

the following inequality holds:

$$\int_{\mathcal{B}} \{\psi(F^*) - \psi(F) - \mathbf{b} \cdot (\mathbf{f}^* - \mathbf{f})\} dm \geq 0. \quad (15.4)$$

Here $\bar{\mathcal{B}}$ is the boundary of \mathcal{B} , and $\psi(F)$ is an abbreviation for

$$\psi(F) = \tilde{\psi}(F(X), \vartheta; X); \quad (15.5)$$

$F^*(X)$ and $F(X)$ are the deformation gradients at X for the configurations \mathbf{f}^* and \mathbf{f} , respectively, both computed relative to the same fixed reference configuration. As in § 14, we do not assume that the body is homogeneous, and hence the function $\tilde{\psi}$ may depend explicitly on X .

We say that $\{\mathbf{f}, \eta\}$ is *strictly IFB stable* if the inequality (15.4) is strict whenever $\{\mathbf{f}^*, \eta^*\}$ obeys (a), (b) and (c) and is such that $\mathbf{f}^* \neq \mathbf{f}$.

Note that the surface tractions do no work if the boundary is fixed and that $-\int_{\mathcal{B}} \mathbf{f}^* \cdot \mathbf{b} dm$ is a potential of the work done by the body forces if these are held at their values $\mathbf{b}(X)$ in the equilibrium state $\{\mathbf{f}, \eta\}$.

The type of stability considered is affected by the prescription of the neighborhood in the requirement (a) of the definition of IFB stability. A global state may be stable with respect to some (small) neighborhood without being stable with respect to other (larger) neighborhoods.

Definition of isothermal stability at fixed surface tractions (IFT stability). An equilibrium state $\{\mathbf{f}, \eta\}$ is called **IFT stable** if $\{\mathbf{f}, \eta\}$ has a uniform temperature ϑ and if for every state $\{\mathbf{f}^*, \eta\}$ which satisfies the following conditions:

- (a) \mathbf{f}^* lies in a prescribed neighborhood of \mathbf{f} ,
- (b) the temperature corresponding to $\{\mathbf{f}^*, \eta\}$ is equal to ϑ for all X in \mathcal{B} ,

the following inequality holds:

$$\Gamma = \int_{\mathcal{B}} \{\psi(F^*) - \psi(F) - \mathbf{b} \cdot (\mathbf{f}^* - \mathbf{f})\} dm - \int_{\bar{\mathcal{B}}} (\mathbf{f}^* - \mathbf{f}) \cdot \mathbf{S} \mathbf{n} dA \geq 0. \quad (15.6)$$

Here $\bar{\mathcal{B}}$ is the boundary surface of the region occupied by \mathcal{B} in the configuration \mathbf{f} ; dA is the element of that surface; and \mathbf{n} is the exterior unit normal.

Note that $-\int_{\bar{\mathcal{B}}} \mathbf{f}^* \cdot \mathbf{S} \mathbf{n} dA$ is a potential of the work done by the surface tractions if they are held at their values in the equilibrium state $\{\mathbf{f}, \eta\}$.

An IFT stable state is always also IFB stable. This follows from the fact that the surface integral in (15.6) gives no contribution if the boundary condition (15.3) holds, so that the inequalities (15.4) and (15.6) become the same in this case.

* A neighborhood of a configuration is defined by the metric

$$\delta(\mathbf{f}, \mathbf{f}^*) = \sup_{X \in \mathcal{B}} \{|\mathbf{f}^*(X) - \mathbf{f}(X)| + |F^{*-1}(X)F(X) - I|\}$$

over the space of all configurations.

If the inequality (15.6) holds for all states which obey items (a) and (b) of the definition of IFT stability and is, furthermore, a strict inequality for all such states for which $F^*(X) \neq F(X)$ for at least one material point X , then we say that $\{\mathbf{f}, \eta\}$ is *strictly IFT stable against deformations and rotations*. For in that case (15.6) can reduce to an equality only if \mathbf{f}^* is related to \mathbf{f} by a simple rigid translation.

To investigate adiabatic mechanical stability, one compares the given equilibrium state $\{\mathbf{f}, \eta\}$ with a class of states which correspond to the same total entropy as $\{\mathbf{f}, \eta\}$. We again consider two special cases.

Definition of adiabatic stability at fixed boundary (AFB stability). *An equilibrium state $\{\mathbf{f}, \eta\}$ is called **AFB stable** if $\{\mathbf{f}, \eta\}$ is thermally stable and if for every state $\{\mathbf{f}^*, \eta^*\}$ which satisfies the following conditions:*

- (a) \mathbf{f}^* lies in a prescribed neighborhood of \mathbf{f} ,
- (b) $\mathbf{f}^*(X) = \mathbf{f}(X)$, when X belongs to $\overline{\mathcal{B}}$,
- (c) $\int_{\mathcal{B}} \eta^*(X) dm = \int_{\mathcal{B}} \eta(X) dm$,

the following inequality holds:

$$\int_{\mathcal{B}} \{ \hat{\varepsilon}[F^*(X), \eta^*(X); X] - \hat{\varepsilon}[F(X), \eta(X); X] - \mathbf{b} \cdot (\mathbf{f}^* - \mathbf{f}) \} dm \geq 0. \quad (15.7)$$

If the inequality in (15.7) is strict for all $\{\mathbf{f}^*, \eta^*\}$ satisfying (a), (b) and (c) and for which $\mathbf{f}^* \neq \mathbf{f}$, then we say that $\{\mathbf{f}, \eta\}$ is *strictly AFB stable*.

Theorem 13. *A thermally stable equilibrium state $\{\mathbf{f}, \eta\}$ is AFB stable if and only if for every state $\{\mathbf{f}^*, \eta^*\}$ which satisfies the following conditions:*

- (a) \mathbf{f}^* lies in a prescribed neighborhood of \mathbf{f} ,
- (b) $\mathbf{f}^*(X) = \mathbf{f}(X)$ when X belongs to $\overline{\mathcal{B}}$,
- (c) $\int_{\mathcal{B}} \{ \hat{\varepsilon}(F^*(X), \eta^*(X); X) - \mathbf{b} \cdot \mathbf{f}^* \} dm = \int_{\mathcal{B}} \{ \hat{\varepsilon}(F(X), \eta(X); X) - \mathbf{b} \cdot \mathbf{f} \} dm$, (15.8)

the following inequality holds:

$$\int_{\mathcal{B}} \eta^*(X) dm \leq \int_{\mathcal{B}} \eta(X) dm. \quad (15.9)$$

Furthermore, $\{\mathbf{f}, \eta\}$ is strictly AFB stable if and only if (15.9) is a strict inequality for every state $\{\mathbf{f}^*, \eta^*\} \neq \{\mathbf{f}, \eta\}$ obeying (a), (b) and (c).

We omit the proof of Theorem 13 because it is analogous to that of Theorem 11. Of course, the validity of Theorem 13 requires the assumption of Postulate II.

Definition of adiabatic stability at fixed surface tractions (AFT stability). *An equilibrium state $\{\mathbf{f}, \eta\}$ is called **AFT stable** if it is thermally stable and if for every state $\{\mathbf{f}^*, \eta^*\}$ which satisfies the following conditions:*

- (a) \mathbf{f}^* is in a prescribed neighborhood of \mathbf{f} ,
- (b) $\int_{\mathcal{B}} \eta^*(X) dm = \int_{\mathcal{B}} \eta(X) dm$,

the following inequality holds:

$$\int_{\mathcal{B}} \{ \hat{\varepsilon}[F^*(X), \eta^*(X); X] - \hat{\varepsilon}[F(X), \eta(X); X] - \mathbf{b} \cdot (\mathbf{f}^* - \mathbf{f}) \} dm - \int_{\mathcal{A}} (\mathbf{f}^* - \mathbf{f}) \cdot \mathbf{S} \mathbf{n} dA \geq 0. \quad (15.10)$$

It will be noticed that a state which is AFT stable is always AFB stable.

If the inequality (15.10) holds for all states which obey (a) and (b) and is a strict inequality for all such states for which $F^*(X) \neq F(X)$ for at least one X , then we say that $\{f, \eta\}$ is *strictly AFT stable against deformations and rotations*.

It is clear that, in analogy to Theorem 13, an alternative, but equivalent, definition of AFT stability can be formulated in which a stable state is defined to be one of maximum entropy among all those states for which (15.10) reduces to an equality.

The definitions of IFB, IFT, AFB and AFT stability given above are applicable only to those physical situations in which the body force field $\mathbf{b} = \mathbf{b}(X)$ is independent of the comparison configuration \mathbf{f}^* . If one is interested in studying cases in which the body force on X depends on X and is also a functional of \mathbf{f}^* , one can modify the definitions of stability by connecting the comparison state \mathbf{f}^* to \mathbf{f} by means of a continuous one-parameter family \mathbf{f}_s , $0 \leq s \leq 1$, $\mathbf{f}_0 = \mathbf{f}$, $\mathbf{f}_1 = \mathbf{f}^*$ and replacing the term

$$-\int_{\mathcal{A}} \mathbf{b} \cdot (\mathbf{f}^* - \mathbf{f}) \, dm$$

in (15.4), (15.6), (15.7), (15.8) and (15.10) by

$$-\int_{\mathcal{A}} \int_0^1 \mathbf{b}(X, \mathbf{f}_s) \cdot \frac{\partial \mathbf{f}_s(X)}{\partial s} \, ds \, dm.$$

If the body force on each material point is derivable from a single-valued potential, then the integral exhibited above is independent of the parametrization, and is simply the difference in the potentials at \mathbf{f} and \mathbf{f}^* .

In the definitions of IFT and AFT stability, we assumed that not only the body forces but also the contact forces at the surface do not depend on the comparison configuration. One can also study, in a way analogous to that outlined above for the body forces, those cases in which the surface tractions depend on the comparison configuration.

Theorem 14. *A state which has isothermal stability of a certain type also has adiabatic stability of the corresponding type.*

Proof. Consider a state $\{f, \eta\}$ which has a uniform temperature ϑ and which has isothermal stability of a particular type. Let \mathbf{f}^* be a configuration which satisfies the boundary conditions, if any, for the appropriate comparison configurations. Define the entropy field η_1 by

$$\eta_1(X) = \tilde{\eta}(F^*(X), \vartheta), \quad (15.11)$$

where F^* is the deformation gradient field corresponding to the configuration \mathbf{f}^* . By (13.1) we have

$$\tilde{\psi}(F^*, \vartheta) - \tilde{\psi}(F, \vartheta) = \hat{\varepsilon}(F^*, \eta_1) - \hat{\varepsilon}(F, \eta) - (\eta_1 - \eta) \vartheta. \quad (15.12)$$

Here F corresponds to \mathbf{f} . Let η^* be any entropy distribution satisfying the condition

$$\int_{\mathcal{A}} \eta^*(X) \, dm = \int_{\mathcal{A}} \eta(X) \, dm, \quad (15.13)$$

which is required for comparison states in adiabatic stability. Define the field β by

$$\beta = \hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F^*, \eta_1) - (\eta^* - \eta_1) \vartheta. \quad (15.14)$$

From (8.5) we get $\beta(X) \geq 0$
 for all X . From (15.12) we have

$$\tilde{\psi}(F^*, \vartheta) - \tilde{\psi}(F, \vartheta) = \hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F, \eta) - \beta - (\eta^* - \eta) \vartheta. \quad (15.15)$$

We integrate (15.15) over \mathcal{B} . According to (15.13) we get no contribution from the term $-(\eta^* - \eta) \vartheta$; hence, since β is non-negative,

$$\int_{\mathcal{B}} [\tilde{\psi}(F^*, \vartheta) - \tilde{\psi}(F, \vartheta)] dm \leq \int_{\mathcal{B}} [\hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F, \eta)] dm. \quad (15.16)$$

Since the work W done by the external forces in going from f to f^* is the same in adiabatic and isothermal stability, it follows from (15.16) that if

$$\int_{\mathcal{B}} [\tilde{\psi}(F^*, \vartheta) - \tilde{\psi}(F, \vartheta)] dm - W \quad (15.17)$$

is non-negative, then

$$\int_{\mathcal{B}} [\hat{\varepsilon}(F^*, \eta^*) - \hat{\varepsilon}(F, \eta)] dm - W \quad (15.18)$$

is non-negative (and strictly positive when (15.17) is strictly positive). Hence, the isothermal stability of $\{f, \eta\}$ implies the corresponding adiabatic stability for $\{f, \eta\}$, *q. e. d.*

Although in writing our proof of Theorem 14 we have used a notation which implies that \mathcal{B} is homogeneous, it is clear that the same argument is valid when \mathcal{B} is not homogeneous.

It appears to us that the converse of Theorem 14 need not be true; *i. e.*, an equilibrium state may have adiabatic stability without being isothermally stable.

16. Gibbs' thermostatistics of fluids

We now consider a type of stability which was proposed by GIBBS* for fluids free from body forces. GIBBS states** that he had in mind a physical situation in which the fluid is "enclosed in a rigid envelop which is non-conducting to heat and impermeable to all the components of the fluid". A body which may be regarded as being in such an envelop is usually called an "isolated system".

Definition of G stability*.** *An equilibrium state $\{f, \eta\}$ of a fluid body \mathcal{B} is called G stable if the following condition is satisfied. Let $\{f^*, \eta^*\}$ be any other state with the same total volume and the same total entropy as $\{f, \eta\}$,*

$$\int_{\mathcal{B}} v^* dm = \int_{\mathcal{B}} v dm, \quad \int_{\mathcal{B}} \eta^* dm = \int_{\mathcal{B}} \eta dm, \quad (16.1)$$

* See the section of [I] which is entitled "Internal stability of homogeneous fluids as indicated by the fundamental equations", (b), pp. 100–115, particularly the subsection entitled "Stability with respect to continuous changes of phase" (b), pp. 105–111.

** [I] (b), p. 100.

*** In this definition we again restrict ourselves to those physical situations in which fluctuations in chemical composition are suppressed. We have in mind situations in which chemical reactions are prohibited and in which the fluid is either homogeneous or does not allow diffusion. For fluids the homogeneous case is the one of practical importance. Situations in which flow is permitted but diffusion is prohibited are rare.

then $\{f, \eta\}$ has a lower total internal energy than $\{f^*, \eta^*\}$,

$$\int_{\mathcal{B}} \bar{\varepsilon}[v^*(X), \eta^*(X); X] dm > \int_{\mathcal{B}} \bar{\varepsilon}[v(X), \eta(X); X] dm, \quad (16.2)$$

unless $v^*(X) = v(X)$ and $\eta^*(X) = \eta(X)$ for all X in \mathcal{B} .

In (16.1) v and v^* denote the specific volume fields for \mathcal{B} corresponding to the configurations f and f^* .

In the following alternative definition f and ε are taken as the independent variables, and the permitted comparison states are such that the total internal energy and total volume of the body are conserved during the variations. This alternative formulation may suggest to the reader why G stability is regarded as being appropriate for discussing the physics of *isolated systems* composed of fluids:

Alternative definition of G stability. *An equilibrium state $\{f, \varepsilon\}$ of a fluid body \mathcal{B} is called G stable if any other state $\{f^*, \varepsilon^*\}$ with the same total volume and the same total internal energy as $\{f, \varepsilon\}$,*

$$\int_{\mathcal{B}} v^* dm = \int_{\mathcal{B}} v dm, \quad \int_{\mathcal{B}} \varepsilon^* dm = \int_{\mathcal{B}} \varepsilon dm, \quad (16.3)$$

has a higher total entropy,

$$\int_{\mathcal{B}} \hat{\eta}[v^*(X), \varepsilon^*(X); X] dm < \int_{\mathcal{B}} \hat{\eta}[v(X), \varepsilon(X); X] dm, \quad (16.4)$$

unless $v^*(X) = v(X)$ and $\varepsilon^*(X) = \varepsilon(X)$ for all X in \mathcal{B} .

The function $\hat{\eta}$ in (16.4) is obtained by solving $\varepsilon = \bar{\varepsilon}(v, \eta; X)$ for η , which is possible in a unique way because $\bar{\varepsilon}$ is strictly increasing in η .

The proof of the equivalence of the two definitions of G stability is analogous to the one given for Theorem 11 of § 14 in the case of thermal stability; one must again use Postulate II of § 8.

The main result of this section is

Theorem 15. *An equilibrium state $\{f, \eta\}$ of a fluid body is G stable if and only if its temperature and pressure are uniform.*

Proof. To prove that the condition is necessary we observe that the functions v, η are solutions of the variational problem

$$\int_{\mathcal{B}} \bar{\varepsilon}(v^*, \eta^*; X) dm = \text{Minimum} \quad (16.5)$$

subject to the constraints (16.1). Therefore, the first variation of

$$\int_{\mathcal{B}} [\bar{\varepsilon}(v^*, \eta^*; X) - \lambda \eta^* - \mu v^*] dm$$

must vanish for $v^* = v$ and $\eta^* = \eta$. Here λ and μ are constant Lagrange parameters. It follows that

$$\bar{\varepsilon}_\eta(v, \eta; X) = \lambda = \text{constant}, \quad \bar{\varepsilon}_v(v, \eta; X) = \mu = \text{constant}. \quad (16.6)$$

Hence, by (8.2) and (11.11), both the temperature, $\vartheta = \bar{\varepsilon}_\eta(v, \eta; X)$, and the pressure, $p = -\bar{\varepsilon}_v(v, \eta; X)$ are uniform over \mathcal{B} .

To prove the sufficiency of the condition of the theorem, we assume that ϑ and p are uniform and that (16.1) holds. From the convexity inequality (11.16),

the inequality (11.9), and the fact that $v = v^3$, $v > 0$, is a convex function of v , one can easily infer that $\bar{\epsilon}(v, \eta)$ must be convex in v and η . Hence, the inequality

$$\bar{\epsilon}(v^*, \eta^*; X) - \bar{\epsilon}(v^*, \eta; X) - (v^* - v) \bar{\epsilon}_v(v, \eta; X) - (\eta^* - \eta) \bar{\epsilon}_\eta(v, \eta; X) \geq 0 \quad (16.7)$$

is valid at all material points X in \mathcal{B} ; (16.7) cannot reduce to an equality for all X unless $v(X) = v^*(X)$ and $\eta(X) = \eta^*(X)$ for all X . Since $p = -\epsilon_v(v, \eta)$ and $\vartheta = \bar{\epsilon}_\eta(v, \eta)$ are independent of X , integration of (16.7) over \mathcal{B} gives

$$\int_{\mathcal{B}} \{ \bar{\epsilon}(v^*, \eta^*; X) - \bar{\epsilon}(v, \eta; X) \} dm + p \int_{\mathcal{B}} (v^* - v) dm - \vartheta \int_{\mathcal{B}} (\eta^* - \eta) dm > 0.$$

The condition (16.1) states that the last two terms vanish and hence that (16.2) holds, q. e. d.

In his discussion of the stability of homogeneous fluids, GIBBS used a definition of stability which is identical to what we have called G stability, except that he did not demand, as we do, that $\{f, \eta\}$ be an equilibrium state*. GIBBS was able to prove that uniform values of $\bar{\epsilon}_\eta(v, \eta)$ and $\bar{\epsilon}_v(v, \eta)$ are necessary for his stability and, furthermore, that the inequality (16.7) is also necessary. He also realized that the constancy of $\bar{\epsilon}_\eta$ and $\bar{\epsilon}_v$ over \mathcal{B} and the validity of (16.7) are sufficient for his stability. If he had gone a step further and postulated that for homogeneous fluids stable states exist for every value of v and η for which $\bar{\epsilon}$ is defined, he would have obtained (16.7) as a property of the function $\bar{\epsilon}$. Such a procedure, however, cannot yield the statements, made in Theorem 6, that $-\bar{\epsilon}_v$ is positive and that $\bar{\epsilon}$ is jointly and strictly convex in v and η .

We conclude with

Theorem 16. *An equilibrium state $\{f, \eta\}$ of a fluid body \mathcal{B} is G stable if and only if both of the following conditions hold:*

- (a) *The temperature corresponding to $\{f, \eta\}$ is uniform.*
- (b) *Any other state $\{f^*, \eta^*\}$ with the same total volume,*

$$\int_{\mathcal{B}} v^* dm = \int_{\mathcal{B}} v dm, \quad (16.8)$$

and the same uniform temperature ϑ has a higher total free energy,

$$\int_{\mathcal{B}} \bar{\psi}(v^*, \vartheta; X) dm > \int_{\mathcal{B}} \bar{\psi}(v, \vartheta; X) dm, \quad (16.9)$$

unless $v^(X) = v(X)$ for all X in \mathcal{B} .*

Proof. The proof that the conditions (a) and (b) are sufficient for the G stability of $\{f, \eta\}$ is completely analogous to the proof of Theorem 14 of § 15.

The necessity of the condition (a) for the G stability of $\{f, \eta\}$ follows from Theorem 15. To prove that (b) is necessary we assume that $\{f, \eta\}$ is stable. We consider another state $\{f^*, \eta^*\}$ which obeys (16.8) and which has the uniform temperature ϑ . Since $v = v^3$ is a convex function of v for $v > 0$, and $\bar{\psi}_v(v, \vartheta) < 0$, the inequality (13.15) implies that

$$\bar{\psi}(v^*, \vartheta; X) - \bar{\psi}(v, \vartheta; X) - (v^* - v) \bar{\psi}_v(v, \vartheta; X) \geq 0; \quad (16.10)$$

* GIBBS does not use either our Postulate I or our definition of (local) thermal equilibrium.

(16.10) cannot reduce to equality for all X unless $v(X) = v^*(X)$ for all X . Now, since we are assuming that $\{f, \eta\}$ is stable, it follows from Theorem 15 and (13.13) that $\bar{\varphi}_v(v, \vartheta; X)$ is independent of X . Thus, by (16.8), if we compute the mass integral of (16.10) over \mathcal{B} , the last term on the left makes no contribution, and we get (16.9). Hence, when $\{f, \eta\}$ is G stable, the condition (b) is valid, q.e.d.

This theorem shows that for G stability of fluids adiabatic and isothermal stability are equivalent.

Acknowledgement. This research was supported in part by the Air Force Office of Scientific Research under Contract AF 49(638)-541 with the Mellon Institute and by the National Science Foundation under Grant NSF-G 5250 to Carnegie Institute of Technology.

References

- [1] (a) GIBBS, J. W.: On the equilibrium of heterogeneous substances. *Trans. Conn. Acad.* **3**, 108—248, 343—524 (1875—1878), or (b) *The Scientific Papers of J. WILLARD GIBBS* **1**, 55—372, particularly 55—62 and 100—115, Longmans, Green, 1906.
- [2] DUHEM, P.: Dissolutions et Mélanges. *Travaux et Mémoires des Facultés de Lille* **3**, No. 2, 1—136 (1893).
- [3] TRUESDELL, C.: Das ungelöste Hauptproblem der endlichen Elastizitätstheorie. *Z. angew. Math. u. Mech.* **36**, 97—103 (1956).
- [4] HADAMARD, J.: *Leçons sur la Propagation des Ondes et les Équations de l'Hydrodynamique*. Paris 1903.
- [5] ERICKSEN, J. L., & R. A. TOUPIN: Implications of Hadamard's condition for elastic stability with respect to uniqueness theorems. *Can. J. Math.* **8**, 432—436 (1956).
- [6] DUHEM, P.: Recherches sur l'élasticité. Troisième partie. La stabilité des milieux élastiques. *Ann. École Norm.* (3) **22**, 143—217 (1905).
- [7] NOLL, W.: A mathematical theory of the mechanical behavior of continuous media. *Arch. Rat. Mech. Anal.* **2**, 197—226 (1958).
- [8] NOLL, W.: The foundations of classical mechanics in the light of recent advances in continuum mechanics. *Proceedings of the Berkeley Symposium on the Axiomatic Method*, 266—281, 1959.
- [9] TRUESDELL, C.: The mechanical foundations of elasticity and fluid dynamics. *J. Rat. Mech. Anal.* **1**, 125—300 (1952); **2**, 593—616 (1953).
- [10] NOLL, W.: On the continuity of the solid and fluid states. *J. Rat. Mech. Anal.* **4**, 3—81 (1955).
- [11] RAMSEY, N. F.: Thermodynamics and statistical mechanics at negative absolute temperatures. *Phys. Rev.* **103**, 20—28 (1956).
- [12] COLEMAN, B. D., & W. NOLL: Conditions for equilibrium at negative absolute temperatures. *Phys. Rev.* **115**, 262—265 (1959).
- [13] BAKER, M., & J. L. ERICKSEN: Inequalities restricting the form of the stress-deformation relations for isotropic elastic solids and Reiner-Rivlin fluids. *J. Washington Acad. Sciences* **44**, 33—35 (1954).
- [14] BARTA, J.: On the non-linear elasticity law. *Acta Tech. Acad. Scient. Hungaricae* **18**, 55—65 (1957).
- [15] HILL, R.: On uniqueness and stability in the theory of finite elastic strain. *J. Mech. Phys. Solids* **5**, 229—241 (1957).

Mellon Institute
Pittsburgh, Pennsylvania
and
Carnegie Institute of Technology
Pittsburgh, Pennsylvania

(Received August 24, 1959)

An Approximation Theorem for Functionals, with Applications in Continuum Mechanics

BERNARD D. COLEMAN & WALTER NOLL

Contents

1. Introduction		355
I. The Approximation Theorem		
2. Influence Functions and Histories		357
3. Retardation		358
4. Memory Functionals		361
5. The Approximation Theorem		363
II. Applications		
6. The Concept of a Simple Fluid		365
7. Approximations of Order n for Simple Fluids		367
8. Simple Materials in General		369
References		370

1. Introduction

We start with a brief discussion of the physical motivation behind the mathematical considerations to be presented in Part I of this paper.

Often in theoretical physics one is concerned with a functional relationship,

$$\sigma = \mathfrak{F} \left(\mathbf{g}(s) \right), \tag{1.1}$$

which states that the present value σ of a physical quantity is determined by the values $\mathbf{g}(s)$ of a second quantity at all times s in the past. In (1.1), $\mathbf{g}(s)$ is the value of the second quantity s time units ago. If the functional \mathfrak{F} is given, the relation (1.1) may be used to predict the present value σ of the first quantity from a knowledge of the “history” \mathbf{g} of the second quantity. The relation (1.1) may be interpreted as expressing the causal nature of a class of physical processes. For definiteness we assume that the values of σ and $\mathbf{g}(s)$ lie in appropriate normed vector spaces, not necessarily finite-dimensional.

In many physical situations, the value σ in (1.1) will be, in some sense, more sensitive to the values of \mathbf{g} for small s than for large s . Intuitively speaking, the “memory” of the system described by (1.1) will “fade away” in time. In order to make this idea precise, we shall introduce a *norm* $\|\mathbf{g}\|$ in the function space of the histories \mathbf{g} for which (1.1) is meaningful. We first choose a number ρ ,

$1 \leq p \leq \infty$, which will be kept fixed, and then define the norm of \mathbf{g} to be

$$\|\mathbf{g}\| = \begin{cases} \sqrt[p]{\int_0^\infty (|\mathbf{g}(s)| h(s))^p ds} & \text{if } 1 \leq p < \infty \\ \sup_{s \geq 0} |\mathbf{g}(s)| h(s) & \text{if } p = \infty, \end{cases} \quad (1.2)$$

where $|\mathbf{g}(s)|$ is the magnitude of $\mathbf{g}(s)$ and where $h(s)$ is a real-valued function which approaches zero rapidly as $s \rightarrow \infty$. Thus, in computing the norm $\|\mathbf{g}\|$ we assign a greater influence to the values of \mathbf{g} for small s (recent past) than for very large s (distant past). We call the function h in (1.2) an *influence function*. The physical idea that the memory of the system is fading corresponds to the mathematical assumption that the functional \mathfrak{F} in (1.1) is *continuous* with respect to convergence in the norm (1.2) of the function space of histories. The influence function h characterizes the rapidity with which the memory is fading.

It is always possible to reduce the relation (1.1) to one which is *normalized* in the sense that the possible histories \mathbf{g} have the value $\mathbf{g}(0) = \mathbf{0}$ at the present instant and that the value of \mathfrak{F} for the zero history $\mathbf{g}(s) \equiv \mathbf{0}$ is zero. Now, physical experience indicates that phenomena which one would expect to be described exactly by a normalized functional relation (1.1) often follow a simpler relation of the form

$$\boldsymbol{\sigma} = \mathbf{I}(\overset{(1)}{\mathbf{g}}), \quad \overset{(1)}{\mathbf{g}} = \frac{d}{ds} \mathbf{g}(s) \Big|_{s=0}, \quad (1.3)$$

where \mathbf{I} is a linear transformation. In particular, it appears that (1.3) accounts well for the observed phenomena in the case of slow processes. This observation leads to the conjecture that, in some mathematically precise sense, the relation (1.3) approximates the normalized relation (1.1) for slow processes. It is the purpose of the present paper to prove an approximation theorem of this kind. Theorem 2 of §5 asserts not only that (1.3) is the first-order approximation to the normalized relation (1.1) for slow processes but shows also the form of the approximations of higher order. The theorem is based on the assumption that the functional \mathfrak{F} of (1.1) is not only continuous but also Fréchet differentiable at the zero history $\mathbf{g}(0) \equiv \mathbf{0}$ in the function space with norm (1.2). A normalized functional \mathfrak{F} satisfying this assumption will be called a *memory functional*. The continuity of \mathfrak{F} for histories \mathbf{g} different from the zero history is not needed to prove the approximation theorem.

In Part II of the paper we apply the approximation theorem to constitutive equations of continuum mechanics. Our main interest there is in the logical status of the theory of Newtonian fluids within the framework of a recently proposed general theory of fluids with memory effects*.

The theory of compressible Newtonian fluids is based on the following constitutive equation for the stress tensor \mathbf{S} :

$$\mathbf{S} = -p \mathbf{I} + 2\eta \mathbf{D} + \lambda(\text{tr } \mathbf{D}) \mathbf{I}; \quad (1.4)$$

here the rate of deformation tensor \mathbf{D} is the symmetric part of the velocity gradient tensor; $p = p(\varrho)$ is the hydrostatic pressure the fluid would be supporting

* Cf. [1], [2] and [3].

if it were at rest at its present density ρ ; η and λ are functions of ρ alone and are called coefficients of viscosity.

Experience shows that for some substances, such as water, the theory of Newtonian fluids accounts remarkably well for experimental measurements over a very wide range of conditions. Other substances, such as molten plastics, definitely do not obey (1.4) exactly, but yet their behavior appears to approximate that of Newtonian fluids in the limit of slow motions.

In §6 we review the mathematical definition of a *simple fluid*. As we have frequently remarked in the past, we believe this definition is capable of covering the behavior of nearly all real fluids including such substances as molten plastics which exhibit “hereditary”, “non-Newtonian” and “viscoelastic” effects. Here we add to the definition of a simple fluid a new requirement: we require that the functionals occurring in the definition be memory functionals in the sense of the definition used in this paper. As we have indicated, this requirement is related to the physical assumption that simple fluids have a fading memory. If this assumption were not satisfied, the term “fluid” would hardly be appropriate.

In §7 we conclude that the theory of Newtonian fluids is indeed the complete first-order approximation to the theory of simple fluids for slow flows. We also indicate what an experimenter should expect to find as second-order corrections to the constitutive equation of a Newtonian fluid. We point out that several special flow problems for incompressible second-order fluids lead to third-order linear partial differential equations.

In §8 we apply our approximation theorem to the theory of the general simple materials defined in [I], Part III.

I. The Approximation Theorem

2. Influence Functions and Histories

An *influence function* h of order $r > 0$ is a real-valued function of a real variable with the following properties:

- (i) $h(s)$ is defined and continuous for $0 \leq s < \infty$.
- (ii) $h(s)$ is positive, $h(s) > 0$.
- (iii) For each $\sigma > 0$, there is a constant M_σ , independent of α , such that

$$\sup_{s \geq \sigma} \frac{h(s/\alpha)}{\alpha^r h(s)} \leq M_\sigma \quad \text{for } 0 < \alpha \leq 1. \tag{2.1}$$

The last condition (iii) means that $h(s)$ must decay to zero at a fast enough rate as $s \rightarrow \infty$. In fact, we have

$$\sup_{s > 0} s^r h(s) = N < \infty, \tag{2.2}$$

which follows from (2.1) by taking $\sigma = 1$ and observing that

$$\left(\frac{1}{\alpha}\right)^r h\left(\frac{1}{\alpha}\right) \leq h(1) \sup_{s \geq 1} \frac{h(s/\alpha)}{\alpha^r h(s)} \leq h(1) M_1.$$

Let a real Banach space \mathcal{S} with norm $|| \cdot ||$ be given. We then define a *history* \mathbf{g} to be a measurable function defined for $0 \leq s < \infty$ with values $\mathbf{g}(s)$ in \mathcal{S} . Two such functions will be regarded as the same if they differ only on a set of measure zero.

For a given influence function h and a given number $p, 1 \leq p \leq \infty$, we define the $\mathcal{L}_{h,p}$ -norm, $\|\mathbf{g}\|_{h,p}$, or simply $\|\mathbf{g}\|$, of a history \mathbf{g} by

$$\|\mathbf{g}\|_{h,p} = \sqrt[p]{\int_0^\infty (|\mathbf{g}(s)| h(s))^p ds} \quad \text{if } 1 \leq p < \infty, \tag{2.3}$$

$$\|\mathbf{g}\|_{h,\infty} = \sup_{s \geq 0} |\mathbf{g}(s)| h(s) \quad \text{if } p = \infty. \tag{2.4}$$

In (2.4) and subsequently $\sup f(s)$ stands for the essential supremum of $f(s)$, *i.e.* for the greatest lower bound of the suprema of all functions which differ from $f(s)$ only on a set of measure zero.

The set of all histories with finite $\mathcal{L}_{h,p}$ -norm forms a Banach space, which we denote by $\mathcal{L}_{h,p}$.

We remark that (2.2) is a necessary but not a sufficient condition for the decay relation (2.1). However, if a function $h(s)$ satisfies (i), (ii), and the limit relation

$$\lim_{s \rightarrow \infty} s^r h(s) = 0 \quad \text{monotonically for large } s, \tag{2.5}$$

then it also satisfies the decay relation (2.1) and hence is an influence function of order r . To prove this fact we consider the expression

$$\frac{h(s/\alpha)}{\alpha^r h(s)} = \frac{(s/\alpha)^r h(s/\alpha)}{s^r h(s)}, \quad 0 < \alpha \leq 1. \tag{2.6}$$

Since $s^r h(s)$ is monotonically decreasing for s larger than some value s_1 , it follows that (2.6) is not greater than 1 for all $s > s_1$. For $0 < \sigma \leq s \leq s_1$ the denominator $s^r h(s)$ of (2.6) has a positive minimum because h satisfies the conditions (i) and (ii). The numerator of (2.6) is bounded by the maximum of $s^r h(s)$, which exists and is finite because of (2.5). It follows that (2.6) is bounded, for $\sigma \leq s < \infty$, by a constant M_σ independent of α , which is the content of (2.1).

The function

$$h(s) = \frac{1}{(s+1)^r}$$

is an influence function of order r . An exponential,

$$h(s) = e^{-\beta s}, \quad \beta > 0,$$

is an influence function of any order.

3. Retardation

The retardation Γ_α with retardation factor $\alpha, 0 < \alpha \leq 1$, is the linear transformation $\mathbf{g} \rightarrow \mathbf{g}_\alpha$ defined, for all histories \mathbf{g} , by

$$(\Gamma_\alpha \mathbf{g})(s) = \mathbf{g}_\alpha(s) = \mathbf{g}(\alpha s). \tag{3.1}$$

We show that Γ_α maps the space $\mathcal{L}_{h,p}$ into itself. For $p \neq \infty$, we have*

$$\begin{aligned} \|\mathbf{g}_\alpha\|^p &= \int_0^\infty (|\mathbf{g}(\alpha s)| h(s))^p ds \\ &= \frac{1}{\alpha} \int_0^\infty (|\mathbf{g}(s)| h(s))^p \left(\frac{h(s/\alpha)}{h(s)}\right)^p ds \leq \|\mathbf{g}\|^p \frac{1}{\alpha} \left(\sup_{s>0} \frac{h(s/\alpha)}{h(s)}\right)^p. \end{aligned} \tag{3.2}$$

* Since no confusion can arise, we omit the indices h, p on the norm.

Since h is continuous and positive,

$$\sup_{\sigma \geq s \geq 0} \frac{h(s/\alpha)}{h(s)} < \infty$$

for any $\sigma > 0$. The decay condition (2.1) implies that

$$\sup_{s \geq \sigma} \frac{h(s/\alpha)}{h(s)} \leq \alpha^r M_\sigma < \infty.$$

Hence,

$$\sup_{s \geq 0} \frac{h(s/\alpha)}{h(s)} = K_\alpha < \infty. \tag{3.3}$$

Combining (3.2) and (3.3) gives

$$\|\mathbf{g}_\alpha\| \leq \alpha^{-\frac{1}{p}} K_\alpha \|\mathbf{g}\|. \tag{3.4}$$

It is easily verified that this inequality remains valid for $p = \infty$ when we put $\frac{1}{\infty} = 1$. It follows from (3.4) that when \mathbf{g} has a finite norm \mathbf{g}_α has a finite norm and is, therefore, in $\mathcal{L}_{h,p}$. Intuitively, retardation replaces a given history by one which is essentially the same, but slower.

If, possibly after a suitable alteration of $\mathbf{g}(s)$ on a set of measure zero, the limits

$$\overset{(0)}{\mathbf{g}} = \lim_{s \rightarrow 0} \mathbf{g}(s), \quad \overset{(h)}{\mathbf{g}} = \lim_{s \rightarrow 0} \frac{h!}{s^k} \left(\mathbf{g}(s) - \sum_{j=0}^{k-1} \frac{s^j}{j!} \overset{(j)}{\mathbf{g}} \right) \tag{3.5}$$

exist for $k = 0, 1, \dots, n$, then we shall say that the history \mathbf{g} has n *generalized derivatives* at $s = 0$. Of course, the existence of ordinary derivatives implies the existence of the corresponding generalized derivatives. Here we shall use the term ‘‘derivative’’ always in the generalized sense of (3.5). Generalized derivatives at $s = 0$ of any order may exist even though $\mathbf{g}(s)$ is not continuous near $s = 0$.

The *Taylor transformation* Π_n is the linear transformation $\mathbf{g} \rightarrow \Pi_n \mathbf{g}$ defined, for all histories \mathbf{g} which are n times differentiable at $s = 0$, by

$$(\Pi_n \mathbf{g})(s) = \sum_{j=0}^n \frac{s^j}{j!} \overset{(j)}{\mathbf{g}}, \tag{3.6}$$

where the $\overset{(j)}{\mathbf{g}}$ are the derivatives (3.5).

This Taylor transformation replaces \mathbf{g} by its n^{th} ‘‘Taylor approximation’’ $\Pi_n \mathbf{g}$. The history $\Pi_n \mathbf{g}$ is a polynomial of degree $\leq n$.

We note that retardation Γ_α preserves the differentiability of a history and that Π_n and Γ_α commute; *i.e.*, for all histories \mathbf{g} having n derivatives at $s = 0$, we have

$$\Pi_n \Gamma_\alpha \mathbf{g} = \Gamma_\alpha \Pi_n \mathbf{g}. \tag{3.7}$$

The set of all histories which have n derivatives at $s = 0$ and which also belong to the Banach space $\mathcal{L}_{h,p}$ will be denoted by \mathcal{D}_n . This set \mathcal{D}_n is a linear subspace of $\mathcal{L}_{h,p}$, but it is not closed in $\mathcal{L}_{h,p}$ and hence not a Banach space.

The following theorem is an analogue of the classical Taylor approximation theorem.

Theorem 1. *Assume that n, p , and the order r of the influence function h satisfy the inequality*

$$n < r - \frac{1}{p} \quad \left(\frac{1}{p} = 0 \quad \text{if} \quad p = \infty \right). \tag{3.8}$$

Then the Taylor transformation Π_n maps the subspace \mathcal{D}_n of $\mathcal{L}_{h,p}$ into itself, and, for all \mathbf{g} in \mathcal{D}_n ,

$$\lim_{\alpha \rightarrow 0} \frac{1}{\alpha^n} \|\Gamma_\alpha(\mathbf{g} - \Pi_n \mathbf{g})\|_{h,p} = 0. \tag{3.9}$$

We can also write (3.9), using (3.7), in the form

$$\Gamma_\alpha \mathbf{g} - \Pi_n \Gamma_\alpha \mathbf{g} = \mathbf{g}_\alpha - \Pi_n \mathbf{g}_\alpha = o(\alpha^n), \tag{3.10}$$

where the order symbol o must be understood in terms of the norm of the function space $\mathcal{L}_{h,p}$. Roughly speaking, the theorem states that in the space $\mathcal{D}_n \in \mathcal{L}_{h,p}$ a slow history is close to its Taylor approximation and that the distance between them is $o(\alpha^n)$.

Proof of Theorem 1. We consider only the case when p is finite. The case when $p = \infty$ can be treated analogously.

First, we show that every polynomial of degree $\leq n$ has a finite norm. For this purpose it is sufficient to prove that

$$\int_0^\infty (s^k h(s))^p ds = L_k < \infty \quad \text{for } k \leq n. \tag{3.11}$$

It follows from (2.2) that

$$\int_1^\infty (s^k h(s))^p ds = \int_1^\infty (s^r h(s))^p s^{p(k-r)} ds \leq N^p \int_1^\infty s^{-p(r-k)} ds. \tag{3.12}$$

The inequalities (3.8) and $k \leq n$ imply that $p(r-k) > 1$ and hence that the integral

$$\int_1^\infty s^{-p(r-k)} ds = \frac{1}{p(r-k)-1}$$

is finite. Since $s^k h(s)$ is continuous, it then follows from (3.12) that the integral (3.11) is finite.

We have shown that any polynomial of degree $\leq n$ belongs to $\mathcal{L}_{h,p}$ and hence to \mathcal{D}_n , because it has n derivatives at $s=0$. Since a Taylor approximation $\Pi_n \mathbf{g}$ is a polynomial of degree $\leq n$, it follows that Π_n maps \mathcal{D}_n into itself.

The definitions (3.5) and (3.6) imply that the history

$$\mathbf{f} = \mathbf{g} - \Pi_n \mathbf{g} \in \mathcal{D}_n \tag{3.13}$$

satisfies the limit relation

$$\lim_{s \rightarrow 0} \frac{|\mathbf{f}(s)|}{s^n} = 0. \tag{3.14}$$

The definitions (2.3), (3.1), and (3.13) show that the assertion (3.9) of the theorem is equivalent to

$$\lim_{\alpha \rightarrow 0} \frac{1}{\alpha^{n p}} \int_0^\infty (|\mathbf{f}(\alpha s)| h(s))^p ds = 0. \tag{3.15}$$

We observe that

$$\frac{1}{\alpha^{n p}} \int_0^\infty (|\mathbf{f}(\alpha s)| h(s))^p ds = \frac{1}{\alpha^{n p + 1}} \int_0^\infty \left(|\mathbf{f}(s)| h\left(\frac{s}{\alpha}\right) \right)^p ds \tag{3.16}$$

and investigate the latter expression. Let $\varepsilon > 0$ be given. It follows from (3.14) that we can find a $\sigma(\varepsilon) > 0$ such that

$$\frac{|f(s)|}{s^n} \leq \varepsilon \quad \text{for } 0 < s \leq \sigma(\varepsilon).$$

Hence,

$$\begin{aligned} \frac{1}{\alpha^{n\beta+1}} \int_0^{\sigma(\varepsilon)} \left(|f(s)| h\left(\frac{s}{\alpha}\right) \right)^\beta ds &\leq \frac{1}{\alpha^{n\beta+1}} \varepsilon^\beta \int_0^{\sigma(\varepsilon)} \left(s^n h\left(\frac{s}{\alpha}\right) \right)^\beta ds \\ &\leq \varepsilon^\beta \int_0^\infty \left(\frac{s^n}{\alpha^n} h\left(\frac{s}{\alpha}\right) \right)^\beta d\left(\frac{s}{\alpha}\right) = \varepsilon^\beta L_n, \end{aligned} \tag{3.17}$$

where L_n is given by (3.11). On the other hand, we have

$$\frac{1}{\alpha^{n\beta+1}} \int_{\sigma(\varepsilon)}^\infty \left(|f(s)| h\left(\frac{s}{\alpha}\right) \right)^\beta ds = \alpha^w \int_{\sigma(\varepsilon)}^\infty \left(|f(s)| h(s) \right)^\beta \left(\frac{h(s/\alpha)}{\alpha^r h(s)} \right)^\beta ds, \tag{3.18}$$

where $w = \beta(r - n) - 1$. Applying the decay condition (2.1) to (3.18), we find that

$$\frac{1}{\alpha^{n\beta+1}} \int_{\sigma(\varepsilon)}^\infty \left(|f(s)| h\left(\frac{s}{\alpha}\right) \right)^\beta ds \leq \alpha^w M_{\sigma(\varepsilon)}^\beta \|f\|^\beta. \tag{3.19}$$

On combining (3.17) and (3.19), we see that

$$\frac{1}{\alpha^{n\beta+1}} \int_0^\infty \left(|f(s)| h\left(\frac{s}{\alpha}\right) \right)^\beta ds \leq \varepsilon^\beta L_n + \alpha^w M_{\sigma(\varepsilon)}^\beta \|f\|^\beta \tag{3.20}$$

holds for all $\varepsilon > 0$ and all $0 < \alpha \leq 1$. The assumed inequality (3.8) insures that $w = \beta(r - n) - 1$ is positive. Therefore, by choosing first ε and then α sufficiently small, we see that the right side of (3.20) can be made as small as desired. It then follows from (3.16) that the limit relation (3.15) holds, which completes the proof of the theorem.

In the special case $n=0$, $r > \frac{1}{\beta}$, $\mathbf{g}^{(0)} = \mathbf{0}$ Theorem 1 states the following. If $\mathbf{g} \in \mathcal{L}_{h,\beta}$ is continuous at $s=0$ with value $\mathbf{0}$, then

$$\lim_{\alpha \rightarrow 0} \|\Gamma_\alpha \mathbf{g}\| = 0. \tag{3.21}$$

If $1 \leq n < r - \frac{1}{\beta}$ and $\mathbf{g}^{(0)} = \mathbf{0}$, instead of (3.21), we have the stronger result

$$\mathbf{g}_\alpha = \Gamma_\alpha \mathbf{g} = O(\alpha). \tag{3.22}$$

4. Memory Functionals

We first recall some definitions* from the theory of functions defined on a real vector space \mathcal{H} with norm $\|\cdot\|$ and having values in another real vector space \mathcal{F} with norm $|\cdot|$.

(1) Suppose $\mathbf{g}_1, \dots, \mathbf{g}_k$ are variables in \mathcal{H} . Then a function $\mathfrak{F}(\mathbf{g}_1, \dots, \mathbf{g}_k)$, defined for all values of the variables \mathbf{g}_i in \mathcal{H} and having values in \mathcal{F} , is called

* These definitions are analogous to those given in Chap. XXVI of [4].

a *bounded k-linear form* if it is linear in each variable \mathbf{g}_i separately and if there is a constant M , independent of \mathbf{g}_i , such that

$$|\mathfrak{P}(\mathbf{g}_1, \dots, \mathbf{g}_k)| \leq M \|\mathbf{g}_1\| \dots \|\mathbf{g}_k\|. \tag{4.1}$$

The form $\mathfrak{P}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ is said to be *symmetric* if any permutation of the variables leaves the value unchanged.

(2) A function $\mathfrak{P}(\mathbf{g})$, defined for all $\mathbf{g} \in \mathcal{H}$ and having values in \mathcal{T} , is called a *bounded homogeneous polynomial of degree k* if there is a bounded symmetric k -linear form $\mathfrak{P}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ such that

$$\mathfrak{P}(\mathbf{g}) = \mathfrak{P}(\mathbf{g}, \dots, \mathbf{g}). \tag{4.2}$$

The symmetric n -linear form $\mathfrak{P}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ is uniquely determined by the homogeneous polynomial $\mathfrak{P}(\mathbf{g})$ and is called the *polar form* of the polynomial.

(3) A function \mathfrak{F} , defined on a neighborhood of $\mathbf{0} \in \mathcal{H}$ and having values in \mathcal{T} is said to be *n times Fréchet-differentiable* at $\mathbf{0} \in \mathcal{H}$ if there are bounded homogeneous polynomials $\delta^k \mathfrak{F}(\mathbf{g})$ of degree $k=0, 1, \dots, n$ such that

$$\mathfrak{F}(\mathbf{g}) = \sum_{k=0}^n \frac{1}{k!} \delta^k \mathfrak{F}(\mathbf{g}) + \|\mathbf{g}\|^n \mathfrak{R}(\mathbf{g}), \tag{4.3}$$

where

$$\lim_{\|\mathbf{g}\| \rightarrow 0} |\mathfrak{R}(\mathbf{g})| = 0. \tag{4.4}$$

The polynomial $\delta^k \mathfrak{F}(\mathbf{g})$ or its polar form $\delta^k \mathfrak{F}(\mathbf{g}_1, \dots, \mathbf{g}_k)$ is called the k^{th} *Fréchet-differential* or the k^{th} *variation* of \mathfrak{F} at $\mathbf{0} \in \mathcal{H}$. The differentials may be obtained recursively by

$$\delta^0 \mathfrak{F}(\mathbf{g}) = \mathfrak{F}(\mathbf{0}), \quad \delta^k \mathfrak{F}(\mathbf{g}) = k! \lim_{\lambda \rightarrow 0} \frac{1}{\lambda^k} \left[\mathfrak{F}(\lambda \mathbf{g}) - \sum_{j=0}^{k-1} \frac{\lambda^j}{j!} \delta^j \mathfrak{F}(\mathbf{g}) \right]. \tag{4.5}$$

We here consider the case in which \mathcal{H} is the function space of all histories \mathbf{g} with the following properties:

- (α) \mathbf{g} has a finite $\mathcal{L}_{h,p}$ -norm,
- (β) \mathbf{g} has n generalized derivatives at $s=0$,
- (γ) \mathbf{g} has a zero limit at $s=0$:

$$\lim_{s \rightarrow 0} \mathbf{g}(s) = \overset{0}{\mathbf{g}} = \mathbf{0}, \tag{4.6}$$

- (δ) n, p , and the order r of the influence function h obey the inequality (3.8).

The conditions (α) and (β) state that $\mathcal{H} \subset \mathcal{D}_n$. Condition (δ) insures that the conclusion (3.9) of Theorem 1 is valid for all $\mathbf{g} \in \mathcal{H}$.

A function \mathfrak{F} defined on a neighborhood in \mathcal{H} of the zero function $\mathbf{0} \in \mathcal{H} \subset \mathcal{L}_{h,p}$ and having values in a real Banach space \mathcal{T} will be called a *memory functional of type (h, n)* if it is n times Fréchet-differentiable at $\mathbf{0} \in \mathcal{H}$ and if it is normalized by

$$\delta^0 \mathfrak{F}(\mathbf{g}) = \mathfrak{F}(\mathbf{0}) = \mathbf{0}. \tag{4.7}$$

In some applications it may be more natural to assume that \mathfrak{F} is defined and Fréchet-differentiable on a neighborhood of zero in the entire space $\mathcal{L}_{h,p}$. However, the approximation theorem of the following section applies only to histories which belong to the subspace \mathcal{H} of $\mathcal{L}_{h,p}$.

5. The Approximation Theorem

Theorem 2. Let \mathfrak{F} be a memory functional of type (h, n) ; the histories \mathbf{g} in the domain space \mathcal{H} of \mathfrak{F} have values in the space \mathcal{S} , and \mathfrak{F} itself has values in \mathcal{T} . Then, for each k -tuple of indices (j_1, j_2, \dots, j_k) such that

$$1 \leq j_1 \leq j_2 \leq \dots \leq j_k \leq n, \quad j_1 + j_2 + \dots + j_n \leq n, \tag{5.1}$$

there exists a bounded k -linear form $\mathbf{l}_{j_1 \dots j_k}$ with variables in \mathcal{S} and values in \mathcal{T} such that, for all $\mathbf{g} \in \mathcal{H}$,

$$\mathfrak{F}(\mathbf{g}_\alpha) = \sum_{(j_1, \dots, j_k)} \mathbf{l}_{j_1 \dots j_k}(\mathbf{g}_\alpha^{(j_1)}, \dots, \mathbf{g}_\alpha^{(j_k)}) + o(\alpha^n), \tag{5.2}$$

where

$$\mathbf{g} = \Gamma_\alpha \mathbf{g}, \quad \text{i.e.,} \quad \mathbf{g}_\alpha(s) = \mathbf{g}(\alpha s), \tag{5.3}$$

and where

$$\mathbf{g}_\alpha^{(j)} = \alpha^j \mathbf{g}^{(j)} \tag{5.4}$$

is the generalized j^{th} derivative at $s=0$ of \mathbf{g}_α , defined according to (3.5). The sum in (5.2) is to be extended over all sets (j_1, \dots, j_k) of indices satisfying (5.1), and the order symbol $o(\alpha^n)$ has the usual sense. The multilinear forms $\mathbf{l}_{j_1 \dots j_k}$ are uniquely determined by \mathfrak{F} .

This theorem permits the asymptotic approximation of a memory functional, for ‘‘slow’’ histories, by a polynomial function of the derivatives at $s=0$ of the argument function of the functional. It is worth noting that the approximating expression

$$\mathfrak{F}_n(\mathbf{g}) = \sum_{(j_1, \dots, j_k)} \mathbf{l}_{j_1 \dots j_k}(\mathbf{g}^{(j_1)}, \dots, \mathbf{g}^{(j_k)}), \tag{5.6}$$

regarded as a functional of \mathbf{g} , is not a memory functional in the sense of the previous section because it is not even continuous at the zero function $\mathbf{g} = \mathbf{0} \in \mathcal{H}$.

For $n = 1$ and $n = 2$, (5.6) reduces to

$$\mathfrak{F}_1(\mathbf{g}) = \mathbf{l}_1(\mathbf{g}^{(1)}), \quad \mathfrak{F}_2(\mathbf{g}) = \mathbf{l}_1(\mathbf{g}^{(1)}) + \mathbf{l}_{11}(\mathbf{g}^{(1)}, \mathbf{g}^{(1)}) + \mathbf{l}_2(\mathbf{g}^{(2)}) \tag{5.7}$$

respectively, where \mathbf{l}_1 and \mathbf{l}_2 are linear and \mathbf{l}_{11} is bilinear.

Proof of Theorem 2. For $n = 0$, the theorem is a trivial consequence of (3.24). For $n \geq 1$, the proof is based on a combination of the Fréchet-differentiability assumption (4.3) and equation (3.40) of Theorem 1 which, in the notation of (5.3), may be written as

$$\mathbf{g}_\alpha = \Pi_n \mathbf{g}_\alpha + o(\alpha^n). \tag{5.8}$$

Since $\mathbf{g}^{(0)} = \mathbf{0}$ by (4.6), the definition (3.6) gives

$$(\Pi_n \mathbf{g}_\alpha)(s) = \sum_{j=1}^n \alpha^j \frac{s^j}{j!} \mathbf{g}^{(j)}. \tag{5.9}$$

The result (3.22) applies to both $\mathbf{g} \in \mathcal{H}$ and $\Pi_n \mathbf{g} \in \mathcal{H}$:

$$\Pi_n \mathbf{g}_\alpha = O(\alpha), \tag{5.10}$$

$$\mathbf{g}_\alpha = O(\alpha). \tag{5.11}$$

Combining (5.11) with (4.3), (4.4), and (4.7), we obtain

$$\mathfrak{F}(\mathbf{g}_\alpha) = \sum_{k=1}^n \frac{1}{k!} \delta^k \mathfrak{F}(\mathbf{g}_\alpha) + o(\alpha^n). \tag{5.12}$$

Consider now a bounded homogeneous polynomial $\mathfrak{P}(\mathbf{g})$ of degree k , $1 \leq k \leq n$, in the sense of (2) of §4. The differentials $\delta^k \mathfrak{F}$, $k > 1$ are such polynomials. The polar form of \mathfrak{P} will also be denoted by \mathfrak{P} . Using the multilinearity of this polar form \mathfrak{P} , we obtain by (4.2), (5.8), and (5.10)

$$\begin{aligned} \mathfrak{P}(\mathbf{g}_\alpha) &= \mathfrak{P}(\Pi_n \mathbf{g}_\alpha + o(\alpha^n), \dots, \Pi_n \mathbf{g}_\alpha + o(\alpha^n)) \\ &= \mathfrak{P}(\Pi_n \mathbf{g}_\alpha, \dots, \Pi_n \mathbf{g}_\alpha) + \sum \mathfrak{P}(O(\alpha), \dots, o(\alpha^n), \dots, O(\alpha)), \end{aligned}$$

where each term of the sum contains at least one variable $o(\alpha^n)$. It follows from the boundedness (4.1) of \mathfrak{P} that the terms in the sum are all $o(\alpha^{n+k-1})$. Since $k \geq 1$, we have

$$\mathfrak{P}(\mathbf{g}_\alpha) = \mathfrak{P}(\Pi_n \mathbf{g}_\alpha) + o(\alpha^n). \tag{5.13}$$

We now investigate $\mathfrak{P}(\Pi_n \mathbf{g}_\alpha)$. Using (5.9) and the multilinearity of the polar form \mathfrak{P} , we obtain

$$\mathfrak{P}(\Pi_n \mathbf{g}_\alpha) = \sum_{j_1=1}^n \dots \sum_{j_k=1}^n \frac{\alpha^{j_1+\dots+j_k}}{j_1! \dots j_k!} \mathfrak{P}(s^{j_1} \mathbf{g}^{(j_1)}, \dots, s^{j_k} \mathbf{g}^{(j_k)}). \tag{5.14}$$

Due to the symmetry of \mathfrak{P} , all terms of (5.14) which differ only in the order of the indices j_1, \dots, j_k are the same. Collecting these terms and separating all terms of order higher than n in α , we get

$$\mathfrak{P}(\Pi_n \mathbf{g}_\alpha) = \sum_{\substack{(j_1, \dots, j_k) \\ k \text{ fixed}}} \alpha^{j_1+\dots+j_k} m_{j_1 \dots j_k} \mathfrak{P}(s^{j_1} \mathbf{g}^{(j_1)}, \dots, s^{j_k} \mathbf{g}^{(j_k)}) + O(\alpha^{n+1}) \tag{5.15}$$

where the sum is to be extended over all k -tuples (j_1, \dots, j_k) satisfying (5.1) and where the $m_{j_1 \dots j_k}$ are positive rational numbers.

Now, the function

$$\mathfrak{p}_{j_1 \dots j_k}(\mathbf{a}_1, \dots, \mathbf{a}_k) = \mathfrak{P}(s^{j_1} \mathbf{a}_1, \dots, s^{j_k} \mathbf{a}_k), \tag{5.16}$$

with variables $\mathbf{a}_i \in \mathcal{S}$ and values in \mathcal{T} is clearly multilinear, because \mathfrak{P} is. Also, $\mathfrak{p}_{j_1 \dots j_k}$ is bounded. Indeed, application of (4.1) to (5.16) yields

$$|\mathfrak{p}_{j_1 \dots j_k}(\mathbf{a}_1, \dots, \mathbf{a}_k)| \leq M \|s^{j_1} \mathbf{a}_1\| \dots \|s^{j_k} \mathbf{a}_k\|,$$

and (3.11) shows

$$\|s^{j_l} \mathbf{a}_l\|_k = |\mathbf{a}_l| \sqrt[j_l]{L_{j_l}},$$

hence

$$|\mathfrak{p}_{j_1 \dots j_k}(\mathbf{a}_1, \dots, \mathbf{a}_k)| \leq \bar{M} |\mathbf{a}_1| \dots |\mathbf{a}_k|. \tag{5.17}$$

Substituting (5.16) into (5.15) and using (5.4) yields

$$\mathfrak{P}(\Pi_n \mathbf{g}_\alpha) = \sum_{\substack{(j_1, \dots, j_k) \\ k \text{ fixed}}} \mathbf{I}_{j_1 \dots j_k}(\mathbf{g}_\alpha^{(j_1)}, \dots, \mathbf{g}_\alpha^{(j_k)}) + O(\alpha^{n+1}) \tag{5.18}$$

where the $\mathbf{I}_{j_1 \dots j_k}$ are bounded k -linear forms.

Applying the results (5.13) and (5.18) to $\mathfrak{F} = \frac{1}{k!} \delta^k \mathfrak{F}$ in (5.12), we obtain formula (5.2).

The uniqueness of the $\mathfrak{I}_{j_1 \dots j_k}$ follows easily from the observation that they are linearly independent and of order $O(\alpha^n)$ or lower.

We remark that the multilinear forms $\mathfrak{I}_{j_1 \dots j_k}$ of (5.2) are not necessarily symmetric.

II. Applications

6. The Concept of a Simple Fluid

The notion of a simple fluid has been given a definition within the framework of a general theory of the mechanical behavior of materials*. This definition is based on the following two physical assumptions**:

(a) The present stress depends on the past history of the first spatial gradient of the displacement function.

(b) A fluid has no preferred configurations.

Using the principle of material objectivity ([I], §11), it was shown in reference [I] that the constitutive equation of a simple fluid can be written in the form (22.12) of [I]. This functional relation, in a slightly different notation, reads

$$\mathbf{S}(t) = \mathfrak{F}_{s=0}^{\infty} (\mathbf{C}_i(t-s); \rho(t)). \tag{6.1}$$

Here $\mathbf{S}(t)$ is the stress and $\rho(t)$ the density at time t . $\mathbf{C}_i(\tau)$ is called the right Cauchy-Green tensor at time τ relative to the configuration at time t . This tensor is defined by

$$\mathbf{C}_i(\tau) = \mathbf{F}_i^T(\tau) \mathbf{F}_i(\tau) \tag{6.2}$$

where

$$\mathbf{F}_i(\tau) = \nabla_{\mathbf{x}} \boldsymbol{\chi}_i(\mathbf{x}, \tau) \tag{6.3}$$

is the gradient the displacement function $\boldsymbol{\xi} = \boldsymbol{\chi}_i(\mathbf{x}, \tau)$ which gives the position at time τ of the material point having the position \mathbf{x} at time t . The stress tensor $\mathbf{S}(t)$ is symmetric. The Cauchy-Green tensor $\mathbf{C}_i(\tau)$ is positive definite and symmetric; for $\tau = t$ it reduces to the unit tensor \mathbf{I} :

$$\mathbf{C}_i(t) = \mathbf{I}. \tag{6.4}$$

The functional \mathfrak{F} in (6.1) is isotropic; *i.e.*, \mathfrak{F} obeys

$$\mathbf{Q} \mathfrak{F}_{s=0}^{\infty} (\mathbf{C}(s); \rho) \mathbf{Q}^T = \mathfrak{F}_{s=0}^{\infty} (\mathbf{Q} \mathbf{C}(s) \mathbf{Q}^T; \rho) \tag{6.5}$$

identically in the history $\mathbf{C}(s) = \mathbf{C}_i(t-s)$ and the orthogonal tensor \mathbf{Q} ; here \mathbf{Q}^T is the transpose of \mathbf{Q} . It follows from (6.5) that the value of \mathfrak{F} for the "rest

* [I], § 21.

** In [3] we give a survey of the theory of simple fluids with emphasis on physical applications. In that paper we anticipate some of the results rigorously derived here. Although the language and the definitions of the present paper are slightly different from those used in §§ 6 and 7 of [3], the arguments presented here prove also the theorems stated there.

history" $\mathbf{C}(s) \equiv \mathbf{I}$ is a scalar multiple $-\rho(\varrho)$ of the unit tensor \mathbf{I} . Defining

$$\mathfrak{F}_{s=0}^{\infty}(\mathbf{G}(s), \varrho) = \rho(\varrho) \mathbf{I} + \mathfrak{F}_{s=0}^{\infty}(\mathbf{I} + \mathbf{G}(s), \varrho), \quad (6.6)$$

we may rewrite (6.1) in the form

$$\mathbf{S}(t) = -\rho(\varrho(t)) \mathbf{I} + \mathfrak{F}_{s=0}^{\infty}(\mathbf{C}_t(t-s) - \mathbf{I}, \varrho(t)). \quad (6.7)$$

The functional $\mathfrak{F}_{s=0}^{\infty}(\mathbf{G}(s); \varrho)$ is defined for functions $\mathbf{G}(s)$ with the property

$$\mathbf{G}(0) = \mathbf{0}, \quad (6.8)$$

and it has the value $\mathbf{0}$ for the zero function $\mathbf{G}(s) \equiv \mathbf{0}$; *i.e.*,

$$\mathfrak{F}_{s=0}^{\infty}(\mathbf{0}, \varrho) = \mathbf{0}. \quad (6.9)$$

It is also isotropic in the sense of (6.5).

We now assume that for each simple fluid defined by a constitutive equation (6.7) there exists an influence function h of an order r such that the functional \mathfrak{F} of (6.7) is a memory functional of type (h, n) in the sense of §4. The domain of \mathfrak{F} is a class \mathcal{H} of histories \mathbf{G} whose values $\mathbf{G}(s)$ are in the space \mathcal{S} of all symmetric tensors. For the norm $|\cdot|$ in \mathcal{S} we use

$$|\mathbf{A}| = \sqrt{\text{tr } \mathbf{A}^2} \quad \text{for } \mathbf{A} \in \mathcal{S}. \quad (6.10)$$

The range space \mathcal{T} of \mathfrak{F} is the same as the range space of the histories \mathbf{G} , *i.e.* the space $\mathcal{T} = \mathcal{S}$ of all symmetric tensors with norm (6.10). Equation (6.9) insures that \mathfrak{F} has the normalization (4.7) required for a memory functional. The density ϱ enters into (6.7) only as a real parameter. The assumption that \mathfrak{F} is a memory functional implies that its domain of definition contains a neighborhood of zero in a function space \mathcal{H} which is defined by the conditions (α) – (δ) of §4. The condition (γ) expresses the assumption that all histories $\mathbf{G} \in \mathcal{H}$ correspond to motions which are continuous at the present instant $s=0$ (*cf.* (6.8)).

If the simple fluid under consideration is *incompressible*, we must make some alterations in our starting assumptions. For every possible motion in such a fluid the density ϱ is constant and the tensor $\mathbf{C}_t(\tau)$ is unimodular. In addition, the stress is determined by the history of the motion only up to a hydrostatic pressure p . Consequently, the equation (6.7) must be replaced by

$$\mathbf{S}(t) = -p \mathbf{I} + \mathfrak{F}_{s=0}^{\infty}(\mathbf{C}_t(t-s) - \mathbf{I}) \quad (6.11)$$

in which the indeterminate pressure p and the functional \mathfrak{F} may be normalized by

$$0 = \text{tr } \mathbf{S}(t) + 3p = \text{tr } \mathfrak{F}_{s=0}^{\infty}(\mathbf{C}_t(t-s) - \mathbf{I}). \quad (6.12)$$

If we were to limit the domain of the functional \mathfrak{F} in (6.11) to kinematically possible histories, this domain would not contain a neighborhood of the zero function in an appropriate function space, \mathcal{H} ; therefore, \mathfrak{F} could not then be a memory functional. We assume, however, that \mathfrak{F} becomes a memory functional

when its domain is extended by putting

$$\mathfrak{F} \Big|_{s=0}^{\infty} (\mathbf{C}(s) - \mathbf{I}) = \mathfrak{F} \Big|_{s=0}^{\infty} ((\det \mathbf{C}(s))^{-1} \mathbf{C}(s) - \mathbf{I}) \tag{6.13}$$

when $\mathbf{C}(s)$ is not unimodular.

Aside from the added properties (6.12) and (6.13), the functional \mathfrak{F} occurring in (6.11) is assumed to be of the same type as that in (6.7) with respect to both isotropy and memory.

7. Approximations of Order n for Simple Fluids

We now apply Theorem 2, §5, to the memory functional occurring in the constitutive equations for simple fluids (6.7) or (6.11).

Suppose a motion with Cauchy-Green tensor $\mathbf{C}_t(\tau)$ is given. The corresponding history \mathbf{G} is defined by

$$\mathbf{G}(s) = \mathbf{C}_t(t - s) - \mathbf{I}. \tag{7.1}$$

If $\mathbf{C}_t(\tau)$ is n times differentiable with respect to τ at $\tau=t$, the k^{th} Rivlin-Ericksen tensor \mathbf{A}_k , $k=1, 2, \dots, n$ is defined as follows:

$$\mathbf{A}_k = \left. \frac{d^k}{d\tau^k} \mathbf{C}_t(\tau) \right|_{\tau=t} = (-1)^k \mathbf{G}^{(k)}, \tag{7.2}$$

where $\mathbf{G}^{(k)}$ is the k^{th} derivative of $\mathbf{G}(s)$ at $s=0$, as in §2.

We now consider histories \mathbf{G}_α obtained from $\mathbf{G} \in \mathcal{H}$ by retardation as in (5.3). The corresponding Rivlin-Ericksen tensors are

$$\mathbf{A}_k^\alpha = \alpha^k \mathbf{A}_k = (-1)^k \mathbf{G}_\alpha^{(k)}. \tag{7.3}$$

They differ only by the inessential factor $(-1)^k$ from the tensors $\mathbf{G}_\alpha^{(k)}$ to be substituted for $\mathbf{g}_\alpha^{(k)}$ in the approximation formula (5.2). This formula, applied to the constitutive equations (6.7) or (6.11), yields the following expression for the stress tensor \mathbf{S}_α corresponding to the retarded history \mathbf{G}_α :

$$\mathbf{S}_\alpha = -p \mathbf{I} + \sum_{(j_1, \dots, j_k)} \mathbf{m}_{j_1 \dots j_k} (\mathbf{A}_{j_1}^\alpha, \dots, \mathbf{A}_{j_k}^\alpha) + o(\alpha^n), \tag{7.4}$$

where the summation is to be extended over all sets of indices (j_1, \dots, j_k) obeying (5.1). The terms $\mathbf{m}_{j_1 \dots j_k} (\mathbf{A}_{j_1}^\alpha, \dots, \mathbf{A}_{j_k}^\alpha)$ are linear in each of the variables. For compressible fluids, it is understood that p and $\mathbf{m}_{j_1 \dots j_k}$ depend on the density ρ .

The equation (7.4) remains valid even when the derivatives shown in (7.2) exist only in the generalized sense of (3.5).

The multilinear forms $\mathbf{m}_{j_1 \dots j_k}$ in (7.4) are isotropic functions, which means that they obey the identities

$$\mathbf{Q} \mathbf{m}_{j_1 \dots j_k} (\mathbf{U}_1, \dots, \mathbf{U}_k) \mathbf{Q}^T = \mathbf{m}_{j_1 \dots j_k} (\mathbf{Q} \mathbf{U}_1 \mathbf{Q}^T, \dots, \mathbf{Q} \mathbf{U}_k \mathbf{Q}^T) \tag{7.5}$$

for all orthogonal \mathbf{Q} and all symmetric tensors $\mathbf{U}_1, \dots, \mathbf{U}_k$. This proposition follows easily from the fact that the memory functional \mathfrak{F} occurring in (6.7) or (6.11) is isotropic in the sense of (6.5), from the observation that the conjugation $\mathbf{G}(s) \rightarrow \mathbf{Q} \mathbf{G}(s) \mathbf{Q}^T$ leaves the norm $\|\mathbf{G}\|$ unchanged, and from the uniqueness of the multilinear forms $\mathbf{m}_{j_1 \dots j_k}$ asserted in Theorem 2, §5.

It follows from known theorems on isotropic functions that each form $\mathfrak{m}_{j_1 \dots j_k}(\mathbf{A}_{j_1}, \dots, \mathbf{A}_{j_k})$, because it is isotropic and multilinear, may be expressed as a sum

$$\mathfrak{m}_{j_1 \dots j_k}(\mathbf{A}_{j_1}, \dots, \mathbf{A}_{j_k}) = \sum_i \zeta_i \tag{7.6}$$

of products of the form

$$\zeta_i = \mu_i \varphi_1 \varphi_2 \dots \varphi_{m_i}(\mathbf{A}_{i_1} \mathbf{A}_{i_2} \dots \mathbf{A}_{i_{q_i}} + \mathbf{A}_{i_{q_i}} \dots \mathbf{A}_{i_2} \mathbf{A}_{i_1}) \tag{7.7}$$

where the φ_j 's are traces of products of some of the tensors \mathbf{A}_r , and are such that each \mathbf{A}_r , $r=1, \dots, k$, occurs precisely once in each term ζ_i . In the case of an incompressible simple fluid the coefficients μ_i are constants, whereas for simple fluids in general the μ_i are functions of the density. Thus, for simple fluids, a finite number of scalar material functions $\check{p}(\rho)$, $\mu_i(\rho)$ suffices to determine the stress \mathbf{S}_α to within terms of order n in α .

The case $n=1$ of (7.4) is of particular interest. With use of (7.6), (7.7), we obtain

$$\mathbf{S}_\alpha = -\check{p} \mathbf{I} + \eta \mathbf{A}_1^\alpha + \frac{\lambda}{2} (\text{tr } \mathbf{A}_1^\alpha) \mathbf{I} + o(\alpha). \tag{7.8}$$

Now, the first Rivlin-Ericksen tensor $\mathbf{A}_1 = 2\mathbf{D}$ differs from the rate of deformation tensor \mathbf{D} only by the factor 2^* . It follows that (7.8) is, to within terms of order $o(\alpha)$, simply the constitutive equation (1.4) of a Newtonian fluid.

When the fluid is incompressible, the case $n=2$ of (7.4) takes a remarkably simple form. The observation that $\text{tr } \mathbf{A}_1 = 0$ for isochoric motions and use of (7.6), (7.7) yield

$$\mathbf{S}_\alpha = -\check{p} \mathbf{I} + \eta \mathbf{A}_1^\alpha + \beta (\mathbf{A}_1^\alpha)^2 + \gamma \mathbf{A}_2^\alpha + o(\alpha^2), \tag{7.9}$$

where η , β , and γ are material constants and where \check{p} is an indeterminate pressure. This pressure \check{p} differs, in general, from the mean pressure \check{p} defined by the normalization (6.12), because it is obtained from \check{p} by absorbing all scalar multiples of \mathbf{I} arising from \mathfrak{m}_{11} and \mathfrak{m}_2 through use of (7.6) and (7.7).

Motivated by (7.9), we can define an *incompressible second-order fluid* by the constitutive equation

$$\mathbf{S} = -\check{p} \mathbf{I} + \eta \mathbf{A}_1 + \beta \mathbf{A}_1^2 + \gamma \mathbf{A}_2. \tag{7.10}$$

Incompressible Newtonian fluids correspond to the special case $\beta = \gamma = 0$ of (7.10).

In some dynamical situations equation (7.10) leads to a linear partial differential equation for the velocity, just as in the Newtonian case. For example, consider a rectilinear shearing flow which, in Cartesian coordinates x, y, z , has a velocity field with components

$$\{v^i\} = \{0, v(x, t), 0\}. \tag{7.11}$$

If the body forces are conservative, substitution of (7.10) and (7.11) into CAUCHY'S dynamical equations leads to the following third-order partial differential equation for $v(x, t)$:

$$\frac{\partial v}{\partial t} = a \frac{\partial^2 v}{\partial x^2} + b \frac{\partial^3 v}{\partial x^2 \partial t}, \quad a = \frac{\eta}{\rho}, \quad b = \frac{\gamma}{\rho}. \tag{7.12}$$

* Cf. [I], (9.7).

A non-steady flow of the Couette type, in cylindrical coordinates r, ϑ, z , has a velocity field with contravariant components

$$\{v^i\} = \{0, \omega(r, t), 0\}. \tag{7.13}$$

For such a flow, instead of (7.12), we get

$$r^3 \frac{\partial \omega}{\partial t} = \frac{\partial}{\partial r} \left[r^3 \left(a \frac{\partial \omega}{\partial r} + b \frac{\partial^2 \omega}{\partial r \partial t} \right) \right]. \tag{7.14}$$

In the Newtonian case, since $b=0$, equations (7.12) and (7.14) reduce to diffusion equations. As for the diffusion equations, many physically interesting solutions of (7.12) and (7.14) may be obtained by separation of variables*. It would be desirable to develop a mathematical theory of third-order partial differential equations of the type (7.12), (7.14).

8. Simple Materials in General

The considerations of the previous two sections may easily be extended to the general simple materials defined in Reference [I], Part III.

The form of the constitutive equation of an *isotropic simple material* may be obtained from (6.1) by replacing the scalar density $\rho(t)$ by the left Cauchy-Green tensor $\mathbf{B}(t)$, taken relative to an undistorted reference state (cf. [I] (22.10)). A consideration analogous to the one which led to (6.7) shows that the constitutive equation of an isotropic simple material may be written in the form

$$\mathbf{S}(t) = \mathfrak{h}(\mathbf{B}(t)) + \mathfrak{F} \left(C_s(t-s) - \mathbf{I}; \mathbf{B}(t) \right). \tag{8.1}$$

Here, the functional \mathfrak{F} depends on a tensor parameter \mathbf{B} , instead of on a scalar parameter ρ as in (6.7). We assume again that there is an influence function h of order n such that the functional \mathfrak{F} in (8.1) is a memory functional of type (h, n) . In place of (7.4), we then obtain the following approximation formula for the stress \mathbf{S}_α produced by a slow motion of an isotropic simple material:

$$\mathbf{S}_\alpha = \mathfrak{h}(\mathbf{B}) + \sum_{(j_1, \dots, j_k)} \mathfrak{m}_{j_1 \dots j_k} (A_{j_1}^\alpha, \dots, A_{j_k}^\alpha; \mathbf{B}) + o(\alpha^n), \tag{8.2}$$

where $\mathfrak{m}_{j_1 \dots j_k}$ is linear in each of its first k tensor variables but not necessarily in the last tensor variable \mathbf{B} .

The formulae corresponding to (8.1) and (8.2) in the case of *anisotropic simple materials* are obtained from (8.1) and (8.2) simply by replacing all tensors \mathbf{T} occurring in these formulae by their conjugates $\mathbf{R}^T \mathbf{T} \mathbf{R}$ with the rotation tensor $\mathbf{R} = \mathbf{R}(t)$ of the displacement from the reference state (cf. [I], (22.8)).

In the case of isotropic materials, the function \mathfrak{h} and the functional \mathfrak{F} in (8.1) are isotropic in the sense that they obey the identities

$$\mathbf{Q} \mathfrak{h}(\mathbf{B}) \mathbf{Q}^T = \mathfrak{h}(\mathbf{Q} \mathbf{B} \mathbf{Q}^T), \tag{8.3}$$

$$\mathfrak{F} \left(\mathbf{Q} \mathbf{G}(s) \mathbf{Q}^T, \mathbf{Q} \mathbf{B} \mathbf{Q}^T \right) = \mathbf{Q} \mathfrak{F} \left(\mathbf{G}(s), \mathbf{B} \right) \mathbf{Q}^T \tag{8.4}$$

* In particular, the sinusoidal vibration problems discussed for Newtonian fluids in §§ 345–346 of LAMB’S treatise [5] are easily solved for second-order fluids. Also, special solutions of (7.14) corresponding to sinusoidal vibrations of a fluid between coaxial cylinders can readily be found.

for all orthogonal tensors \mathbf{Q} . As in §7, it follows that the $\mathbf{m}_{j_1 \dots j_k}$ in (8.2) are isotropic functions of all their variables; *i.e.*, they obey the identities

$$\mathbf{Q} \mathbf{m}_{j_1 \dots j_k}(\mathbf{U}_1, \dots, \mathbf{U}_k; \mathbf{B}) \mathbf{Q}^T = \mathbf{m}_{j_1 \dots j_k}(\mathbf{Q} \mathbf{U}_1 \mathbf{Q}^T, \dots, \mathbf{Q} \mathbf{U}_k \mathbf{Q}^T; \mathbf{Q} \mathbf{B} \mathbf{Q}^T) \quad (8.5)$$

for all orthogonal \mathbf{Q} and all symmetric tensors $\mathbf{U}_1, \dots, \mathbf{U}_k, \mathbf{B}$ for which $\mathbf{m}_{j_1 \dots j_k}$ is defined. The methods developed by SPENCER & RIVLIN ([6], [7] and [8]) may be used to derive explicit representations for the $\mathbf{m}_{j_1 \dots j_k}$ of a type analogous to but more complicated than (7.6), (7.7). Using such a representation, one can show that, in the case $n=1$, the approximation formula (8.2) reduces to

$$\mathbf{S}_\alpha = \mathfrak{h}(\mathbf{B}) + \mathbf{A}_1^\alpha \mathfrak{t}_1(\mathbf{B}) + \mathfrak{t}_1(\mathbf{B}) \mathbf{A}_1^\alpha + \text{tr}(\mathbf{A}_1^\alpha \mathfrak{t}_2(\mathbf{B})) \mathfrak{t}_3(\mathbf{B}), + o(\alpha), \quad (8.6)$$

where \mathfrak{h} , \mathfrak{t}_1 , \mathfrak{t}_2 and \mathfrak{t}_3 are isotropic functions of the one variable \mathbf{B} and hence have representations of the form

$$\mathfrak{h}(\mathbf{B}) = \beta_0 \mathbf{I} + \beta_1 \mathbf{B} + \beta_2 \mathbf{B}^2, \quad (8.7)$$

in which β_0 , β_1 and β_2 are scalar functions of the three principal invariants of \mathbf{B} .

The first term $\mathfrak{h}(\mathbf{B})$ in the expression (8.2) for the stress \mathbf{S}_α corresponds to purely elastic response. The sum in (8.2) may be interpreted as representing the internal friction for slow motions.

Acknowledgement. The research leading to this paper was supported by the Air Force Office of Scientific Research under Contract AF 49(638)541 with Mellon Institute and by the National Science Foundation under Grant NSF-G 5250 to Carnegie Institute of Technology.

References

- [1] NOLL, W.: A mathematical theory of the mechanical behavior of continuous media. Arch. Rational Mech. Anal. **2**, 197–266 (1958).
- [2] COLEMAN, B. D., & W. NOLL: On certain steady flows of general fluids. Arch. Rational Mech. Anal. **4**, 289–303 (1959).
- [3] COLEMAN, B. D., & W. NOLL: Recent results in the continuum theory of viscoelastic fluids. To appear in Ann. New York Acad. Sci.
- [4] HILLE, E., & R. S. PHILLIPS: Functional Analysis and Semi-Groups. American Mathematical Society Colloquium Publications, Vol. XXXI. New York 1957.
- [5] LAMB, H.: Hydrodynamics, Chap. XI, 6th Edit. Cambridge: Univ. Press 1932.
- [6] SPENCER, A. J. M., & R. S. RIVLIN: The theory of matrix polynomials and its application to the mechanics of isotropic continua. Arch. Rational Mech. Anal. **3**, 309–336 (1959).
- [7] SPENCER, A. J. M., & R. S. RIVLIN: Finite integrity bases for five or fewer symmetric 3×3 matrices. Arch. Rational Mech. Anal. **3**, 435–446 (1959).
- [8] SPENCER, A. J. M., & R. S. RIVLIN: Further results in the theory of matrix polynomials. Arch. Rational Mech. Anal. **4**, 214–230 (1960).

Mellon Institute
Pittsburgh, Pennsylvania
and
Mathematics Department
Carnegie Institute of Technology
Pittsburgh, Pennsylvania

(Received August 20, 1960)

Foundations of Linear Viscoelasticity*

BERNARD D. COLEMAN

AND

WALTER NOLL

1. INTRODUCTION

THE classical linear theory of viscoelasticity was apparently first formulated by Boltzmann¹ in 1874. His original presentation covered the three-dimensional case, but was restricted to isotropic materials. The extension of the theory to anisotropic materials is, however, almost immediately evident on reading Boltzmann's paper, and the basic hypotheses of the theory have not changed since 1874. Since that date, much work has been done on the following aspects of linear viscoelasticity: solution of special boundary value problems,^{2a} reformulation^{3,4} of the one-dimensional version of the theory in terms of new material functions (such as "creep functions" and frequency-dependent complex "impedances") which appear to be directly accessible to measurement, experimental determination^{2b} of the material functions for those materials for which the theory appears useful, prediction of the form of the material functions from molecular models, and, recently, axiomatization^{5,6} of the theory. In this article, instead of being concerned with these matters, we reexamine the fundamental hypotheses of linear viscoelasticity in the light of recent advances in nonlinear continuum mechanics.

The basic assumption of the classical linear theory of viscosity is a constitutive equation relating the stress tensor $T(t)$ at time t to the history of the infinitesimal strain tensor $E(t-s)$, $0 \leq s < \infty$. This assumption asserts that if $E(t-s)$, taken relative to a natural reference configuration corresponding to zero equilibrium stress, is small in magnitude for all s , then

$$T(t) = \Omega\{E(t)\} + \Phi(0)\{E(t)\} + \int_0^\infty \Phi(s)\{E(t-s)\}ds, \quad (1.1)$$

* This work was supported by the Air Force Office of Scientific Research under contract and by the National Science Foundation under Grant NSF-G5250.

¹ S. Boltzmann, Sitzber. Akad. Wiss. Wien, Math. naturw. Kl. 70, 275 (1874); Pogg. Ann. Phys. 7, 624 (1876).

² (a) E. H. Lee, in *Viscoelasticity*, edited by J. T. Bergen (Academic Press, Inc., New York, 1960), p. 1; (b) J. D. Ferry and K. Ninomiya, *ibid.*, p. 55.

³ B. Gross, *Mathematical Structure of the Theories of Viscoelasticity* (Hermann & Cie., Paris, 1953). A summary of relationships between those material functions which occur in the one-dimensional formulation of the theory is given in this reference and in reference 4.

⁴ H. Leaderman, Trans. Soc. Rheol. 1, 213 (1957).

⁵ E. R. Love, Australian J. Phys. 9, 1 (1956).

⁶ H. König and J. Meixner, Math. Nachr. 19, 265 (1958).

where

$$\dot{\Phi}(s) = (d/ds)\Phi(s), \quad (1.2)$$

and Φ is such that

$$\lim_{s \rightarrow \infty} \Phi(s) = 0. \quad (1.3)$$

Here $\Phi(s)\{ \}$ (for each s) and $\Omega\{ \}$ are linear transformations of the space of symmetric tensors into itself. As a function of time, Φ has a simple physical significance and is called the "stress relaxation function." For if we consider a deformation history such that the material is kept in its natural reference configuration ($E=0$) for all times $t < 0$ and has the strain E^* for all times $t \geq 0$, then for such a history Eq. (1.1) yields

$$T(t) = \begin{cases} 0 & \text{if } t < 0 \\ \Phi(t)\{E^*\} + \Omega\{E^*\} & \text{if } t \geq 0. \end{cases} \quad (1.4)$$

In the familiar special case of an isotropic material, $\Phi(t)$ is completely determined by two scalar-valued functions of time: the stress relaxation functions for shear and dilatation. The linear transformation Ω characterizes the "linear equilibrium stress-strain law" of infinitesimal elasticity theory; i.e., Eqs. (1.3) and (1.4) yield

$$\lim_{t \rightarrow \infty} T(t) = \Omega\{E^*\}. \quad (1.5)$$

For an isotropic solid, Ω is determined by the two Lamé constants.

We refer to the classical linear theory based on Eqs. (1.1)–(1.3) as *infinitesimal viscoelasticity* because, roughly speaking, it can be applicable only to those situations in which the strain is small at all times.

In Sec. 4 we show how Eq. (1.1) must be modified when the reference configuration is arbitrary and not necessarily one in which the equilibrium stress is zero. In particular, in the case of a fluid, $T(t)$ should be replaced by $T(t) + p_r I$, where p_r is the equilibrium hydrostatic pressure corresponding to the reference configuration, and I denotes the unit (or identity) tensor. For a fluid Ω is determined by the equilibrium compressibility.

It is often claimed that the theory of infinitesimal viscoelasticity can be derived from an assumption that on a microscopic level matter can be regarded as composed of "linear viscous elements" (also called

“dashpots”) and “linear elastic elements” (called “springs”) connected together in intricate “networks.”⁷ The motivation behind some of the recent work on spring and dashpot networks appears to be the hope that the consideration of such readily visualized models will suggest a formalism for immersing viscoelasticity in a general thermodynamical theory of irreversible processes.

We feel that the physicist’s confidence in the usefulness of the theory of infinitesimal viscoelasticity does not stem from a belief that the materials to which the theory is applied are really composed of microscopic networks of springs and dashpots, but comes rather from other considerations. First, there is the observation that the theory works for many real materials. But second, and perhaps more important to theoreticians, is the fact that the theory looks plausible because it seems to be a mathematization of little more than certain intuitive prejudices about smoothness in macroscopic phenomena. It is natural to assume that the dependence of the stress on the history of the deformation should be, in some sense, a *smooth dependence*. (Smoothness assumptions are usually so “natural” to physicists that they are seldom made explicit.) Since we know that in small neighborhoods smooth dependences are approximately *linear*, it is felt that if only small deformations are considered, the stress should be given by a linear functional of the deformation history, and that this functional should yield the form exhibited in Eqs. (1.1)–(1.3).

This article tries to make precise these observations about smoothness, and in so doing seeks to obtain a mathematical derivation of infinitesimal viscoelasticity from plausible macroscopic assumptions. To do this one must first presume a nonlinear theory of the mechanical behavior of materials with memory, and, if the undertaking is to be at all worthwhile, the presumed nonlinear theory must rest on constitutive equations based only on very general physical principles. Our development starts with the recently formulated general theory⁸ of “simple materials” (i.e., materials for which the stress depends in an arbitrary way on the history of the first spatial gradient of the displacement). The theory of simple materials is outlined in Sec. 3.

To make precise the notion of smoothness we must introduce a topology into the space of functions characterizing the history of the deformation; i.e., we must have a way of knowing when two histories are close to each other. We do this by defining a *norm*. The particular norm used here is one of those considered in our paper on memory functionals.⁹ This norm has two important properties: first, it makes our space of histories a Hilbert space; second, it places greater

emphasis on the deformations which occurred in the recent past than on those which occurred in the distant past. We believe that this second property is essential if one is to formulate a smoothness assumption for macroscopic phenomena that is compatible with the everyday observation that memories are imperfect. The memory of a macroscopic object for its past deformations fades in the sense that deformations which occurred in the distant past have a smaller effect on the present forces than have more recent deformations.

We mathematize the notion of smoothness by assuming that the constitutive functionals which give the stress in a simple material are *Fréchet differentiable* in our Hilbert space of histories.

In considering finite deformations in simple materials, it is often convenient to take the present configuration as the reference configuration for describing the history of the deformation. Indeed, when dealing with a fluid, this is the natural thing to do, because a fluid has no preferred configurations. However, we can do this even for solids, provided we maintain in the constitutive equations a tensor parameter which tells how the present configuration is related to a preferred configuration.

The function space norm which we use has the property that the norm of a history is small if the deformations have been small at all times in the past; indeed, our derivation of infinitesimal viscoelasticity is a combination of this fact with our differentiability assumption. However, when one takes the present state as a reference, the deformation at the present time is zero, and if one further notes that the distant past is of little importance, it becomes clear that there are several ways in which a history can be small in norm. In particular, any history for which the motion has been *slow* in the recent past has a small norm. This observation has suggested to us the consideration of a new linear approximation for the general constitutive functionals of simple materials. We call the theory based on this new approximation *finite linear viscoelasticity*; it includes the classical infinitesimal theory as a special case, but has the advantage of being meaningful in situations involving finite deformations. The arguments presented in Secs. 3 and 5 show that finite linear viscoelasticity furnishes a complete first-order approximation to the theory of simple materials in the limit in which the history of the deformation, taken relative to the present configuration, is small in norm.

The smoothness considerations presented can be extended to obtain higher order approximations to the general constitutive equations of simple materials. In Sec. 6 we discuss a second-order theory of viscoelasticity for incompressible simple fluids.

2. KINEMATICS

We present a brief outline of the kinematics required for a discussion of simple materials. For a more complete

⁷ D. R. Bland, *The Theory of Viscoelasticity* (Pergamon Press, New York, 1960), Chap. 2.

⁸ W. Noll, *Arch. Ratl. Mech. Anal.* 2, 197 (1958).

⁹ B. D. Coleman and W. Noll, *Arch. Ratl. Mech. Anal.* 6, 356 (1960).

presentation which goes back to first principles, see Noll.⁸

Consider a particular material point X of a body \mathfrak{B} . Suppose that X occupies the position \mathbf{X} in Euclidean space \mathcal{E} when \mathfrak{B} is in a reference configuration. Let ξ be the position of X in \mathcal{E} at time τ . For the dependence of ξ on \mathbf{X} and τ , we write

$$\xi = \chi(\mathbf{X}, \tau). \tag{2.1}$$

The gradient $F(\tau)$ of $\chi(\mathbf{X}, \tau)$ with respect to \mathbf{X} ,

$$F(\tau) = \nabla_{\mathbf{X}} \chi(\mathbf{X}, \tau), \tag{2.2}$$

is called the *deformation gradient* at the material point X at time τ . It is a tensor which possesses an inverse $F(\tau)^{-1}$. (Here the term "tensor" is used a synonym for "linear transformation of the three-dimensional Euclidean vector space into itself.") The value of $F(\tau)$ at each point of \mathfrak{B} is affected not only by the configuration of \mathfrak{B} at time τ but also by our choice of a reference configuration for \mathfrak{B} . This reference configuration may be chosen for convenience and need not necessarily be a configuration actually occupied by the body during its motion.

It is often useful to employ the configuration at the present time t , rather than a fixed configuration, as the reference. The corresponding deformation gradient is denoted by $F_t(\tau)$ and called the *relative deformation gradient*. The deformation gradients enjoy the following important property, which is a direct consequence of the chain rule for the differentiation of composite vector-valued functions:

$$F(\tau) = F_t(\tau)F(t), \tag{2.3}$$

where the indicated multiplication is the usual composition of linear transformations (matrix product).

An immediate consequence of the definition of $F_t(t)$ is that

$$F_t(t) = I, \tag{2.4}$$

where I is the unit (or identity) tensor. From Eq. (2.3) we obtain the relation

$$F_t(\tau) = F(\tau)F(t)^{-1}. \tag{2.5}$$

Let $\rho(\tau)$ give the mass density at X as a function of τ ; it follows from a theorem of kinematics that

$$\det F_t(\tau) = \rho(t)/\rho(\tau). \tag{2.6}$$

If $F(\tau)$ is independent of X , we say that the configuration of \mathfrak{B} at time τ and the reference configuration of \mathfrak{B} are related by a *homogeneous deformation*. If $F = F(\tau)$ is orthogonal, i.e., if

$$F^T F = F F^T = I, \tag{2.7}$$

in which F^T denotes the transpose of F , then this "homogeneous deformation" represents a rigid rotation of the body. If F is symmetric positive-definite, then the body has been subjected to a *pure stretch*; in this case the proper vectors of F give the principal direc-

tions of stretch and the proper numbers of F are the principal stretch ratios.

A theorem of algebra, called the *polar decomposition theorem*, states that any invertible tensor F can be written in two ways as the product of a symmetric positive-definite tensor and an orthogonal tensor:

$$F = RU, \tag{2.8}$$

$$F = VR. \tag{2.9}$$

Furthermore, the orthogonal tensor R and the symmetric positive-definite tensors U and V in these decompositions are uniquely determined by F and obey the following relations:

$$U^2 = F^T F \equiv C, \tag{2.10}$$

$$V^2 = F F^T \equiv B, \tag{2.11}$$

$$U = R^T V R. \tag{2.12}$$

Equations (2.8) and (2.9) have the following significance in kinematics: Any homogeneous deformation with deformation gradient F may be regarded as being the result of a pure stretch U followed by a rigid rotation R , or a rigid rotation R followed by a pure stretch V . These interpretations uniquely determine the pairs R , U and R , V . The rigid rotations entering these two interpretations are the same; however, the pure stretches U and V can be different. It follows from Eq. (2.12) that although these stretches may have different principal directions, they must yield the same stretch ratios. We call the tensor R the *rotation tensor* and the tensors U and V , respectively, the *right and left stretch tensors*. The symmetric positive-definite tensors C and B , defined by Eqs. (2.10) and (2.11), are called, respectively, the *right and left Cauchy-Green tensors*; they obviously contain the same information as the corresponding stretch tensors, and their components are often easier to compute.

The rotation tensor, the stretch tensors, and the Cauchy-Green tensors computed from the relative deformation gradient F_t are denoted by R_t , U_t , V_t , C_t , and B_t . The modifier *relative* is used to indicate that the present configuration (time t) is used as the reference. For example, $C_t(\tau)$, is called the *relative right Cauchy-Green tensor*.

The following formulas are consequences of Eq. (2.4):

$$U_t(t) = V_t(t) = C_t(t) = B_t(t) = R_t(t) = I. \tag{2.13}$$

For simplicity we have emphasized the interpretation for *homogeneous deformations* of the tensors defined by Eqs. (2.8)–(2.11). These definitions obviously apply also to nonhomogeneous deformations, and similar interpretations can be given to them in the nonhomogeneous case if one merely first observes that the deformations considered in continuum mechanics are sufficiently smooth to be approximately homogeneous in small regions of \mathfrak{B} .

We note that there is no unique way to measure "the strain" corresponding to an arbitrary finite deformation.

We now establish the connection between the kinematics of *finite deformations* sketched in the foregoing and the more familiar kinematics of *infinitesimal deformations*.

The *magnitude* $|A|$ of a tensor A is defined by

$$|A|^2 = \text{Tr}(AA^T), \tag{2.14}$$

where Tr denotes the trace of a tensor. If Cartesian coordinates are used, then $|A|^2$ is the sum of the squares of the elements of the 3×3 matrix corresponding to A . We also use the definition (2.14) or magnitude when A is replaced by a linear transformation Γ of the six-dimensional space of symmetric tensors. In this case, the square of the magnitude $|\Gamma|$ of Γ is the sum of the squares of the 6×6 matrix corresponding to Γ .

Let a motion with deformation gradient $F = F(\tau)$ be given. We put

$$H = F - I \tag{2.15}$$

and

$$\epsilon = \sup_{\tau} |H(\tau)|. \tag{2.16}$$

H is the gradient of the displacement vector field. We say that the deformation corresponding to $F(\tau)$ is *infinitesimal at all times* τ if

$$\epsilon \ll 1. \tag{2.17}$$

The *infinitesimal strain tensor* $E = E(\tau)$ is defined by

$$E = \frac{1}{2}(H + H^T). \tag{2.18}$$

In the following we consider functions of τ which are determined by $H(\tau)$ and which have the property that for each τ their magnitude is less than $K\epsilon^n$, where K is a number independent of τ , the function $H(\tau)$, and ϵ . Any such function is denoted by the order symbol $O(\epsilon^n)$; i.e.,

$$|O(\epsilon^n)| < K\epsilon^n. \tag{2.19}$$

It is easy to show that

$$F = I + H = I + O(\epsilon), \tag{2.20}$$

$$F^{-1} = I - H + O(\epsilon^2) = I + O(\epsilon). \tag{2.21}$$

Also, it is not difficult to establish the following relations between the stretch tensors U, V and Cauchy-Green tensors C, B , on the one hand, and the infinitesimal strain tensor E , on the other hand:

$$U - I = E + O(\epsilon^2) = O(\epsilon), \tag{2.22}$$

$$V - I = E + O(\epsilon^2) = O(\epsilon), \tag{2.23}$$

$$C - I = 2E + O(\epsilon^2) = O(\epsilon), \tag{2.24}$$

$$B - I = 2E + O(\epsilon^2) = O(\epsilon). \tag{2.25}$$

Thus, if terms of order $O(\epsilon^2)$ can be neglected, the stretch tensors U, V and Cauchy-Green tensors C, B can be expressed in terms of E . For finite deformations,

however, the infinitesimal strain tensor E is devoid of kinematical significance.

Finally, we note the following relations between the *infinitesimal rotation tensor* W , defined by

$$W = \frac{1}{2}(H - H^T), \tag{2.26}$$

and the finite rotation tensor R :

$$R = I + W + O(\epsilon^2) = I + O(\epsilon),$$

$$R^T = R^{-1} = I - W + O(\epsilon^2) = I + O(\epsilon). \tag{2.27}$$

In order to find an expression for the relative Cauchy-Green tensor $C_t(\tau)$, we first substitute Eqs. (2.20) and (2.21) into Eq. (2.5) and obtain

$$F_t(\tau) = I + H(\tau) - H(t) + O(\epsilon^2). \tag{2.28}$$

Equation (2.10), written for the relative tensors F_t and C_t , reads

$$C_t(\tau) = F_t(\tau)^T F_t(\tau). \tag{2.29}$$

Substitution of Eq. (2.28) into Eq. (2.29) and use of Eq. (2.18) yield

$$C_t(\tau) = I + 2[E(\tau) - E(t)] + O(\epsilon^2) = I + O(\epsilon). \tag{2.30}$$

For finite deformations there is no simple relation between $C_t(\tau), C(\tau)$, and $C(t)$.

3. FADING MEMORY

The theory of simple materials is based on the following physical assumption: *The present stress is given by a functional of the past history of the deformation gradient.*

Suppose the deformation gradient $F(\tau)$ is given (for all $\tau \leq t$) computed relative to a fixed reference configuration. The right Cauchy-Green tensor $C(t)$ and rotation tensor $R(t)$ corresponding to $F(t)$ are determined by Eqs. (2.8) and (2.10). On using Eqs. (2.5) and (2.29) we can compute the relative Cauchy-Green tensors $C_t(\tau)$ for all $\tau \leq t$. We now put

$$\bar{C}_t(\tau) = R^T(t) C_t(\tau) R(t). \tag{3.1}$$

If the material has always been at rest, we have, by Eqs. (2.13) and (3.1),

$$\bar{C}_t(\tau) \equiv I \quad \text{for } \tau \leq t. \tag{3.2}$$

The principle of material objectivity, which states that the properties of a material should appear the same to all observers, can be used to show that the general constitutive equation for simple materials reduces to the form

$$\bar{T}(t) \equiv R^T(t) T(t) R(t) = \mathfrak{F} \left(\bar{C}_t(t-s); C(t) \right), \tag{3.3}$$

where $T(t)$ is the stress tensor at time t and the symbol \mathfrak{F} denotes a functional. [This may be compared with reference 8, Eq. (22.8). Here we use a somewhat more suggestive notation.]

It is useful to put Eq. (3.3) into a slightly different form by writing the right-hand side as the sum of an "equilibrium term" $\mathfrak{H}(C(t))$ and a term which vanishes when the material has always been at rest, i.e., when Eq. (3.2) holds:

$$\bar{T}(t) = \mathfrak{H}(C(t)) + \int_{s=0}^{\infty} (\bar{C}_t(t-s) - I; C(t)), \quad (3.4)$$

$$\int_{s=0}^{\infty} \mathfrak{F}(0; C(t)) = 0. \quad (3.5)$$

For present purposes it is sufficient to regard the constitutive equation (3.4) as the definition of a simple material.

We now add a new physical assumption: *The memory of a simple material fades in time.*

There is no unique way to give this statement a precise meaning. We consider a particular mathematical interpretation of it. For this purpose we first introduce the concept of an influence function which is used to characterize the rate at which the memory fades. (This definition of an influence function is slightly different and somewhat less technical than the one we gave in reference 9.) A function h is called an *influence function* of order $r > 0$ if it satisfies the following conditions:

(a) $h(s)$ is defined for $0 \leq s < \infty$ and has positive real values: $h(s) > 0$.

(b) $h(s)$ decays to zero according to

$$\lim_{s \rightarrow \infty} s^r h(s) = 0 \quad (3.6)$$

monotonically for large s . For example,

$$h(s) = (s+1)^{-p}$$

is an influence function of order r if $r < p$. An exponential

$$h(s) = e^{-\beta s}, \quad \beta > 0$$

is an influence function of any order.

Any function $G(s)$, defined for $s \geq 0$ and with values which are symmetric tensors, is called a *history*. The argument function $G(s) = \bar{C}_t(t-s) - I$ of the functional \mathfrak{F} of Eq. (3.4) is a history. The tensor $C(t)$ in Eq. (3.4) plays the role of a parameter.

Let an influence function $h(s)$ be given. We then define the norm $\|G(s)\|$ of a history $G(s)$ by

$$\|G(s)\|^2 = \int_0^{\infty} |G(s)|^2 h(s)^2 ds, \quad (3.7)$$

where $|G(s)|$ is the magnitude of the tensor $G(s)$ defined by Eq. (2.14). The influence function $h(s)$ determines the influence assigned to the values of $G(s)$ in computing the norm $\|G(s)\|$. Since $h(s) \rightarrow 0$ as $s \rightarrow \infty$, the values of $G(s)$ for small s (recent past) have a greater weight than the values for large s (distant past).

The collection of all histories with finite norm (3.7)

forms a Hilbert space \mathfrak{H} . A history $G(s)$ belongs to the space \mathfrak{H} if it does not grow too fast as $s \rightarrow \infty$.

Consider now an influence function h and a functional

$$\int_{s=0}^{\infty} \mathfrak{F}(G(s))$$

which is defined on a neighborhood of the zero history in the Hilbert space \mathfrak{H} corresponding to h and whose values are symmetric tensors. Assume that the value of \mathfrak{F} for the zero history is zero, i.e., that

$$\int_{s=0}^{\infty} \mathfrak{F}(0) = 0. \quad (3.8)$$

We say that \mathfrak{F} is *Fréchet-differentiable* at the zero history if there is a continuous linear functional $\delta\mathfrak{F}$ such that

$$\int_{s=0}^{\infty} \mathfrak{F}(G(s)) = \delta\mathfrak{F}(G(s)) + \mathfrak{R}(G(s)). \quad (3.9)$$

where the "remainder" \mathfrak{R} is of order $o(\|G(s)\|)$ in the sense that

$$\lim_{\|G(s)\| \rightarrow 0} \|G(s)\|^{-1} \int_{s=0}^{\infty} \mathfrak{R}(G(s)) = 0. \quad (3.10)$$

The linear functional $\delta\mathfrak{F}$ is called the *first variation* or *Fréchet differential* of \mathfrak{F} at the zero history.

We now translate our physical assumption of fading memory into the following mathematical requirement:

(F) *There exists an influence function $h(s)$ of an order $r > \frac{1}{2}$ such that, for each value of the tensor parameter C , the functional \mathfrak{F} of the constitutive equation (3.4) is Fréchet-differentiable at the zero history in the Hilbert space \mathfrak{H} corresponding to $h(s)$.*

If we indicate the dependence on the tensor parameter C , Eq. (3.9) becomes

$$\int_{s=0}^{\infty} \mathfrak{F}(G(s); C) = \delta\mathfrak{F}(G(s); C) + \mathfrak{R}(G(s); C). \quad (3.11)$$

We now invoke the theorem of the theory of Hilbert spaces which states that every continuous linear functional may be written as an inner product. It follows from this theorem that the first variation $\delta\mathfrak{F}$ has an integral representation of the form

$$\delta\mathfrak{F}(G(s); C) = \int_0^{\infty} \mathbf{\Gamma}(s; C) \{G(s)\} ds. \quad (3.12)$$

Here $\mathbf{\Gamma}(s; C) \{ \}$, for each s and each C , is a linear transformation of the space of symmetric tensors into itself with the property that

$$\int_0^{\infty} |\mathbf{\Gamma}(s; C)|^2 h(s)^{-2} ds < \infty, \quad (3.13)$$

where $|\Gamma(s; C)|$ is the magnitude of $\Gamma(s; C)$ as defined by Eq. (2.14). The property (3.13) shows that $\Gamma(s; C)$ must approach zero at a faster rate than the influence function $h(s)$ as $s \rightarrow \infty$. Substitution of Eqs. (3.12) and (3.11) into Eq. (3.4) yields

$$\bar{T} = \mathfrak{H}(C) + \int_0^\infty \Gamma(s; C)\{G(s)\}ds + \overset{\infty}{\mathfrak{H}}(G(s); C), \quad (3.14)$$

where

$$G(s) = \bar{C}_t(t-s) - I. \quad (3.15)$$

It is understood that the variables \bar{T} , C , and $G(s)$ depend on the present time t .

It seems natural to add to the requirement (F) the following two assumptions:

(F') *The Fréchet-differentiability of \mathfrak{F} postulated in (F) is uniform in the tensor parameter C .*

(D) *The tensor function $\mathfrak{H}(C)$ of (3.14) is continuously differentiable.*

By the assumption (F') we mean that the first variation

$$\overset{\infty}{\delta\mathfrak{F}}(G(s); C)$$

depends continuously on C in the strong sense and that the convergence in Eq. (3.10) is uniform in C .

4. INFINITESIMAL VISCOELASTICITY

We first remark that any function of order $O(\epsilon^n)$ in the sense of Eq. (2.19) is also a function of order $O(\epsilon^n)$ with respect to the Hilbert-space norm (3.7); i.e., there is a constant \bar{K} , independent of ϵ , such that

$$\|O(\epsilon^n)\| < \bar{K}\epsilon^n. \quad (4.1)$$

In order to prove this inequality we substitute $O(\epsilon^n)$ for $G(s)$ in the definition (3.7) of the norm and use the inequality (2.19):

$$\|O(\epsilon^n)\|^2 = \int_0^\infty |O(\epsilon^n)|^2 h(s)^2 ds < (K\epsilon^n)^2 \int_0^\infty h(s)^2 ds. \quad (4.2)$$

The requirement (F) of Sec. 3 ensures that the number r of Eq. (3.6) is greater than $\frac{1}{2}$. It follows that the integral $\int_0^\infty h^2(s) ds$ is finite and hence that the inequality (4.1) holds with

$$\bar{K} = K \left(\int_0^\infty h(s)^2 ds \right)^{\frac{1}{2}}. \quad (4.3)$$

This remark shows that the order symbols in Eqs. (2.20)–(2.30) may be interpreted in terms of the convergence in the Hilbert space of histories defined in Sec. 3. This interpretation must be used to justify most of the subsequent considerations.

By combining Eqs. (3.1), (2.30), and (2.27), we find the following expression for the history

$$G(s) = \bar{C}_t(t-s) - I$$

which enters the constitutive equation (3.4) of a simple material:

$$G(s) = 2[E(t-s) - E(t)] + O(\epsilon^2) = O(\epsilon). \quad (4.4)$$

On substituting Eq. (4.4) into Eq. (3.11) and using Eq. (3.10) and the linearity and continuity of the first variation $\delta\mathfrak{F}$, we obtain

$$\overset{\infty}{\delta\mathfrak{F}}(G(s); C) = 2\delta\overset{\infty}{\mathfrak{F}}(E(t-s) - E(t); C) + o(\epsilon). \quad (4.5)$$

where the order symbol $o(\epsilon)$ is used in the sense that

$$\lim_{\epsilon \rightarrow 0} \epsilon^{-1}|o(\epsilon)| = 0. \quad (4.6)$$

It is not difficult to prove that the uniformity assumption (F') of Sec. 3 implies that Eq. (4.5) remains valid if, on the right-hand side, the tensor $C = I + O(\epsilon)$ is replaced by the unit tensor I :

$$\overset{\infty}{\delta\mathfrak{F}}(G(s); C) = 2\delta\overset{\infty}{\mathfrak{F}}(E(t-s) - E(t); I) + o(\epsilon). \quad (4.7)$$

We now substitute the integral representation (3.12), for $C = I$, into Eq. (4.7), and obtain

$$\overset{\infty}{\delta\mathfrak{F}}(G(s); C) = \int_0^\infty 2\Gamma(s)\{E(t-s)\}ds - \int_0^\infty 2\Gamma(s)ds\{E(t)\} + o(\epsilon). \quad (4.8)$$

On defining $\Phi(s)$ by

$$\Phi(s) = -2 \int_s^\infty \Gamma(\sigma) d\sigma, \quad \dot{\Phi}(s) = \frac{d}{ds}\Phi(s) = 2\Gamma(s), \quad (4.9)$$

we may rewrite Eq. (4.8) in the form

$$\overset{\infty}{\delta\mathfrak{F}}(G(s); C) = \Phi(0)\{E(t)\} + \int_0^\infty \dot{\Phi}(s)\{E(t-s)\}ds + o(\epsilon), \quad (4.10)$$

where

$$\lim_{\epsilon \rightarrow 0} \Phi(s) = 0. \quad (4.11)$$

Assumption (D) of Sec. 3 and Eq. (2.24) imply that the equilibrium term $\mathfrak{H}(C)$ of Eq. (3.4) has the form

$$\mathfrak{H}(C(t)) = T_r + \Omega\{E(t)\} + o(\epsilon). \quad (4.12)$$

Here, the linear transformation $\Omega\{ \}$ of the space of symmetric tensors is the gradient of the tensor function $\mathfrak{H}(C)$ at $C = I$. The tensor

$$T_r = \mathfrak{H}(I) \quad (4.13)$$

is the *residual stress*, i.e., the stress the material would sustain if it had been held in the reference configuration at all times in the past.

Substitution of Eqs. (4.10) and (4.12) into the constitutive equation (3.4) yields

$$\bar{T}(t) = T_r + [\Omega + \Phi(0)]\{E(t)\} + \int_0^\infty \dot{\Phi}(s)\{E(t-s)\}ds + o(\epsilon). \quad (4.14)$$

Finally, going back to the definition (3.3) of \bar{T} and using Eqs. (2.27), we obtain the following expression for the stress tensor $T(t)$:

$$T(t) - T_r = W(t)T_r - T_rW(t) + [\Omega + \Phi(0)]\{E(t)\} + \int_0^\infty \dot{\Phi}(s)\{E(t-s)\}ds + o(\epsilon). \quad (4.15)$$

When ϵ , given by Eq. (2.16), is small enough, the remainder term $o(\epsilon)$ can be neglected in comparison with the other terms on the right-hand side of Eq. (4.15), which are of order $O(\epsilon)$. Thus, the constitutive equation of infinitesimal viscoelasticity reads

$$T(t) - T_r = W(t)T_r - T_rW(t) + [\Omega + \Phi(0)]\{E(t)\} + \int_0^\infty \dot{\Phi}(s)\{E(t-s)\}ds. \quad (4.16)$$

When the reference configuration is a natural state, we have $T_r = 0$, and Eq. (4.16) reduces to the classical equation (1.1). Equation (4.16), with $T_r \neq 0$, applies to infinitesimal deformations superposed on a large deformation from an unstressed natural state. In this case, the reference configuration is not the natural state but the deformed state with equilibrium stress T_r . If T_r is a hydrostatic pressure $T_r = -pI$, the terms involving $W(t)$ in Eq. (4.16), cancel. The stress relaxation function $\Phi(s)$ depends not only on the material but also on the configuration which has been taken as the reference.

We remark that the special case $\Phi(s) \equiv 0$ of Eq. (4.16) corresponds to the theory of infinitesimal elastic deformations superposed on large deformations. The special case $\Phi(s) \equiv 0$ and $T_r = 0$ corresponds to the classical theory of infinitesimal elasticity.

5. FINITE LINEAR VISCOELASTICITY

Motivation

Let us return to Eq. (3.14), which, under our hypothesis (F), is equivalent to the fundamental constitutive equation (3.4). It follows from Eq. (3.10) that the remainder term of Eq. (3.14) is small compared to the term involving the integral, provided the history $G(s) = \bar{C}_t(t-s) - I$ has a small Hilbert-space norm. Thus, the equation

$$\bar{T} = \mathfrak{h}(C) + \int_0^\infty \Gamma(s; C)\{G(s)\}ds \quad (5.1)$$

approximates the general constitutive equation of a

simple material in the limit

$$\|G(s)\| \rightarrow 0, \quad (5.2)$$

and the error approaches zero faster than $\|G(s)\|$. We call the theory based on Eq. (5.1) *finite linear viscoelasticity*.

One way of achieving the limit (5.2) is to let ϵ , defined by Eq. (2.16), go to zero. The discussion of Sec. 4 shows that, in this case, Eq. (5.1) reduces to the constitutive equation (4.16) of infinitesimal viscoelasticity.

When we consider, however, the definition (3.7) of the norm $\|G(s)\|$, we see that the limit (5.2) may be achieved even when ϵ does not approach zero. In order for $\|G(s)\|$ to be small, it is not necessary that the deformation (relative to the configuration at the present time t) be small at *all* past times $\tau < t$, but only that the deformation be small in the *recent* past. In particular, $\|G(s)\|$ is small for "slow" motions. To make this remark precise we consider a history $G(s)$ which has finite norm and corresponds to a deformation which makes no jump at the present, so that

$$\lim_{\epsilon \rightarrow 0} G(s) = 0. \quad (5.3)$$

We then construct for each $\alpha, 0 < \alpha \leq 1$, a "retarded" history

$$G_\alpha(s) = G(\alpha s). \quad (5.4)$$

It follows from Eq. (3.21) of reference 9 that

$$\lim_{\alpha \rightarrow 0} \|G_\alpha(s)\| = 0, \quad (5.5)$$

i.e., that the limit (5.2) may be achieved by retardation of a given process.

Aside from the fact that the finite theory based on Eq. (5.1) applies to a much larger class of problems than the infinitesimal theory, there is a fundamental difference between the two theories. The infinitesimal theory is physically meaningless for finite deformations because it does not have the invariance properties required by the principle of material objectivity. The finite linear theory, on the other hand, enjoys the correct invariance. Thus, it is conceivable that there exists *some* material which obeys Eq. (5.1) for *arbitrary* finite deformations. The infinitesimal theory cannot possibly apply to *any* material when finite deformations are considered.

Finally, we remark that in the derivation of Eq. (5.1) no assumption has been made about the magnitude of the tensor parameter C . Hence, the finite theory based on Eq. (5.1) is applicable when the present and the reference configuration are related by an arbitrary large deformation.

Isotropic Materials

When dealing with isotropic materials it is convenient to take the reference configuration to be *undis-*

torted. (A precise definition of this term is given in reference 8.) Then the equilibrium stress is hydrostatic. Furthermore, the results in Sec. 22 of reference 8 show that the constitutive equation (5.1) reduces to

$$T = \mathfrak{h}(B) + \int_0^\infty \Gamma(s; B) \{J(s)\} ds, \tag{5.6}$$

where $B = B(t)$ is the left Cauchy-Green tensor, defined by Eq. (2.11), and the history $J(s)$ is given by

$$J(s) = C_t(t-s) - I. \tag{5.7}$$

Furthermore, the tensor function \mathfrak{h} and the linear functional given by the integral in Eq. (5.6) are *isotropic* in the sense that they obey the identities

$$Q\mathfrak{h}(B)Q^T = \mathfrak{h}(QBQ^T), \tag{5.8}$$

$$Q \int_0^\infty \Gamma(s; B) \{J(s)\} ds Q^T = \int_0^\infty \Gamma(s; QBQ^T) \{QJ(s)Q^T\} ds \tag{5.9}$$

for all orthogonal tensors Q . A fundamental theorem of the theory of isotropic tensor functions (for an elegant recent proof see reference 10, Sec. 59) states that \mathfrak{h} has a representation

$$\mathfrak{h}(B) = h_0 I + h_1 B + h_2 B^2, \tag{5.10}$$

where $h_0, h_1,$ and h_2 are scalar invariants of B . Also, it can be shown that the identity (5.9) implies the following representation for Γ :

$$\begin{aligned} \Gamma(s; B) \{J(s)\} &= \mathfrak{f}_1(s; B) J(s) + J(s) \mathfrak{f}_1(s; B) \\ &+ \text{Tr}[J(s) \mathfrak{f}_2(s; B)] I + \text{Tr}[J(s) \mathfrak{f}_3(s; B)] B \\ &+ \text{Tr}[J(s) \mathfrak{f}_4(s; B)] B^2. \end{aligned} \tag{5.11}$$

Here, for each s , the tensor functions $\mathfrak{f}_i(s; B)$ are isotropic in the sense of Eq. (5.8) and hence have representations of the form (5.10). The proof of this result is too technical to be included here. Equations (5.10) and (5.11) and the representations for the \mathfrak{f}_i may be used to render the constitutive equation (5.6) explicit. The resulting formula shows that, in the finite theory of linear viscoelasticity, the behavior of an isotropic material is determined by 11 independent scalar material functions; three of these depend on three variables and the remaining eight on four variables. The assumption of isotropy alone yields no further simplification. The special case $\Gamma \equiv 0$ of Eq. (5.6) corresponds to the theory of finite (nonlinear) isotropic elasticity.

Fluids

We now consider materials which not only obey a constitutive equation of the form (5.1) but which are also *simple fluids* in the sense of the definition given in

¹⁰ J. Serrin, "Mathematical principles of classical fluid mechanics," in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1959), Vol. VIII/1.

reference 8. (Coleman and Noll¹¹ give a summary of the general theory of simple fluids with emphasis on physical applications.) Such materials are isotropic, and hence Eqs. (5.6)–(5.9) apply. Moreover, the functions $\mathfrak{h}(B)$ and $\Gamma(s; B)$ in Eq. (5.6) depend on B only through the determinant of B or, equivalently, only through the present density $\rho = \rho(t)$. Thus, for a fluid, Eq. (5.6) becomes

$$T = \mathfrak{h}(\rho) + \int_0^\infty \Gamma(s; \rho) \{J(s)\} ds. \tag{5.12}$$

The isotropy identities (5.8) and (5.9) may be written in the form

$$Q\mathfrak{h}(\rho)Q^T = \mathfrak{h}(\rho), \tag{5.13}$$

$$\begin{aligned} \int_0^\infty Q[\Gamma(s; \rho) \{J(s)\}] Q^T \\ - \Gamma(s; \rho) \{QJ(s)Q^T\} ds = 0. \end{aligned} \tag{5.14}$$

Since Eq. (5.13) is valid for all orthogonal tensors Q , it follows that $\mathfrak{h}(\rho)$ must reduce to a scalar multiple of the unit tensor:

$$\mathfrak{h}(\rho) = -p(\rho)I. \tag{5.15}$$

We call $p(\rho)$ the *equilibrium pressure*; it is the pressure the fluid would be supporting if it had remained at rest in its present configuration at all times in the past.

Equation (5.14) is valid for all orthogonal Q and for all possible histories $J(s)$ belonging to the Hilbert space \mathfrak{H} . The only element of a Hilbert space which is orthogonal to all elements of the space is the zero element. This fact implies that the integrand in Eq. (5.14) must be identically zero. Hence, the transformation $\Gamma(s; \rho) \{ \}$ satisfies the identity

$$Q[\Gamma(s; \rho) \{J\}] Q^T = \Gamma(s; \rho) \{QJQ^T\}. \tag{5.16}$$

for all orthogonal tensors Q and all symmetric tensors J . In other words, for each s and ρ , $\Gamma(s; \rho) \{ \}$ is an *isotropic* linear transformation of the space of symmetric tensors. The representation theorem for such isotropic transformations [special case of the theorem embodied in Eq. (5.10) (see reference 10, Sec. 59)] states that $\Gamma(s; \rho) \{J(s)\}$ must be of the form

$$\Gamma(s; \rho) \{J(s)\} = \mu(s; \rho) J(s) + \lambda(s; \rho) (\text{Tr } J(s)) I, \tag{5.17}$$

where $\mu(s; \rho)$ and $\lambda(s; \rho)$ are scalar functions of the time lapse s and the present density ρ . On substituting Eqs. (5.15) and (5.17) into Eq. (5.12), we obtain the following *constitutive equation of a simple fluid in the theory of finite linear viscoelasticity*:

$$\begin{aligned} T = -p(\rho)I + \int_0^\infty \mu(s; \rho) J(s) ds \\ + \left[\int_0^\infty \lambda(s; \rho) \text{Tr } J(s) ds \right] I. \end{aligned} \tag{5.18}$$

¹¹ B. D. Coleman and W. Noll, *Ann. N. Y. Acad. Sci.* **89**, 672 (1961).

In this theory, the mechanical behavior of a fluid is determined by the three scalar material functions $p(\rho)$, $\mu(s; \rho)$, and $\lambda(s; \rho)$.

If the fluid under consideration is *incompressible*, certain modifications must be made in this analysis. In incompressible materials, the motion determines the stress only up to a hydrostatic pressure. In other words, the constitutive equation gives only the *extra stress*

$$T_e = T + pI, \tag{5.19}$$

where p is an indeterminate pressure. In the incompressible case, the two terms in Eq. (5.18) which are scalar multiples of the unit tensor I may be absorbed into the indeterminate pressure term pI . From these remarks we see that in finite linear viscoelasticity the stress in an incompressible fluid is given by the remarkably simple equation

$$T_e = T + pI = \int_0^\infty \mu(s) J(s) ds, \tag{5.20}$$

where, since the density is constant, $\mu(s)$ is a function of time lapse s .

The "relaxation function" $\phi(s)$ determined by geologists from measurements of the decay of shearing motions for simple (infinitesimal) shear in incompressible fluids is related to the material function $\mu(s)$ as follows:

$$\phi(s) = -2 \int_s^\infty \mu(\sigma) d\sigma, \quad \mu(s) = \frac{1}{2} (d/ds) \phi(s). \tag{5.21}$$

Thus, the relaxation function $\phi(s)$ is sufficient to determine the mechanical behavior of incompressible fluids in the theory of finite linear viscoelasticity.

For simple fluids, the property (3.13) is equivalent to the conditions

$$\int_0^\infty |\mu(s; \rho)|^2 h(s)^{-2} ds < \infty, \tag{5.22}$$

$$\int_0^\infty |\lambda(s; \rho)|^2 h(s)^{-2} ds < \infty.$$

These conditions relate the rate of decay of the influence function to the rate of decay of the material functions $\mu(s; \rho)$ and $\lambda(s; \rho)$ as $s \rightarrow \infty$.

6. SECOND-ORDER VISCOELASTICITY

In Sec. 3 we showed, on the basis of our assumption (F), that the (nonlinear) functional \mathfrak{F} giving the stress of a simple material may be approximated by a *linear* functional. The error in this approximation approaches zero faster than the Hilbert-space norm $\|G(s)\|$ of the history (3.15). The analysis of Sec. 3 may be generalized if the assumption (F) is replaced by a stronger assumption

which requires that the functional \mathfrak{F} be not just once but n times Fréchet differentiable at the zero history. It is then possible to approximate \mathfrak{F} by a *polynomial* functional of degree n with an error that approaches zero faster than the n th power of the norm $\|G(s)\|$. For example, when $n=2$, we find the following generalization of Eq. (3.14):

$$\bar{T} = \mathfrak{h}(C) + \int_0^\infty \Gamma(s; C) \{G(s)\} ds + \mathfrak{D}(G(s); C) + \mathfrak{R}'(G(s); C). \tag{6.1}$$

Here, \mathfrak{D} is a continuous *quadratic* functional depending on the tensor parameter C ; the remainder \mathfrak{R}' is of order $o(\|G(s)\|^2)$, i.e.,

$$\lim_{\|G(s)\| \rightarrow 0} \|\mathfrak{R}'(G(s); C)\| = 0. \tag{6.2}$$

Relation (6.1) shows that the equation

$$\bar{T} = \mathfrak{h}(C) + \int_0^\infty \Gamma(s; C) \{G(s)\} ds + \mathfrak{D}(G(s); C) \tag{6.3}$$

approximates the general constitutive equation of a simple material in the limit $\|G(s)\| \rightarrow 0$, and the error approaches zero faster than $\|G(s)\|^2$. We call the theory based on Eq. (6.3) *second-order viscoelasticity*.

The quadratic functional \mathfrak{D} of Eq. (6.3) may be expressed in terms of a bounded symmetric operator on the Hilbert space of histories. It is not possible, in general, to represent \mathfrak{D} by integrals. However, an integral representation does exist if the operator corresponding to \mathfrak{D} is completely continuous. We consider only this special case.

Explicit forms of the constitutive equations for isotropic materials and for simple fluids in second-order viscoelasticity may be obtained by an analysis similar to the one given in Sec. 5 in finite linear viscoelasticity. The resulting formulas are too complicated to be included here in full. Without giving the details of the derivation, we state the *constitutive equation of an incompressible fluid in the second-order theory of viscoelasticity*:

$$T + pI = \int_0^\infty \mu(s) J(s) ds + \int_0^\infty \int_0^\infty [\alpha(s_1, s_2) J(s_1) J(s_2) + \beta(s_1, s_2) \{ \text{Tr } J(s_1) \} J(s_2)] ds_1 ds_2. \tag{6.4}$$

Here, p is an indeterminate pressure, $J(s)$ is the history given by Eq. (5.7), and $\mu(s)$, $\alpha(s_1, s_2)$ and $\beta(s_1, s_2)$ are scalar material functions. The function $\mu(s)$ is the same as in Eq. (5.23). The functions α and β are symmetric,

i.e.,

$$\alpha(s_1, s_2) = \alpha(s_2, s_1), \quad \beta(s_1, s_2) = \beta(s_2, s_1). \quad (6.5)$$

In order to illustrate the behavior predicted by Eq. (6.4), we consider a class of motions called *simple shearing motions*. These motions are defined by the property that the velocity field $v(x) = \{v_x, v_y, v_z\}$, in some Cartesian coordinate system x, y, z , has the components

$$v_x = 0, \quad v_y = v(x, t), \quad v_z = 0. \quad (6.6)$$

It follows from Eqs. (5.6), (5.8), and (5.10) of reference 11 that the matrix function corresponding to the history $J(s)$ defined by Eq. (5.7) has the form

$$[J(s)] = \lambda_t(s) \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} + \lambda_t(s)^2 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad (6.7)$$

where

$$\lambda_t(s) = \int_0^s \frac{d}{dx} v(x, t - \sigma) d\sigma. \quad (6.8)$$

In order to obtain the components T_{xx}, T_{xy} , etc., of the stress tensor T , we substitute Eq. (6.7) into Eq. (6.4). After a simple calculation, we find

$$T_{xy} = \int_0^\infty \mu(s) \lambda_t(s) ds + \int_0^\infty \int_0^\infty \gamma(s_1, s_2) \times \lambda_t(s_1)^2 \lambda_t(s_2) ds_1 ds_2, \quad (6.9)$$

$$T_{xx} - T_{yy} = \int_0^\infty \mu(s) \lambda_t^2(s) ds + \int_0^\infty \int_0^\infty \gamma(s_1, s_2) \times \lambda_t(s_1)^2 \lambda_t(s_2)^2 ds_1 ds_2, \quad (6.10)$$

$$T_{yy} - T_{zz} = \int_0^\infty \int_0^\infty \alpha(s_1, s_2) \lambda_t(s_1) \lambda_t(s_2) ds_1 ds_2, \quad (6.11)$$

where

$$\gamma(s_1, s_2) = \alpha(s_1, s_2) + \beta(s_1, s_2). \quad (6.12)$$

Equations (6.9)–(6.11), together with Cauchy’s equations of motion, lead to a rather complicated system of integro-differential equations.

We now consider the special case when

$$\epsilon = \sup_{s>0} |\lambda_t(s)| \quad (6.13)$$

is small. Physically, this case corresponds to shearing motions with the property that the configuration of the fluid at *all* past times differs from the present configuration only by a small deformation. Shearing vibrations of small amplitude have this property. It is clear from Eqs. (6.13) and (6.7) that the Hilbert space norm $\|J(s)\|$ is of order $O(\epsilon^2)$. But the terms involving double integrals in Eqs. (6.9) and (6.10) are of order

$O(\epsilon^3)$ and $O(\epsilon^4)$, respectively. Therefore, for small ϵ , Eqs. (6.9) and (6.10) reduce to

$$T_{xy} = \int_0^\infty \mu(s) \lambda_t(s) ds, \quad (6.14)$$

$$T_{xx} - T_{yy} = \int_0^\infty \mu(s) \lambda_t^2(s) ds. \quad (6.15)$$

Equation (6.14) for the shearing stress T_{xy} is the same as the corresponding equation in the theory of infinitesimal viscoelasticity. The normal stress differences given by Eqs. (6.15) and (6.11), zero in the infinitesimal theory, do not vanish in the second-order theory. Equations (6.11) and (6.15) may be used, for example, for the interpretation of data on normal stresses obtained in experiments involving shearing vibrations of small amplitude. It is remarkable that the normal stress difference (6.15) depends only on the material function $\mu(s)$ or, equivalently, the shear relaxation function $\phi(s)$ given by Eq. (5.24).

These results on simple shearing motions can easily be generalized to motions that have a form similar to (6.6) in an appropriate curvilinear orthogonal coordinate system. (The method to be employed is analogous to the one used in Sec. 2 of reference 12.)

7. FINAL REMARKS

In our considerations in Secs. 3–6 we have used the relative right Cauchy-Green tensor C_t as a measure of strain. As we remarked at the end of Sec. 2, there is no unique “strain tensor” when finite deformations are considered. Instead of C_t we could also have used the relative right stretch tensor $U_t = (C_t)^{1/2}$, the inverse C_t^{-1} , $\log C_t$, or any other tensor related to C_t by a smooth one-to-one transformation. To different choices of the measure of strain correspond different theories of finite linear viscoelasticity. However, the difference of the stresses computed using two different such theories is of order $o(\|G(s)\|)$. Hence, since any finite linear theory can be expected to be accurate only when terms of order $o(\|G(s)\|)$ can be neglected, we can say that the various theories corresponding to the various measures of strain are equivalent.

To different choices of the measure of strain also correspond different theories of second-order viscoelasticity. These different theories are equivalent in the sense that the corresponding stresses differ only by terms of order $o(\|G(s)\|^2)$.

On the basis of a molecular model for certain incompressible fluids, Lodge¹³ has derived a constitutive equation corresponding to Eq. (5.23) when $J(s)$ is

¹² B. D. Coleman and W. Noll, Arch. Ratl. Mech. Anal. 4, 289 (1959).

¹³ A. S. Lodge, Trans. Faraday Soc. 52, 120 (1956).

nputed using C_t^{-1} rather than C_t as a measure of strain. Our analysis shows that any other molecular model must give the same or an equivalent result, provided only that terms of order $o(\|G(s)\|)$ may be neglected.

As we have remarked in Sec. 5, the norm $\|G(s)\|$ is all in particular for "slow" motions, and hence the finite linear theory applies in this case. For slow flows in simple fluids, the finite linear theory is actually equivalent to the classical theory of Newtonian fluids, provided that the influence function h satisfies the condition (6.3) with $r > \frac{3}{2}$. This fact and analogous results for fluids of higher order are proved in reference 9.

Rivlin and his co-workers¹⁴ in recent years have developed memory theories involving multiple integrals similar to the second-order theory proposed in Sec. 6. The emphasis in their work has been on the representation theorems following from material objectivity and symmetry. In particular, the representations mentioned here in Secs. 5 and 6 can be derived using their results. An investigation of higher order theories of viscoelasticity based on the existence and complete continuity of Fréchet differentials of order > 2 would make much more use of such representation theorems.

¹⁴A. J. M. Spencer and R. S. Rivlin, Arch. Ratl. Mech. Anal. 4, 214 (1960).

Those items in the **Erratum** published in
 REVIEW OF MODERN PHYSICS, Vol. 36, No. 4, 1103, October 1964,
 which have not been inserted in the text above

Page 239. In the second line of the second paragraph, place of "viscosity" read "viscoelasticity."

Page 239. Reference 1 should read:

¹L. Boltzmann, Sitzber. Kaiserlich, Akad. Wiss. Wien), Math.-Naturwiss. Kl. 70, Sect. II, 275-306 (874).

Page 246. The sentence starting seven lines below eq. (5.11) and reading "The resulting formula shows at . . ." should be deleted and replaced by:

The resulting formula shows that in finite linear

viscoelasticity the behavior of an isotropic material is determined by 15 independent scalar-valued material functions.

Pages 247 and 248. The sentence containing Eq. (6.5) should be replaced by:

The function α is uniquely determined if and only if it is chosen to be symmetric, i.e.,

$$\alpha(s_1, s_2) = \alpha(s_2, s_1); \tag{6.5}$$

the function β need not be symmetric.

Motions with Constant Stretch History

WALTER NOLL

Introduction

The class of motions considered in the present paper has recently been introduced by COLEMAN* [1], who noticed that it includes almost all steady flows that have ever been studied in connection with non-classical fluids. Examples are: Simple shearing flow, laminar flow through pipes (circular or not), Couette flow, helical flow, torsional flow, and circular flow between coaxial cones or concentric spheres. All these examples are actually special cases of what is called here steady curvilinear flow (see Section 3). The curvilinear flows belong to a class of motions which have been called viscometric flows (see Section 3 of the present paper or Section 3 of [1]).

The main result of this paper is contained in Section 2, where a representation theorem is proved which characterizes all motions with constant stretch history. It turns out that the nature of such a motion is governed by one constant tensor \mathbf{M} .

Section 3 deals mainly with conditions which insure that \mathbf{M} is nilpotent. In particular, the viscometric flows are shown to correspond to the case when $\mathbf{M}^2 = \mathbf{0}$. For curvilinear flows it is shown how \mathbf{M} can be computed explicitly. Some examples of motions with constant stretch history that do not reduce to viscometric flows are mentioned.

In Section 4 the representation theorem is used to obtain information about the behavior of incompressible simple fluids¹ in motions with constant stretch history. It is shown that this behavior is described by an isotropic tensor function of \mathbf{M} . This description leads, in particular, to a very efficient general treatment of viscometric flows of incompressible simple fluids.

The kinematical results of COLEMAN ([1], Section 6) are corollaries of the Theorem 3 presented here.²

1. Preliminaries

We collect for easy reference some definitions and results from the kinematics of continuous media. For more details we refer to [4] or [2].

Consider a continuous medium in smooth motion. Let ξ be the location at time τ of that material point which, at time t , is located at x . Denote the

* He used the term "substantially stagnant motion".

¹ The concept of a simple fluid was introduced by the author in [2]. For details on the theory of simple fluids see [3] or [1] and the literature quoted there.

² The result derived here shows, moreover, that the condition (β) of COLEMAN's Theorems 1 or 3 may be dropped and the condition (γ) may be weakened by replacing A_3 by A_4 .

dependence of ξ on t , \mathbf{x} , and τ by

$$\xi = \chi_t(\mathbf{x}, \tau). \quad (1.1)$$

The gradient

$$\mathbf{F}_t(\tau) = \text{grad}_{\mathbf{x}} \chi_t(\mathbf{x}, \tau) \quad (1.2)$$

of ξ with respect to \mathbf{x} is called the *relative deformation gradient*. It depends, of course, not only on t and τ but also on the material point. We shall fix our attention on only one particular material point throughout.

In addition to the variable times t and τ , we consider also a fixed time, taken to be 0. It is often convenient to think of t as the variable present time and of τ as a variable past time. The various values of the relative deformation gradient are related by

$$\mathbf{F}_0(\tau) = \mathbf{F}_t(\tau) \mathbf{F}_0(t), \quad \mathbf{F}_t(t) = \mathbf{I}, \quad (1.3)$$

where \mathbf{I} denotes the unit tensor.³

The *relative right Cauchy-Green tensor* is defined by

$$\mathbf{C}_t(\tau) = \mathbf{F}_t^T(\tau) \mathbf{F}_t(\tau), \quad (1.4)$$

where the superscript T denotes transposition. The tensor function $\mathbf{C}_t^s(s)$, $s \geq 0$, defined by

$$\mathbf{C}_t^s(s) = \mathbf{C}_t(t-s) = \mathbf{F}_t^T(t-s) \mathbf{F}_t(t-s) \quad (1.5)$$

is a measure of the deformation history taking t as the present time. The proper numbers of $\mathbf{C}_t^s(s)$ are the squares of the principal stretches of the deformation carrying the present configuration (time t) into the past configuration s time units ago. The proper vectors of $\mathbf{C}_t^s(s)$ determine the directions of these principal stretches.

The following relations follow from (1.3) and (1.4):

$$\mathbf{C}_0(\tau) = \mathbf{F}_0^T(t) \mathbf{C}_t(\tau) \mathbf{F}_0(t), \quad \mathbf{C}_t(t) = \mathbf{C}_t^t(0) = \mathbf{I}. \quad (1.6)$$

Differentiations, denoted by superimposed dots, are assumed not to affect parameters indicated by indices. For example:

$$\dot{\mathbf{F}}_t(\tau) = \frac{d}{d\tau} \mathbf{F}_t(\tau), \quad \dot{\mathbf{F}}_t(t) = \frac{d}{d\tau} \mathbf{F}_t(\tau) \Big|_{\tau=t}. \quad (1.7)$$

The *velocity gradient tensor* $\mathbf{L}(t)$ is given by

$$\mathbf{L}(t) = \dot{\mathbf{F}}_t(t) = \mathbf{D}(t) + \mathbf{W}(t); \quad (1.8)$$

its symmetric part $\mathbf{D}(t) = \frac{1}{2}(\mathbf{L}(t) + \mathbf{L}^T(t))$ is called the *stretching tensor*, and its skew part $\mathbf{W}(t)$ is called the *spin tensor*. The n^{th} Rivlin-Ericksen tensor $\mathbf{A}_n(t)$ is defined by

$$\mathbf{A}_n(t) = \overset{(n)}{\mathbf{C}}_t(t) = (-1)^n \overset{(n)}{\mathbf{C}}_t^t(0), \quad \mathbf{A}_1(t) = 2\mathbf{D}(t). \quad (1.9)$$

If $\mathbf{A}_n(t) = \mathbf{0}$ for all t , then $\mathbf{C}_t^s(s)$ is a polynomial of degree $< n$ in s .

The following results on exponentials of tensors will be needed. The exponential $e^{\mathbf{M}}$ of an arbitrary tensor \mathbf{M} is defined by the convergent power series

$$e^{\mathbf{M}} = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{M}^n. \quad (1.10)$$

³ The term "tensor" is used as a synonym for linear transformation.

The following rules are valid:

$$(e^{\mathbf{M}})^T = e^{\mathbf{M}^T}, \quad (e^{\mathbf{M}})^{-1} = e^{-\mathbf{M}}, \quad e^{\mathbf{0}} = \mathbf{I}, \quad (1.11)$$

$$e^{\mathbf{M}+\mathbf{N}} = e^{\mathbf{M}}e^{\mathbf{N}} \quad \text{if } \mathbf{MN} = \mathbf{NM},^4 \quad (1.12)$$

$$\mathbf{Q}e^{\mathbf{M}}\mathbf{Q}^T = e^{\mathbf{QM}\mathbf{Q}^T} \quad \text{if } \mathbf{Q} \text{ is orthogonal,} \quad (1.13)$$

$$\frac{d}{d\tau} e^{\tau\mathbf{M}} = \mathbf{M}e^{\tau\mathbf{M}} = e^{\tau\mathbf{M}}\mathbf{M}. \quad (1.14)$$

A tensor \mathbf{M} is called *nilpotent* if $\mathbf{M}^n = \mathbf{0}$ for some integer $n \geq 0$. In three dimensions, if \mathbf{M} is nilpotent, then $\mathbf{M}^3 = \mathbf{0}$. Moreover, the matrix $[\mathbf{M}]$ of \mathbf{M} , relative to an appropriate orthonormal basis, has the form

$$[\mathbf{M}] = \begin{vmatrix} 0 & 0 & 0 \\ \varkappa & 0 & 0 \\ \lambda & \nu & 0 \end{vmatrix}. \quad (1.15)$$

If $\mathbf{M}^2 = \mathbf{0}$, then the basis may be chosen such that

$$[\mathbf{M}] = \varkappa \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}. \quad (1.16)$$

The exponential $e^{\tau\mathbf{M}}$ is a polynomial in τ if and only if \mathbf{M} is nilpotent. In this case,

$$e^{\tau\mathbf{M}} = \mathbf{I} + \tau\mathbf{M} + \frac{1}{2}\tau^2\mathbf{M}^2. \quad (1.17)$$

2. The representation theorem

Definition. We say that a motion, along the path of a material point, has *constant stretch history* if there is an orthogonal tensor function $\mathbf{P}(t)$ such that the histories $\mathbf{C}_i^t(s)$ and $\mathbf{C}_0^0(s)$, for $s \geq 0$, are related by

$$\mathbf{C}_i^t(s) = \mathbf{P}^T(t)\mathbf{C}_0^0(s)\mathbf{P}(t), \quad s \geq 0. \quad (2.1)$$

The condition (2.1) means physically that for an observer moving with the material point both the magnitudes of the principal stretches and the changes of direction of the principal axes of strain depend only on the time lapse s and not on the present time t .

Theorem 1. *A motion has constant stretch history if and only if the deformation gradient $\mathbf{F}_0(\tau)$ relative to some fixed time 0 has the representation*

$$\mathbf{F}_0(\tau) = \mathbf{Q}(\tau)e^{\tau\mathbf{M}}, \quad \mathbf{Q}(0) = \mathbf{I}, \quad (2.2)$$

where \mathbf{M} is a constant tensor and $\mathbf{Q}(\tau)$ is an orthogonal tensor function.

Before proving this theorem, we give a list of formulas which follow directly from (2.2) and the results given in the previous section. We use the abbreviations

$$\mathbf{M}_i = \mathbf{Q}(t)\mathbf{M}\mathbf{Q}^T(t), \quad \mathbf{Z}(t) = \dot{\mathbf{Q}}(t)\mathbf{Q}^T(t). \quad (2.3)$$

Clearly, $\mathbf{Z}(t)$ is skew, i.e., $\mathbf{Z}(t)^T = -\mathbf{Z}(t)$.

⁴ Note that $e^{\mathbf{M}+\mathbf{N}} \neq e^{\mathbf{M}}e^{\mathbf{N}}$ if \mathbf{M} and \mathbf{N} do not commute.

(a) Relative deformation gradient $\mathbf{F}_t(\tau)$:

$$\mathbf{F}_t(\tau) = \mathbf{Q}(\tau) e^{(\tau-t)\mathbf{M}} \mathbf{Q}^T(t). \quad (2.4)$$

(b) Relative right Cauchy-Green tensor $\mathbf{C}_t^t(s) = \mathbf{C}_t(t-s)$:

$$\mathbf{C}_t^t(s) = e^{-s\mathbf{M}_t^T} e^{-s\mathbf{M}_t} = \mathbf{Q}(t) e^{-s\mathbf{M}^T} e^{-s\mathbf{M}} \mathbf{Q}^T(t). \quad (2.5)$$

(c) Velocity gradient $\mathbf{L}(t)$:

$$\mathbf{L}(t) = \mathbf{M}_t + \mathbf{Z}(t), \quad (2.6)$$

$$\dot{\mathbf{L}}(t) = \dot{\mathbf{Z}}(t) + \mathbf{Z}(t)\mathbf{L}(t) - \mathbf{L}(t)\mathbf{Z}(t). \quad (2.7)$$

(d) Stretching $\mathbf{D}(t)$ and spin $\mathbf{W}(t)$:

$$\mathbf{D}(t) = \frac{1}{2} (\mathbf{M}_t + \mathbf{M}_t^T), \quad \mathbf{W}(t) = \frac{1}{2} (\mathbf{M}_t - \mathbf{M}_t^T) + \mathbf{Z}(t). \quad (2.8)$$

(e) Rivlin-Ericksen tensors $\mathbf{A}_n(t)$:

$$\mathbf{A}_n(t) = \sum_{k=0}^n \binom{n}{k} (\mathbf{M}_t^k)^T \mathbf{M}_t^{n-k} = \mathbf{Q}(t) \mathbf{A}_n(0) \mathbf{Q}^T(t). \quad (2.9)$$

It is often useful to consider an orthonormal basis $\mathbf{b}_1(t)$, $\mathbf{b}_2(t)$, $\mathbf{b}_3(t)$ which is attached to the material point and which rotates, as the point moves, according to the law

$$\dot{\mathbf{b}}_k(t) = \mathbf{Q}(t) \mathbf{b}_k(0). \quad (2.10)$$

The matrix $[\mathbf{M}]$ of \mathbf{M}_t , relative to this rotating basis, is then independent of t . It follows from (2.5), (2.8)₁, and (2.9), that the matrices of $\mathbf{C}_t^t(s)$, $\mathbf{D}(t)$ and $\mathbf{A}_n(t)$ are also independent of t . If we define the matrix $[\mathbf{F}_t(\tau)]$ of $\mathbf{F}_t(\tau)$ by

$$\|\mathbf{b}_k(\tau) \cdot \mathbf{F}_t(\tau) \mathbf{b}_m(t)\| = [\mathbf{F}_t(\tau)], \quad (2.11)$$

then (2.4) is equivalent to

$$[\mathbf{F}_t(t-s)] = e^{-s[\mathbf{M}]}. \quad (2.12)$$

Proof of Theorem 1. Assume that the motion has constant stretch history. The function $\mathbf{P}(t)$ is not uniquely determined by (2.1), but in view of the assumed smoothness of the motion $\mathbf{P}(t)$ may be chosen to be a smooth function of t which satisfies $\mathbf{P}(0) = \mathbf{I}$. We abbreviate

$$\mathbf{E}(t) = \mathbf{P}(t) \mathbf{E}_0(t), \quad \mathbf{H}(\tau) = \mathbf{C}_0(-\tau) = \mathbf{C}_0^0(\tau). \quad (2.13)$$

Note that $\mathbf{E}(0) = \mathbf{I}$, $\mathbf{H}(0) = \mathbf{I}$. We infer from (2.1), (1.6), and (2.13) that

$$\mathbf{H}(s-t) = \mathbf{C}_0(t-s) = \mathbf{E}^T(t) \mathbf{H}(s) \mathbf{E}(t), \quad s \geq 0. \quad (2.14)$$

We differentiate (2.14) with respect to t and then put $t=0$, obtaining

$$-\dot{\mathbf{H}}(s) = \mathbf{M}^T \mathbf{H}(s) + \mathbf{H}(s) \mathbf{M}, \quad s \geq 0, \quad (2.15)$$

where $\mathbf{M} = \dot{\mathbf{E}}(0)$. Now, (2.15) is a differential equation for $\mathbf{H}(s)$ which must have a unique solution satisfying the initial condition $\mathbf{H}(0) = \mathbf{I}$. This solution is given by

$$\mathbf{H}(s) = e^{-s\mathbf{M}^T} e^{-s\mathbf{M}}, \quad s \geq 0, \quad (2.16)$$

as is easily verified with the help of (1.14). Substituting (2.16) back into (2.14) and putting $s - t = \tau$ yields

$$\mathbf{H}(\tau) = \mathbf{E}^T(t) e^{-t\mathbf{M}^T} e^{-\tau\mathbf{M}^T} e^{-t\mathbf{M}} e^{-\tau\mathbf{M}} \mathbf{E}(t), \quad \tau \geq -t, \quad (2.17)$$

which shows that $\mathbf{H}(\tau)$ depends analytically on τ for $\tau \geq -t$. Since t is arbitrary, it follows that $\mathbf{H}(\tau)$ is an analytic function of τ for all values of τ . From (2.16) and the principle of analytic continuation it follows that

$$\mathbf{H}(-\tau) = e^{\tau\mathbf{M}^T} e^{\tau\mathbf{M}} = \mathbf{C}_0(\tau) = \mathbf{F}_0^T(\tau) \mathbf{F}_0(\tau) \quad (2.18)$$

holds for all τ . (2.18) is equivalent to

$$(\mathbf{F}_0(\tau) e^{-\tau\mathbf{M}})^T (\mathbf{F}_0(\tau) e^{-\tau\mathbf{M}}) = \mathbf{I}, \quad (2.19)$$

which shows that $\mathbf{Q}(\tau) = \mathbf{F}_0(\tau) e^{-\tau\mathbf{M}}$ is orthogonal and hence that (2.2) holds.

If, conversely, (2.2) is valid, then (2.1) follows directly from (2.5), with the choice $\mathbf{P}(t) = \mathbf{Q}(t)^T$. Q.E.D.

Remark. In certain degenerate cases the tensor \mathbf{M} of (2.2) is not uniquely determined by the motion. Consider, for example, the case when $\mathbf{H}(s) = \mathbf{C}_0^0(s) \equiv \mathbf{I}$ for all s , which corresponds to a motion that is locally rigid along the path line under consideration. In this case, $\mathbf{F}_0(\tau)$ is itself orthogonal, and hence (2.2) holds when \mathbf{M} is an arbitrary skew tensor, for $e^{\tau\mathbf{M}}$ is orthogonal when \mathbf{M} is skew.

If, for some value of τ , the proper numbers of $\mathbf{H}(\tau) = \mathbf{C}_0^0(\tau)$ are distinct, then \mathbf{M} is uniquely determined by the motion, as may be seen from the following argument. If

$$\mathbf{H}(\tau) = e^{-\tau\mathbf{M}_1^T} e^{-\tau\mathbf{M}_1} = e^{-\tau\mathbf{M}_2^T} e^{-\tau\mathbf{M}_2},$$

then $\mathbf{H}(\tau)$ satisfies

$$-\dot{\mathbf{H}}(\tau) = \mathbf{M}_i^T \mathbf{H}(\tau) + \mathbf{H}(\tau) \mathbf{M}_i, \quad i = 1, 2. \quad (2.20)$$

Taking the difference of the two equations (2.20) gives

$$\mathbf{A}^T \mathbf{H}(\tau) + \mathbf{H}(\tau) \mathbf{A} = \mathbf{0}, \quad \mathbf{A} = \mathbf{M}_1 - \mathbf{M}_2. \quad (2.21)$$

For $\tau = 0$ we obtain $\mathbf{A}^T + \mathbf{A} = \mathbf{0}$, which means that \mathbf{A} must be skew. Fix the value of τ , and denote the proper numbers of $\mathbf{H}(\tau)$ by h_1, h_2, h_3 . Let $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ be a corresponding orthonormal basis of proper vectors, and let A_{km} be the components of \mathbf{A} relative to the basis \mathbf{f}_k . Equation (2.21) is then equivalent to

$$0 = A_{mk} h_m + h_k A_{km} = A_{km} (h_k - h_m). \quad (2.22)$$

If the proper numbers h_k are distinct, it follows from (2.22) that $A_{km} = 0$ for $k \neq m$. Since \mathbf{A} is skew, we then must have $\mathbf{A} = \mathbf{M}_1 - \mathbf{M}_2 = \mathbf{0}$, i.e., $\mathbf{M}_1 = \mathbf{M}_2$.

3. Special cases

Theorem 2. Consider a motion with constant stretch history. Then the following conditions are equivalent:

- (a) One of the Rivlin-Ericksen tensors vanishes for all t .
- (b) In the representation (2.2) the tensor \mathbf{M} is nilpotent and hence satisfies $\mathbf{M}^3 = \mathbf{0}$.

(c) The history $\mathbf{C}_i^t(s)$ has the form

$$\mathbf{C}_i^t(s) = (\mathbf{I} - s\mathbf{M}_i^T + \frac{1}{2}s^2\mathbf{M}_i^{2T}) (\mathbf{I} - s\mathbf{M}_i + \frac{1}{2}s^2\mathbf{M}_i^2), \quad (3.1)$$

where $\mathbf{M}_i = \mathbf{Q}(t)\mathbf{M}\mathbf{Q}(t)^T$.

(d) The motion is isochoric, and $\mathbf{A}_m(t) \equiv \mathbf{0}$ for $m \geq 5$ and all t .

Proof. $\mathbf{A}_n(t) \equiv \mathbf{0}$ for some n implies that $\mathbf{C}_0^0(s)$ must be a polynomial in s . By (2.5), $\mathbf{C}_0^0(s)$ is the product of the two entire functions $e^{-s\mathbf{M}^T}$ and $e^{-s\mathbf{M}}$. It follows that $e^{-s\mathbf{M}}$ must be a polynomial in s . But this can be the case only when \mathbf{M} is nilpotent. (3.1) follows then from (1.17) and (2.5). It is clear from (2.9) that $\mathbf{A}_m(t) \equiv \mathbf{0}$ for $m \geq 5$. Since \mathbf{M} is nilpotent, its trace is zero. Hence, by (2.8), $\text{tr} \mathbf{D}(t) \equiv \mathbf{0}$, which shows that the motion must be isochoric. Q.E.D.

The following theorem is a corollary of Theorem 2:

Theorem 3. Consider a motion with constant stretch history. Then the following conditions are equivalent:

(a) $\mathbf{A}_4(t) = \mathbf{0}$ for all t .

(b) The tensor \mathbf{M} of the representation (2.2) satisfies $\mathbf{M}^2 = \mathbf{0}$.

(c) The history $\mathbf{C}_i^t(s)$ is of the form

$$\mathbf{C}_i^t(s) = (\mathbf{I} - s\mathbf{M}_i^T) (\mathbf{I} - s\mathbf{M}_i) = \mathbf{I} - s(\mathbf{M}_i + \mathbf{M}_i^T) + s^2\mathbf{M}_i^T\mathbf{M}_i, \quad (3.2)$$

where $\mathbf{M}_i = \mathbf{Q}(t)\mathbf{M}\mathbf{Q}(t)^T$.

(d) The motion is isochoric, and $\mathbf{A}_m(t) \equiv \mathbf{0}$ for $n \geq 3$ and all t .

The motions of the type described by Theorem 3 are called *locally viscometric motions*. A *viscometric flow* is a flow which is locally viscometric at every material point of the flowing medium.

Definition. We call *steady curvilinear flow* a motion whose velocity field $\mathbf{v}(\mathbf{x})$ has the contravariant components

$$v^1 = 0, \quad v^2 = u(x^1), \quad v^3 = w(x^1) \quad (3.3)$$

in an orthogonal curvilinear coordinate system x^k whose metric components g_{kk} are constant along the curves $\xi = \xi(s)$ defined by

$$\xi^1 = x^1, \quad \xi^2 = x^2 + su(x^1), \quad \xi^3 = x^3 + sw(x^1). \quad (3.4)$$

The condition on the g_{kk} is satisfied in particular when the g_{kk} depend only on x^1 or when $w(x^1) \equiv 0$ and the g_{kk} depend only on x^1 and x^3 .

Theorem 4. Every steady curvilinear flow is viscometric. Let $\tilde{\mathbf{e}}_k = \tilde{\mathbf{e}}_k(\mathbf{x})$ be the orthonormal basis of the unit vectors in the direction of the coordinate lines, and let the orthonormal basis $\mathbf{b}_k = \mathbf{b}_k(\mathbf{x})$ be defined by

$$\begin{aligned} \mathbf{b}_1 &= \tilde{\mathbf{e}}_1, \\ \mathbf{b}_2 &= \alpha \tilde{\mathbf{e}}_2 + \beta \tilde{\mathbf{e}}_3, \\ \mathbf{b}_3 &= -\beta \tilde{\mathbf{e}}_2 + \alpha \tilde{\mathbf{e}}_3, \end{aligned} \quad (3.5)$$

where

$$\alpha = \frac{u'}{\kappa} \sqrt{\frac{g_{22}}{g_{11}}}, \quad \beta = \frac{w'}{\kappa} \sqrt{\frac{g_{33}}{g_{11}}}, \quad \kappa = \sqrt{\frac{u'^2 g_{22} + w'^2 g_{33}}{g_{11}}} \quad (3.6)$$

in which u' and w' denote the derivatives of $u(x^1)$ and $w(x^1)$ with respect to x^1 .

The matrix $[\mathbf{M}]$ of the tensor \mathbf{M}_t of (2.5) relative to the basis (3.5) is then given by

$$[\mathbf{M}] = \kappa \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}. \tag{3.7}$$

Proof. The deformation function (1.1) is obtained by solving the differential equations

$$\dot{\xi}^1 = 0, \quad \dot{\xi}^2 = u(\xi^1), \quad \dot{\xi}^3 = w(\xi^1) \tag{3.8}$$

with the initial conditions $\xi^k|_{\tau=t} = x^k$. The solution is given by

$$\xi^1 = x^1, \quad \xi^2 = x^2 + (\tau - t)u(x^1), \quad \xi^3 = x^3 + (\tau - t)w(x^1). \tag{3.9}$$

Let $\mathbf{e}_k = \mathbf{e}_k(\mathbf{x})$ be the natural basis of the coordinate system used, and let $\mathbf{e}^k = \mathbf{e}^k(\mathbf{x})$ be the dual of \mathbf{e}_k (\mathbf{e}^k is characterized by $\mathbf{e}^k \cdot \mathbf{e}_m = \delta_m^k$). By (3.9) and (1.2) we then have

$$\|\mathbf{e}^k(\boldsymbol{\xi}) \cdot \mathbf{F}_t(\tau) \mathbf{e}_m(\mathbf{x})\| = \left\| \frac{\partial \xi^k}{\partial x^m} \right\| = \begin{vmatrix} 1 & 0 & 0 \\ (\tau - t)u' & 1 & 0 \\ (\tau - t)w' & 0 & 1 \end{vmatrix}. \tag{3.10}$$

The unit vectors $\tilde{\mathbf{e}}_k$ in the direction of the coordinate lines are related to the \mathbf{e}_k and \mathbf{e}^k by

$$\mathbf{e}_k = \sqrt{g_{kk}} \tilde{\mathbf{e}}_k, \quad \mathbf{e}^k = \sqrt{g_{kk}^{-1}} \tilde{\mathbf{e}}_k. \tag{3.11}$$

It follows from (3.9) and our hypothesis on the g_{kk} that the g_{kk} are constant along the path lines of the motion and hence that $g_{kk}(\boldsymbol{\xi}) = g_{kk}(\mathbf{x}) = g_{kk}$. Therefore, (3.10) can be rewritten, with use of (3.11), in the form

$$\|\tilde{\mathbf{e}}_k(\boldsymbol{\xi}) \cdot \mathbf{F}_t(t-s) \tilde{\mathbf{e}}_m(\mathbf{x})\| = [\mathbf{I}] - s \begin{vmatrix} 0 & 0 & 0 \\ \sqrt{g_{22} g_{11}^{-1}} u' & 0 & 0 \\ \sqrt{g_{33} g_{11}^{-1}} w' & 0 & 0 \end{vmatrix} \tag{3.12}$$

where $[\mathbf{I}]$ is the unit matrix. Finally, the matrix $[\mathbf{F}_t(t-s)]$ in the sense of (2.11), for the basis defined by (3.5) and (3.6), is easily computed to be

$$[\mathbf{F}_t(t-s)] = [\mathbf{I}] - s\kappa \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}. \tag{3.13}$$

It follows that (2.12) is satisfied when $[\mathbf{M}]$ is given by (3.7). Since $\mathbf{M}^2 = \mathbf{0}$, the motion is a viscometric flow. Q.E.D.

Consider a flow whose velocity field $\mathbf{v}(\mathbf{x})$ in Cartesian coordinates x, y, z has the components

$$u = 0, \quad v = \kappa x, \quad w = \lambda x + \nu y, \tag{3.14}$$

where $\kappa, \lambda,$ and ν are constants $\neq 0$. The components ξ, η, ζ of the deformation function (1.1) corresponding to (3.14) are easily computed:

$$\xi = x, \quad \eta = (\tau - t)\kappa x + y, \quad \zeta = (\tau - t)(\lambda x + \nu y) + \frac{1}{2}(\tau - t)^2 \nu \kappa x + z. \tag{3.15}$$

The derivatives of ξ, η, ζ with respect to x, y, z are the components of the relative deformation gradient $\mathbf{F}_i(\tau)$. Thus, the matrix $[\mathbf{F}_i(\tau)]$ of $\mathbf{F}_i(\tau)$ is given by

$$[\mathbf{F}_i(\tau)] = [\mathbf{I}] + (\tau - t) \begin{vmatrix} 0 & 0 & 0 \\ \kappa & 0 & 0 \\ \lambda & \nu & 0 \end{vmatrix} + \frac{1}{2}(\tau - t)^2 \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \kappa\nu & 0 & 0 \end{vmatrix}, \quad (3.16)$$

which shows that $\mathbf{F}_i(\tau) = e^{(\tau-t)\mathbf{M}}$, where the matrix $[\mathbf{M}]$ of \mathbf{M} is given by (1.15). We see that (3.14) describes a flow with constant stretch history such that $\mathbf{M}^3 = \mathbf{0}$ but $\mathbf{M}^2 \neq \mathbf{0}$. Hence, this flow is of the type described by Theorem 2 at all its material points, but it is not viscometric.

An example of a motion which corresponds to the case when the tensor \mathbf{M} in the representation (2.2) is not nilpotent but symmetric is *steady extension*, which is treated in a separate paper [5].

4. Simple Fluids

Flows with constant stretch history are very important for the theory of incompressible simple fluids. Such fluids obey the constitutive equation

$$\mathbf{T} + p\mathbf{I} = \mathfrak{F} \left(\mathbf{C}_i^t(s) \right) \quad (4.1)$$

where \mathbf{T} is the stress tensor and where p is an indeterminate pressure, which may be normalized so that $p = -\frac{1}{3}\text{tr}\mathbf{T}$. The response functional \mathfrak{F} satisfies the isotropy relation

$$\mathbf{Q}^T \mathfrak{F} \left(\mathbf{H}(s) \right) \mathbf{Q} = \mathfrak{F} \left(\mathbf{Q}^T \mathbf{H}(s) \mathbf{Q} \right) \quad (4.2)$$

for all functions $\mathbf{H}(s)$ in a suitable class and for all orthogonal tensors \mathbf{Q} .

Consider a flow with constant stretch history. The formula (2.5) then applies and, in conjunction with (4.1) and (4.2), gives

$$\mathbf{Q}(t) \mathbf{T}(t) \mathbf{Q}^T(t) + p\mathbf{I} = \mathbf{g}(\mathbf{M}), \quad (4.3)$$

where

$$\mathbf{g}(\mathbf{M}) = \mathfrak{F} \left(e^{-\mathbf{M}^T s} e^{-\mathbf{M} s} \right). \quad (4.4)$$

The values $\mathbf{g}(\mathbf{M})$ are symmetric tensors. The argument \mathbf{M} need not be symmetric, but it must satisfy $\text{tr}\mathbf{M} = 0$ in order that the motion be isochoric. (4.2) implies that $\mathbf{g}(\mathbf{M})$ must be isotropic in the sense that

$$\mathbf{Q}^T \mathbf{g}(\mathbf{M}) \mathbf{Q} = \mathbf{g}(\mathbf{Q}^T \mathbf{M} \mathbf{Q}) \quad (4.5)$$

holds for all \mathbf{M} and all orthogonal \mathbf{Q} . Consider a rotating basis $\mathbf{b}_k(t)$ which satisfies (2.10), and let $[\mathbf{T}] = \|T_{\langle k m \rangle}\|$ be the matrix of the stress tensor $\mathbf{T}(t)$ relative to that basis. Then (4.3) is equivalent to

$$\|T_{\langle k m \rangle}\| = [\mathbf{T}] = -p[\mathbf{I}] + \hat{\mathbf{g}}([\mathbf{M}]), \quad (4.6)$$

where $\hat{\mathbf{g}}$ is the matrix function which corresponds to the tensor function \mathbf{g} relative to any orthonormal basis. It follows that the behavior of an incompressible simple fluid for flows with constant stretch history is completely determined by the matrix function $\hat{\mathbf{g}}([\mathbf{M}])$.

For a viscometric flow, which is characterized by $\mathbf{M}^2 = \mathbf{0}$, we may assume that $[\mathbf{M}]$ has the form (1.16), *i.e.*

$$[\mathbf{M}] = \kappa \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}. \tag{4.7}$$

In this case, the components of $[\mathbf{T}] + \rho[\mathbf{I}]$ are determined by κ . This means that the off-diagonal components of $[\mathbf{T}]$ and the differences of the diagonal components of $[\mathbf{T}]$ are functions of κ . We write

$$T^{(12)} = \tau(\kappa), \quad T^{(11)} - T^{(33)} = \sigma_1(\kappa), \quad T^{(22)} - T^{(33)} = \sigma_2(\kappa). \tag{4.8}$$

The other two off-diagonal terms $T^{(13)}$ and $T^{(23)}$ must be zero, as is shown by the following argument. The orthogonal matrix

$$[\mathbf{Q}] = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix} \tag{4.9}$$

commutes with (4.7). It follows from (4.5) and (4.6) that $[\mathbf{T}]$ must commute with (4.9). It is easily seen that this can be the case only if $T^{(13)} = T^{(23)} = 0$.

Making use of (4.5) with the choice

$$[\mathbf{Q}] = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{4.10}$$

one can readily prove that the shear stress functions $\tau(\kappa)$ must be odd while the normal stress functions $\sigma_1(\kappa)$ and $\sigma_2(\kappa)$ must be even in κ .

The behavior of incompressible simple fluids in steady extension is discussed in [5].

Acknowledgement. This research was supported by the National Science Foundation under Grant NSF-G 16745 to Carnegie Institute of Technology.

References

[1] COLEMAN, B. D.: Kinematical concepts with applications in the mechanics and thermodynamics of incompressible viscoelastic fluids. *Arch. Rational Mech. Anal.* **9**, 273–300 (1961).
 [2] NOLL, W.: A mathematical theory of the mechanical behavior of continuous media. *Arch. Rational Mech. Anal.* **2**, 197–226 (1958).
 [3] COLEMAN, B. D., & W. NOLL: Simple fluids with fading memory. *Symposium on Second Order Effects in Elasticity, Plasticity, and Fluid Dynamics*. Haifa, Israel 1962.
 [4] TRUESDELL, C., & R. A. TOUPIN: *The Classical Field Theories*. In: *Encyclopedia of Physics*, vol. III/1, edited by S. FLÜGGE. Berlin-Göttingen-Heidelberg: Springer 1960.
 [5] COLEMAN, B. D., & W. NOLL: Steady extension of incompressible simple fluids. *Phys. of Fluids* **5**, 840–843 (1962).

The Johns Hopkins University
 Baltimore, Maryland

(Received July 23, 1962)

LA MÉCANIQUE CLASSIQUE, BASÉE SUR UN AXIOME D'OBJECTIVITÉ ⁽¹⁾

par Walter NOLL

1. Introduction.

Il y a plusieurs points de vue sur l'axiomatisation d'une théorie physique. On peut les schématiser en parlant du point de vue du logicien, du point de vue du mathématicien et du point de vue du physicien.

Le logicien entreprend de donner des formalisations complètes en utilisant le symbolisme de la logique formelle. Il est intéressé plus dans la structure formelle d'une théorie que dans son contenu physique. L'accent du logicien est concentré sur les méta-théorèmes et non sur les théorèmes de la théorie.

Le physicien, au contraire, s'il fait l'axiomatique, entreprend d'ordonner les phénomènes d'expérience d'une façon assez logique sans exclure son intuition et sans être trop préoccupé des questions de rigueur. L'accent du physicien est concentré sur le contenu physique d'une théorie et non sur sa structure formelle.

Le mathématicien a une position intermédiaire. Il entreprend d'être rigoureux comme le logicien, mais son accent, comme celui du physicien, est concentré sur les théorèmes d'une théorie et non sur sa forme logique. Le but du mathématicien est de créer et d'analyser des structures complexes. Il emploie la logique et la théorie des ensembles d'un point de vue naïf ; mais, la théorie des ensembles admise, le reste doit être rigoureux. L'objectif de l'axiomatique en physique est d'achever la clarté et la précision des

⁽¹⁾ Ce rapport est basé sur des recherches supportées par la National Science Foundation, Grant NSF-G 5250 to Carnegie Institute of Technology.

concepts. C'est la théorie mathématique qui fournit la précision définitive des concepts de la physique.

Dans ce rapport je me propose de traiter la mécanique classique comme une structure mathématique. L'objectif de ce traitement est de donner des définitions mathématiques précises des concepts généraux de la mécanique classique en extrayant tout ce qui est commun à ses branches diverses : la mécanique des masses ponctuelles, la mécanique des corps rigides, la mécanique des fluides et des milieux élastiques, et la mécanique moderne des matériaux qui ne sont ni fluides ni élastiques dans le sens classique (voir [3] et les références données dans ce papier). Je crois que c'est une telle axiomatisation que Hilbert eut en vue quand il posa le sixième de ses célèbres problèmes en 1900 [1]. Je ne crois pas que la partie du sixième problème de Hilbert qui concerne la mécanique a trouvé une solution satisfaisante.

Dans une autre communication [2] j'ai présenté une tentative d'axiomatisation de la mécanique classique générale. Les axiomes fondamentaux étaient les lois de la balance des actions et des moments et ce que j'ai appelé le principe d'objectivité matérielle, qui demande l'invariance des équations constitutives par changement de repère arbitraire.

Récemment, j'ai observé qu'on peut achever une axiomatisation plus simple et plus naturelle en remplaçant les axiomes de la balance par un autre axiome d'objectivité qui demande l'invariance du travail par changement de repère arbitraire. Cet axiome est compatible avec les réalités physiques si l'on regarde les forces d'inertie comme des forces véritables qui sont les interactions entre les corps dans notre système solaire et la totalité des objets dans le reste de l'univers. Les repères inertiels n'entrent plus dans la partie générale de la nouvelle axiomatisation. La loi d'inertie est regardée comme un postulat constitutif. Les repères privilégiés sont liés explicitement à la position du reste de l'univers.

Dans la suite j'esquisserai les idées essentielles de la nouvelle axiomatisation sans donner tous les détails techniques.

2. Univers matériel.

Soit \mathcal{E} un espace euclidéen à trois dimensions et \mathcal{U} l'espace vectoriel associé. Un *univers matériel* est un ensemble \mathcal{U} muni d'une structure par un ensemble Φ d'applications biunivoques de \mathcal{U}

sur \mathcal{E} et par une mesure positive finie m sur \mathcal{U} . On demande de Φ des propriétés convenables telles qu'on peut définir, à partir de Φ , sur \mathcal{U} une structure de variété continûment différentiable isomorphe à \mathcal{E} .

Les éléments $X \in \mathcal{U}$ sont appelés les *points matériels* de l'univers, les applications $\chi \in \Phi$, les *configurations* de l'univers et la valeur $m(\mathcal{B})$ de la mesure m pour une partie \mathcal{B} de \mathcal{U} la *masse* de la partie \mathcal{B} .

Un *corps* \mathcal{B} est défini comme un sous-ensemble fermé de \mathcal{U} qui est assez régulier et qui n'a pas de sous-ensembles stricts fermés avec la même masse que \mathcal{B} . Deux corps \mathcal{B} et \mathcal{C} sont dits *séparés* si la masse de leur intersection est nulle, $m(\mathcal{B} \cap \mathcal{C}) = 0$. On demande que, pour tout corps \mathcal{B} , il existe un autre corps $\overset{\circ}{\mathcal{B}}$ séparé de \mathcal{B} telle que $m(\mathcal{B} \cup \overset{\circ}{\mathcal{B}}) = m(\mathcal{U})$. Le corps \mathcal{B} est appelé l'*extérieur* de \mathcal{B} .

3. Systèmes de forces.

Un *système de forces* pour un univers matériel \mathcal{U} est défini comme une fonction f à valeurs vectorielles de deux corps séparés dans \mathcal{U} . La valeur $f(\mathcal{B}, \mathcal{C})$ est appelée la *force exercée sur le corps \mathcal{B} par le corps \mathcal{C}* . On demande que f ait les propriétés suivantes :

a) Pour tout corps \mathcal{B} il existe une fonction $f_{\mathcal{B}}$ vectorielle complètement additive définie sur des sous-ensembles de \mathcal{B} et telle que :

$$(1) \quad f_{\mathcal{B}}(\mathcal{C}) = f(\mathcal{C}, \overset{\circ}{\mathcal{B}})$$

si \mathcal{C} est un corps contenu dans \mathcal{B} .

b) Pour tout corps \mathcal{B} il existe une fonction $f_{\overset{\circ}{\mathcal{B}}}$ vectorielle complètement additive définie sur des sous-ensembles de l'extérieur $\overset{\circ}{\mathcal{B}}$ et telle que :

$$(2) \quad f_{\overset{\circ}{\mathcal{B}}}(\mathcal{C}) = f(\mathcal{B}, \mathcal{C})$$

si \mathcal{C} est un corps dans l'extérieur $\overset{\circ}{\mathcal{B}}$ de \mathcal{B} .

En vertu de ces propriétés on peut parler de l'intégrale de Stieltjes $\int_{\mathcal{B}} g df_{\mathcal{B}}$ d'une fonction g continue définie sur le corps \mathcal{B} .

On appelle :

$$(3) \quad f_{\mathcal{B}}(\mathcal{B}) = \int_{\mathcal{B}} df_{\mathcal{B}} = f(\mathcal{B}, \overset{\circ}{\mathcal{B}})$$

la *force résultante* exercée sur \mathcal{B} .

Un système de forces f peut être une combinaison de *forces de contact* et de *forces de distance*. Pour un système de masses ponctuelles, c'est-à-dire si m est une mesure discrète, on ne peut avoir que des forces de distance. Les forces de contact sont essentielles dans les milieux continus.

4. Processus dynamiques.

Étant donné un univers matériel \mathcal{U} , on considère une paire $\{\chi_t, f_t\}$, $-\infty < t < +\infty$, d'une famille de configurations χ_t de \mathcal{U} et d'une famille de systèmes de forces f_t pour \mathcal{U} . Le paramètre réel t de ces familles est appelé le *temps*. On emploie la notation $\chi_t(X) = \chi(X, t)$ et on demande que la *vitesse* :

$$(4) \quad v(X, t) = \frac{d}{dt} \chi(X, t)$$

existe et soit assez régulière. La famille χ_t ou la fonction $\chi(X, t)$ est appelée un *mouvement* de \mathcal{U} . Le point spatial $x = \chi(X, t)$ est la *position* du point matériel X au temps t .

Le *travail* par unité de temps des forces exercées sur un corps \mathcal{B} est défini par

$$(5) \quad w(\mathcal{B}, t) = \int_{\mathcal{B}} v(X, t) \cdot df_{t, \mathcal{B}},$$

où $f_{t, \mathcal{B}}$ est la mesure vectorielle associée à f_t par (1).

On dit que la paire $\{\chi_t, f_t\}$ définit un *processus dynamique* si l'axiome fondamental du § 6 ci-dessous est satisfait. Pour énoncer cet axiome il est d'abord nécessaire d'analyser la notion de changement de repère.

5. Changements de repère.

Un *événement* est une paire (x, t) d'un point x de l'espace \mathcal{E} , la *position* de l'événement, et d'un nombre réel t , le *temps* de l'événement. La totalité des événements est appelée l'*espace-temps*. On définit un *changement de repère* comme un automorphisme $(x, t) \leftrightarrow (x', t')$ de l'espace-temps qui préserve l'intervalle temporel de toute paire d'événements et la distance spatiale de toute paire d'événements simultanés. La géométrie analytique nous apprend que tout changement de repère a une représentation analytique de la forme :

$$(6) \quad \begin{aligned} x' &= c(t) + Q(t)x, \\ t' &= t + a, \end{aligned}$$

où les points $x, x', c(t)$ sont confondus avec leurs vecteurs de position par rapport à une origine arbitraire. $c(t)$ est une fonction ponctuelle de t , $Q(t)$ est une fonction de t à valeurs transformations orthogonales, et a est une constante réelle. On demande que $c(t)$ et $Q(t)$ soient continûment différentiables. Un changement de repère est déterminé par les données $c(t), Q(t), a$.

Si l'on demande que les relations entre les points, les scalaires, les vecteurs et les tenseurs soient préservées par changement de repère on obtient des lois définitives de transformation pour ces quantités. Par exemple, les scalaires ne changent pas par changement de repère et les vecteurs u subissent une transformation de la forme :

$$(7) \quad u' = Q(t)u.$$

Deux processus dynamiques $\{x_i, f_i\}$ et $\{x'_i, f'_i\}$ sont dits *équivalents* s'ils sont liés par un changement de repère de la forme :

$$(8) \quad x'_i = c(t) + Q(t)x_i,$$

$$(9) \quad f'_i = Q(t)f_i,$$

où $t' = t + a$.

L'interprétation physique d'un changement de repère c'est d'un changement de l'observateur seulement. Des processus équivalents ne rendent que des descriptions différentes d'une même réalité physique.

Une fonction définie à partir d'un processus dynamique est dite *objective* si elle transforme en accord avec les lois de changement de repère. Autrement elle est dite *relative*. Les systèmes de forces sont objectifs par définition (9). Par contre, les vitesses et les accélérations sont relatives parce qu'elles ne transforment pas en accord avec la loi (7). Il est aisé à voir que la partie symétrique du tenseur gradient de la vitesse est objective. Ce tenseur est une mesure du changement de déformation.

6. L'axiome fondamental.

En général, le travail (5) n'est pas invariant si l'on transforme x_i et f_i par (8) et (9). Autrement dit, le travail n'est pas objectif pour un pair $\{x_i, f_i\}$ arbitraire. L'axiome fondamental c'est la restriction des considérations de la mécanique aux paires $\{x_i, f_i\}$ pour lesquelles le travail est invariant.

AXIOME FONDAMENTAL. — *Pour un processus dynamique le travail est objectif. C'est-à-dire, pour tout corps \mathcal{B} le travail $w(\mathcal{B}, t)$ défini par (5) est invariant par changement de repère arbitraire (8) (9).*

On peut regarder cet axiome comme la définition d'un processus dynamique.

7. Les lois fondamentales.

C'est une conséquence simple de (8) et (4) que la loi de transformation de la vitesse par changement de repère a la forme (1) :

$$(10) \quad \mathbf{v}' = \mathbf{Q}(\mathbf{v} + \mathbf{v}_0 + \mathbf{W}\mathbf{x})$$

où :

$$(11) \quad \mathbf{v}_0 = \mathbf{Q}^T \dot{\mathbf{c}}, \quad \mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t), \quad \mathbf{W} = \mathbf{Q}^T \dot{\mathbf{Q}}.$$

Ici, \mathbf{Q}^T dénote la transposée de la transformation orthogonale \mathbf{Q} et le point superposé dénote la dérivée par rapport au temps t . La transformation linéaire \mathbf{W} , produit des transformations $\dot{\mathbf{Q}}$ et \mathbf{Q}^T , est antisymétrique.

Si l'on substitue (9) et (10) dans la formule (5) pour le travail, on obtient :

$$w'(\mathcal{B}) = \int_{\mathcal{B}} \mathbf{v}' \cdot d\mathbf{f}'_{\mathcal{B}} = \int_{\mathcal{B}} \mathbf{Q}(\mathbf{v} + \mathbf{v}_0 + \mathbf{W}\mathbf{x}) \cdot \mathbf{Q} d\mathbf{f}_{\mathcal{B}}.$$

L'invariance du produit intérieur de deux vecteurs par transformation orthogonale donne :

$$w'(\mathcal{B}) = \int_{\mathcal{B}} (\mathbf{v} + \mathbf{v}_0 + \mathbf{W}\mathbf{x}) \cdot d\mathbf{f}_{\mathcal{B}},$$

d'où :

$$(12) \quad w'(\mathcal{B}) = w(\mathcal{B}) + \mathbf{v}_0 \cdot \int_{\mathcal{B}} d\mathbf{f}_{\mathcal{B}} + tr \left[\mathbf{W} \left(\int_{\mathcal{B}} \mathbf{x} \otimes d\mathbf{f}_{\mathcal{B}} \right) \right],$$

où tr dénote la trace d'une transformation linéaire et \otimes le produit tensoriel de deux vecteurs (2).

L'axiome fondamental affirme que $w'(\mathcal{B}) = w(\mathcal{B})$ pour tout \mathcal{B} et tout changement de repère. En vertu de (12) il résulte que :

$$\mathbf{v}_0 \cdot \int_{\mathcal{B}} d\mathbf{f}_{\mathcal{B}} + tr \left[\mathbf{W} \left(\int_{\mathcal{B}} \mathbf{x} \otimes d\mathbf{f}_{\mathcal{B}} \right) \right] = 0$$

(1) Dans ce qui suit, la dépendance des variables d'un processus dynamique du temps t et du point matériel \mathbf{X} est entendu.

(2) Le produit tensoriel $\mathbf{a} \otimes \mathbf{b}$ est confondu avec la transformation linéaire définie par $(\mathbf{a} \otimes \mathbf{b})\mathbf{u} = \mathbf{a}(\mathbf{b} \cdot \mathbf{u})$ pour tout vecteur \mathbf{u} .

pour tout vecteur v_0 et toute transformation antisymétrique \mathbf{W} .
Il s'ensuit que :

$$(13) \quad \int_{\mathcal{B}} df_{\mathcal{B}} = f_{\mathcal{B}}(\mathcal{B}) = f(\mathcal{B}, \overset{\circ}{\mathcal{B}}) = 0$$

et :

$$(14) \quad \int_{\mathcal{B}} \mathbf{x} \cdot \wedge df_{\mathcal{B}} = 0,$$

où \wedge dénote le produit extérieur de deux vecteurs.

Les équations (13) et (14) sont le contenu des deux lois fondamentales de la mécanique :

LOI DE LA BALANCE DES ACTIONS. — *La force résultante exercée sur tout corps est nulle.*

LOI DE LA BALANCE DES MOMENTS. — *Le moment résultant des forces exercées sur tout corps est nul.*

Il faut se rappeler que nous n'avons pas dissocié les forces d'inertie des forces ordinaires. On obtient les énoncés conventionnels des lois fondamentales après avoir introduit une telle dissociation.

Les deux lois fondamentales ne sont pas seulement nécessaires, mais aussi suffisantes pour la validité de l'axiome fondamental. Toutes les autres lois générales de la dynamique classique sont des conséquences des deux lois fondamentales. En particulier, on a le théorème de l'action et de la réaction : pour deux corps séparés \mathcal{B} et \mathcal{C} assez réguliers, la force exercée sur \mathcal{B} par \mathcal{C} est l'opposée de la force exercée sur \mathcal{C} par \mathcal{B} :

$$(15) \quad f(\mathcal{B}, \mathcal{C}) = -f(\mathcal{C}, \mathcal{B}).$$

8. Les postulats constitutifs.

La structure mathématique développée jusqu'ici n'est suffisante pour la description d'aucune situation physique particulière. Elle n'introduit qu'une langue générale qui est très commode dans la plupart des situations de la mécanique classique. Pour délimiter la nature particulière d'une situation mécanique il faut introduire des *postulats constitutifs*.

D'abord, on ne peut jamais traiter l'univers total. On concentre l'attention sur une partie \mathcal{S} de l'univers \mathcal{U} et on ne considère que le mouvement de \mathcal{S} et les forces exercées sur les corps \mathcal{B}

dans \mathcal{S} . On appelle *système matériel* une telle partie \mathcal{S} de l'univers matériel.

Un *postulat constitutif de la première espèce* spécifie l'action de l'extérieur $\overset{\circ}{\mathcal{S}}$ du système \mathcal{S} sur les corps \mathcal{B} dans \mathcal{S} . On suppose que les forces $f(\mathcal{B}, \overset{\circ}{\mathcal{S}}) = f_{\mathcal{S}}(\mathcal{B})$ exercées sur les corps \mathcal{B} dans \mathcal{S} par le reste de l'univers $\overset{\circ}{\mathcal{S}}$ dépendent d'une manière simple du mouvement de \mathcal{B} relatif à un repère défini à partir de $\overset{\circ}{\mathcal{S}}$. Ce repère est appelé le *repère privilégié*.

Un exemple de postulat constitutif de la première espèce est le *postulat d'inertie*. Pour tout corps \mathcal{B} dans le système \mathcal{S} , la force exercée sur \mathcal{B} par le reste de l'univers $\overset{\circ}{\mathcal{S}}$ est donnée, relatif au repère privilégié, par :

$$(16) \quad f(\mathcal{B}, \overset{\circ}{\mathcal{S}}) = f_{\mathcal{S}}(\mathcal{B}) = - \frac{d}{dt} \int_{\mathcal{B}} v dm.$$

On adopte ce postulat, en particulier, dans les situations physiques où \mathcal{S} est le système solaire. En ce cas, le repère privilégié est le repère dans lequel les étoiles fixes sont au repos en moyenne.

Des postulats constitutifs de la première espèce plus généraux résultent, si l'on augmente le côté droit de (16) par une fonction prescrite du temps, des points matériels de \mathcal{B} et de leurs positions et vitesses.

Un *postulat constitutif de la deuxième espèce* spécifie l'action mutuelle des corps contenus dans le système matériel \mathcal{S} . La forme générale d'un tel postulat est une relation fonctionnelle entre les restrictions des fonctions $\chi(X, t)$ et $f(\mathcal{B}, \mathcal{C})$ à $X \in \mathcal{S}$ et $\mathcal{B} \subset \mathcal{S}$, $\mathcal{C} \subset \mathcal{S}$. On appelle une telle relation une *équation constitutive*.

Un exemple d'une équation constitutive dans la mécanique des masses ponctuelles est donné par :

$$(17) \quad f_i(X, Y) = h(\chi(X, t), \chi(Y, t))$$

où h est une fonction vectorielle prescrite de deux variables ponctuelles. En ce cas, il n'est pas nécessaire de distinguer les corps à un seul point matériel et les points matériels X, Y de masses positives.

Pour une théorie générale des équations constitutives dans la mécanique des milieux continus voir [3].

9. Le principe de l'objectivité matérielle.

Les postulats constitutifs ne peuvent pas être complètement arbitraires parce qu'on ne change pas une situation physique si l'on ne change que l'observateur.

PRINCIPE DE L'OBJECTIVITÉ MATÉRIELLE. — *Un postulat constitutif doit être objectif, c'est-à-dire invariant par changement de repère.*

Le contenu précis de ce principe est le suivant : si un postulat constitutif est satisfait pour un processus dynamique $\{ \chi_t, f_t \}$ le même postulat constitutif doit être satisfait pour tout processus dynamique $\{ \chi'_t, f'_t \}$ équivalent à $\{ \chi_t, f_t \}$. Du point de vue formel, le principe est une sorte de méta-axiome. Il dissocie les postulats constitutifs admissibles des postulats non admissibles.

Pour les postulats constitutifs de la première espèce, le principe d'objectivité matérielle permet le calcul des forces extérieures dans les repères non privilégiés. Par exemple, le principe affirme qu'on peut obtenir la forme du postulat d'inertie dans un repère arbitraire de la manière suivante : supposant que f' et v' correspondent au repère privilégié on a, d'après (16) :

$$f'_s(\mathcal{B}) = - \frac{d}{dt'} \int_{\mathcal{B}} v' dm.$$

Après y substituer (9) et (10) on obtient la *forme générale du postulat d'inertie* :

$$(18) \quad f'_s(\mathcal{B}) = - \mathbf{Q}^r \frac{d}{dt} \left\{ \mathbf{Q} \int_{\mathcal{B}} (v + v_0 + \mathbf{W}x) dm \right\}.$$

Ici, v_0 est la vitesse de l'origine, \mathbf{Q} est l'orientation et \mathbf{W} est la vitesse angulaire du repère considéré relatif au repère privilégié.

Il est aisé de dériver de (18) la formule classique pour les forces d'inertie dans un repère non inerte. Le principe d'objectivité matérielle est nécessaire pour justifier cette formule. Pour les physiciens, le principe est trop évident pour prendre la peine de l'énoncer.

Les conséquences du principe de l'objectivité matérielle sont moins évidentes pour les postulats constitutifs de la deuxième espèce. Le principe affirme que la forme d'une équation constitutive doit être la même dans tout repère, privilégié ou non. Par

exemple, on montre aisément que l'équation constitutive (17) n'est admissible que si la fonction vectorielle \mathbf{h} a la forme spéciale :

$$(19) \quad \mathbf{h}(\mathbf{x}, \mathbf{y}) = \overrightarrow{\mathbf{x}\mathbf{y}} h(r),$$

où $h(r)$ est une fonction scalaire de la distance r des points \mathbf{x} et \mathbf{y} . Dans le cas particulier $h(r) = Gr^{-3}$ on a l'équation constitutive de la gravitation.

Pour des applications moins triviales du principe de l'objectivité matérielle dans la mécanique des milieux continus voir [3].

RÉFÉRENCES

- [1] D. HILBERT. — *Mathematische Probleme*. Göttinger Nachr., 1900, 253-297.
- [2] W. NOLL. — *The foundations of classical mechanics in the light of recent advances in continuum mechanics. Proceedings of the Berkeley Symposium on the Axiomatic Method*, 1959, 266-281.
- [3] W. NOLL. — A mathematical theory of the mechanical behavior of continuous media. *Arch. Rat. Mech. Anal.*, 1958, 2, 197-226.

The Thermodynamics of Elastic Materials with Heat Conduction and Viscosity

BERNARD D. COLEMAN & WALTER NOLL

1. Introduction

The basic physical concepts of classical continuum mechanics are *body*, *configuration* of a body, and *force system* acting on a body. In a formal rational development of the subject, one first tries to state precisely what mathematical entities represent these physical concepts: a body is regarded to be a smooth manifold whose elements are the material points; a configuration is defined as a mapping of the body into a three-dimensional Euclidean space, and a force system is defined to be a vector-valued function defined for pairs of bodies¹. Once these concepts are made precise one can proceed to the statement of *general principles*, such as the principle of objectivity or the law of balance of linear momentum, and to the statement of specific *constitutive assumptions*, such as the assertion that a force system can be resolved into body forces with a mass density and contact forces with a surface density, or the assertion that the contact forces at a material point depend on certain local properties of the configuration at the point. While the general principles are the same for all work in classical continuum mechanics, the constitutive assumptions vary with the application in mind and serve to define the *material* under consideration. When one has stated the mathematical nature of bodies, configurations and forces, and has laid down the ways in which these concepts occur in the general principles and the constitutive assumptions, then the properties of these concepts are fixed, and one can present rigorous arguments without recourse to “operational definitions” and other metaphysical paraphernalia, which may be of some use in deciding on the applicability of a theory to a specific physical situation but seem to have no place in its mathematical development.

To discuss the *thermodynamics* of continua, it appears that to the concepts of continuum mechanics one must add five new basic concepts: these are *temperature*, specific *internal energy*², specific *entropy*³, *heat flux*, and *heat supply*⁴ (due to radiation). Once mechanics is axiomatized, it is easy to give the mathematical entities representing the thermodynamic concepts: temperature, specific internal energy, specific entropy and heat supply are scalar fields defined over

¹ For more extensive discussions of the foundations of continuum mechanics see references [1]–[4].

² Sometimes called “internal energy density”.

³ Sometimes called “entropy density”.

⁴ Sometimes called “density of absorbed radiation”.

the body, while heat flux is a vector field over the body. We believe that in presenting thermodynamics one should retain all the general principles of mechanics but add to them two new principles: the *first law* of thermodynamics, *i.e.* the law of balance of energy¹, and the *second law*, which for continua takes the form of the Clausius-Duhem inequality². The constitutive assumptions of our present work are the following:

(1) That there exists a caloric equation of state relating the specific internal energy to the "strain"³ and the specific entropy.

(2) That there exists a temperature equation relating the temperature to the strain and specific entropy.

(3) That the stress tensor of mechanics is the sum of two terms, one of which, the "elastic" term, depends on only the strain and the specific entropy, and the other, the "viscous" term, depends on both these variables and, in addition, has a linear dependence on the "rate of strain"⁴.

(4) That the heat flux depends on only the strain, the specific entropy and the spatial gradient of the temperature. (We assume smoothness but not linearity for this dependence.)

We allow the heat supply to be assignable in any way compatible with the general principles, just as body forces are often left assignable in mechanics.

The constitutive assumptions considered here are not the most general imaginable⁵, but they are sufficiently general to cover many applications; in particular, they include the starting assumptions of thermoelasticity theory and the classical theory of viscous fluids with heat conduction.

We feel that a statement of the role of the thermodynamic fields in the general laws of thermodynamics and in specific constitutive assumptions should serve to fix the mathematics of thermodynamics and permit a rational development of the subject without requiring the introduction of non-mathematical concepts, such as Heat Reservoirs, Perpetual Motion Machines, and Reversibly Added Heat⁶.

Here we proceed as follows. We first give a precise meaning to the term *thermodynamic process*; such a process is defined to be a time-dependent set of configurations, force systems, and temperature, internal energy, entropy, heat supply and heat flux fields, compatible with the principles of mechanics and the law of conservation of energy. A thermodynamic process is said to be *admissible* if it is compatible with the constitutive assumptions under consideration. We then demand that the Clausius-Duhem inequality hold for *all*

¹ Cf. §§ 241 and 242 of [4].

² Cf. § 257 of [4].

³ More precisely, the deformation gradient.

⁴ More precisely, the velocity gradient.

⁵ For example, they do not allow for all the long range memory effects covered in the purely mechanical theory of simple materials; cf. [5], [2].

⁶ We do not believe it impossible to assign a mathematical meaning to Heat Reservoirs, Perpetual Motion Machines, and Reversibly Added Heat, but we feel that this has not yet been done, and we doubt its necessity. In several of the extant "axiomatizations" of thermodynamics, reference to these vague concepts appears to disguise the absence of, rather than to furnish, a mathematical justification for the conclusions drawn.

admissible processes. This requirement places restrictions on the constitutive assumptions.

Our main interest here is in the rigorous derivation of a set of necessary and sufficient restrictions on the assumptions (1)–(4) for the Clausius-Duhem inequality to hold for all admissible processes. Each of the restrictions we find is, by itself, a familiar formula or inequality. What is new here is the logical connection.

In a future article we shall apply the general principles of the present framework to more general constitutive assumptions and in so doing obtain a thermodynamic theory compatible with long range memory and non-linear viscoelastic effects, such as stress relaxation.

2. Thermodynamic processes

Consider a body consisting of material points X . A thermodynamic process for this body is described by eight functions of X and the time t , with physical interpretations as follows:

(1) The spatial *position* $\mathbf{x} = \boldsymbol{\chi}(X, t)$; here the function $\boldsymbol{\chi}$, called the *deformation function*, describes a *motion* of the body.

(2) The symmetric *stress tensor* $\mathbf{T} = \mathbf{T}(X, t)$.

(3) The *body force* $\mathbf{b} = \mathbf{b}(X, t)$ per unit mass (exerted on the body by the external world).

(4) The specific *internal energy* $\varepsilon = \varepsilon(X, t)$.

(5) The *heat flux* vector $\mathbf{q} = \mathbf{q}(X, t)$.

(6) The *heat supply* $r = r(X, t)$ per unit mass and unit time (absorbed by the material and furnished by radiation from the external world).

(7) The specific *entropy* $\eta = \eta(X, t)$.

(8) The local *temperature* $\vartheta = \vartheta(X, t)$, which is assumed to be always positive, $\vartheta > 0$.

We say that such a set of eight functions is a *thermodynamic process* if the following two conservation laws¹ are satisfied not only for the body but also for each of its parts \mathcal{B} :

(A) *The law of balance of linear momentum:*

$$\int_{\mathcal{B}} \dot{\mathbf{x}} \, dm = \int_{\mathcal{B}} \mathbf{b} \, dm + \int_{\partial\mathcal{B}} \mathbf{T}\mathbf{n} \, ds. \quad (2.1)$$

(B) *The law of balance of energy*

$$\frac{1}{2} \frac{d}{dt} \int_{\mathcal{B}} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \, dm + \int_{\mathcal{B}} \dot{\varepsilon} \, dm = \int_{\mathcal{B}} (\dot{\mathbf{x}} \cdot \mathbf{b} + r) \, dm + \int_{\partial\mathcal{B}} (\dot{\mathbf{x}} \cdot \mathbf{T}\mathbf{n} - \mathbf{q} \cdot \mathbf{n}) \, ds. \quad (2.2)$$

In (2.1) and (2.2), dm denotes the element of mass in the body, $\partial\mathcal{B}$ the surface of \mathcal{B} , ds the element of surface area in the configuration at time t , and \mathbf{n} the exterior unit normal vector to $\partial\mathcal{B}$ in the configuration at time t ; superimposed dots denote time-derivatives.

¹ A thorough discussion of these conservation laws is given in [4], §§ 196–205, 240, 241.

The assumed symmetry of the stress tensor \mathbf{T} insures that the moment of momentum is automatically balanced. Couple stresses, body couples and other mechanical interactions not included in \mathbf{T} or \mathbf{b} are assumed to be absent.

Under suitable smoothness assumptions the balance equations (2.1) and (2.2) in integral form are equivalent to the following two balance equations in differential form¹:

$$\operatorname{div} \mathbf{T} - \rho \ddot{\mathbf{x}} = -\rho \mathbf{b}, \quad (2.3)$$

$$\operatorname{tr}\{\mathbf{T}\mathbf{L}\} - \operatorname{div} \mathbf{q} - \rho \dot{\varepsilon} = -\rho r; \quad (2.4)$$

here, ρ denotes the mass density and \mathbf{L} the velocity gradient.

In order to define a thermodynamic process it is sufficient to prescribe the six functions χ , \mathbf{T} , ε , \mathbf{q} , η , and ϑ . The two remaining functions \mathbf{b} and r are then uniquely determined by (2.3) and (2.4).

3. The constitutive equations

A *material* is defined by a *constitutive assumption*, which is a restriction on the processes that are admissible in a body consisting of the material. An *elastic material with heat conduction and viscosity* is defined by five *response functions* $\hat{\varepsilon}$, $\hat{\vartheta}$, $\hat{\mathbf{T}}$, \mathbf{I} and $\hat{\mathbf{q}}$. A process is said to be *admissible* in a homogeneous body consisting of such a material if the following constitutive equations hold at each material point X and at all times t :

$$\varepsilon = \hat{\varepsilon}(\mathbf{F}, \eta), \quad (3.1)$$

$$\vartheta = \hat{\vartheta}(\mathbf{F}, \eta), \quad (3.2)$$

$$\mathbf{T} = \hat{\mathbf{T}}(\mathbf{F}, \eta) + \mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}], \quad (3.3)$$

$$\mathbf{q} = \hat{\mathbf{q}}(\mathbf{F}, \eta, \operatorname{grad} \vartheta). \quad (3.4)$$

Here \mathbf{F} denotes the deformation of gradient, at X and t , computed relative to a fixed homogeneous reference configuration, and \mathbf{L} denotes, as before, the velocity gradient; *i.e.*, if we identify the material point X with its position \mathbf{X} in the reference configuration, we have

$$\mathbf{F} = \nabla_{\mathbf{X}} \chi(\mathbf{X}, t) \quad (3.5 a)$$

and

$$\mathbf{L} = \dot{\mathbf{F}}\mathbf{F}^{-1}. \quad (3.5 b)$$

The value $\mathbf{I}(\mathbf{F}, \eta)$ of the response function \mathbf{I} is a linear transformation over the nine-dimensional space of tensors, and the square brackets in (3.3) indicate that this transformation operates on the tensor \mathbf{L} .² The response functions $\hat{\varepsilon}$, $\hat{\vartheta}$, $\hat{\mathbf{T}}$, \mathbf{I} and $\hat{\mathbf{q}}$ depend on the choice of the reference configuration.

¹ See the sections of [4] cited above.

² The assumed linear dependence of the "viscous stress" $\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}]$ on \mathbf{L} is not essential to our present arguments; for example, for our derivation of (5.5) it suffices that $\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}] \rightarrow 0$ as $\mathbf{L} \rightarrow 0$. Our work on materials with fading memory [6] suggests, however, that if the extra-stress depends non-linearly on the velocity gradient then it should depend also on acceleration gradients. Since we here do not allow for an effect of acceleration gradients on the viscous stress, we prefer to let $\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}]$ be linear in \mathbf{L} and thus stay within the theory of "linearly viscous materials", a theory which gives, in the sense of [6], a complete first-order approximation to the non-linear theory of materials with fading memory.

To specify an admissible process, it is sufficient to prescribe the deformation function $\boldsymbol{\chi}$ and the entropy distribution η as functions of time. The deformation gradient \mathbf{F} and the velocity gradient \mathbf{L} can be computed from $\boldsymbol{\chi}$, and then ε , ϑ , and \mathbf{T} can be determined from (3.1)–(3.3). With the temperature distribution ϑ thus obtained, the heat flux \mathbf{q} can be calculated from (3.4). Finally \mathbf{b} and \mathbf{r} can be chosen so that (2.3) and (2.4) hold¹. In Sections 4 and 5 we shall make frequent use of this observation that $\boldsymbol{\chi}$ and η can be chosen *independently* and that to every such choice corresponds a unique admissible thermodynamic process.

The constitutive equations (3.1)–(3.4) must obey the *principle of material objectivity*², which states that an admissible process must remain admissible after a change of frame or observer. Such a change of frame is defined by a time-dependent orthogonal tensor \mathbf{Q} . The scalars ε , η , and ϑ remain unaffected by a change of frame, but \mathbf{F} , \mathbf{T} , \mathbf{L} , and $\text{grad } \vartheta$ transform as follows:

$$\begin{aligned} \mathbf{F} &\rightarrow \mathbf{Q}\mathbf{F}, \\ \mathbf{T} &\rightarrow \mathbf{Q}\mathbf{T}\mathbf{Q}^T, \\ \mathbf{L} &\rightarrow \mathbf{Q}\mathbf{L}\mathbf{Q}^T + \dot{\mathbf{Q}}\mathbf{Q}^T, \\ \mathbf{q} &\rightarrow \mathbf{Q}\mathbf{q}, \\ \text{grad } \vartheta &\rightarrow \mathbf{Q} \text{ grad } \vartheta. \end{aligned} \tag{3.6}$$

Using the methods of reference [7], it is easy to show that the principle of material objectivity is satisfied if and only if the response functions obey the following identities:

$$\begin{aligned} \hat{\varepsilon}(\mathbf{F}, \eta) &= \hat{\varepsilon}(\mathbf{Q}\mathbf{F}, \eta), \\ \hat{\vartheta}(\mathbf{F}, \eta) &= \hat{\vartheta}(\mathbf{Q}\mathbf{F}, \eta), \\ \mathbf{Q}\hat{\mathbf{T}}(\mathbf{F}, \eta)\mathbf{Q}^T &= \hat{\mathbf{T}}(\mathbf{Q}\mathbf{F}, \eta), \\ \mathbf{Q}\mathfrak{I}(\mathbf{F}, \eta)[\mathbf{L}]\mathbf{Q}^T &= \mathfrak{I}(\mathbf{Q}\mathbf{F}, \eta)[\mathbf{Q}\mathbf{L}\mathbf{Q}^T + \mathbf{W}], \\ \mathbf{Q}\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}) &= \hat{\mathbf{q}}(\mathbf{Q}\mathbf{F}, \eta, \mathbf{Q}\mathbf{v}) \end{aligned} \tag{3.7}$$

for all scalars η , all vectors \mathbf{v} , all tensors \mathbf{F} and \mathbf{L} , all orthogonal tensors \mathbf{Q} , and all skew tensors \mathbf{W} .

¹ In many treatments of continuum thermo-mechanics \mathbf{b} and \mathbf{r} are regarded as assigned *a priori*. We do not follow this practice here. Nor are we disturbed by the fact that it might be difficult to control \mathbf{b} and \mathbf{r} experimentally. Physical considerations suggest, however, that for each set of assigned values of \mathbf{F} , \mathbf{L} , and η the heat supply \mathbf{r} should have a negative lower bound $r_0 < 0$ whose magnitude $|r_0|$ is the radiation the body would emit into an environment at zero absolute temperature.

To account explicitly for this lower bound it would suffice to restrict the admissible fields $\boldsymbol{\chi}$ and η to those which give rise to values of \mathbf{T} , \mathbf{L} , $\text{div } \mathbf{q}$, ϱ , and $\dot{\varepsilon}$ such that the left side of (2.4) is greater than ϱr_0 . Our arguments would not be affected much by this restriction.

² See [1]–[3].

The identities (3.7) can be used to derive reduced forms for the constitutive equations (3.1)–(3.4). One set of such reduced constitutive equations is

$$\begin{aligned}\varepsilon &= \hat{\varepsilon}(\mathbf{U}, \eta), \\ \vartheta &= \hat{\vartheta}(\mathbf{U}, \eta), \\ \mathbf{R}^T \mathbf{T} \mathbf{R} &= \hat{\mathbf{T}}(\mathbf{U}, \eta) + \mathbf{I}(\mathbf{U}, \eta) [\mathbf{R}^T \mathbf{D} \mathbf{R}], \\ \mathbf{q} &= \mathbf{R} \hat{\mathbf{q}}(\mathbf{U}, \eta, \mathbf{R}^T \text{grad } \vartheta).\end{aligned}\tag{3.8}$$

Here \mathbf{U} is the (symmetric and positive-definite) right stretch tensor and \mathbf{R} the (orthogonal) rotation tensor, determined by the polar decomposition $\mathbf{F} = \mathbf{R} \mathbf{U}$, and $\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T)$ is the stretching tensor.

An alternative set of reduced constitutive equations is

$$\begin{aligned}\varepsilon &= \hat{\varepsilon}(\mathbf{C}, \eta), \\ \vartheta &= \hat{\vartheta}(\mathbf{C}, \eta), \\ \mathbf{F}^T \mathbf{T} \mathbf{F} &= \tilde{\mathbf{T}}(\mathbf{C}, \eta) + \tilde{\mathbf{I}}(\mathbf{C}, \eta) [\mathbf{F}^T \mathbf{D} \mathbf{F}], \\ \mathbf{q} &= \mathbf{F} \hat{\mathbf{q}}(\mathbf{C}, \eta, \mathbf{F}^T \text{grad } \vartheta),\end{aligned}\tag{3.9}$$

where $\mathbf{C} = \mathbf{U}^2 = \mathbf{F}^T \mathbf{F}$ is the right Cauchy-Green tensor. Noting that

$$\mathbf{F}^T \text{grad } \vartheta = \nabla_{\mathbf{X}} \vartheta,\tag{3.10}$$

where $\nabla_{\mathbf{X}} \vartheta$ is the gradient of the temperature with respect to the position \mathbf{X} in the reference configuration, we find that the last equation of (3.9) has the form¹

$$\mathbf{q} = \mathbf{F} \hat{\mathbf{q}}(\mathbf{C}, \eta, \nabla_{\mathbf{X}} \vartheta).\tag{3.11}$$

In rough terms, the *isotropy group* of a material is the set of density-preserving changes of reference configuration which leave the response of the material unaltered. A more formal definition is the following². The isotropy group \mathcal{I} of an elastic material with heat conduction and viscosity is the set of all unimodular tensors \mathbf{H} for which the following identities hold:

$$\begin{aligned}\hat{\varepsilon}(\mathbf{F}, \eta) &= \hat{\varepsilon}(\mathbf{F} \mathbf{H}, \eta), \\ \hat{\vartheta}(\mathbf{F}, \eta) &= \hat{\vartheta}(\mathbf{F} \mathbf{H}, \eta), \\ \hat{\mathbf{T}}(\mathbf{F}, \eta) &= \hat{\mathbf{T}}(\mathbf{F} \mathbf{H}, \eta), \\ \mathbf{I}(\mathbf{F}, \eta) [\mathbf{L}] &= \mathbf{I}(\mathbf{F} \mathbf{H}, \eta) [\mathbf{L}], \\ \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}) &= \hat{\mathbf{q}}(\mathbf{F} \mathbf{H}, \eta, \mathbf{v}),\end{aligned}\tag{3.12}$$

for all scalars η , all vectors \mathbf{v} and all tensors \mathbf{F} and \mathbf{L} . It is easy to show that the tensors \mathbf{H} in \mathcal{I} form a group. Of course, the group \mathcal{I} depends not only on the material under consideration but also on the choice of reference configuration. It can be shown, however, that the isotropy groups corresponding

¹ This is the form derived by PIPKIN & RIVLIN ([8], equation (17.2)), and by GREEN & ADKINS ([9], equation (8.5.17)).

² This definition is analogous to that given in [2] and discussed in detail for elastic materials (without heat conduction or viscosity) in [10].

to two different reference configurations of the same material are conjugate and hence isomorphic.

The unit tensor \mathbf{I} always belongs to \mathcal{I} . Suppose now that the inversion $-\mathbf{I}$ is in \mathcal{I} . (We note that if $-\mathbf{I}$ is in the isotropy group corresponding to one reference configuration of a material, then $-\mathbf{I}$ is in the isotropy groups corresponding to all other reference configurations of the same material.) In this case (3.12)₅ for $\mathbf{H} = -\mathbf{I}$ yields the identity

$$\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}) = \hat{\mathbf{q}}(-\mathbf{F}, \eta, \mathbf{v}). \tag{3.13}$$

On the other hand, (3.7)₅ with $\mathbf{Q} = -\mathbf{I}$ and \mathbf{v} replaced by $-\mathbf{v}$, gives

$$-\hat{\mathbf{q}}(\mathbf{F}, \eta, -\mathbf{v}) = \hat{\mathbf{q}}(-\mathbf{F}, \eta, \mathbf{v}). \tag{3.14}$$

Combining (3.13) and (3.14), we obtain the identity

$$\hat{\mathbf{q}}(\mathbf{F}, \eta, -\mathbf{v}) = -\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}). \tag{3.15}$$

In words: if $-\mathbf{I}$ is in \mathcal{I} , then, the heat flux must be given by an *odd* function of the temperature gradient. If we put $\mathbf{v} = \mathbf{0}$ in (3.15), it follows that

$$\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{0}) = \mathbf{0}. \tag{3.16}$$

Of course, (3.16) holds whenever $\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v})$ is linear in \mathbf{v} , *i.e.*, when Fourier's law is assumed. We have shown here that even if $\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v})$ is not linear in \mathbf{v} , the presence of the inversion $-\mathbf{I}$ in the isotropy group implies that \mathbf{q} vanishes when $\text{grad } \vartheta$ vanishes¹.

PIPKIN & RIVLIN [β], § 18, have referred to the equation (3.16) as expressing the "non-existence of a piezo-caloric effect".

4. The Clausius-Duhem inequality

We regard \mathbf{q}/ϑ to be a vectorial *flux of entropy* and r/ϑ to be a scalar *supply of entropy*. In other words, for each process we define the *production of entropy* in the body \mathcal{B} to be

$$\Gamma = \frac{d}{dt} \int_{\mathcal{B}} \eta \, dm - \int_{\mathcal{B}} \frac{1}{\vartheta} r \, dm + \int_{\partial \mathcal{B}} \frac{1}{\vartheta} \mathbf{q} \cdot \mathbf{n} \, ds. \tag{4.1}$$

Under suitable smoothness assumptions one can write

$$\Gamma = \int_{\mathcal{B}} \gamma \, dm, \tag{4.2}$$

where

$$\begin{aligned} \gamma &= \dot{\eta} - \frac{r}{\vartheta} + \varrho^{-1} \text{div}(\mathbf{q}/\vartheta) \\ &= \dot{\eta} - \frac{r}{\vartheta} + \frac{1}{\varrho \vartheta} \text{div} \mathbf{q} - \frac{1}{\varrho \vartheta^2} \mathbf{q} \cdot \text{grad } \vartheta \end{aligned} \tag{4.3}$$

is the *specific production of entropy*.

¹ There are real materials, such as crystals in the pedial class, whose material symmetry does not imply that $-\mathbf{I}$ be in \mathcal{I} .

We now lay down the following

Postulate. For every process admissible in a body consisting of a given material and for every part \mathcal{B} of this body the inequality

$$\Gamma \geq 0 \quad (4.4)$$

is valid.

The inequality (4.4) is called the *Clausius-Duhem inequality*, and the postulate above is a statement of the second law of thermodynamics within the framework presented here. The main purpose of the present paper is to derive logical consequences from the postulate.

In order that the inequality (4.4) hold for *all* parts \mathcal{B} of the body, it is necessary and sufficient that the specific entropy production (4.3) be non-negative at all material points X and at all times t . Using (2.4) and $\vartheta > 0$, we see that this condition is equivalent to

$$\rho \vartheta \gamma = \rho (\vartheta \dot{\eta} - \dot{\varepsilon}) + \text{tr}\{\mathbf{T}\mathbf{L}\} - \frac{1}{\vartheta} \mathbf{q} \cdot \text{grad } \vartheta \geq 0. \quad (4.5)$$

It follows from (3.1) and the chain rule for the differentiation of composite functions that

$$\dot{\varepsilon} = \text{tr}\{\hat{\varepsilon}_{\mathbf{F}}(\mathbf{F}, \eta) \dot{\mathbf{F}}\} + \hat{\varepsilon}_{\eta}(\mathbf{F}, \eta) \dot{\eta}, \quad (4.6)$$

where $\hat{\varepsilon}_{\mathbf{F}}$ is the gradient¹ of ε with respect to the tensor \mathbf{F} , and $\hat{\varepsilon}_{\eta}$ the derivative of ε with respect to the scalar η . Substituting (4.6) into (4.5) and using the relation $\dot{\mathbf{F}} = \mathbf{L}\mathbf{F}$, we obtain

$$\rho [\vartheta - \hat{\varepsilon}_{\eta}(\mathbf{F}, \eta)] \dot{\eta} + \text{tr}\{[\mathbf{T} - \rho \mathbf{F} \hat{\varepsilon}_{\mathbf{F}}(\mathbf{F}, \eta)] \mathbf{L}\} - \frac{1}{\vartheta} \mathbf{q} \cdot \text{grad } \vartheta \geq 0. \quad (4.7)$$

5. Temperature relation, stress relation, dissipation inequality

Our postulate states that the inequality (4.7) must hold for *every* admissible process. We consider now processes that are defined by *homogeneous deformations* and *homogeneous entropy distributions*. For such processes, \mathbf{F} and η are independent of the material point and depend on only the time. By (3.2), the temperature distribution ϑ is also homogeneous, and hence $\text{grad } \vartheta = \mathbf{0}$. The inequality (4.7) then becomes, after substitution of (3.2) and (3.3),

$$\rho [\hat{\vartheta}(\mathbf{F}, \eta) - \hat{\varepsilon}_{\eta}(\mathbf{F}, \eta)] \dot{\eta} + \text{tr}\{[\hat{\mathbf{T}}(\mathbf{F}, \eta) - \rho \mathbf{F} \hat{\varepsilon}_{\mathbf{F}}(\mathbf{F}, \eta)] \mathbf{L}\} + \text{tr}\{\mathbf{I}(\mathbf{F}, \eta) [\mathbf{L}] \mathbf{L}\} \geq 0. \quad (5.1)$$

As functions of time, \mathbf{F} and η can be chosen arbitrarily. It follows that for a *particular* time, the values of \mathbf{F} , η , $\mathbf{L} = \dot{\mathbf{F}}\mathbf{F}^{-1}$, and $\dot{\eta}$ can be assigned arbitrarily. Taking $\mathbf{L} = \mathbf{0}$ in (5.1) shows that $[\hat{\vartheta}(\mathbf{F}, \eta) - \hat{\varepsilon}_{\eta}(\mathbf{F}, \eta)] \dot{\eta} \geq 0$ must hold for *every* value of $\dot{\eta}$; this can be the case only if the *temperature relation*

$$\vartheta = \hat{\vartheta}(\mathbf{F}, \eta) = \hat{\varepsilon}_{\eta}(\mathbf{F}, \eta) \quad (5.2)$$

holds.

Choosing $\dot{\eta} = 0$ and substituting $\alpha \mathbf{L}$ for \mathbf{L} in (5.1), where α is an arbitrary number, we obtain

$$\alpha \text{tr}\{[\hat{\mathbf{T}}(\mathbf{F}, \eta) - \rho \mathbf{F} \hat{\varepsilon}_{\mathbf{F}}(\mathbf{F}, \eta)] \mathbf{L}\} + \alpha^2 \text{tr}\{\mathbf{L} \mathbf{I}(\mathbf{F}, \eta) [\mathbf{L}]\} \geq 0. \quad (5.3)$$

¹ The definition of the gradient used here is that given in reference [7], equations (C. 11) and (C. 11c). In components: $[\hat{\varepsilon}_{\mathbf{F}}(\mathbf{F}, \eta)]_{ij} = \frac{\partial \varepsilon}{\partial F_{ji}}$.

This polynomial inequality in α must be satisfied for all values of α when \mathbf{F} , η , and \mathbf{L} are kept fixed. It is clear that this is the case if and only if the coefficient of α vanishes and the coefficient of α^2 is non-negative. Thus we obtain the *stress relation*

$$\hat{\mathbf{T}}(\mathbf{F}, \eta) = \rho \mathbf{F} \hat{\boldsymbol{\varepsilon}}_{\mathbf{F}}(\mathbf{F}, \eta) \tag{5.4}$$

and the *dissipation inequality*

$$\text{tr}\{\mathbf{L}\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}]\} \geq 0 \tag{5.5}$$

which must hold for all tensors \mathbf{L} .

Let $\mathbf{F} = \mathbf{R}\mathbf{U}$, where \mathbf{R} is the rotation tensor and \mathbf{U} the right stretch tensor. Writing (3.7)₄ with the choice $\mathbf{Q} = \mathbf{R}^T$ and $\mathbf{W} = \frac{1}{2}\mathbf{R}^T(\mathbf{L}^T - \mathbf{L})\mathbf{R}$, multiplying the result by $\mathbf{R}^T\mathbf{L}\mathbf{R}$, and then taking the trace, we obtain

$$\text{tr}\{\mathbf{L}\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}]\} = \text{tr}\{\mathbf{S}\mathbf{I}(\mathbf{U}, \eta)[\mathbf{S}]\}, \tag{5.6}$$

where $\mathbf{S} = \frac{1}{2}\mathbf{R}^T(\mathbf{L} + \mathbf{L}^T)\mathbf{R}$. It is evident that \mathbf{L} can always be chosen so that \mathbf{S} is an arbitrarily prescribed symmetric tensor. Therefore, by (5.6), the dissipation inequality (5.5) is equivalent to the requirement that

$$\text{tr}\{\mathbf{S}\mathbf{I}(\mathbf{U}, \eta)[\mathbf{S}]\} \geq 0 \tag{5.7}$$

hold for all symmetric tensors \mathbf{S} .

6. Heat conduction inequality. Summary

Let us assume now that the temperature relation (5.2) can be solved for the entropy, so that (5.2) is equivalent to

$$\eta = \bar{\eta}(\mathbf{F}, \vartheta). \tag{6.1}$$

The invertibility of (5.2) is assured, in particular, when for each fixed \mathbf{F} , ε is a convex function of η .

Continuing to keep the homogeneous reference configuration fixed, let us consider deformations $\boldsymbol{\chi}$ that are both homogeneous and constant in time. Each such deformation is characterised by a unique value of \mathbf{F} , and, of course, we have $\mathbf{L} = \mathbf{0}$. Let us also consider time-independent temperature distributions $\vartheta = \vartheta(X)$. By (6.1), for each pair $\vartheta = \vartheta(X)$, $\mathbf{F} = \text{const.}$, a corresponding time-independent entropy distribution $\eta = \eta(X)$ is uniquely determined. Thus, by the last paragraph of Section 2, a prescribed constant \mathbf{F} and a prescribed function $\vartheta(X)$ uniquely determine an admissible process. For such a process, the inequality (4.7) reduces to

$$-\mathbf{q} \cdot \text{grad } \vartheta \geq 0. \tag{6.2}$$

Since $\vartheta = \vartheta(X)$ can be so chosen that, at a particular material point, both ϑ and $\text{grad } \vartheta$ have arbitrarily assigned values, it follows from (6.2) and (3.4) that the *heat conduction inequality*

$$\mathbf{v} \cdot \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}) \leq 0 \tag{6.3}$$

must hold for all vectors \mathbf{v} , all tensors \mathbf{F} and all scalars η .

Keeping \mathbf{F} and η fixed, let us define a scalar function of the vector \mathbf{v} by

$$\varphi(\mathbf{v}) = \mathbf{v} \cdot \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v}). \quad (6.4)$$

The heat conduction inequality (6.3) states that $\varphi(\mathbf{v})$ has the maximum 0 at $\mathbf{v} = \mathbf{0}$. A rule of calculus tells us then that the gradient of $\varphi(\mathbf{v})$ with respect to \mathbf{v} must vanish for $\mathbf{v} = \mathbf{0}$. Using the rules for the differentiation of products, we find from (6.4) that

$$\{[\nabla_{\mathbf{v}} \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v})]^T \mathbf{v} + \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v})\}|_{\mathbf{v}=\mathbf{0}} = \hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{0}) = \mathbf{0}. \quad (6.5)$$

Therefore, the condition (3.15), expressing the non-existence of a piezo-caloric effect, is also a consequence of the heat conduction inequality and hence can be expected to hold even if $-\mathbf{I}$ does not belong to the isotropy group of the material. This is, in fact, a result already obtained by PIPKIN & RIVLIN [8], § 18.

We summarize the results obtained as a consequence of our postulate: *An elastic material with heat conduction and viscosity is determined by the three response functions $\hat{\varepsilon}$, \mathbf{I} , and $\hat{\mathbf{q}}$. The temperature and stress are given by the relations (5.2) and (5.4). The response function \mathbf{I} is restricted by the dissipation inequality (5.5). If the temperature relation is invertible, the response function $\hat{\mathbf{q}}$ is restricted by the heat conduction inequality (6.3) and hence also by (3.15).*

Using (5.2), (3.3), and (5.4), we can rewrite the inequality (4.5) as follows:

$$\rho \vartheta \gamma = \text{tr}\{\mathbf{L}\mathbf{I}(\mathbf{F}, \eta)[\mathbf{L}]\} - \frac{1}{\vartheta} \mathbf{q} \cdot \text{grad } \vartheta \geq 0. \quad (6.6)$$

It is clear from (6.6) that the temperature relation (5.2), the stress relation (5.4), the dissipation inequality (5.5), and the heat conduction inequality (6.3) are *sufficient* that the Clausius-Duhem inequality hold for all admissible processes.

7. Linearly viscous fluids

We say that a material is a *fluid* if its isotropy group \mathcal{I} is the group \mathcal{U} of all unimodular tensors¹. If $\mathcal{I} = \mathcal{U}$ for one reference configuration, then $\mathcal{I} = \mathcal{U}$ for all reference configurations of the same material; hence, the property of being a fluid is intrinsic to the material.

For an *elastic fluid with heat conduction and viscosity* we have the identities (3.12) for all unimodular tensors \mathbf{H} , and, therefore, $\hat{\varepsilon}(\mathbf{F}, \eta)$, $\mathbf{I}(\mathbf{F}, \eta)$ and $\hat{\mathbf{q}}(\mathbf{F}, \eta, \mathbf{v})$ can depend on \mathbf{F} only through $|\det \mathbf{F}|$ or, equivalently, through the specific volume $v = \frac{1}{\rho}$. The temperature and stress relations now reduce to

$$\vartheta = \hat{\varepsilon}_\eta(v, \eta), \quad \hat{\mathbf{T}}(v, \eta) = \hat{\varepsilon}_v(v, \eta) \mathbf{I}. \quad (7.1)$$

The identity (3.7)₄ implies that

$$\mathbf{Q} \mathbf{I}(v, \eta) [\mathbf{D}] \mathbf{Q}^T = \mathbf{I}(v, \eta) [\mathbf{Q} \mathbf{D} \mathbf{Q}^T] \quad (7.2)$$

holds for all symmetric tensors \mathbf{D} and all orthogonal tensors \mathbf{Q} . It is well known² that the identity (7.2) holds if and only if \mathbf{I} is of the form

$$\mathbf{I}(v, \eta) [\mathbf{D}] = 2\mu \mathbf{D} + \lambda (\text{tr } \mathbf{D}) \mathbf{I}, \quad (7.3)$$

¹ Cf. [2].

² Cf. [4], § 298.

where μ and λ are scalar functions of v and η . Substituting (7.1)₂ and (7.3) into (3.8)₃, we find

$$\mathbf{T} = -p\mathbf{I} + 2\mu\mathbf{D} + \lambda(\text{tr}\mathbf{D})\mathbf{I} \tag{7.4}$$

where

$$p = -\hat{\varepsilon}_v(v, \eta). \tag{7.5}$$

Of course, (7.4) is the familiar constitutive equation for linearly viscous fluids. It is significant, however, that the pressure p is given by the same formula (7.5) which holds for perfect (inviscid) fluids. The dissipation inequality (5.7) states that

$$2\mu \text{tr}\mathbf{S}^2 + \lambda(\text{tr}\mathbf{S})^2 \geq 0 \tag{7.6}$$

for all symmetric tensors \mathbf{S} . It is known² that the inequalities

$$\mu \geq 0, \quad \lambda + \frac{2}{3}\mu \geq 0 \tag{7.7}$$

are equivalent to (7.6). They state that the shear viscosity μ and the bulk viscosity $\lambda + \frac{2}{3}\mu$ must both be non-negative.

For fluids, the identity (3.7)₅ reduces to

$$\mathbf{Q} \hat{\mathbf{q}}(v, \eta, \mathbf{v}) = \hat{\mathbf{q}}(v, \eta, \mathbf{Q}\mathbf{v}). \tag{7.8}$$

This identity holds if and only if $\hat{\mathbf{q}}$ is of the form

$$\hat{\mathbf{q}}(v, \eta, \mathbf{v}) = -\kappa \mathbf{v} \tag{7.9}$$

where κ is a scalar function of v, η , and $\mathbf{v} \cdot \mathbf{v}$. By (7.9), the heat conduction inequality (6.3) reduces to

$$\kappa \geq 0, \tag{7.10}$$

which states that the heat conductivity κ cannot be negative.

Acknowledgements. We are grateful to Drs. J. L. ERICKSEN, A. E. GREEN, R. A. TOUPIN and C. TRUESDELL for stimulating discussions.

This research was supported by the Air Force Office of Scientific Research under Contract AF 49(638)541 with Mellon Institute and by the National Science Foundation under Grant G 11316 to The Johns Hopkins University.

References

- [1] NOLL, W.: The foundations of classical mechanics in the light of recent advances in continuum mechanics. In: *The Axiomatic Method with Special Reference to Geometry and Physics*. Pp. 266—281. Amsterdam: North Holland Co. 1959.
- [2] NOLL, W.: A mathematical theory of the mechanical behavior of continuous media. *Arch. Rational Mech. Anal.* **2**, 197—226 (1958).
- [3] NOLL, W.: La mécanique classique, basée sur un axiome d'objectivité. A paper read at the Colloque Internationale sur la Méthode Axiomatique in Mécanique Classique et Moderne, Paris, 1959 (to be published by Gauthier-Villars, Paris).
- [4] TRUESDELL, C., & R. A. TOUPIN: The Classical Field Theories. In: *Encyclopedia of Physics*, Vol. III/1, edited by S. FLÜGGE. Berlin-Göttingen-Heidelberg: Springer 1960.
- [5] GREEN, A. E., & R. S. RIVLIN: The mechanics of non-linear materials with memory. *Arch. Rational Mech. Anal.* **1**, 1—34 (1957).
- [6] COLEMAN, B. D., & W. NOLL: An approximation theorem for functionals, with applications in continuum mechanics. *Arch. Rational Mech. Anal.* **6**, 355—370 (1960).

- [7] NOLL, W.: On the continuity of the solid and fluid states. *J. Rational Mech. Anal.* **4**, 3—81 (1955).
- [8] PIPKIN, A. C., & R. S. RIVLIN: The formulation of constitutive equations in continuum physics. Technical Report DA 4531/4 to the Department of the Army, Ordnance Corps (1958).
- [9] GREEN, A. E., & J. E. ADKINS: *Large Elastic Deformations and Non-Linear Continuum Mechanics*. Oxford: Clarendon Press 1960.
- [10] COLEMAN, B. D., & W. NOLL,: Material symmetry and thermostatic inequalities in finite elastic deformations, forthcoming.

Mellon Institute
Pittsburgh, Pennsylvania
and
The Johns Hopkins University
Baltimore, Maryland

(Received February 11, 1963)

Material Symmetry and Thermostatic Inequalities in Finite Elastic Deformations

BERNARD D. COLEMAN & WALTER NOLL

Contents

Introduction	87
Applicability of the Theory	87
Thermostatic Inequalities	88
On Notation	90
1. Mechanical Preliminaries	90
Kinematics	91
Stress	93
Changes of Frame	93
2. The Response Function and the Isotropy Group	93
3. Undistorted States of Aeolotropic Solids	97
4. Thermostatic Inequalities	101
5. States of Pressure	104
6. Differential Inequalities	107
Isotropy	109
Cubic Symmetry	110

Introduction

This paper is concerned with *elastic materials*; these are substances for which the present stress S on a material point depends on only the present local configuration¹ M of that point:

$$S = \mathbf{S}(M). \tag{1}$$

There are several highly developed branches of physics which rest on special cases of the constitutive assumption (1): Euler's hydrodynamics of perfect fluids, the classical theory of infinitesimal elastic deformations, and the modern theory of finite elastic deformations. Our present interest is not in the solution of special problems in these subjects but rather in the theory of the form of the function \mathbf{S} itself; here we investigate the general limitations placed on \mathbf{S} by material symmetry and thermodynamic considerations.

Applicability of the Theory

Let us consider the situations in which we expect a theory based on (1) to be useful. It is known that there is a large, albeit not all-inclusive, class of

¹ In Sections 1 and 2 we explain more precisely some of the mechanical and kinematical terms used in this Introduction with only abbreviated and heuristic descriptions.

substances, called *simple materials*, whose behavior in general mechanical processes is described by the assertion that the stress at a material point is determined by a function of the history of the local configuration of the point:

$$S = \mathfrak{S}_{s=0}^{\infty}(M^s). \quad (2)$$

Here M^s denotes the local configuration s seconds ago, and \mathfrak{S} is a functional whose argument is the function M^s , $0 \leq s < \infty$, and whose value is the symmetric tensor S . Most of the literature on continuum mechanics deals with simple materials. For examples we have the theory of linear viscoelasticity, the hydrodynamics of perfect, viscous and non-Newtonian fluids, and also recent general theories of materials with nonlinear memory.^{2,3}

Now, when we consider a material point which has been in its present local configuration at all times in the past, the function M^s reduces to a constant,

$$M^s \equiv M = \text{const.}, \quad 0 \leq s < \infty, \quad (3)$$

and (2) yields the result that S depends only on M :

$$S = \mathbf{S}(M). \quad (4)$$

Thus, the theory of elastic materials describes simple materials which have always been at rest.

Suppose that the material described by the functional \mathfrak{S} in (2) has a "fading memory", *i.e.*, that \mathfrak{S} is continuous in such a way that S is "more sensitive" to changes in the local configuration M^s at small s (recent past) than to changes in the local configuration at large s (distant past).⁴ Then, given any history M^s , we can construct *retarded histories* $M^{\alpha s}$, $0 < \alpha < 1$, and prove that⁷

$$\lim_{\alpha \rightarrow 0} \mathfrak{S}_{s=0}^{\infty}(M^{\alpha s}) = \mathbf{S}(M^0). \quad (5)$$

The function \mathbf{S} in Eqs. (4) and (5) is the same function.

Equation (5) tells us that for every history M^s there exists a retarded history $M^{\alpha s}$, "essentially the same as M^s but slower", such that the present stress corresponding to $M^{\alpha s}$ is given, to as good an approximation as we wish, by $\mathbf{S}(M^0)$. This result justifies, within the framework of the theory of simple materials, the use of the static stress-strain function \mathbf{S} in the discussion of slow processes; *i.e.*, it gives a motivation to the *dynamic* (as distinguished from the static) theories of perfect fluids and perfectly elastic solids.

Thermostatic Inequalities

In the theory of *infinitesimal* elastic deformations from a natural state (*i.e.* a state with zero stress), it is usually assumed that the stress-strain function \mathbf{S} is determined by the gradient of a *positive-definite* strain-energy function σ .

² GREEN, A. E., & R. S. RIVLIN, Arch. Rational Mech. Anal. **1**, 1 (1957).

³ NOLL, W., Arch. Rational Mech. Anal. **2**, 197 (1958).

⁴ This concept of "fading memory" is made precise in reference 5 and exploited further in reference 6.

⁵ COLEMAN, B. D., & W. NOLL, Arch. Rational Mech. Anal. **6**, 355 (1960).

⁶ COLEMAN, B. D., & W. NOLL, Rev. Mod. Phys. **33**, 239 (1961).

⁷ The crucial step in proving (5) is Eq. (3.21) of reference 5.

It has often been supposed that the positive-definiteness of σ should be justifiable in terms of fundamental principles in classical thermodynamics, but rigorous arguments to this effect have not come forth. Nor has classical thermodynamics told us how to extend this assumption on σ (and hence on \mathbf{S}) for infinitesimal elasticity to an assertion about \mathbf{S} that is in accord with experience in *finite* elasticity.⁸

Yet, Eq. (3) tells us that the function \mathbf{S} must be appropriate for those physical situations in which a simple material can be regarded as having been forever at rest, and \mathbf{S} must therefore describe "equilibrium states." Furthermore, our thermodynamic intuition suggests that an equilibrium state is one which maximizes an entropy or minimizes a free energy with respect to an appropriate class of disturbances, and this, in turn, suggests that the function \mathbf{S} should be subject to certain inequalities. The failure of classical thermodynamics to yield these inequalities is, in part, due to the difficulty in deciding what to include in the "class of disturbances".

In 1959^{9,10} we proposed a theory of thermostatics which led to a general inequality on \mathbf{S} . We refer to that inequality as the TI (Thermostatic Inequality). In references 9 and 11 consequences of the TI were obtained for various types of materials. In the theory of infinitesimal deformations from a natural state, the TI reduces to the classical assumption of the positive-definiteness of σ . The consequences of the TI for finite elastic deformations seem to be in accord with all experience in solids.

For example, consider a homogeneous deformation of an undistorted isotropic elastic cube in the direction of its three axes. The TI then implies that the greater stretch occurs always in the direction of the greater applied force.⁹

For fluids, the TI is equivalent to the assertion that the pressure be positive and that the compression modulus be greater than two-thirds of the pressure. That the pressure is positive seems to be in accord with experience. However, for a fluid whose critical point occurs at a high pressure there can be a range of densities at which the compression modulus is less than two-thirds of the pressure.

Our TI is equivalent to a requirement of stability against homogeneous disturbances at fixed surface forces. Such a requirement should be appropriate for solids, because it is the surface forces that are controlled in most measurements on solids. For fluids, however, it is usually pressure, and not surface force, that is controlled. When the pressure is fixed during a deformation, the surface forces do not remain fixed but change their direction so as to stay normal to the surface and change their magnitude so as to compensate for the variation of the surface area. The familiar statement that the compression modulus is positive can be derived from a requirement of stability against variations in volume at fixed pressure. Such a requirement is suitable for fluids but does not yield adequate restrictions on the behavior of solids; it does not even yield the positivity of the shear modulus for isotropic infinitesimal elasticity.

⁸ TRUESDELL, C., *Z. angew. Math. u. Mech.* **36**, 97 (1956).

⁹ COLEMAN, B. D., & W. NOLL, *Arch. Rational Mech. Anal.* **4**, 97 (1959).

¹⁰ A different formulation of the 1959 theory is given in reference 11.

¹¹ COLEMAN, B. D., *Arch. Rational Mech. Anal.* **9**, 172 (1962).

It is still an open question whether it is possible to find a single inequality from which one can derive all thermodynamic restrictions on the static behavior of elastic materials, whether they be fluids, solids, or neither.

In this article we make a detailed study of implications of the TI, stating, without proof, some of the known theorems and presenting several new theorems. Proofs are supplied for all new theorems stated.

Throughout the present discussion we ignore the dependence of the stress on a thermodynamic parameter, such as the temperature or the entropy density. Furthermore, we follow the procedure of a recent article by TRUESDELL & TOUPIN¹², and, whenever possible, we work with a generalization of the TI, called the WTI (Weakened Thermostatic Inequality) which can be expressed directly in terms of the stress-strain function \mathbf{S} without recourse to mention of internal energy or free-energy functions. This deliberate suppression of mention of thermodynamic variables is done to emphasize the mechanical significance of our thermodynamic considerations.

Before going into the theory of the TI, we discuss the characterization of material symmetry in terms of stress-strain functions, giving simplified proofs of some known propositions and presenting several new results.

On Notation

The term *tensor* is used here as a synonym for linear transformation of a vector space into itself. We denote tensors by Latin majuscules: F, U, Q, R, I . We reserve the symbol X , however, to represent material points of a physical body. For the trace of a tensor F we write $\text{tr } F$, and for the determinant of F we write $\det F$. We say that F is invertible if F has an inverse F^{-1} , which is the case if and only if $\det F \neq 0$. The transpose of F is denoted by F^T ; F is symmetric if $F^T = F$ and skew if $F^T = -F$. The "unit tensor" (*i.e.* the identity transformation) is denoted by I . If $Q Q^T = Q^T Q = I$, then Q is orthogonal. If Q is orthogonal and such that $\det Q = +1$, then Q is proper orthogonal. If the inner product $\mathbf{v} \cdot U \mathbf{v}$ is strictly greater than zero for all nonzero vectors \mathbf{v} , then we say that the tensor U is positive-definite. If U is symmetric, then a necessary and sufficient condition that U be positive-definite is that all the proper numbers of U be positive.

1. Mechanical Preliminaries

For convenience, we briefly summarize some definitions and general results of continuum mechanics which are prerequisite to the present paper. Since detailed axiomatic treatments of this material have been given elsewhere,^{3,13,14} the results stated here are asserted without proof.

¹² TRUESDELL, C., & R. A. TOUPIN, Arch. Rational Mech. Anal. **12**, 1 (1963). The inequality which we here call the "WTI" is called the "GCN condition" by TRUESDELL & TOUPIN.

¹³ NOLL, W., in: Proceedings of the Berkeley Symposium on the Axiomatic Method, pp. 266–281. Amsterdam: North Holland.

¹⁴ Another brief summary is given in §§ 1 and 5 of reference 9. For a thorough exposition of the foundations of continuum mechanics see C. TRUESDELL & R. A. TOUPIN, in: Encyclopedia of Physics, Edited by S. FLÜGGE, Vol. III/1. Berlin-Göttingen-Heidelberg: Springer 1959.

Kinematics

A body \mathcal{B} is a smooth manifold whose elements X are called *material points*. A *configuration* \mathbf{f} of \mathcal{B} is a one-to-one mapping of \mathcal{B} onto a region \mathcal{R} in a three-dimensional Euclidean point space \mathcal{E} . The point $\mathbf{x}=\mathbf{f}(X)$ is called the *position* of X in the configuration \mathbf{f} , and $\mathcal{R}=\mathbf{f}(\mathcal{B})$ is called the *region* occupied by \mathcal{B} in \mathbf{f} . The *mass distribution* m of \mathcal{B} is a measure defined on subsets of \mathcal{B} . To each configuration \mathbf{f} of \mathcal{B} corresponds a mass density ρ .

If we have two configurations \mathbf{f}_1 and \mathbf{f}_2 of the same body \mathcal{B} , we can consider the composition $\mathbf{g}_{2,1}$ of \mathbf{f}_2 with the inverse \mathbf{f}_1^{-1} of \mathbf{f}_1 ; $\mathbf{g}_{2,1}$ is defined by

$$\mathbf{g}_{2,1}(\mathbf{x})=\mathbf{f}_2(\mathbf{f}_1^{-1}(\mathbf{x})).$$

Clearly, $\mathbf{g}_{2,1}$ is a mapping of the region \mathcal{R}_1 occupied by \mathcal{B} in the configuration \mathbf{f}_1 onto the region \mathcal{R}_2 occupied by \mathcal{B} in the configuration \mathbf{f}_2 . The class of admissible configurations \mathbf{f} is assumed to be such that for any two configurations \mathbf{f}_1 and \mathbf{f}_2 , $\mathbf{g}_{2,1}$ is a smooth homeomorphism¹⁵ of \mathcal{R}_1 onto \mathcal{R}_2 . Now, since $\mathbf{g}_{2,1}$ is a mapping of one region of ordinary Euclidean space onto another, there is no difficulty in computing the gradient of $\mathbf{g}_{2,1}$, $\text{grad } \mathbf{g}_{2,1}$. Because \mathbf{f}_1 is one-to-one, $\text{grad } \mathbf{g}_{2,1}$ can be regarded as either a field over \mathcal{R}_1 or as a field over the manifold \mathcal{B} . Here it is convenient for us to take the latter point of view. We denote the values of $\text{grad } \mathbf{g}_{2,1}$ by $F_{2,1}(X)$. Of course $F_{2,1}(X)$ is a tensor; it is called the *deformation gradient* (at the material point X) of the configuration \mathbf{f}_2 computed taking the configuration \mathbf{f}_1 as reference. Because the function $\mathbf{g}_{2,1}$ is one-to-one, the tensors $F_{2,1}(X)$ are invertible.

If we have three configurations $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ of the same body \mathcal{B} and compute the three deformation gradients $F_{3,1}(X)$, $F_{3,2}(X)$ and $F_{3,1}(X)$, then it follows from the chain rule for differentials of vector functions that

$$F_{3,1}(X)=F_{3,2}(X) F_{2,1}(X). \quad (1.1)$$

If $F_{2,1}$ is independent of X , we say that \mathbf{f}_1 and \mathbf{f}_2 are related by a homogeneous deformation. If, in this case, $F_{2,1}$ is a proper orthogonal tensor, then $F_{2,1}$ is said to characterize a *rigid rotation*. If $F_{2,1}$ is orthogonal (with no restriction on the sign of $\det F_{2,1}$), then $F_{2,1}$ is said to correspond to an *orthogonal transformation*, i.e. a combination of rigid rotations and reflections. If $F_{2,1}$ has the form αQ , where α is a scalar and Q is orthogonal, then $F_{2,1}$ describes a *similarity transformation*, i.e. a combination of an orthogonal transformation and a uniform expansion by the factor α . If $F_{2,1}$ is a positive definite symmetric tensor, then \mathbf{f}_2 is said to be obtained from \mathbf{f}_1 by a *pure stretch*; the proper vectors of $F_{2,1}$ then give the *principal directions of stretch*, and the proper numbers of $F_{2,1}$ are the *principal stretches*.

Let us return to deformations which are not necessarily homogeneous; such deformations can still be regarded, in a sense which we shall now make precise, as homogeneous on a "local level".

If, at a particular material point X , $F_{2,1}(X)$ is the unit tensor I , then we say that \mathbf{f}_1 and \mathbf{f}_2 give rise to the same local configuration at X . To state this

¹⁵ A smooth homeomorphism is a one-to-one function \mathbf{g} such that both \mathbf{g} and its inverse \mathbf{g}^{-1} are continuously differentiable.

more formally, we observe that for each fixed X we can define an equivalence relation " \sim " among all the configurations \mathbf{f} of \mathcal{B} by the condition that $\mathbf{f}_1 \sim \mathbf{f}_2$ if $F_{2,1}(X) = I$; the resulting equivalence classes of configurations are the *local configurations* M of X . If M_1 is the equivalence class of \mathbf{f}_1 at X and if M_2 is the equivalence class of \mathbf{f}_2 at X , then $F_{2,1}(X)$ depends only on M_1 and M_2 ; it is also true that M_2 is uniquely determined by $F_{2,1}(X)$ and M_1 . Hence we can call $F_{2,1}(X)$ simply *the deformation gradient from M_1 to M_2* . Dropping the symbol X in $F_{2,1}(X)$, we write

$$M_2 = F_{2,1} M_1. \quad (1.2)$$

If we denote the mass densities at X corresponding to the local configurations M_1 and M_2 by ϱ_1 and ϱ_2 , respectively, then

$$\varrho_2 = \frac{1}{|\det F_{2,1}|} \varrho_1. \quad (1.3)$$

It is often convenient to employ a fixed *local reference configuration* M_o and to characterize all other local configurations

$$M = F M_o \quad (1.4)$$

by their deformation gradients F from M_o . It follows from (1.4) that if

$$M_1 = F_1 M_o, \quad M_2 = F_2 M_o, \quad (1.5)$$

then the deformation gradient H from M_1 to M_2 is given by

$$H = F_2 F_1^{-1}, \quad F_2 = H F_1. \quad (1.6)$$

A theorem of algebra, called the *polar decomposition theorem*, states that any invertible tensor F can be written in two ways as the product of a symmetric positive-definite tensor and an orthogonal tensor:

$$F = R U, \quad (1.7)$$

$$F = V R. \quad (1.8)$$

Furthermore, the orthogonal tensor R and the symmetric positive-definite tensors U and V in these decompositions are uniquely determined by F and obey the following relations:

$$U^2 = F^T F, \quad (1.9)$$

$$V^2 = F F^T, \quad (1.10)$$

$$U = R^T V R. \quad (1.11)$$

Equations (1.7) and (1.8) have the following significance in kinematics: Any deformation gradient F may be regarded as being the result of a pure stretch U followed by an orthogonal transformation R , or an orthogonal transformation R followed by a pure stretch V . These interpretations uniquely determine the pairs R, U and R, V . The orthogonal transformations entering these two interpretations are the same; however, the pure stretches U and V can be different. It follows from (1.11) that although U and V may have different principal directions, they must have the same proper numbers u_i . We call the orthogonal

tensor R the *rotation tensor*¹⁶ and the tensors U and V , respectively, the *right* and *left stretch tensors*. If $M_2 = FM_1$, then we call U and V , respectively, the *right* and *left stretch tensors from M_1 to M_2* , and the proper numbers u_i of U and V are called the *principal stretches from M_1 to M_2* .

Stress

One distinguishes between contact forces and body forces. On assuming the principle of linear momentum and the principle of angular momentum, one can show, after laying down some strong continuity assumptions, that the contact force $d\mathbf{c}$ across an oriented surface element with unit normal \mathbf{n} and area dA is given by

$$d\mathbf{c} = \mathbf{S}\mathbf{n}dA, \quad (1.12)$$

where S is a symmetric tensor called the *stress tensor*.¹⁷

A proper vector of S is said to determine a *principal axis of stress*. Proper numbers of S are called *principal stresses*.

Since we are here interested in only "local phenomena", we ignore body forces throughout our discussion.

Changes of Frame

A *change of frame* is the mathematical embodiment of the physical concept of a change of observer.

In this paper we assume that the stress tensor S at a material point X is determined by the present value of the local configuration M at that point. It follows that, for our present purposes, we can define a change of frame to be the following simultaneous transformation of the local configuration M_X and the stress S_X at each material point X of a body:

$$M_X \rightarrow Q M_X, \quad (1.13a)$$

$$S_X \rightarrow Q S_X Q^{-1}; \quad (1.13b)$$

here Q is an orthogonal tensor independent of X . The change of frame (1.13) is said to be characterized by Q .

In other words, a change of frame is a simultaneous orthogonal transformation of both the present configuration of \mathcal{B} and the present force system acting on \mathcal{B} .

The *principle of material objectivity* is a mathematization, for classical mechanics, of the physical idea that the behavior of a material should be independent of the observer. In our present theory this principle reduces to the statement that the dependence of the stress on the local configuration must be invariant under changes of frame; *i.e.*, this dependence must be such that if M is changed from M to QM , where Q is any orthogonal tensor, then the stress S corresponding to M must change to QSQ^{-1} .

2. The Response Function and the Isotropy Group

We are now prepared to state formally our

Fundamental Constitutive Assumption. To each local configuration M there corresponds a unique value of the stress tensor S .

¹⁶ In applications R is usually proper orthogonal.

¹⁷ For details, see reference 13.

Such a correspondence between S and M is called an *elastic material* (or, simply, "a material"). Since the behavior of a material should be independent of the observer, we also lay down an

Assumption of Material Objectivity. The dependence of the stress S on the local configuration M is such that if $M_2 = Q M_1$ where Q is any orthogonal tensor, then $S_2 = Q S_1 Q^T$ where S_1 and S_2 are the stresses corresponding to M_1 and M_2 , respectively.

We follow standard procedure and express the dependence of S on M by picking a local reference configuration M_o , characterizing all local configurations by their deformation gradients F from M_o in accordance with (1.4), and then regarding S as a function g of F :

$$S = g(F). \quad (2.1)$$

The function g is called the *response function*, of our material, *taken relative to M_o* .

It follows from our Assumption of Material Objectivity that g must obey the identity

$$g(QF) = Q g(F) Q^{-1} \quad (2.2)$$

for all tensors F and all orthogonal tensors Q . Using Eqs. (2.2) and (1.7), we infer that Eq. (2.1) is equivalent to

$$S = R g(U) R^{-1}. \quad (2.3)$$

The function g depends on the choice of the reference configuration M_o . To exhibit the nature of this dependence of g on M_o , we note that it follows from (1.5) and (1.6) that if F and F' are two deformation gradients characterizing the same local configuration M with respect to two different local reference configurations M_o and M'_o , *i.e.*

$$M = F M_o = F' M'_o,$$

and if

$$M'_o = G M_o,$$

then

$$F = F' G.$$

If the response functions relative to M_o and M'_o are denoted by g and g' , respectively, we infer

$$S = g(F) = g'(F') = g'(F G^{-1}).$$

This proves

Proposition 1. *If g and g' are the response functions for the same material but taken relative to different local configurations M_o and M'_o , with $M'_o = G M_o$, then*

$$g(F) = g'(F G^{-1}) \quad (2.4)$$

for all F .

For a given material it may turn out that g remains the same function if the local reference configuration M_o is changed to another local reference configuration $M'_o = H M_o$ with the same density, *i.e.*, that

$$g(F) = g'(F) \quad (2.5)$$

for all F . (The physical interpretation of the identity (2.5) is that the two configurations M_o and M'_o must be equal in their response to equal deformations, *i.e.*, that the classes of reference configurations for which the identity (2.5) holds must characterize the isotropy of the material under consideration.) Now, it follows from (2.4) that if (2.5) holds for M_o and $M'_o = HM_o$, then g satisfies the identity

$$g(F) = g(FH^{-1}) \quad (2.6)$$

for all F . Since we are here assuming that M'_o and M_o have the same density, it is clear that $|\det H| = 1$; *i.e.*, H is here a unimodular transformation. If the identity (2.6) holds for H^{-1} , it holds also for H ; furthermore, the set of unimodular transformations H for which the identity (2.6) holds forms a group. Let us now state the following important formal definition whose physical motivation should, we feel, now be clear.

Definition. The group \mathcal{I} of unimodular tensors H for which the identity

$$g(FH) = g(F) \quad (2.7)$$

holds for all F , where g is the response function relative to the reference configuration M_o , is called the *isotropy group relative to M_o* .

\mathcal{I} depends on M_o , but we shall see that for a given material the isotropy groups \mathcal{I} and \mathcal{I}' relative to two distinct local reference configurations M_o and M'_o are conjugate and hence isomorphic. In fact, combination of (2.4) and (2.7) shows that

$$g'(FG^{-1}) = g'(FHG^{-1}) \quad (2.8)$$

for all F . If we put $H' = GHG^{-1}$ and $F' = FG^{-1}$, then $FHG^{-1} = F'H'$, and (2.8) becomes the assertion that

$$g'(F') = g'(F'H') \quad (2.9)$$

for all F' , but this simply says that H' is in \mathcal{I}' . Hence we have

Proposition 2. *If G is the deformation gradient from M_o to M'_o , then H belongs to \mathcal{I} if and only if GHG^{-1} belongs to \mathcal{I}' , where \mathcal{I} and \mathcal{I}' are, respectively, the isotropy groups relative to M_o and M'_o .*

$\mathcal{I} = \mathcal{I}'$ would mean that the change of reference configuration G does not affect the isotropy group. By Proposition 2 this is the case if and only if $G^{-1}HG$ belongs to \mathcal{I} . In the language of group theory, this condition states that G must belong to the *normalizer* group of \mathcal{I} within the full linear group \mathcal{L} . Thus we have

Proposition 3. *A change of reference configuration from M_o to GM_o leaves the isotropy group \mathcal{I} unchanged if and only if G belongs to normalizer of \mathcal{I} in \mathcal{L} .*

Suppose Q is orthogonal and belongs to \mathcal{I} ; then we can combine the identity (2.6), with $H = Q$, and the identity (2.2), to obtain

Proposition 4. *An orthogonal tensor Q belongs to the isotropy group \mathcal{I} if and only if*

$$Qg(F)Q^{-1} = g(QFQ^{-1}) \quad (2.10)$$

for all tensors F .

Definitions. If an elastic material has a local configuration M_0 such that the corresponding isotropy group is either the orthogonal group \mathcal{O} or a group which contains \mathcal{O} as a subgroup, then we say that the material is *isotropic*, and M_0 is called an *undistorted* local configuration. If a material is such that for some M_0 , \mathcal{I} is a subgroup (which need not be proper) of \mathcal{O} , then we say that the material is a *solid*, and we again call M_0 *undistorted*. Hence for the undistorted configurations M_0 of an *isotropic solid* $\mathcal{I} = \mathcal{O}$. In general, we say that a configuration M_0 is *undistorted* if \mathcal{I} is comparable to \mathcal{O} ; i.e. if \mathcal{I} either is itself a subgroup of \mathcal{O} or contains \mathcal{O} as a subgroup. We say that a material is a *fluid* if \mathcal{I} is the full unimodular group \mathcal{U} . Since \mathcal{O} is a subgroup of \mathcal{U} , an elastic fluid is isotropic. It follows from Proposition 2 that if $\mathcal{I} = \mathcal{U}$ for one local configuration M_0 , then $\mathcal{I}' = \mathcal{U}$ for every other local configuration M'_0 ; hence, every configuration of a fluid is undistorted.

Note. Continuum mechanics does not yet have a standard terminology for the various mathematical concepts behind the intuitive notion of "isotropy". The particular definitions we give here are specializations to elastic materials of the definitions given by NOLL⁸ for general simple materials. The reader is cautioned that, whereas we use *isotropic* as a quality of a *material*, other writers regard *isotropic* as a quality of both a *material and a configuration*. When these writers say that "the material is isotropic in the configuration M ", we say that " M is an undistorted configuration of the isotropic material".

The following proposition is an immediate consequence of Proposition 4.

Proposition 5. *The response function g of an isotropic material, when taken relative to an undistorted state, obeys the following identity for all tensors F and all orthogonal tensors Q :*

$$g(QFQ^{-1}) = Qg(F)Q^{-1}. \quad (2.11)$$

Tensor-valued functions obeying (2.11) are called *isotropic functions*.

The next three propositions illustrate some of the physical motivation behind our formal definitions.

Proposition 6. *The stress on an undistorted state of an isotropic material is always a hydrostatic pressure.*

This proposition states that if \mathcal{I} contains \mathcal{O} as a subgroup (properly or improperly), then $g(I) = -pI$.

Proof. By hypothesis and Proposition 5 we have

$$Qg(I)Q^{-1} = g(QIQ^{-1}) = g(I)$$

for all orthogonal tensors Q ; i.e., $g(I)$ must commute with all orthogonal tensors. This is possible only if $g(I)$ is a scalar multiple ($-pI$) of the unit tensor I , *q.e.d.*

Proposition 7. *The stress on an elastic fluid is always a hydrostatic pressure $-pI$ which depends on only the density ϱ .*

In other words, for a simple fluid,

$$g(F) = -p(\varrho)I \quad (2.12)$$

where

$$\varrho = \frac{1}{|\det F|} \varrho_0, \quad (2.13)$$

ρ being the density of the present local configuration FM and ρ_0 the density of the local reference configuration M_0 .

Proof. For any given tensor F , the tensor $H = |\det F| F^{-1}$ is obviously unimodular and hence belongs to the isotropy group $\mathcal{I} = \mathcal{U}$ of the fluid. Therefore, by (2.7),

$$g(F) = g(FH) = g(|\det F| I)$$

depends only on $|\det F|$ or, equivalently, on the density ρ . That $g(F)$ is a hydrostatic pressure for all M and F is a consequence of Proposition 6 and the fact that a simple fluid is an isotropic material all of whose configurations are undistorted.

Proposition 8. *Two undistorted local configurations of an isotropic solid can differ by only a similarity transformation.*

Proof. By definition, the isotropy group of an isotropic solid, relative to an undistorted reference configuration, is the orthogonal group \mathcal{O} . Hence, by Proposition 3, a change G from one undistorted reference configuration to another must belong to the normalizer of \mathcal{O} in \mathcal{L} . But it is a known result in group theory that the normalizer of \mathcal{O} is the group of all similarity transformations,¹⁸ *q.e.d.*

3. Undistorted States of Aeolotropic Solids

Proposition 6 of Section 2 shows that the stress on an undistorted local configuration of an isotropic material (solid or not) must be a hydrostatic pressure. Here we investigate the limitations material symmetry imposes upon the stress on undistorted configurations of various types of aeolotropic solids. We also extend Proposition 8 of Section 2 to obtain limitations on the possible strains which can relate undistorted states of general solids.

As far as purely mechanical behavior is concerned, the symmetry of an elastic solid is characterized by a corresponding "type" of isotropy group \mathcal{I} . By our definition of the term *solid*, \mathcal{I} will be a subgroup of the orthogonal group \mathcal{O} if an undistorted local configuration M_0 is used as a reference. If $M'_0 = GM_0$ is another undistorted reference configuration, then the isotropy group \mathcal{I}' corresponding to M'_0 will also be a subgroup of the orthogonal group. Moreover, according to Proposition 3 of Section 2, \mathcal{I}' will be the conjugate $\mathcal{I}' = G\mathcal{I}G^{-1}$ of \mathcal{I} under G ; *i.e.*, every transformation Q' in \mathcal{I}' will be of the form

$$Q' = GQG^{-1} \quad (3.1)$$

where Q is a member of \mathcal{I} . Let

$$G = RU \quad (3.2)$$

be the polar decomposition of G , so that U is the right stretch tensor and R the rotation tensor of the deformation carrying M_0 into M'_0 . Combining Eqs. (3.1) and (3.2), we find

$$\begin{aligned} Q'G &= GQ, \\ Q'RU &= RUQ = (RQ)(Q^{-1}UQ). \end{aligned} \quad (3.3)$$

¹⁸ Cf. H. WEYL, *The Classical Groups*, p. 22. Princeton 1946.

Since both $Q'R$ and RQ are orthogonal tensors while U and $Q^{-1}UQ$ are both positive-definite and symmetric, (3.3) gives two polar decompositions of the one tensor $Q'G$. Therefore, the uniqueness of a polar decomposition implies that

$$Q'R = RQ \quad \text{and} \quad U = Q^{-1}UQ. \quad (3.4)$$

The first of the Eqs. (3.4) shows that every member Q' of \mathcal{S}' is of the form

$$Q' = RQR^{-1}, \quad (3.5)$$

i.e., that \mathcal{S}' is actually a *conjugate* $\mathcal{S}' = R\mathcal{S}R^{-1}$ of \mathcal{S} within the orthogonal group. We say that two subgroups of the orthogonal group are of the *same type* if they are conjugate in this sense. The isotropy group of a solid depends on the choice of reference configuration, even when this configuration is restricted to be undistorted. The type to which this group belongs, however, represents an intrinsic property of the material.

Crystalline solids are classified into thirty-two classes, each of which is characterized by a certain type of symmetry group. The symmetry groups are finite subgroups of the orthogonal group. On putting $Q = -I$ in (2.2) we see that

$$g(F(-I)) = (-I)g(F)(-I) = g(F), \quad (3.6)$$

which shows that the inversion $-I$ is always a member of the isotropy group. This inversion, however, does not belong to the symmetry group of some of the crystal classes.¹⁹ The following assumption seems to be natural:

Hypothesis on Crystalline Solids. Consider a crystal which, relative to some undistorted state M_o , has the crystallographic symmetry group (*i.e.* point group) \mathcal{C} . The isotropy group \mathcal{S} of this crystal, relative to M_o , is assumed to be the group generated by \mathcal{C} and the inversion $-I$.

The 32 types of symmetry groups²⁰ give rise to only 11 types of isotropy groups. In describing a group \mathcal{S} we need not list all the elements of \mathcal{S} but only a set of *generators* of \mathcal{S} , *i.e.* a set of members of \mathcal{S} which, when they and their inverses are multiplied among themselves in various combinations, yield all the elements of \mathcal{S} . For an isotropy group relative to an undistorted state of a solid it is always possible to find a list of generators $-I, Q_1, \dots, Q_m$ such that each Q_i is *proper orthogonal*. We denote by $R_{\mathbf{n}}^{\varphi}$ the right-handed rotation through the angle φ , $0 < \varphi < 2\pi$, about an axis in the direction of the unit vector \mathbf{n} . Each proper orthogonal transformation $\neq I$ is a rotation of the form $R_{\mathbf{n}}^{\varphi}$. Table 1 gives a description of the 11 types of isotropy groups for crystals. In this table, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ denotes a right handed orthonormal basis and $\mathbf{d} = \frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$.

A change from one group to another of the same type corresponds merely to a change of the orthonormal basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

¹⁹ If we were dealing here with vector phenomena, as in electromagnetic theory, then we should not automatically have $-I$ in \mathcal{S} , and we should have instead $\mathcal{S} = \mathcal{C}$. In other words, it is because our present theory involves only tensors of order two that our isotropy groups are sometimes bigger than the crystallographic point groups.

²⁰ Complete lists of members for all these groups are given by G. F. SMITH & R. S. RIVLIN, *Trans. Am. Math. Soc.* **88**, 175 (1958). The names we employ for the various crystal systems and classes are those used by SMITH & RIVLIN, who, in turn, state that they come from the 1952 Edition of DANA's *Textbook of Mineralogy*, revised by C. S. HURLBUT (New York: John Wiley).

Table 1. *Isotropy Groups for the Various Crystal Classes*

Ref. No.	Crystal Class	Proper Orthogonal Generators of \mathcal{I}	Number of Elements in \mathcal{I}
	<i>Triclinic System</i>		
1.	all classes	I	2
	<i>Monoclinic System</i>		
2.	all classes	$R_{\mathbf{k}}^{\pi}$	4
	<i>Rhombic System</i>		
3.	all classes	$R_{\mathbf{i}}^{\pi}, R_{\mathbf{j}}^{\pi}$	8
	<i>Tetragonal System</i>		
4.	{ tetragonal-disphenoidal tetragonal-pyramidal tetragonal-dipyramidal }	$R_{\mathbf{k}}^{2\pi/2}$	8
5.	{ tetragonal-scalenohedral ditetragonal-pyramidal tetragonal-trapezohedral ditetragonal-dipyramidal }	$R_{\mathbf{k}}^{2\pi/2}, R_{\mathbf{i}}^{\pi}$	16
	<i>Cubic System</i>		
6.	{ tetratoidal } { diploidal }	$R_{\mathbf{i}}^{\pi}, R_{\mathbf{j}}^{\pi}, R_{\mathbf{d}}^{2\pi/3}$	24
7.	{ hextetrahedral } { gyroidal } { hexoctahedral }	$R_{\mathbf{i}}^{2\pi/2}, R_{\mathbf{j}}^{2\pi/2}$	48
	<i>Hexagonal System</i>		
8.	{ trigonal-pyramidal } { rhombohedral }	$R_{\mathbf{k}}^{2\pi/3}$	6
9.	{ ditrigonal-pyramidal } { trigonal-trapezohedral } { hexagonal-scalenohedral }	$R_{\mathbf{k}}^{2\pi/3}, R_{\mathbf{i}}^{\pi}$	12
10.	{ trigonal-dipyramidal } { hexagonal-pyramidal } { hexagonal-dipyramidal }	$R_{\mathbf{k}}^{2\pi/3}$	12
11.	{ ditrigonal-dipyramidal } { dihexagonal-pyramidal } { hexagonal-trapezohedral } { dihexagonal-dipyramidal }	$R_{\mathbf{k}}^{2\pi/3}, R_{\mathbf{i}}^{\pi}$	24

With respect to their elastic behavior alone, crystals can show only the 11 types of symmetry characterized by the 11 types of isotropy groups.

A type of symmetry appropriate not to crystals but to materials with a bundled or laminated structure or to drawn fibers is *transverse isotropy*. It is defined by the assumption that the isotropy group consists of $\pm I$ and $\pm R_{\mathbf{k}}^{\varphi}$, for some fixed unit vector \mathbf{k} , and all angles φ , $0 < \varphi < 2\pi$.

Consider an elastic solid and a particular undistorted reference configuration M_o . Denote the corresponding response function by g and the corresponding isotropy group by \mathcal{I} . The stress on M_o is given by

$$S_o = g(I). \quad (3.7)$$

The following is an immediate consequence of Eq. (2.10) for $F=I$:

Proposition 1. *The stress S_o on an undistorted local configuration M_o of a solid commutes with every member of the isotropy group \mathcal{I} for M_o ; i.e.*

$$QS_o = S_oQ \quad (3.8)$$

holds for all Q in \mathcal{I} .

It is easily seen that (3.8) holds for all Q in \mathcal{I} if it holds for all rotations Q_i , $i=1, \dots, m$ in a list $-I, Q_1, \dots, Q_m$ of generators of \mathcal{I} . Therefore, the problem of finding the possible stresses S_o on an undistorted state reduces to the problem of finding those symmetric tensors that commute with a certain set of rotations of the form $R_{\mathbf{n}}^{\varphi}$. For each of the 11 cases given in Table 1 and for the case of transverse isotropy, this problem can easily be solved with the help of the following

Commutation Theorem²¹. *The symmetric tensor S commutes with the orthogonal tensor Q if and only if Q leaves each of the characteristic spaces of S invariant, i.e., if it maps each characteristic space into itself.*

Here, a *characteristic space* of S is defined to be a maximal subspace (of the ordinary three-dimensional vector space) consisting of only proper vectors of S . If S has three distinct proper numbers, then it has three one-dimensional characteristic spaces. If S has only two distinct proper numbers, then it has two characteristic spaces, one of which is one-dimensional and the other two-dimensional. In this case, S must be of the form²²

$$S = -pI + q\mathbf{n} \otimes \mathbf{n}, \quad (3.9)$$

where $-p$ and $q-p$ are the proper numbers of S and \mathbf{n} is a unit vector which generates the one-dimensional characteristic space. If the three proper numbers of S coincide, then the entire three-dimensional vector space is the only characteristic space of S , and S is of the form

$$S = -pI. \quad (3.10)$$

Now if $\varphi \neq \pi$, the only spaces left invariant by the rotation $R_{\mathbf{n}}^{\varphi}$ are the one-dimensional space of all multiples of \mathbf{n} the two-dimensional space of all vectors perpendicular to \mathbf{n} , and the entire three-dimensional vector space. The rotation $R_{\mathbf{n}}^{\pi}$ leaves invariant, in addition, each one-dimensional space generated by a vector perpendicular to \mathbf{n} .

Proposition 1, Table 1, and the results from geometrical linear algebra just described enable one to establish easily the results collected in Table 2. In this table, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is the same orthonormal basis as is used for the isotropy groups in Table 1. Recall that a proper vector of the stress S_o determines a principal axis of stress.

Let us now return to Eqs. (3.4). It follows from the second of these equations that $QU = UQ$, which is the content of

²¹ This theorem is a corollary to Theorem 2, p. 77, and Theorem 3, p. 157, of P. R. HALMOS, *Finite-Dimensional Vector Spaces*. Princeton: Van Nostrand, 2nd ed., 1958.

²² The symbol \otimes denotes a tensor product; i.e., $\mathbf{n} \otimes \mathbf{n}$ is the tensor with the property that $(\mathbf{n} \otimes \mathbf{n})\mathbf{v} = \mathbf{n}(\mathbf{n} \cdot \mathbf{v})$ for all vectors \mathbf{v} .

Proposition 2. *Let M_0 and M'_0 be two local reference configurations of a solid, and let \mathcal{I} be the isotropy group relative to M_0 . If M_0 is undistorted, then M'_0 is undistorted if and only if the right stretch tensor U from M_0 to M'_0 commutes with every member of \mathcal{I} .*

Table 2. *Restrictions on the Stress for Undistorted States of Aeolotropic Solids*

Type of Isotropy	Ref. No. in Table 1	Restriction on S_0
Triclinic system	1	no restriction
Monoclinic system	2	\mathbf{k} is a proper vector of S_0
Rhombic system	3	$\mathbf{i}, \mathbf{j}, \mathbf{k}$ are proper vectors of S_0
{ Tetragonal system Hexagonal system Transverse isotropy	{ 4, 5, 8, 9, 10, 11 }	$S_0 = -pI + q\mathbf{k} \otimes \mathbf{k}$

It follows from this proposition that if M_0 is an undistorted configuration of a solid with the symmetry listed in the first column of Table 2, then a necessary and sufficient condition that M'_0 be an undistorted configuration of that solid is that the right stretch tensor U relating M'_0 to M_0 obey the restrictions listed for S_0 in the second column.

4. Thermoelastic Inequalities

As we mentioned in the Introduction, there are reasons, whose origins lie outside of mechanics, for believing that stress-strain functions used in physical applications should be subjected to restrictions beyond those following from Material Objectivity. These restrictions should have the form of inequalities and, we believe, should follow from a properly formulated theory of the thermodynamics of continuous media.

The inequalities we wish to consider here can be most simply expressed through use of the first Piola-Kirchhoff stress tensor, T , defined by

$$T = |\det F| (F^{-1} S)^T = \frac{\rho_0}{\rho} S F^{T^{-1}}, \quad (4.1)$$

where S is the ordinary stress and F the deformation gradient. The definition of T depends on the choice of the reference configuration. T need not be symmetric. We denote the relation between T and F by

$$T = \mathfrak{h}(F), \quad (4.2)$$

where the response function \mathfrak{h} is related to the response function g of Eq. (2.1) by

$$\mathfrak{h}(F) = |\det F| g(F) F^{T^{-1}}. \quad (4.3)$$

When the equation of material objectivity (2.2) is expressed in terms of \mathfrak{h} , it reads

$$\mathfrak{h}(QF) = Q \mathfrak{h}(F). \quad (4.4)$$

Using Eqs. (4.4) and (1.7), we see that Eq. (4.2) is equivalent to

$$T = R \mathfrak{h}(U), \quad (4.5)$$

where R is the rotation tensor and U the right stretch tensor corresponding to F .

In a work on the foundations of thermostatics⁹ we have laid down postulates which, in our present theory, are equivalent to the following

Thermostatic Inequality (TI). There exists a (scalar-valued) *energy function* σ such that

$$\sigma(F^*) - \sigma(F) - \text{tr}\{(F^* - F) \mathfrak{h}(F)^T\} > 0 \quad (4.6)$$

for all pairs of deformation gradients F, F^* such that $F^* \neq F$ and F^*F^{-1} is positive-definite and symmetric.

The energy function $\sigma(F)$ in (4.6) can be interpreted as either the internal energy density (per unit volume in the reference configuration) or the Helmholtz free energy density, depending on whether one regards the entropy density or the temperature as the thermodynamic quantity being held fixed, and then suppressed, in defining σ .

In reference 9 it is shown that (4.6) implies

$$\mathfrak{h}(F)^T = \text{grad}_F \sigma(F). \quad (4.7)$$

Once (4.7) is established, our assumption of Material Objectivity, *i.e.* (2.2), is equivalent to the assertion that σ in (4.6) obey the identity

$$\sigma(QF) = \sigma(F) \quad (4.8)$$

for all F and all orthogonal Q .²³

An alternative formulation of our TI is the following assertion.²⁴ Consider the Class \mathfrak{C} of continuous rectifiable curves F_t in the space of all invertible tensors F , and let the parameter t for these curves vary from 0 to 1. Let \mathfrak{C}' be the set of all curves F_t in \mathfrak{C} for which $F_1 \neq F_0$ and $F_1 F_0^{-1}$ is both symmetric and positive-definite. Then, the following inequality must hold for curves F_t in \mathfrak{C}' :

$$\int_{F_0}^{F_1} \text{tr}\{\mathfrak{h}(F_t)^T dF_t\} > \text{tr}\{(F_1 - F_0) \mathfrak{h}(F_0)^T\}. \quad (4.9)$$

The integral on the left in (4.9) is to be interpreted as a line integral along the curve F_t from $t=0$ to $t=1$.

The quantities appearing on each side of (4.9) represent work, per unit volume in the reference configuration, done against contact forces at a material point X as the local configuration of X is deformed along path $F_t M_0$, $0 \leq t \leq 1$. The quantity on the left in (4.9) gives the "true" work done, *i.e.* the work done assuming that at each t the contact forces on each material surface at X are those which one calculates using the stress tensor $S = \mathfrak{g}(F_t)$ and the actual

²³ In essence, the proof of (4.8) is given in Theorem I of W. NOLL, J. Rational Mech. Anal. 4, 3 (1955); see also § 1 of reference 11.

²⁴ The proof that (4.9) is equivalent to the assertion that σ exists and obeys (4.5) is given in reference 11. In writing (4.9) we have made use of the fact that $\mathfrak{g}(F)$ is symmetric. This enables us to simplify Eqs. (1.4) and the inequality (2.3) of reference 11.

configuration of the surface at time t ; the quantity on the right in (4.9), however, gives the work which would be done along the path $F_i M_o$ if the contact forces were to remain fixed at their initial values.²⁵ Thus, (4.9) states that \mathfrak{h} must be such that contact forces always change in a process which results in a pure stretch, and, furthermore, they always change in such a way that the work done against them is greater than that which would have been done had they remained fixed.

Let us now return to (4.6). Of course, when F^*F^{-1} is symmetric and positive definite, so is FF^{*-1} , and we can interchange F and F^* in (4.6). If we do this and add the resulting inequality to (4.6), the terms involving σ cancel, and we get, as do TRUESDELL & TOUPIN¹²,

$$\text{tr}\{(F^* - F)[\mathfrak{h}(F^*) - \mathfrak{h}(F)]^T\} > 0. \quad (4.10)$$

We now state the

Weakened Thermostatic Inequality WTI. The response function \mathfrak{h} is such that, for all pairs of tensors F^*, F for which $F^* \neq F$ and F^*F^{-1} is positive-definite and symmetric, we have the inequality (4.10).

The WTI is equivalent to demanding that

$$\text{tr}\{(U - I)F[\mathfrak{h}(UF) - \mathfrak{h}(F)]^T\} > 0 \quad (4.11)$$

for all F and all positive-definite symmetric $U \neq I$. Since the reference configuration can always be chosen such that $F = I$, the WTI is also equivalent to the requirement that

$$\text{tr}\{(U - I)[\mathfrak{h}(U) - \mathfrak{h}(I)]\} > 0 \quad (4.12)$$

for all positive-definite symmetric U , no matter what reference configuration is used in defining the response function \mathfrak{h} .

Of course the TI implies the WTI; the WTI does not imply TI. It is possible to find (theoretical) materials obeying our WTI but yet such that \mathfrak{h} is not derivable from an energy function σ through (4.7).

Remark. To say that a material obeys the TI is equivalent to the assertion that the response function \mathfrak{h} is such that (4.9) holds for *all* curves F_i in \mathfrak{C} . To say that a material obeys the WTI, however, is to assert only that \mathfrak{h} is such that (4.9) holds for those curves F_i in \mathfrak{C}' which are straight lines.²⁶

It follows from these observations that any special result derived from the WTI inequality (4.10) will also hold for materials obeying the TI inequality (4.6). Since most of the known implications of the TI, and also the implications of the TI which we wish to report here, can hold also under the weaker (*i.e.* more general) assumption of the WTI, we shall base our present discussion of thermostatics upon the WTI. Throughout the rest of this paper, if the WTI is not mentioned in the statement of a proposition, it is to be understood to be present as an axiom.

²⁵ Note that, when the configuration of a surface is changing, keeping contact forces fixed is not equivalent to keeping the stress tensor S fixed; rather, it is equivalent to keeping the Piola-Kirchhoff tensor T fixed.

²⁶ BRAGG, L. E., & B. D. COLEMAN, *J. Math. Phys.* **4**, 1074 (1963).

5. States of Pressure

Here we obtain some results which show that the WTI places limitations on the properties of local configurations which give rise to a hydrostatic pressure. Our main new result is Theorem 3, which tells us, among other things, that the stress tensor can vanish only in an undistorted configuration of a solid.

Lemma. *Consider two local configurations M_1 and M_2 of the same material, and suppose that they both correspond to hydrostatic pressures:*

$$S_1 = -p_1 I, \quad (5.1a)$$

$$S_2 = -p_2 I. \quad (5.1b)$$

Here p_1 may or may not equal p_2 . Consider the right stretch U from M_1 to M_2 ; i.e., put

$$M_2 = F M_1, \quad F = R U, \quad (5.2)$$

where R is orthogonal and U positive-definite and symmetric. Then, the following inequality must hold whenever $U \neq I$:

$$p_1(\operatorname{tr} U - 3) + p_2(\det U)(\operatorname{tr} U^{-1} - 3) > 0. \quad (5.3)$$

Proof. We use $M_1 = M_0$ as the reference configuration. By Eqs. (4.1), the Piola-Kirchhoff tensors corresponding to M_1 and M_2 are

$$T_1 = S_1 = -p_1 I, \quad (5.4a)$$

$$T_2 = |\det F| (F^{-1} S_2)^T = -p_2(\det U) R U^{-1}. \quad (5.4b)$$

Hence, by (4.5), we obtain

$$\mathfrak{h}(I) = -p_1 I, \quad \mathfrak{h}(U) = -p_2(\det U) U^{-1}. \quad (5.5)$$

Substitution of Eqs. (5.5) into the form (4.12) of the WTI gives

$$\operatorname{tr}\{(U - I)[(-p_2)(\det U) U^{-1} + p_1 I]\} > 0, \quad (5.6)$$

whenever $U \neq I$; (5.6) is equivalent to (5.3), *q.e.d.*

Theorem 1. *Suppose that two local configurations M_1 and M_2 of a material correspond to the same hydrostatic stress $-pI$ with $p > 0$. Suppose further that M_1 and M_2 differ by more than an orthogonal transformation, i.e., that the principal stretches u_i , $i = 1, 2, 3$ from M_1 to M_2 are not all 1. Then the following two situations are both impossible:*

$$u_i \geq 1 \quad \text{for all } i, \quad (5.7a)$$

$$u_i \leq 1 \quad \text{for all } i. \quad (5.7b)$$

Proof. We use the Lemma, putting $p_1 = p_2 = p$ in (5.1). Of course, the u_i mentioned here are just the proper numbers of U defined in (5.2). Now the inequality (5.3) states that if $U \neq I$, i.e. if the u_i are not all 1, then

$$p g > 0, \quad (5.8)$$

where

$$\begin{aligned} g &= (\operatorname{tr} U - 3) + (\det U)(\operatorname{tr} U^{-1} - 3) \\ &= (u_1 + u_2 + u_3 - 3) + (u_2 u_3 + u_1 u_3 + u_2 u_3 - 3 u_1 u_2 u_3). \end{aligned} \quad (5.9)$$

Since, by hypothesis, $p > 0$, the inequality (5.8) implies

$$g > 0. \quad (5.10)$$

Differentiation of Eq. (5.9) yields equations of the type

$$\frac{\partial g}{\partial u_j} = u_k + u_l + 1 - 3u_k u_l, \quad (5.11)$$

where j, k , and l are all distinct. Suppose now that the inequalities (5.7a) hold. We then also have inequalities of the form

$$u_k u_l \geq \frac{u_k + u_l + 1}{3}, \quad k \neq l, \quad (5.12)$$

and hence, by Eq. (5.11),

$$\frac{\partial g}{\partial u_j} \leq 0, \quad j = 1, 2, 3. \quad (5.13)$$

Since g reduces to 0 when $u_1 = u_2 = u_3 = 1$, it follows from (5.13) that $g \leq 0$ when (5.7a) holds, which contradicts (5.10). Hence the inequalities (5.7a) are impossible. To show that the inequalities (5.7b) are also impossible, we need only interchange the roles of the configurations M_1 and M_2 , *q.e.d.*

Corollary. *Suppose two local configurations M_1 and M_2 give rise to the same positive hydrostatic pressure p and differ only by a similarity transformation $F = \alpha Q$, Q orthogonal, $\alpha > 0$. Then $\alpha = 1$; i.e., M_1 and M_2 must have the same density and can differ by only a rotation or reflection.*

Theorem 2. *Suppose that, for a given material, M_1 and M_2 are two local configurations which have the same density and give rise to hydrostatic pressures p_1 and p_2 , respectively, so that Eqs. (5.1) hold. If $p_1 \leq 0$ and $p_2 \leq 0$, then M_1 and M_2 can differ by only an orthogonal transformation, and we must have $p_1 = p_2$.*

Proof. Since here M_1 and M_2 have the same density, the right stretch tensor U of (5.2) satisfies

$$\det U = \det U^{-1} = 1. \quad (5.14)$$

Thus, if $U \neq I$, the inequality (5.3) becomes

$$p_1(\operatorname{tr} U - 3) + p_2(\operatorname{tr} U^{-1} - 3) > 0. \quad (5.15)$$

The arithmetic mean of a set of positive numbers is strictly greater than the geometric mean unless the numbers are all equal to one. Applied to the proper numbers of a positive-definite, symmetric tensor A , this observation states that

$$\sqrt[3]{\det A} < \frac{\operatorname{tr} A}{3} \quad \text{if } A \neq I. \quad (5.16)$$

On putting $A = U$ and $A = U^{-1}$ in (5.16) and using (5.14), we obtain

$$\operatorname{tr} U - 3 > 0, \quad \operatorname{tr} U^{-1} - 3 > 0, \quad (5.17)$$

whenever $U \neq I$. If $p_1 \leq 0$ and $p_2 \leq 0$, the inequalities (5.15) and (5.17) are inconsistent. Hence $U = I$, which means that M_1 and M_2 can differ only by an orthogonal transformation $F = R$ (see Eq. (5.2)). Moreover, using Eqs. (2.1)

and (2.2) with M_1 as the reference configuration, $F=I$ and $Q=R$, we find

$$S_1 = -p_1 I = g(I), \quad (5.18a)$$

$$S_2 = -p_2 I = g(R) = Rg(I)R^{-1} = R(-p_1 I)R^{-1} = -p_1 I; \quad (5.18b)$$

i.e. $p_1 = p_2$, q.e.d.

Corollary. *If $g(H) = g(I) = -pI$ where H is unimodular but non-orthogonal, then $p > 0$.*

Theorem 3. *If there exists a local configuration M_o such that the corresponding stress S_o vanishes or is a negative hydrostatic pressure, then the material under consideration is a solid, and M_o is an undistorted configuration.*

The theorem states that if $S_o = -pI$, with $p \leq 0$, then the isotropy group \mathcal{I} relative to M_o is a subgroup of \mathcal{O} .

Proof. By the definition of \mathcal{I} , if H is in \mathcal{I} , then $|\det H| = 1$ and $g(H) = g(I) = S_o$. Since we here have $S_o = -pI$ with $p \leq 0$, the Corollary to Theorem 2 shows that H must be orthogonal, q.e.d.

Let us now consider elastic fluids. On turning back to Proposition 7 of Section 2 and observing that the stress in a fluid is characterized by a scalar-valued function p of a scalar ϱ , we see that our WTI is equivalent to our TI for a fluid.²⁷ It follows from Theorem 3 that $p(\varrho)$ in (2.12) is always positive. Furthermore, it follows from Theorem 1 that $p(\varrho)$ is an invertible function. These observations are sharpened in the following theorem.¹¹

Theorem 4. *For an elastic fluid, a necessary and sufficient condition for the WTI (and hence also for the TI) is that the function $p(\varrho)$ in (2.12) obey the following inequalities for all ϱ :*

$$p(\varrho) > 0, \quad (5.19)$$

$$\varrho \frac{d p(\varrho)}{d \varrho} \geq \frac{2}{3} p(\varrho). \quad (5.20)$$

In (5.20) it is to be understood that equality holds on only a nowhere dense set of values of ϱ .

Theorem 4 has the same content as Theorem 6 of reference 9, as is easily seen by observing that (5.20) is equivalent to the statement that the internal energy density in a fluid is a convex function of the cube root of the specific volume.²⁸

Various implications of the WTI for isotropic materials are derived by COLEMAN & NOLL²⁹, TRUESDELL & TOUPIN¹², BRAGG & COLEMAN³⁰, and NOLL & TRUESDELL^{31, 32}.

²⁷ This is not the case when thermodynamic variables such as the temperature or entropy density are considered.

²⁸ An extension of the inequality (5.20) to a statement about the mean pressure on arbitrary states of strain in general elastic materials is given in reference 26.

²⁹ Section 12 of reference 9.

³⁰ BRAGG, L. E., & B. D. COLEMAN, *J. Math. Phys.* **4**, 424 (1963).

³¹ NOLL, W., & C. TRUESDELL, in: *Encyclopedia of Physics*, Edited by S. FLÜGGE, Vol. III/3. Berlin-Göttingen-Heidelberg: Springer (forthcoming).

³² For its discussion of inequalities in isotropic materials the article of TRUESDELL & TOUPIN¹² is the most exhaustive currently in print. It also contains some interesting remarks on uniqueness questions for boundary value problems in materials of arbitrary symmetry.

6. Differential Inequalities

We assume now that the response function g of Eq. (2.1) or, equivalently, the response function \mathfrak{h} of Eq. (4.2) is continuously differentiable and investigate the restriction that the WTI imposes on the gradients of g and \mathfrak{h} .

Let D be an arbitrary non-zero symmetric tensor, and consider

$$U = I + \tau D, \quad (6.1)$$

where τ is a real parameter $\neq 0$. It is clear that when $|\tau|$ is sufficiently small the tensor U given by Eq. (6.1) is not only symmetric but also positive-definite and $\neq I$. Let us make the abbreviation

$$f(\tau) = \text{tr}\{D\mathfrak{h}(I + \tau D)^T\}. \quad (6.2)$$

Starting with (4.12) and letting U be given by (6.1), we find that the WTI may be written

$$\tau[f(\tau) - f(0)] > 0. \quad (6.3)$$

Dividing (6.3) by τ^2 and taking the limit $\tau \rightarrow 0$, we obtain

$$\dot{f}(0) \geq 0, \quad (6.4)$$

where the superimposed dot denotes the derivative. The directional derivative $\nabla\mathfrak{h}[D]$ of \mathfrak{h} at I , in the direction of the tensor D , is defined by

$$\left. \frac{d}{d\tau} \mathfrak{h}(I + \tau D) \right|_{\tau=0} = \nabla\mathfrak{h}[D]. \quad (6.5)$$

Here $\nabla\mathfrak{h}$, the *gradient* of \mathfrak{h} at I , is a linear transformation of the six-dimensional space of symmetric tensors D into the nine-dimensional space of arbitrary tensors. We now define a quadratic form \mathfrak{Q} on the six-dimensional space of all symmetric tensors D by

$$\mathfrak{Q}(D) = \text{tr}\{D(\nabla\mathfrak{h}[D])^T\}. \quad (6.6)$$

Equations (6.2), (6.5), and (6.6) show that the inequality (6.4) states that $\mathfrak{Q}(D) \geq 0$. Thus we have

Theorem 1. *In order for the WTI to hold it is necessary that the quadratic form \mathfrak{Q} defined by (6.6) be positive-semidefinite, i.e. that*

$$\mathfrak{Q}(D) \geq 0 \quad (6.7)$$

for all symmetric tensors D .

Of course, the response function \mathfrak{h} and hence the quadratic form \mathfrak{Q} depend on the choice of the reference configuration; \mathfrak{Q} must be positive semi-definite for every such choice.

To cast (6.7) into a more transparent form we consider a smooth one-parameter family of deformation gradients $F(\tau)$ such that $F(0) = I$. The corresponding rotation tensors are denoted by $R(\tau)$ and the corresponding right stretch tensors by $U(\tau)$. We then have $R(0) = U(0) = I$. If we put

$$\dot{R}(0) = W, \quad \dot{U}(0) = D, \quad (6.8)$$

we find, by differentiating Eq. (1.7) with respect to τ and then putting $\tau = 0$, that

$$\dot{F}(0) = W + D. \quad (6.9)$$

Moreover, it is clear that D is symmetric and it is easy to show that W is skew (by differentiating $R^T R = I$).

Let $S(\tau)$ and $T(\tau)$ be the stress and the Piola-Kirchhoff tensor, respectively, corresponding to $F(\tau)$. Equation (4.1) shows that $T(0) = S(0)$. We use the notation

$$S(0) = T(0) = S_o, \quad \dot{S}(0) = \dot{S}_o, \quad \dot{T}(0) = \dot{T}_o. \quad (6.10)$$

On differentiating (4.5) we find

$$\dot{T}_o = W \mathfrak{h}(I) + \nabla \mathfrak{h}[D], \quad (6.11)$$

where $\nabla \mathfrak{h}$ is the gradient of \mathfrak{h} at I and where the notations (6.9) and (6.10) have been used. Since $S_o = \mathfrak{h}(I)$, Eq. (6.11) is equivalent to

$$\dot{T}_o - W S_o = \nabla \mathfrak{h}[D]. \quad (6.12)$$

Let us denote the density of the local configuration corresponding to $F(\tau)$ by $\varrho(\tau)$ and use the notation $\varrho_o = \varrho(0)$, $\dot{\varrho}_o = \dot{\varrho}(0)$. If we differentiate Eq. (4.1) with respect to τ and then put $\tau = 0$, we find, using (6.8)–(6.10),

$$\varrho_o \dot{S}_o = \dot{\varrho}_o S_o + \varrho_o S_o (D - W) + \varrho_o \dot{T}_o. \quad (6.13)$$

The equation of continuity³³ implies $\dot{\varrho}_o = -(\text{tr } D) \varrho_o$, hence (6.13) is equivalent to

$$\dot{S}_o = -S_o (\text{tr } D) + S_o (D - W) + \dot{T}_o. \quad (6.14)$$

Combining Eqs. (6.6), (6.12), and (6.14), we find

$$\mathbf{\Omega}(D) = \text{tr} \{ (\dot{S}_o + S_o W - W S_o) D \} + \text{tr } D \text{tr} (S_o D) - \text{tr} (S_o D^2). \quad (6.15)$$

If we differentiate Eq. (2.3) with respect to τ and then put $\tau = 0$, we find

$$\dot{S}_o + S_o W - W S_o = \nabla \mathfrak{g}[D], \quad (6.16)$$

where $\nabla \mathfrak{g}$ is the gradient of the response function \mathfrak{g} at the tensor I . We may regard $\nabla \mathfrak{g}$ as a linear transformation of the six-dimensional space of symmetric tensors into itself. When components are used, the matrix of $\nabla \mathfrak{g}$ has 36 components, which are the elastic coefficients for an infinitesimal deformation from the reference configuration. Of course, these elastic coefficients depend on the choice of the reference configuration.

It follows from Eqs. (6.15) and (6.16) that the quadratic form $\mathbf{\Omega}(D)$ may be expressed in terms of $\nabla \mathfrak{g}$ by

$$\mathbf{\Omega}(D) = \text{tr} \{ D \nabla \mathfrak{g}[D] \} + (\text{tr } D) \text{tr} (S_o D) - \text{tr} (S_o D^2). \quad (6.17)$$

We call a one-parameter family $M = M(t)$ of local configurations a *deformation path*. If we define $F(\tau)$ by $M(t + \tau) = F(\tau) M(t)$, the tensors D and W are called the *stretching* and the *spin* of the deformation path. In the notation of Eqs. (6.10)–(6.17) the stress corresponding to $M(t)$ is S_o . We now omit the subscript o and note that Eq. (6.15) and Theorem 1 yield

Theorem 2. *In order for the WTI to hold it is necessary that for every deformation path*

$$\text{tr} (S^* D) \geq 0, \quad (6.18)$$

³³ Or, equivalently, (1.3).

where D is the stretching, and S^* is the invariant stress rate defined by

$$S^* = \dot{S} + SW - WS + S \operatorname{tr} D - \frac{1}{2}(SD + DS). \quad (6.19)$$

In (6.19) S is the stress and W the spin.

In his "Recherches sur l'élasticité," P. DUHEM³⁴ makes the following assertion.

"Considérons d'abord un système défini par un certain nombre de variables normales $\alpha, \beta, \dots, \lambda$, hors la température absolue T . Supposons qu'à une certaine température T , le système prenne un état d'équilibre lorsqu'on le soumet aux actions extérieures A, B, \dots, L , et que cet état d'équilibre varie d'une manière continue lorsque, sans faire varier la température T , on fait varier les valeurs A, B, \dots, L des actions extérieures. Si les actions A, B, \dots, L éprouvent des variations infiniment petites dA, dB, \dots, dL , que nous nommerons des actions perturbatrices, les valeurs des variables $\alpha, \beta, \dots, \lambda$ qui conviennent à l'équilibre éprouvent des variations $d\alpha, d\beta, \dots, d\lambda$, que nous nommerons des perturbations; l'expression

$$dA d\alpha + dB d\beta + \dots + dL d\lambda$$

sera nommée le travail perturbateur isothermique.

"Dire qu'un travail perturbateur est positif, c'est dire, sous une forme mathématique précise, que la perturbation se produit dans le sens vers lequel tendent les actions perturbatrices. Il est clair que les systèmes que la nature nous offre seront tels, en général, que tout travail perturbateur isothermique, accompli à partir d'un état d'équilibre, soit positif. C'est ce que nous experimerons en disant qu'ils sont soumis à la loi du déplacement isothermique de l'équilibre."

Our Theorem 2 above furnishes a precise mathematization, for elastic materials, of DUHEM's "law of isothermal displacement of equilibrium".

Isotropy

We now seek the limitations which the WTI places on the elastic coefficients characterizing the response of an isotropic material to an infinitesimal deformation from an undistorted state.

Taking the gradient of Eq. (2.11) with respect to F at $F=I$, we find that

$$\nabla_{\mathfrak{g}}[QDQ^{-1}] = Q\nabla_{\mathfrak{g}}[D]Q^{-1} \quad (6.20)$$

holds for all orthogonal tensors Q in the isotropy group and all symmetric tensors D . Here we assume that \mathfrak{g} is the response function for an undistorted configuration of an isotropic material; it then follows that $\nabla_{\mathfrak{g}}[D]$ is a linear isotropic tensor function of D . It is a consequence of this remark that $\nabla_{\mathfrak{g}}$ has the representation³⁵

$$\nabla_{\mathfrak{g}}[D] = \lambda(\operatorname{tr} D)I + 2\mu D, \quad (6.21)$$

where λ and μ (the Lamé coefficients) are elastic coefficients which depend on the reference configuration. By Proposition 6 of Section 2, the stress S_0 on the undistorted reference configuration must be of the form

$$S_0 = -pI. \quad (6.22)$$

Substitution of Eqs. (6.21) and (6.22) into (6.17) gives

$$\Omega(D) = (\lambda - p)(\operatorname{tr} D)^2 + (2\mu + p)\operatorname{tr} D^2. \quad (6.23)$$

³⁴ DUHEM, P., Ann. École Norm. **22**, 143–217 (1905), p. 193.

³⁵ Cf. A. E. H. LOVE, A Treatise on the Mathematical Theory of Elasticity (Cambridge, 4th Edition, 1927), §§ 106–110. LOVE's assumption of the existence of a strain energy function does not affect the generality of his results for isotropic materials or cubic crystals.

Now, quadratic forms of the type

$$q = a(d_1 + d_2 + d_3)^2 + b(d_1^2 + d_2^2 + d_3^2) \tag{6.24}$$

are well known. A necessary and sufficient condition that q be positive semi-definite in the triple (d_1, d_2, d_3) is that both

$$b \geq 0 \quad \text{and} \quad 3a + 2b \geq 0. \tag{6.25}$$

If we take $a = \lambda - p$, $b = 2\mu + p$, and d_i to be the proper numbers of D , the quadratic form (6.24) reduces to (6.23). Therefore, a necessary and sufficient condition that $\Omega(D)$ be positive semi-definite is that

$$2\mu + p \geq 0 \quad \text{and} \quad 3(\lambda - p) + 2(2\mu + p) \geq 0. \tag{6.26}$$

From Theorem 1 we now obtain the following result.

Theorem 3. *The WTI implies that for each undistorted configuration of an isotropic material the following inequalities hold:*

$$\mu \geq -\frac{1}{2}p, \tag{6.27a}$$

$$\kappa \geq \frac{2}{3}p; \tag{6.27b}$$

here p is the pressure, μ the shear modulus, and $\kappa = \lambda + \frac{2}{3}\mu$ the compression modulus; of course p , μ , and κ can all depend on the choice of the reference configuration.

If $p = 0$, i.e., if the reference configuration is a natural state of an isotropic solid, then Theorem 3 yields the familiar assertion that

$$\mu \geq 0 \quad \text{and} \quad \kappa \geq 0. \tag{6.28}$$

For an elastic fluid we have $\mu = 0$ and $\kappa = \rho \, d\phi(\rho)/d\rho$, and hence Theorem 3 yields

$$p \geq 0, \quad \rho \frac{d\phi(\rho)}{d\rho} \geq \frac{2}{3}p.$$

This condition is just a bit weaker than the necessary and sufficient condition for the WTI stated in Theorem 4 of Section 5, for our present Theorem 3 does not contain the precious result that $p = 0$ is *impossible* in a fluid. Yet we think Theorem 3 to be not without interest. Through its conclusion (6.27a) it yields a condition on μ for a particular negative pressure p to be possible: If the configuration is varied among undistorted configurations in such a way that p decreases and becomes negative, μ must always remain greater than $-\frac{1}{2}p$, otherwise the material will not support the negative pressure.³⁶

Cubic Symmetry

When the material has cubic symmetry and the reference configuration is undistorted, it follows from (6.20) that there exists an orthonormal basis \mathbf{h}_i such that the components of $\text{Vg}[D]$ and D are related by³⁷

$$\text{Vg}[D]_{ii} = \alpha D_{ii} + \beta(D_{jj} + D_{kk}), \tag{6.29a}$$

$$\text{Vg}[D]_{ij} = 2\gamma D_{ij}. \tag{6.29b}$$

³⁶ We take this to mean that cavitation will occur, as seems to be the case with fluids.

³⁷ The result is essentially given in LOVE's treatise.³⁵ The three coefficients α, β, γ , are related to those on p. 160 of reference 35 by $C_{11} = \alpha, C_{12} = \beta, C_{44} = \gamma$.

No summation convention is used here, and i, j, k is any permutation of 1, 2, 3. It follows from the last entry in Table 2 that the stress S_o in the reference configuration must be a hydrostatic pressure:

$$S_o = -pI. \quad (6.30)$$

Substitution of (6.29) and (6.30) into (6.17) yields

$$\mathbf{\Omega}(D) = \alpha \sum_{i=1}^3 D_{ii}^2 + 2(\beta - p) \sum_{\substack{i,j=1 \\ j>i}}^3 D_{ii} D_{jj} + (4\gamma + 2p) \sum_{\substack{i,j=1 \\ j>i}}^3 D_{ij}^2. \quad (6.31)$$

Thus $\mathbf{\Omega}(D)$ is the sum of the two quadratic forms

$$\mathbf{\Omega}_1(D) = \alpha \sum_{i=1}^3 D_{ii}^2 + (2\beta - p) \sum_{\substack{i,j=1 \\ j>i}}^3 D_{ii} D_{jj}, \quad (6.32a)$$

and

$$\mathbf{\Omega}_2(D) = (4\gamma + 2p) \sum_{\substack{i,j=1 \\ j>i}}^3 D_{ij}^2. \quad (6.32b)$$

The form $\mathbf{\Omega}_1(D)$ involves the three diagonal components of D , which can be chosen at will, while the form $\mathbf{\Omega}_2(D)$ involves only the three off-diagonal components of D , which also can be chosen at will. Hence, in order for $\mathbf{\Omega}(D)$ to be positive-semidefinite it is necessary and sufficient that both of the forms (6.32) be positive-semidefinite.

It is clear that the form $\mathbf{\Omega}_2(D)$ of (6.32b) is positive-semidefinite if and only if $4\gamma + 2p \geq 0$. The form $\mathbf{\Omega}_1(D)$ of (6.32a) is of the type (6.24) when we put $a + b = \alpha$, $2a = 2\beta - p$, and $d_i = D_{ii}$, $i = 1, 2, 3$. Using (6.25), it follows that the form $\mathbf{\Omega}(D)$ of (6.31) is positive-semidefinite if and only if the three inequalities

$$2\gamma + p \geq 0, \quad \alpha - \beta + \frac{1}{2}p \geq 0, \quad \beta + 2\alpha - \frac{1}{2}p \geq 0 \quad (6.33)$$

hold. We summarize:

Theorem 4. *The WTI implies that for each undistorted configuration of a material with cubic symmetry, the elastic coefficients α, β, γ and the pressure p obey the inequalities*

$$\gamma \geq -\frac{1}{2}p, \quad \alpha - \beta \geq -\frac{1}{2}p, \quad \beta + 2\alpha \geq \frac{1}{2}p. \quad (6.34)$$

Acknowledgement. This research was supported by the Air Force Office of Scientific Research under Contract AF 49(638)541 with Mellon Institute and by the National Science Foundation under Grant NSF-G 16745 to Carnegie Institute of Technology.

Mellon Institute
Pittsburgh
and
Department of Mathematics
Carnegie Institute of Technology
Pittsburgh

(Received August 16, 1963)

EUCLIDEAN GEOMETRY AND MINKOWSKIAN CHRONOMETRY

WALTER NOLL

1. Introduction. The term "Minkowskian Chronometry" is used here for the study of the structure of space-time appropriate to Einstein's special relativity. This term suggests a parallel with Euclidean geometry. Originally, geometry was an empirical science which dealt with measuring of distances on earth. The Greeks transformed this science into a beautiful mathematical discipline. Chronometry is the science of the measurement of time intervals. Einstein's first paper was a critical study of time measurements. It was Minkowski, however, who made a geometrical discipline out of Einstein's chronometry. More recently, Synge ([2], [3]) has been an able advocate of the geometrical point of view in relativity.

The present paper is an attempt to axiomatize Minkowskian chronometry using direct coordinate-free methods. I believe that the coordinate-free approach fosters the cultivation of intuition, a scarce commodity in relativity because the phenomena this theory is intended to describe are as yet rather remote from our daily experience. I hope, moreover, that Minkowskian chronometry will become, as Euclidean geometry did, a branch of mathematics that is of interest purely for its esthetic value.

Section 2 contains a collection of definitions and results on vector spaces with an inner product that is not necessarily positive definite. We give a new proof of the inertia theorem of Sylvester, a proof that remains valid for infinite dimensional spaces.

In Section 3 we present certain inequalities valid in inner product spaces of index one. The most notable of these are the "reversed Schwarz inequality" and the "reversed triangle inequality."

Section 4 consists of an axiomatic introduction to pseudo-Euclidean geometry, based on the fact that the structure of a pseudo-Euclidean space is determined by its "separation function," which in the Euclidean case is identical to the square of the metric. Associated with each pseudo-Euclidean space is a unique "translation space," which is a vector space with inner product.

“Inner product” is just another term for “nondegenerate bilinear form.” Usually the term is used, however, only when the form is positive definite. But indefinite inner products can be treated in the same geometric spirit as definite inner products usually are, and we adopt here the terminology that has become standard so far only for the case of definite inner products (see, e.g. Halmos [5]). The theory of bilinear forms is extensive, much of it going back to the nineteenth century. In most textbooks, however, the subject is treated in analytical language with components. A notable exception is the treatise of Bourbaki [1], to which we refer for further information on the topics treated in Sections 2 and 4.

When the index of the translation space is zero, pseudo-Euclidean geometry reduces to Euclidean geometry, and when the index is one, it reduces to Minkowskian chronometry. From a physical point of view it is reasonable to use distance as a primitive notion in Euclidean geometry, because there are yardsticks to measure distances. In this case, the separation of two points is just the square of their distance. There are no “separation meters,” however, to measure the separation of two arbitrary events in relativistic space-time. In Section 5 we give an axiomatic introduction to Minkowskian chronometry based on primitive notions that have a more direct physical meaning: observers, clock-readings, signals. We show—and this is the least trivial part of the paper—that these primitive data determine the separation function uniquely. The fact that the index of the translation space is one is not an assumption of the theory but comes out as a theorem.

Section 6 deals with temporal order, i.e. with the possibility of distinguishing future from past. It turns out that such order, in Minkowskian chronometry, is related to the distinction between emission and reception of signals. We note that in classical space-time, such distinction is not sufficient to determine temporal order. The axioms given here may answer a problem proposed by Suppes ([4], Sect. 4).

We do not impose anywhere a restriction on the dimension, which may even be infinite. Of course, in the presently known physical applications of Minkowskian chronometry the dimension is 4.

2. Inner product spaces. (For details see [1].) Let \mathcal{U} be a real vector space. A nondegenerate symmetric bilinear form on \mathcal{U} will be called an inner product. The inner product of $\mathbf{u}, \mathbf{v} \in \mathcal{U}$ will be denoted by $\mathbf{u} \cdot \mathbf{v}$, and we will abbreviate $\mathbf{u} \cdot \mathbf{u} = \mathbf{u}^2$. Nondegeneracy means that

$$(2.1) \quad \mathbf{u} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \in \mathcal{U} \text{ implies } \mathbf{u} = \mathbf{0}.$$

A *quadratic form* Φ on \mathcal{U} is a function $\Phi: \mathcal{U} \rightarrow \mathcal{R}$ (\mathcal{R} = set of real numbers) such that $\Phi(\mathbf{v}) = \mathbf{v} \cdot \mathbf{v}$ for some inner product. The quadratic form determines the inner product uniquely.

An *inner product space* \mathcal{U} is a vector space with an additional structure defined by an inner product. If \mathcal{U} is a subspace of \mathcal{U} , then $\mathcal{U}^\perp = \{\mathbf{v} \mid \mathbf{v} \cdot \mathbf{u} = 0 \text{ for all}$

$u \in \mathfrak{u}$ } is called the *orthogonal complement* of \mathfrak{u} . If $\dim \mathfrak{u}$ is finite, then $\mathfrak{u}^{\perp\perp} = \mathfrak{u}$ and $\dim \mathfrak{u} + \dim \mathfrak{u}^{\perp} = \dim \mathfrak{V}$. A subspace \mathfrak{u} is called *regular* if $\mathfrak{u} \cap \mathfrak{u}^{\perp} = \{0\}$ and *singular* if $\mathfrak{u} \cap \mathfrak{u}^{\perp} \neq \{0\}$. If $\dim \mathfrak{u}$ is finite and \mathfrak{u} is regular, then \mathfrak{V} has the direct decomposition $\mathfrak{V} = \mathfrak{u} \oplus \mathfrak{u}^{\perp}$.

A linear transformation $Q: \mathfrak{V} \rightarrow \mathfrak{V}$ is called *orthogonal* if it preserves the inner product. In view of (2.1), Q is then also one-to-one and hence an automorphism of the inner product space \mathfrak{V} .

The following terminology is suggested by the physical applications. The sets of vectors

$$(2.2) \quad \begin{aligned} \mathfrak{V}_+ &= \{ \mathbf{v} \mid \mathbf{v}^2 > 0 \text{ or } \mathbf{v} = \mathbf{0} \}, \\ \mathfrak{V}_- &= \{ \mathbf{v} \mid \mathbf{v}^2 < 0 \text{ or } \mathbf{v} = \mathbf{0} \}, \\ \mathfrak{V}_0 &= \{ \mathbf{v} \mid \mathbf{v}^2 = 0 \} \end{aligned}$$

will be called the *space-cone*, the *time-cone*, and the *signal-cone*, respectively. These cones have only the zero vector $\mathbf{0}$ in common. A vector \mathbf{v} is said to be *space-like*, *time-like*, or a *signal vector* depending on whether $\mathbf{v} \in \mathfrak{V}_+$, $\mathbf{v} \in \mathfrak{V}_-$, or $\mathbf{v} \in \mathfrak{V}_0$. The maximal dimension of the time-like subspaces of \mathfrak{V} , (i.e. subspaces contained in \mathfrak{V}_-) will be called the *index* of \mathfrak{V} ; it will be denoted by $i = \text{ind } \mathfrak{V}$. (The customary definition is $\text{ind } \mathfrak{V} = \text{maximal dimension of the signal subspaces of } \mathfrak{V}$. Our definition will give the same value if the sign of the inner product is properly adjusted, as is shown in Theorem 2 below.)

THEOREM 1. *If \mathfrak{u} is a time-like subspace of maximal dimension $i = \text{ind } \mathfrak{V}$ and if $i < \infty$, then the complement \mathfrak{u}^{\perp} is space-like and \mathfrak{V} has the direct decomposition*

$$(2.3) \quad \mathfrak{V} = \mathfrak{u} \oplus \mathfrak{u}^{\perp}, \quad \mathfrak{u} \subset \mathfrak{V}_-, \quad \mathfrak{u}^{\perp} \subset \mathfrak{V}_+.$$

In addition, for any decomposition of the type (2.3), we have $\dim \mathfrak{u} = i$.

Proof. Let $\mathfrak{u} \subset \mathfrak{V}_-$, $\dim \mathfrak{u} = i$. \mathfrak{u} is then regular and \mathfrak{V} has the direct decomposition $\mathfrak{V} = \mathfrak{u} \oplus \mathfrak{u}^{\perp}$. Assume that \mathfrak{u}^{\perp} is not space-like. Then there is a vector $\mathbf{w} \in \mathfrak{u}^{\perp}$ such that $\mathbf{w}^2 \leq 0$, $\mathbf{w} \neq \mathbf{0}$. By (2.1), there is a vector $\mathbf{v} \in \mathfrak{V}$ such that $\mathbf{w} \cdot \mathbf{v} = \alpha \neq 0$. Let

$$\mathbf{v} = \mathbf{u} + \mathbf{z}, \quad \mathbf{u} \in \mathfrak{u}, \quad \mathbf{z} \in \mathfrak{u}^{\perp}$$

be the decomposition of \mathbf{v} . We have $\alpha = \mathbf{w} \cdot \mathbf{v} = \mathbf{w} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{z} = \mathbf{w} \cdot \mathbf{z}$. Since $\mathbf{w}, \mathbf{z} \in \mathfrak{u}^{\perp}$, the vector $\mathbf{p} = \mathbf{z} + \beta \mathbf{w}$ also belongs to \mathfrak{u}^{\perp} for any choice of β . Now,

$$\mathbf{p}^2 = \mathbf{z}^2 + 2\beta \mathbf{w} \cdot \mathbf{z} + \beta^2 \mathbf{w}^2 \leq \mathbf{z}^2 + 2\beta \alpha.$$

Since $\alpha \neq 0$, we can adjust β such that $\mathbf{p}^2 < 0$. If $\mathbf{x} \in \mathfrak{u}$ and $\mathbf{y} = \mathbf{x} + \lambda \mathbf{p}$, then $\mathbf{y}^2 = \mathbf{x}^2 + \lambda^2 \mathbf{p}^2 < 0$, unless $\mathbf{y} = \mathbf{0}$. Therefore, the space \mathfrak{u}' spanned by \mathfrak{u} and \mathbf{p} is time-like and of dimension $i+1$, which contradicts the assumption that \mathfrak{u} is of maximal dimension.

Consider now an arbitrary decomposition of the form (2.3) and assume that $\hat{\mathfrak{u}}$ is a time-like subspace of maximal dimension i . Every $\hat{\mathbf{u}} \in \hat{\mathfrak{u}}$ has a unique

decomposition $\hat{u} = u + z, u \in \mathfrak{u}, z \in \mathfrak{u}^\perp$. This decomposition defines a linear transformation $\hat{u} \rightarrow u = L\hat{u}$ of $\hat{\mathfrak{u}}$ into \mathfrak{u} . If $u = 0$, then $\hat{u} = 0 + z \in \mathfrak{u}^\perp \subset \mathfrak{V}_+$, and hence $\hat{u} = 0$. Thus L is one-to-one, which implies $i = \dim \hat{\mathfrak{u}} \leq \dim \mathfrak{u}$. On the other hand, $\dim \mathfrak{u} \leq \dim \hat{\mathfrak{u}} = i$ by definition of i . Q.E.D.

In the case when $\dim \mathfrak{V}$ is finite, Theorem 1 reduces to a version of the classical inertia theorem of Sylvester (cf. [1], Chapter 7, no. 2).

Lincoln E. Bragg has disclosed to me a counterexample which shows that the conclusion of Theorem 1 may be false when $i = \infty$. From now on we assume that the index i of \mathfrak{V} is finite ($\dim \mathfrak{V}$ need not be finite).

COROLLARY. *Let a decomposition of the type (2.3) be given. Then any other decomposition of the same type has the form*

$$(2.4) \quad \mathfrak{V} = Q(\mathfrak{u}) \oplus Q(\mathfrak{u}^\perp),$$

where Q is orthogonal.

THEOREM 2. *If $\dim \mathfrak{V} - i \geq i$, then the maximal dimension of the signal subspaces of \mathfrak{V} is also given by i .*

Proof. Let \mathfrak{s} be a signal subspace of maximal dimension i' , and consider a decomposition of the form (2.3). By Theorem 1 we have $\dim \mathfrak{u} = i, \dim \mathfrak{u}^\perp = \dim \mathfrak{V} - i \geq i$. The unique decomposition $\mathfrak{s} = u + z, u \in \mathfrak{u}, z \in \mathfrak{u}^\perp$ defines two linear transformations $\mathfrak{s} \rightarrow u = L\mathfrak{s}$ and $\mathfrak{s} \rightarrow z = N\mathfrak{s}$ of \mathfrak{s} into \mathfrak{u} and \mathfrak{u}^\perp , respectively. It is easily seen that both L and N are one-to-one. It follows that $i' = \dim \mathfrak{s} \leq \dim \mathfrak{u} = i$. If $i' < i$, $L(\mathfrak{s})$ and $N(\mathfrak{s})$ must be proper subspaces of \mathfrak{u} and \mathfrak{u}^\perp , respectively. It is then possible to find $x \in \mathfrak{u}$ and $y \in \mathfrak{u}^\perp$ such that x is orthogonal to $L(\mathfrak{s})$ and y is orthogonal to $N(\mathfrak{s})$. If x and y are normalized such that $x^2 = -1, y^2 = +1$, the space spanned by \mathfrak{s} and $x + y$ is easily seen to be a signal subspace of dimension $i' + 1$, which contradicts the definition of i' . Q.E.D.

3. Spaces of index one. From now on we assume that \mathfrak{V} is an inner product space of index one. The following two theorems are corollaries of the results of the previous section.

THEOREM 1. *Let l be a time-like unit vector ($l^2 = -1$). Then every vector $v \in \mathfrak{V}$ has unique decomposition of the form*

$$(3.1) \quad v = \xi l + w, \quad w \cdot l = 0, \quad w \in \mathfrak{V}_+.$$

THEOREM 2. *If a vector is orthogonal to a nonzero time-like vector then it must be space-like.*

In an indefinite inner product space ($\text{ind } \mathfrak{V} \neq 0$), Schwarz's inequality is of course not valid. When $\text{ind } \mathfrak{V} = 1$ the following two theorems partially replace the Schwarz inequality:

THEOREM 3 (reversed Schwarz inequality). *If \mathbf{u} and \mathbf{v} are both time-like, then*

$$(3.2) \quad (\mathbf{u} \cdot \mathbf{v})^2 \geq \mathbf{u}^2 \mathbf{v}^2,$$

and equality holds only when \mathbf{u} and \mathbf{v} are linearly dependent.

THEOREM 4. *For any three nonzero time-like vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} ,*

$$(3.3) \quad (\mathbf{u} \cdot \mathbf{v})(\mathbf{v} \cdot \mathbf{w})(\mathbf{w} \cdot \mathbf{u}) < 0.$$

Proof. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}_-$, all $\neq \mathbf{0}$. Consider

$$\mathbf{z} = \alpha \mathbf{u} - \beta \mathbf{v}, \quad \alpha = \mathbf{v} \cdot \mathbf{w}, \quad \beta = \mathbf{u} \cdot \mathbf{w}.$$

We have $\mathbf{z} \cdot \mathbf{w} = \alpha \beta - \beta \alpha = 0$. It follows from Theorem 2 that

$$0 \leq \mathbf{z}^2 = \alpha^2 \mathbf{u}^2 + \beta^2 \mathbf{v}^2 - 2\alpha\beta \mathbf{u} \cdot \mathbf{v},$$

i.e.

$$(3.4) \quad 2(\mathbf{v} \cdot \mathbf{w})(\mathbf{u} \cdot \mathbf{w})(\mathbf{u} \cdot \mathbf{v}) \leq \alpha^2 \mathbf{u}^2 + \beta^2 \mathbf{v}^2.$$

Here equality can hold only when $\mathbf{z} = \mathbf{0}$, i.e. when \mathbf{u} and \mathbf{v} are linearly dependent. Since \mathbf{w} is time-like and $\neq \mathbf{0}$, Theorem 2 shows that $\alpha = \mathbf{v} \cdot \mathbf{w}$ and $\beta = \mathbf{u} \cdot \mathbf{w}$ cannot be zero. Therefore, the right-hand side of (3.4) is negative, which proves Theorem 4. Theorem 3 is trivial when $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$. Otherwise it follows from (3.4) when we put $\mathbf{w} = \mathbf{u}$ and observe that $\mathbf{u}^2 < 0$. Q.E.D.

Consider the relation $\mathbf{u} \cdot \mathbf{v} < 0$ within the set $\hat{\mathcal{V}}_-$ of all nonzero time-like vectors. Theorem 4 shows that this relation is transitive. Clearly, it is also reflexive and symmetric. Hence $\mathbf{u} \cdot \mathbf{v} < 0$ is an equivalence relation in $\hat{\mathcal{V}}_-$. The vectors $\mathbf{u} \in \hat{\mathcal{V}}_-$ and $-\mathbf{u} \in \hat{\mathcal{V}}_-$ belong to different equivalence classes, hence there are at least two such classes. But there are no more than two classes, because if $\mathbf{u}, \mathbf{v}, \mathbf{w}$ belonged to three different classes, then the inner products on the left side of (3.3) would be all positive, which contradicts (3.3). We adjoin the zero vector $\mathbf{0}$ to each of the two equivalence classes and denote the resulting sets by \mathcal{V}_-^1 and \mathcal{V}_-^2 , respectively. It is easily verified that \mathcal{V}_-^1 and \mathcal{V}_-^2 are in fact convex cones. We summarize:

THEOREM 5. *The time cone \mathcal{V}_- is the union of two convex cones \mathcal{V}_-^1 and \mathcal{V}_-^2 which have only the zero vector in common. Two time-like vectors \mathbf{u}, \mathbf{v} belong to the same cone if and only if $\mathbf{u} \cdot \mathbf{v} \leq 0$. If $\mathbf{u} \in \mathcal{V}_-$ belongs to one of the cones, then $-\mathbf{u}$ belongs to the other, i.e., $\mathcal{V}_-^1 = -\mathcal{V}_-^2$.*

We call \mathcal{V}_-^1 and \mathcal{V}_-^2 the two *directed time cones*.

If \mathbf{v} is time-like, we call $\tau(\mathbf{v}) = \sqrt{-\mathbf{v}^2}$ the *duration* of \mathbf{v} . We have $\tau(\mathbf{v}) \geq 0$, and $\tau(\mathbf{v}) = 0$ only if $\mathbf{v} = \mathbf{0}$. The following "reversed triangle inequality" is a consequence of Theorems 3 and 5. The proof is analogous to that of the ordinary triangle inequality.

THEOREM 6 (reversed triangle inequality). *If the two vectors \mathbf{u} and \mathbf{v} both belong to one of the directed time cones ($\mathbf{u} \cdot \mathbf{v} \leq 0$), then*

$$(3.5) \quad \tau(\mathbf{u} + \mathbf{v}) \geq \tau(\mathbf{u}) + \tau(\mathbf{v}),$$

and equality holds only when \mathbf{u} and \mathbf{v} are linearly dependent.

Remark. The reversed triangle inequality expresses what is often misleadingly called “the relativistic clock paradox.”

4. Pseudo-Euclidean Geometry. Let a set \mathcal{E} of points x, y, \dots and a function

$$(4.1) \quad \sigma: \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{R}$$

be given, where \mathcal{R} denotes the real line. The function σ defines a certain structure on \mathcal{E} . The automorphisms of this structure are the one-to-one mappings α of \mathcal{E} onto itself which satisfy

$$(4.2) \quad \sigma(\alpha(x), \alpha(y)) = \sigma(x, y)$$

for all $x, y \in \mathcal{E}$. We denote by \mathcal{G} the group of all these automorphisms. We impose restrictive conditions on σ by assuming that \mathcal{G} contains a subgroup \mathcal{U} which satisfies the axioms (E₁)–(E₄) stated below.

(E₁) \mathcal{U} is commutative.

(E₂) \mathcal{U} is transitive.

(E₃) If $\mathbf{v} \in \mathcal{U}$ maps some point $x \in \mathcal{E}$ onto itself, then \mathbf{v} is the identity mapping.

We write the group operation in \mathcal{U} additively and denote the identity mapping by $\mathbf{0}$. We write $x + \mathbf{v} \in \mathcal{E}$ for the image $\mathbf{v}(x)$ of $x \in \mathcal{E}$ under $\mathbf{v} \in \mathcal{U}$. It follows from (E₂) and (E₃) that for any two points $x, y \in \mathcal{E}$ there is a unique $\mathbf{v} \in \mathcal{U}$ which maps x onto y . The mapping \mathbf{v} determined in this way will be denoted by $\mathbf{v} = y - x$. The “sum” $x + \mathbf{v}$ and “point-difference” $y - x$ thus defined obey the rules suggested by the notation.

The following proposition is easily established: The value $\sigma(x, y)$ depends only on the point difference $y - x$; i.e., there is a function $\Phi: \mathcal{U} \rightarrow \mathcal{R}$ such that

$$(4.3) \quad \Phi(y - x) = \sigma(x, y)$$

for all $x, y \in \mathcal{E}$.

The last condition required of \mathcal{U} is the following.

(E₄) \mathcal{U} is the underlying additive group of a real vector space and Φ is a non-degenerate quadratic form on \mathcal{U} ; i.e., \mathcal{U} can be given the structure of an inner product space such that

$$(4.4) \quad (y - x)^2 = (y - x) \cdot (y - x) = \sigma(x, y).$$

In view of (E₄) we refer to the automorphism \mathbf{v} in \mathcal{U} as *vectors*.

UNIQUENESS THEOREM. *There is at most one subgroup \mathcal{U} of the automorphism group \mathcal{G} such that \mathcal{U} satisfies the axioms (E₁)–(E₄). If such a subgroup exists then its structure as an inner product space, as required for (E₄), is unique.*

Proof. Assume that \mathcal{U} and $\hat{\mathcal{U}}$ are subgroups of \mathcal{G} which both satisfy (E_1) – (E_4) . If two points $x, y \in \mathcal{E}$ are given, we denote by $y-x$ the unique vector in \mathcal{U} which maps x onto y and by $y \hat{\triangle} x$ the unique vector in $\hat{\mathcal{U}}$ which maps x onto y .

We choose a fixed point $q \in \mathcal{E}$ and define a mapping f of \mathcal{U} into $\hat{\mathcal{U}}$ by

$$(4.5) \quad \hat{\mathbf{v}} = f(\mathbf{v}) = (q + \mathbf{v}) \hat{\triangle} q,$$

so that

$$(4.6) \quad q + \mathbf{v} = q + f(\mathbf{v}) = q + \hat{\mathbf{v}}$$

for all $\mathbf{v} \in \mathcal{U}$. It is clear that f is one-to-one and onto, and that $f(\mathbf{0}) = \hat{\mathbf{0}} = \mathbf{0}$.

By (4.4) we have

$$(4.7) \quad \sigma(x, y) = (y - x)^2 = (y \hat{\triangle} x)^2.$$

Substituting $x = q + \mathbf{u}$ and $y = q + \mathbf{v}$ into (4.7) we find that

$$(4.8) \quad (\mathbf{v} - \mathbf{u})^2 = (\hat{\mathbf{v}} - \hat{\mathbf{u}})^2$$

holds for all $\mathbf{u}, \mathbf{v} \in \mathcal{U}$. In particular, when $\mathbf{u} = \mathbf{0} = \hat{\mathbf{u}} = \hat{\mathbf{0}}$, (4.8) shows that $\mathbf{v}^2 = \hat{\mathbf{v}}^2$ for all $\mathbf{v} \in \mathcal{U}$. Hence, if we expand (4.8),

$$(4.9) \quad \mathbf{v}^2 - 2\mathbf{v} \cdot \mathbf{u} + \mathbf{u}^2 = \hat{\mathbf{v}}^2 - 2\hat{\mathbf{v}} \cdot \hat{\mathbf{u}} + \hat{\mathbf{u}}^2,$$

the square terms cancel and we obtain

$$(4.10) \quad \mathbf{v} \cdot \mathbf{u} = \hat{\mathbf{v}} \cdot \hat{\mathbf{u}} = f(\mathbf{v}) \cdot \hat{\mathbf{u}}.$$

Repeated use of (4.10) shows that

$$\begin{aligned} f(\alpha\mathbf{v} + \beta\mathbf{w}) \cdot \hat{\mathbf{u}} &= (\alpha\mathbf{v} + \beta\mathbf{w}) \cdot \mathbf{u} = \alpha(\mathbf{v} \cdot \mathbf{u}) + \beta(\mathbf{w} \cdot \mathbf{u}) \\ &= [\alpha f(\mathbf{v}) + \beta f(\mathbf{w})] \cdot \hat{\mathbf{u}}. \end{aligned}$$

Hence

$$[\alpha f(\mathbf{v}) + \beta f(\mathbf{w}) - f(\alpha\mathbf{v} + \beta\mathbf{w})] \cdot \hat{\mathbf{u}} = 0,$$

which holds for all $\hat{\mathbf{u}} \in \hat{\mathcal{U}}$. Since (E_4) requires that the inner product is non-degenerate it follows from (2.1) that

$$(4.11) \quad \alpha f(\mathbf{v}) + \beta f(\mathbf{w}) = f(\alpha\mathbf{v} + \beta\mathbf{w})$$

is valid for all $\mathbf{v}, \mathbf{w} \in \mathcal{U}$ and all real α, β . It is the content of (4.10), (4.11) and the remark after (4.6) that f is an inner product space isomorphism of \mathcal{U} onto $\hat{\mathcal{U}}$.

Let $x \in \mathcal{E}$ and $\mathbf{v} \in \mathcal{U}$ be given. Put $\mathbf{u} = x - q$ so that $x = q + \mathbf{u}$. Using (4.6) twice and (4.11) once we derive

$$x + \mathbf{v} = q + \mathbf{u} + \mathbf{v} = q + f(\mathbf{u} + \mathbf{v}) = q + f(\mathbf{u}) + f(\mathbf{v}) = q + \mathbf{u} + f(\mathbf{v}) = x + f(\mathbf{v}).$$

Thus, $x + \mathbf{v} = x + f(\mathbf{v})$ is valid for all $x \in \mathcal{E}$, which shows that the mappings \mathbf{v} and $f(\mathbf{v})$ of \mathcal{E} onto \mathcal{E} are the same. Therefore, f is the identity mapping of \mathcal{U} onto itself, which completes the proof that \mathcal{U} and $\hat{\mathcal{U}}$ coincide as inner product spaces.

DEFINITION. A set \mathcal{E} which is endowed with a structure defined by a real valued function σ on $\mathcal{E} \times \mathcal{E}$ is called a pseudo-Euclidean space if the axioms (E_1) – (E_4) are satisfied. The function σ will be called the separation function of \mathcal{E} . The inner product space \mathcal{V} determined by σ is called the translation space of \mathcal{E} . (The definition given here differs from that of Bourbaki ([1] Chapter 6, no. 6) in that the translation space \mathcal{V} is *not* regarded as part of the defining structure.)

The uniqueness theorem insures that the translation space is well defined.

If the separation function σ is nonnegative and if the translation space \mathcal{V} is finite-dimensional, then \mathcal{E} may be regarded as a Euclidean space in the ordinary sense.

The following is a corollary of the uniqueness theorem.

REPRESENTATION THEOREM. Let \mathcal{E} be a pseudo-Euclidean space and q a point in \mathcal{E} . Then every automorphism α of \mathcal{E} has a unique representation of the form

$$(4.12) \quad \alpha(x) = \alpha(q) + Q(x - q),$$

where Q is an orthogonal transformation of the translation space \mathcal{V} .

Proof. The mapping Q defined by

$$(4.13) \quad Q(\mathbf{v}) = \alpha \circ \mathbf{v} \circ \alpha^{-1}$$

is an isomorphism of the subgroup \mathcal{V} of \mathcal{A} onto its conjugate $\mathcal{V}^* = \alpha \circ \mathcal{V} \circ \alpha^{-1}$ in \mathcal{A} . The mapping Q may be used to transport the inner product space structure of \mathcal{V} to \mathcal{V}^* . It is clear that \mathcal{V}^* then satisfies the conditions (E_1) – (E_4) . Hence, by the uniqueness theorem, it must coincide with \mathcal{V} , and Q must be an automorphism of the inner product space \mathcal{V} . Since we use the notation $\mathbf{u}(q) = q + \mathbf{u}$ when $\mathbf{u} \in \mathcal{V}$, the image of $q \in \mathcal{E}$ under the mapping $Q\mathbf{v} \circ \alpha = \alpha \circ \mathbf{v}$ (see (4.13)) is given by

$$(4.14) \quad \alpha(q) + Q\mathbf{v} = \alpha(q + \mathbf{v}).$$

We obtain (4.12) by substituting $\mathbf{v} = x - q$ into (4.14).

Remark I. In the Euclidean case (4.12) is the well-known formula for rigid displacements. Many textbooks, however, derive this formula under unnecessary a priori assumptions of smoothness or even linearity of α ; a theorem that does not require such assumptions appears as Exercise 21a, chapter 6, in Bourbaki [1]. The uniqueness theorem and the representation theorem, including proofs, remain valid when the field of real numbers is replaced by an arbitrary commutative ring \mathcal{R} of characteristic $\neq 2$. In this case, the translation space is a module over \mathcal{R} .

Remark II. One may be tempted to define a subspace \mathcal{F} of a pseudo-Euclidean space \mathcal{E} by the condition that the restriction $\sigma_{\mathcal{F}}$ to \mathcal{F} of the separation function σ of \mathcal{E} satisfy the axioms (E_1) – (E_4) and hence give \mathcal{F} the structure of a pseudo-Euclidean space. Unfortunately, it may happen that such a “subspace” \mathcal{F} is

not a subspace in the customary sense, i.e., a set of the form

$$(4.15) \quad \mathcal{F} = \{x \mid x = p + u, u \in \mathfrak{U}\},$$

where \mathfrak{U} is a subspace of the translation space of \mathcal{E} . For example, let $\mathcal{E} = \mathcal{R} \times \mathcal{R} \times \mathcal{R}$ with σ defined by

$$(4.16) \quad \sigma(x, y) = (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2.$$

The set \mathcal{F} of all triples of the form $(\phi(\xi), \phi(\xi), \xi)$, where ϕ is an arbitrary function $\phi: \mathcal{R} \rightarrow \mathcal{R}$, is a "subspace," but not of the type (4.15) when ϕ is not linear. But for this fact Theorem 1 of the following section would be trivial.

Remark III. The uniqueness theorem given in this Section shows that the complete structure of a pseudo-Euclidean space, including the translation space, is determined by a knowledge of the separation function σ alone. One may ask whether σ itself may not be uniquely determined by the prescription of even less information. A result in this direction was found by Suppes [2]. Under the assumption that the translation space has index 1 and dimension 4, he showed that σ is uniquely determined by its values $\sigma(x, y)$ for all pairs (x, y) such that $\sigma(x, y) < 0$. This result is easily obtained by an adaptation of the reasoning leading from (5.10) to (5.14) given in the following section. The uniqueness theorem of the following section is a different result of the same type.

5. Minkowskian Chronometry. We assume that the following data are given:

- (a) a set \mathcal{E} , whose elements x, y, \dots will be called *events*;
- (b) a family Ω of subsets of \mathcal{E} which covers \mathcal{E} . The members \mathcal{L} of Ω will be called *observers*;
- (c) for each observer $\mathcal{L} \in \Omega$, a nonpositive separation function $\sigma_{\mathcal{L}}$ on \mathcal{L} which gives \mathcal{L} the structure of a one-dimensional pseudo-Euclidean space;
- (d) a symmetric binary relation \sim on \mathcal{E} with the following property: Given any observer \mathcal{L} and any event $x \notin \mathcal{L}$, there are at least two events y_1 and y_2 in \mathcal{L} that are related to x . The relation \sim will be called the *signal relation* and a pair (x, y) of events related by \sim will be called a *signal*.

Remarks on physical interpretation: $x \in \mathcal{L}$ means physically that x is an event which is experienced by the observer \mathcal{L} . We imagine that each observer is equipped with a clock. If $\tau_{\mathcal{L}}(x, y)$ is the time-difference of the two clock readings at the events x and y of \mathcal{L} , then the separation $\sigma_{\mathcal{L}}(x, y)$ of $x, y \in \mathcal{L}$ is assumed to be given by $\sigma_{\mathcal{L}}(x, y) = -(\tau_{\mathcal{L}}(x, y))^2$. The two events of a signal (x, y) are interpreted to be the emission and reception of a light, radio, or other electromagnetic signal. If $x \notin \mathcal{L}$, we may imagine x to be the event of reflection of a signal which is sent out by \mathcal{L} at y_1 and returns to \mathcal{L} at y_2 .

The data described under (a)–(d) define a certain structure on \mathcal{E} . We impose restrictions on this structure by assuming that there exists a separation function σ on all of \mathcal{E} which endows \mathcal{E} with the structure of a pseudo-Euclidean space and which satisfies the following two axioms:

(M₁) σ is an extension of the separation function $\sigma_{\mathcal{L}}$ for each observer $\mathcal{L} \in \Omega$. In other words,

$$(5.1) \quad \sigma(x, y) = \sigma_{\mathcal{L}}(x, y)$$

holds whenever $x, y \in \mathcal{L}$.

(M₂) The pair (x, y) is a signal if and only if the separation of x and y is zero, i.e. $\sigma(x, y) = 0$ if and only if $y \sim x$.

UNIQUENESS THEOREM. *There is at most one separation function σ which satisfies the axioms (M₁) and (M₂).*

We first prove a number of preliminary theorems, assuming that some separation function σ satisfying (M₁) and (M₂) on \mathcal{E} is given. The corresponding translation space is denoted by \mathcal{U} , as in Section 4.

THEOREM 1. *Every observer \mathcal{L} is a time-like straight line in \mathcal{E} . More precisely: There is a vector $l \in \mathcal{U}$ with the following properties*

- (i) l is a time-like unit vector, i.e. $l^2 = -1$.
- (ii) if $q \in \mathcal{L}$ is given, then $x \in \mathcal{L}$ if and only if $x = q + \xi l$, and $\xi \in \mathcal{R}$.

A vector l with the properties (i) and (ii) will be called a *direction vector* of the observer \mathcal{L} . It is clear that if l is a direction vector, then $-l$ is also one and there can be no others.

Proof. Let \mathcal{U}' be the one-dimensional translation space corresponding to $\sigma_{\mathcal{L}}$. The event-difference in \mathcal{U}' of the two events $x, y \in \mathcal{L}$ will be denoted by $y \overset{\mathcal{L}}{\sim} x$. This difference must be carefully distinguished from the event difference $y - x$ in \mathcal{U} , which corresponds to the separation function σ . It follows from axiom (M₁) and from (4.3) that

$$(5.2) \quad (y - x)^2 = (y \overset{\mathcal{L}}{\sim} x)^2 = \sigma(x, y) = \sigma_{\mathcal{L}}(x, y)$$

for $x, y \in \mathcal{L}$. Since \mathcal{U}' is assumed to be one-dimensional and $\sigma_{\mathcal{L}}$ nonpositive, \mathcal{L} can be represented in the form

$$\mathcal{L} = \{x \mid x = q + \xi l', \xi \in \mathcal{R}\},$$

where q is a fixed event in \mathcal{L} and $l' \in \mathcal{U}'$ is such that $l'^2 = -1$.

We now define a mapping f of \mathcal{R} into \mathcal{U} by $f(\xi) = (q + \xi l') - q$. \mathcal{L} is then the set of events x that are of the form

$$(5.3) \quad x = q + \xi l' = q + f(\xi), \quad \xi \in \mathcal{R}.$$

The same argument as the one that led from (4.5) to (4.10) shows that for all $\xi, \eta \in \mathcal{R}$ we must have

$$(5.4) \quad -\xi\eta = f(\xi) \cdot f(\eta).$$

We fix ξ and η and consider the vector $s \in \mathcal{U}$ given by

$$(5.5) \quad s = \eta f(\xi) - \xi f(\eta).$$

It follows from (5.4) that

$$\mathbf{s}^2 = \eta^2(-\xi^2) - 2\xi\eta(-\xi\eta) + \xi^2(-\eta^2) = 0,$$

i.e. that \mathbf{s} is a signal vector. Consider the event $z = q + \mathbf{s}$. By (5.2) we have

$$(5.6) \quad \mathbf{s}^2 = (z - q)^2 = \sigma(q, z) = 0.$$

If $z \in \mathcal{L}$ then $\mathbf{s} = z - q$ must be of the form $\mathbf{s} = \mathbf{f}(\zeta)$. By (5.4) it then follows that $0 = \mathbf{s}^2 = -\zeta^2$ and hence that $\mathbf{s} = \mathbf{f}(0) = \mathbf{0}$.

Assume now that $\mathbf{s} \neq \mathbf{0}$, in which case $z \notin \mathcal{L}$. Axiom (M_2) and (5.6) imply that $q \in \mathcal{L}$ and $z \notin \mathcal{L}$ must be related by a signal. The signal relation has the property that there must be another event $p \in \mathcal{L}$, $p \neq q$, which is also related to z . Using axiom (M_2) again, we find

$$(5.7) \quad 0 = \sigma(p, z) = (z - p)^2 = (\mathbf{s} + \mathbf{v})^2 = 2\mathbf{s} \cdot \mathbf{v} + \mathbf{v}^2,$$

where $\mathbf{v} = q - p$. Since $p \in \mathcal{L}$, \mathbf{v} must be of the form $\mathbf{v} = \mathbf{f}(\lambda)$. By (5.4) and (5.5), we obtain

$$\mathbf{s} \cdot \mathbf{v} = \eta(-\xi\lambda) - \xi(-\eta\lambda) = 0$$

and hence, by (5.7), $0 = \mathbf{v}^2 = -\lambda^2$. Consequently, $\mathbf{v} = \mathbf{f}(0) = \mathbf{0} = q - p$, i.e. $p = q$, which contradicts $p \neq q$.

We conclude that always $\mathbf{s} = \mathbf{0}$ and hence, by (5.5), that

$$(5.8) \quad \eta \mathbf{f}(\xi) = \xi \mathbf{f}(\eta)$$

holds for all real ξ and η . Putting $\eta = 1$ and $\mathbf{f}(1) = \mathbf{l} \in \mathcal{U}$, we see that (5.8) gives $\mathbf{f}(\xi) = \xi \mathbf{l}$. It follows from (5.3) that \mathcal{L} is the set of events of the form $x = q + \xi \mathbf{l}$. The relation $\mathbf{l}^2 = -1$ is a consequence of (5.4). Hence \mathbf{l} has the properties (i) and (ii). Q.E.D.

THEOREM 2. *Let \mathcal{L} be an observer with direction vector \mathbf{l} and let $q \in \mathcal{L}$, $x \in \mathcal{E}$. An event $y = q + \eta \mathbf{l} \in \mathcal{L}$ is then related to x by a signal if and only if η is a root of the equation*

$$(5.9) \quad \eta^2 + 2\mathbf{l} \cdot (x - q)\eta + (x - q)^2 = 0.$$

When $x \notin \mathcal{L}$, there are exactly two events

$$(5.10) \quad y_1 = q + \eta_1 \mathbf{l}, \quad y_2 = q + \eta_2 \mathbf{l}$$

in \mathcal{L} that are related to x by a signal and we have

$$(5.11) \quad \sigma(q, x) = (x - q)^2 = -\eta_1 \eta_2, \quad (x - q) \cdot \mathbf{l} = -(\eta_1 + \eta_2),$$

$$(5.12) \quad [\mathbf{l} \cdot (x - q)]^2 + 4(x - q)^2 > 0.$$

Proof. By axiom (M_2) , $y = q + \eta \mathbf{l}$ is related to x by a signal if and only if

$$\sigma(x, y) = (y - x)^2 = (q - x + \eta \mathbf{l})^2 = -\eta^2 - 2\mathbf{l} \cdot (x - q)\eta + (x - q)^2 = 0,$$

which proves the first part of the theorem.

When $x \notin \mathcal{L}$, the property of the signal relation described in (d) requires that the roots of (5.9) must be real and distinct. (5.12) is a necessary and sufficient condition for this requirement. The roots of (5.9) always satisfy (5.11).

Proof of the uniqueness theorem. Let q and x be any two events in \mathcal{E} . Since the family of all observers covers \mathcal{E} , there is at least one observer such that $q \in \mathcal{L}$. If $x \in \mathcal{L}$ also, then by axiom (M_1)

$$(5.13) \quad \sigma(q, x) = \sigma_{\mathcal{L}}(q, x).$$

If $x \notin \mathcal{L}$, we consider the events $y_1, y_2 \in \mathcal{L}$, determined as in Theorem 2. With the notation of Theorem 2, we have

$$\begin{aligned} \sigma(q, y_i) &= (y_i - q)^2 = \eta_i^2 l^2 = -\eta_i^2 & i = 1, 2 \\ \sigma(y_1, y_2) &= (y_2 - y_1)^2 = (\eta_2 - \eta_1)^2 l^2 = -\eta_2^2 + 2\eta_1\eta_2 - \eta_1^2. \end{aligned}$$

It follows that

$$\eta_1\eta_2 = \frac{1}{2}[\sigma(y_1, y_2) - \sigma(q, y_1) - \sigma(q, y_2)].$$

Noting that $y_1, y_2, q \in \mathcal{L}$, we infer from axiom (M_1) and $(5.11)_1$ that

$$(5.14) \quad \sigma(q, x) = -\frac{1}{2}[\sigma_{\mathcal{L}}(y_1, y_2) - \sigma_{\mathcal{L}}(q, y_1) - \sigma_{\mathcal{L}}(q, y_2)].$$

Equations (5.13) and (5.14) show that $\sigma(q, x)$ is uniquely determined when $\sigma_{\mathcal{L}}$ is given. Q.E.D.

DEFINITION. A set \mathcal{E} which is endowed with a structure defined by $\Omega, \{\sigma_{\mathcal{L}} | \mathcal{L} \in \Omega\}$, and \sim as described under (a)–(d) is called a Minkowskian domain if the axioms (M_1) and (M_2) are satisfied.

The uniqueness theorem shows that a Minkowskian domain is also endowed, in a canonical way, with the structure of a pseudo-Euclidean space.

THEOREM 3. The translation space \mathcal{V} of a Minkowskian domain has index 1.

Proof. Let \mathcal{L} be an observer with direction vector l , and let $\mathcal{U} = \{u | u = \xi l, \xi \in \mathcal{R}\}$ be the one-dimensional subspace of \mathcal{V} generated by l . Consider a nonzero vector $v \in \mathcal{U}^\perp$, which then satisfies $l \cdot v = 0$. Choose an event $q \in \mathcal{L}$ and put $x = q + v$. It is clear that $x \notin \mathcal{L}$. Since $l \cdot v = l \cdot (x - q) = 0$, (5.12) states that $(x - q)^2 = v^2 > 0$, which shows that v is space-like. It follows that $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^\perp$ is a direct decomposition with the property that $\mathcal{U} \subset \mathcal{V}_-, \mathcal{U}^\perp \subset \mathcal{V}_+$, i.e., it is a decomposition of the type (2.3). Theorem 1 of Section 2 implies that $\text{ind } \mathcal{V} = \dim \mathcal{U} = 1$.

Remark. As is shown by Theorem 1, the family Ω of observers is a congruence of straight lines in \mathcal{E} . Since Ω covers \mathcal{E} , there must be at least one line through each event. For example, Ω could be the set of all straight lines parallel to a given time-like direction. Unless Ω contains “very many” observers, a knowledge of $\sigma(x, y)$ for all pairs (x, y) such that x, y belong to the same observer is *not* sufficient to determine σ . It is the interconnection of σ with the signal relation, as provided by axiom (M_2) , which renders σ unique.

6. Temporal order. As in the previous section we assume that data as described under (a), (b), and (c) are given. Instead of (d), we suppose that we have (d') a binary relation \rightarrow on \mathcal{E} with the following property: Given any observer \mathcal{L} and any event $x \in \mathcal{E}$, there is an event $y_1 \in \mathcal{L}$ such that $y_1 \rightarrow x$ and there is an event $y_2 \in \mathcal{L}$ such that $x \rightarrow y_2$. Moreover, $x \rightarrow y$ and $y \rightarrow x$ can both be valid only if $x = y$. The relation \rightarrow will be called the *directed signal relation* and the relation defined by

$$(6.1) \quad x \sim y \text{ if and only if } x \rightarrow y \text{ or } y \rightarrow x$$

the (undirected) signal relation associated with \rightarrow .

It is clear that \sim has the property described under (d) of the previous section.

Remark on physical interpretation. The change from (d) to (d') corresponds to introducing the possibility of distinguishing the emission x from the reception y of a signal $x \rightarrow y$.

We require here the existence of a separation function σ on \mathcal{E} which satisfies not only the axioms (M₁) and (M₂) but also

(M₃) *If \mathcal{L} is an observer and if y_1, y_2 and z_1, z_2 are events in \mathcal{L} such that*

$$(6.2) \quad y_1 \rightarrow p \rightarrow y_2, \quad z_1 \rightarrow q \rightarrow z_2$$

for some $p, q \in \mathcal{E}$, then

$$(6.3) \quad \sigma_{\mathcal{L}}(y_1, z_1) + \sigma_{\mathcal{L}}(y_2, z_2) - \sigma_{\mathcal{L}}(y_1, z_2) - \sigma_{\mathcal{L}}(y_2, z_1) \geq 0.$$

DEFINITION. A set \mathcal{E} which is endowed with a structure defined by $\Omega, \{\sigma_{\mathcal{L}} | \mathcal{L} \in \Omega\}$, and \rightarrow as described under (a)–(c) and (d') is called a *directed Minkowskian domain* if the axioms (M₁), (M₂), and (M₃) are satisfied.

From now on we shall exclude the trivial case when the domain \mathcal{E} coincides with one of the observers. The dimension of the translation space \mathcal{U} of \mathcal{E} is then at least two.

DEFINITION. We say that the event $x \in \mathcal{E}$ is *earlier than* the event $y \in \mathcal{E}$, and we write $x < y$, if there is an event p such that $x \rightarrow p \rightarrow y$.

THEOREM 1. Every observer \mathcal{L} has a unique direction vector l with the following property: For any two events $x, y \in \mathcal{E}$ the relation $x < y$ holds if and only if

$$(6.4) \quad (y - x)^2 \leq 0, \quad (y - x) \cdot l \leq 0.$$

The direction vector l determined by the condition (6.4) will be called the *proper direction vector* of \mathcal{L} .

LEMMA. Let \mathcal{L} be an observer with direction vector l and let $x_1, x_2 \in \mathcal{E}$. Assume that $q \in \mathcal{L}$ and $z_i = q + \xi_i l$ are the two events related to x_i by a signal ($i = 1, 2$). Then

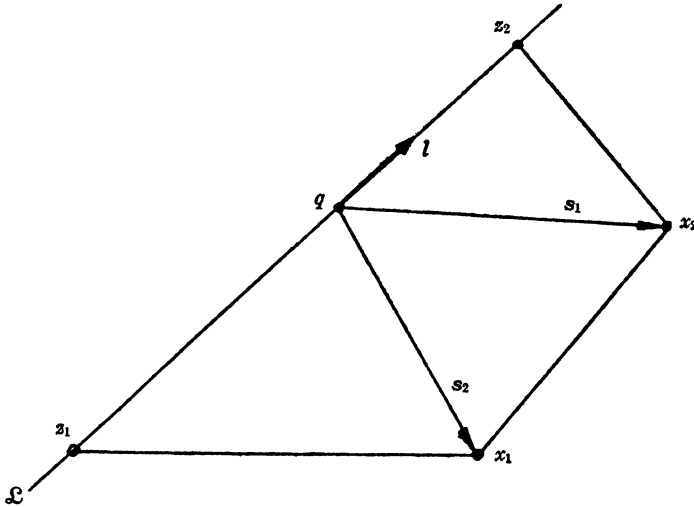
$$(6.5) \quad (x_2 - x_1) \cdot l = (\xi_1 - \xi_2), \quad \xi_1 \xi_2 (x_2 - x_1)^2 \geq 0,$$

and (6.5)₂ can reduce to equality only if $(x_2 - x_1)^2 = 0$.

Proof. Theorem 2 of Section 5 applies when we replace there $q, x, y_1, y_2, \eta_1, \eta_2$ by $q, x_i, q, z_i, 0, \xi_i$, respectively. $(5.11)_2$ then yields

$$(6.6) \quad \mathbf{s}_i \cdot \mathbf{l} = -\xi_i,$$

where $\mathbf{s}_i = x_i - q$. Hence $(x_2 - x_1) \cdot \mathbf{l} = (\mathbf{s}_2 - \mathbf{s}_1) \cdot \mathbf{l} = -(\xi_2 - \xi_1)$, which proves $(6.5)_1$.



Consider the vector $\mathbf{u} = \xi_2 \mathbf{s}_1 - \xi_1 \mathbf{s}_2$. It follows from (6.6) that $\mathbf{u} \cdot \mathbf{l} = 0$. Therefore, by Theorem 3 of Section 5 and Theorem 2 of Section 3, \mathbf{u} must be space-like. Since $\mathbf{s}_i^2 = 0$, we obtain

$$(6.7) \quad \mathbf{u}^2 = -2\xi_1 \xi_2 \mathbf{s}_1 \cdot \mathbf{s}_2 \geq 0.$$

On the other hand we have

$$(x_2 - x_1)^2 = (\mathbf{s}_2 - \mathbf{s}_1)^2 = -2\mathbf{s}_1 \cdot \mathbf{s}_2,$$

which shows that (6.7) is equivalent to $(6.5)_2$. Equality can hold in (6.7) only if $\mathbf{u} = \xi_2 \mathbf{s}_1 - \xi_1 \mathbf{s}_2 = \mathbf{0}$, in which case $x_2 - x_1 = \mathbf{s}_2 - \mathbf{s}_1$ must be a signal vector. Q.E.D.

Proof of Theorem 1. The proof will be carried out in four steps.

I. Assume first that $x, y \in \mathcal{L}$. Let \mathbf{l}' be one of the two direction vectors of \mathcal{L} (see Theorem 1 of the previous section) and let $y - x = \alpha \mathbf{l}'$. Since the translation space \mathcal{U} of \mathcal{E} is such that $\text{ind } \mathcal{U} = 1, \dim \mathcal{U} \geq 2$, there is a space-like unit vector orthogonal to \mathbf{l}' , i.e. a vector \mathbf{u} such that $\mathbf{u} \cdot \mathbf{l}' = 0, \mathbf{u}^2 = 1$. Put $q = x + \frac{1}{2}\alpha \mathbf{l}' \in \mathcal{L}$ and $z = q + \frac{1}{2}\alpha \mathbf{u}$, so that

$$x = q - \frac{\alpha}{2} \mathbf{l}', \quad y = q + \frac{\alpha}{2} \mathbf{l}', \quad z - q = \frac{\alpha}{2} \mathbf{u}.$$

The two roots of the equation

$$\eta^2 + \mathbf{l}' \cdot (z - q) \eta - (z - q)^2 = \eta^2 - \left(\frac{\alpha}{2}\right)^2$$

are $\eta = \pm \frac{1}{2}\alpha$. Hence, by Theorem 2 of Section 5, x and y are the events in \mathcal{L} that are related to z by a signal. Therefore x and y , in some order, must be the two events in \mathcal{L} that correspond to z as described in (d'), and we must have either $x \rightarrow z \rightarrow y$, i.e. $x < y$, or $y \rightarrow z \rightarrow x$, i.e. $y < x$. Both $x < y$ and $y < x$ can hold only when $x = y = z$.

II. We choose two events $z_1, z_2 \in \mathcal{L}$ such that $(z_2 - z_1)^2 = -1$. By the result of I we may assume that z_1, z_2 are arranged such that $z_1 < z_2$. We put $l = z_2 - z_1$. Let $x, y \in \mathcal{L}$ and assume $x < y$. An easy calculation, starting from (6.3) with y_1, y_2 replaced by x, y , shows that

$$\sigma_{\mathcal{L}}(x, z_1) + \sigma_{\mathcal{L}}(y, z_2) - \sigma_{\mathcal{L}}(x, z_2) - \sigma_{\mathcal{L}}(y, z_1) = -2(y - x) \cdot l \leq 0.$$

Under the restrictive hypothesis $x, y \in \mathcal{L}$ the conclusion of the theorem now follows. We note that if $x, y \in \mathcal{L}$, then

$$(6.8) \quad y - x = \xi l, \quad \text{where} \quad \begin{aligned} \xi &\geq 0 \text{ if } x < y \\ \xi &\leq 0 \text{ if } y < x. \end{aligned}$$

(6.8) characterizes the proper direction vector of \mathcal{L} .

III. Assume $x < y$ holds for two events $x, y \in \mathcal{E}$, i.e. $x \rightarrow q \rightarrow y$ for some $q \in \mathcal{E}$. Let \mathcal{L}' be an observer passing through q and let l' be the proper direction vector of \mathcal{L}' . According to the requirements described in (d'), there are events $z_1, z_2 \in \mathcal{L}'$ such that

$$z_1 \rightarrow x \rightarrow q \quad \text{and} \quad q \rightarrow y \rightarrow z_2,$$

i.e. such that $z_1 < q < z_2$. By II, (6.8) applies with x, y replaced by z_1, q or q, z_2 . Hence

$$(6.9) \quad z_i = q + \xi_i l', \quad \xi_1 \leq 0, \quad \xi_2 \geq 0.$$

The lemma and (6.9) show that

$$(6.10) \quad (y - x) \cdot l' \leq 0, \quad (y - x)^2 \leq 0.$$

Therefore the conclusion (6.4) holds for the particular observer \mathcal{L}' .

Let \mathcal{L} be another observer and let $q' \in \mathcal{L}'$ be chosen arbitrarily. It is clear that there are events $x_1, x_2 \in \mathcal{L}$ and $z'_1, z'_2 \in \mathcal{L}'$ such that $z'_1 \rightarrow x_1 \rightarrow q' \rightarrow x_2 \rightarrow z'_2$. If $\mathcal{L}' \neq \mathcal{L}$ we must have $x_1 \neq x_2$ and hence

$$(6.11) \quad x_2 - x_1 = \alpha l, \quad \alpha > 0,$$

where l is the proper direction vector of \mathcal{L} . Also, (6.9) remains valid when q, z_i, ξ_i are replaced by q', z'_i, ξ'_i . Thus, the lemma applies again and we find, using (6.11), that

$$(6.12) \quad (x_2 - x_1) \cdot l' = \alpha l \cdot l' \leq 0, \quad l \cdot l' \leq 0.$$

If $y \neq x$, it follows from (6.10), (6.12) and Theorem 2 of Section 3 that actually

$$(y - x)^2 \leq 0, \quad (y - x) \cdot l' < 0, \quad l \cdot l' < 0.$$

Theorem 4 of Section 3 shows that we must have $(y-x) \cdot l < 0$. If $y=x$, $(y-x) \cdot l = 0$ is trivially true. Hence (6.4) holds for the proper direction vector l of an arbitrary observer.

IV. Assume that $(y-x)^2 \leq 0$. It is not hard to see that it is then possible to choose an observer \mathcal{L} and events $q, z_1, z_2 \in \mathcal{L}$ such that the hypotheses of the lemma are satisfied for $x_2=y, x_1=x$. If $(y-x)^2 < 0$, (6.5)₂ implies $\xi_1 \xi_2 < 0$, which states that ξ_1 and ξ_2 have opposite sign. From part II we conclude that either $z_1 \rightarrow x \rightarrow q \rightarrow y \rightarrow z_2$ or $z_2 \rightarrow y \rightarrow q \rightarrow x \rightarrow z_1$. Hence, we must have either $x < y$ or $y < x$. If $(y-x)^2 = 0$, axiom (M₂) shows that x and y are related by a signal and hence that $x < y$ or $y < x$, trivially. Thus, x and y are comparable. This observation, together with the result of III, completes the proof of Theorem 1. Q.E.D.

Remark. The number $-(y-x) \cdot l$ is the "time-difference" of the events x and y relative to the observer \mathcal{L} . Theorem 1 states, therefore, that x is earlier than y if and only if it is "earlier" for every observer.

An argument based on Theorem 1 and Theorem 4 of Section 3 will prove

THEOREM 2. *The relation $<$ is a partial order on \mathcal{E} , which has the property that x and y are comparable if and only if $\sigma(x, y) \leq 0$.*

One of the two directed time-cones described in Theorem 5 of Section 3, say \mathcal{V}_- , is singled out by the property that $y-x \in \mathcal{V}_-$ implies $x < y$. We may call \mathcal{V}_- the *future time-cone* and $\mathcal{V}_+ = -\mathcal{V}_-$ the *past time-cone*.

The research leading to this paper was done under Grant NSF-G16745 by the National Science Foundation to Carnegie Institute of Technology.

References

1. N. Bourbaki, *Algèbre*, Chap. 9, Hermann, Paris, 1959.
2. J. L. Synge, *Relativity, The Special Theory*, North Holland, Amsterdam, 1956.
3. ———, *Relativity, The General Theory*, North Holland, Amsterdam, 1960.
4. P. Suppes, *Axioms for Relativistic Kinematics with or without Parity*. Symposium on the Axiomatic Method, with special reference to geometry and physics, North Holland, Amsterdam, 1959, p. 291–307.
5. P. R. Halmos, *Finite Dimensional Vector Spaces*, Van Nostrand, Princeton, 1958.

*Proof of the Maximality of the Orthogonal
Group in the Unimodular Group*

WALTER NOLL

Proof of the Maximality of the Orthogonal Group in the Unimodular Group

WALTER NOLL

Let \mathcal{D} be a real inner product space of dimension n . The inner product is assumed to be positive definite. The *unimodular group* \mathcal{U} and the *orthogonal group* \mathcal{O} of \mathcal{D} are defined by

$$\mathcal{U} = \{\mathbf{H} \mid \mathbf{H} \in \mathcal{L}, |\det \mathbf{H}| = 1\} \quad (1)$$

and

$$\mathcal{O} = \{\mathbf{Q} \mid \mathbf{Q} \in \mathcal{L}, \mathbf{Q}^T = \mathbf{Q}^{-1}\}, \quad (2)$$

where \mathcal{L} is the algebra of all linear transformations of \mathcal{D} into itself. It is the purpose of this note to give a simple proof of the fact that \mathcal{O} is a *maximal subgroup* of \mathcal{U} .

This theorem has an important corollary in the theory of simple materials. In the terminology introduced in [1], the theorem implies that every isotropic simple material is either a solid or a fluid.

I convinced myself of the validity of the theorem about two years ago, on the basis of a tedious and ugly argument. Since I could not believe that the theorem was new, I asked a number of distinguished mathematicians for references, but I had no success¹. I still believe that the theorem must have been discovered previously, but it is certainly not *well* known.

The polar decomposition theorem states that every invertible linear transformation is the product of an orthogonal transformation and one that is positive definite and symmetric. It follows that every group \mathcal{G} containing \mathcal{O} is generated by its positive definite and symmetric members. It is therefore sufficient to prove the

Theorem. *Assume that $\mathbf{S} \in \mathcal{U}$ is positive definite and symmetric and has at least two distinct proper numbers s and t , $s > t$. Then every positive definite and symmetric $\mathbf{H} \in \mathcal{U}$ belongs to the group \mathcal{G} generated by \mathcal{O} and \mathbf{S} .*

The proof is based on the following

Lemma. *For every ξ such that*

$$\left(\frac{t}{s}\right)^2 \leq \xi \leq \left(\frac{s}{t}\right)^2 \quad (3)$$

¹ The proof of BRAUER [2], which came to my attention after this note was written, is being published in response to my inquiries.

there exists an $\mathbf{R} \in \mathcal{O}$ such that the symmetric transformation

$$\mathbf{T} = \mathbf{S}^{-1} \mathbf{R} \mathbf{S}^2 \mathbf{R}^{-1} \mathbf{S}^{-1} \tag{4}$$

has the proper numbers $(\xi, \xi^{-1}, 1, \dots, 1)$.

Proof. Let \mathbf{g} and \mathbf{f} be unit proper vectors of \mathbf{S} corresponding to the proper numbers s and t , respectively. Let \mathcal{W} be the $(n - 2)$ -dimensional subspace of \mathcal{D} that is orthogonal to both \mathbf{g} and \mathbf{f} . We define $\mathbf{R}_\vartheta \in \mathcal{O}$ by

$$\begin{aligned} \mathbf{R}_\vartheta \mathbf{w} &= \mathbf{w} \quad \text{if } \mathbf{w} \in \mathcal{W}, \\ \mathbf{R}_\vartheta \mathbf{g} &= \cos \vartheta \mathbf{g} + \sin \vartheta \mathbf{f}, \\ \mathbf{R}_\vartheta \mathbf{f} &= -\sin \vartheta \mathbf{g} + \cos \vartheta \mathbf{f}, \end{aligned}$$

and we put

$$\mathbf{T}_\vartheta = \mathbf{S}^{-1} \mathbf{R}_\vartheta \mathbf{S}^2 \mathbf{R}_\vartheta^{-1} \mathbf{S}^{-1}.$$

It is easily seen that $\mathbf{T}_\vartheta \mathbf{w} = \mathbf{w}$ if $\mathbf{w} \in \mathcal{W}$ and

$$\mathbf{T}_{(\pi/2)} \mathbf{g} = \begin{pmatrix} t \\ s \end{pmatrix}^2 \mathbf{g}, \quad \mathbf{T}_{(\pi/2)} \mathbf{f} = \begin{pmatrix} s \\ t \end{pmatrix}^2 \mathbf{f}.$$

Hence $\mathbf{T}_{(\pi/2)}$ has the proper numbers $\left(\frac{t}{s}\right)^2, \left(\frac{t}{s}\right)^{-2}, 1, \dots, 1$. Since $\mathbf{T}_0 = \mathbf{1}$ and since \mathbf{T}_ϑ depends continuously on ϑ , it follows that there exists a ϑ such that the assertion of the lemma holds for $\mathbf{T} = \mathbf{T}_\vartheta$, QED.

Proof of the theorem. Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be an orthonormal basis of proper vectors of \mathbf{H} , with corresponding proper numbers h_1, \dots, h_n . Let \mathbf{H}_k be the symmetric tensor that has the same proper vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ as \mathbf{H} , but with corresponding proper numbers, $1, \dots, 1, \eta_k, \eta_k^{-1}, 1, \dots, 1$, where η_k corresponds to \mathbf{e}_k and η_k^{-1} to \mathbf{e}_{k+1} . If the η_k are chosen such that

$$\eta_1 = h_1, \quad \eta_k = \eta_{k-1} h_k, \quad k = 1, 2, \dots, n - 1,$$

then, observing that $h_1 h_2 \dots h_n = 1$, we infer

$$\mathbf{H} = \mathbf{H}_1 \mathbf{H}_2 \dots \mathbf{H}_{n-1}. \tag{5}$$

We now fix k and choose m large enough that $\xi = \sqrt[m]{\eta_k}$ satisfies the inequality (3). We then determine \mathbf{T} according to the lemma. Let $\bar{\mathbf{g}}$ and $\bar{\mathbf{f}}$ be unit proper vectors of \mathbf{T} which correspond to the proper numbers ξ and ξ^{-1} , respectively. We can find an orthogonal transformation \mathbf{Q} which maps the basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ onto an orthonormal basis of proper vectors of \mathbf{T} in such a way that

$$\mathbf{Q} \mathbf{e}_k = \bar{\mathbf{g}}, \quad \mathbf{Q} \mathbf{e}_{k+1} = \bar{\mathbf{f}}.$$

With ξ, \mathbf{T} , and $\mathbf{Q} \in \mathcal{O}$ chosen in this manner, we easily see that

$$\mathbf{H}_k = \mathbf{Q}^{-1} \mathbf{T}^m \mathbf{Q}. \tag{6}$$

Of course, \mathbf{T}, \mathbf{Q} , and m depend on k .

If we substitute (4) into (6) and the result into (5), and if we do so for each of the \mathbf{H}_k , we see that \mathbf{H} is indeed generated by \mathbf{S} and orthogonal transformations. QED.

Acknowledgement. This work was supported by the National Science Foundation under Grant NSF-G 16745 to Carnegie Institute of Technology.

References

- [1] NOLL, W., A mathematical theory of the mechanical behavior of continuous media. Arch. Rational Mech. Anal. **2**, 197–226 (1958).
- [2] BRAUER, R., On the relation between the orthogonal group and the unimodular group. Arch. Rational Mech. Anal. **18**, 97 (1965).

Carnegie Institute of Technology
Pittsburgh, Pennsylvania

(Received August 4, 1964)

Space-Time Structures in Classical Mechanics

WALTER NOLL

Space-Time Structures in Classical Mechanics

WALTER NOLL

1. Introduction

The English language contains many words that denote spatial or temporal concepts: 'now', 'later', 'soon', 'simultaneous', 'here', 'there', 'far', 'location', 'equidistant', etc. The grammar is in part organized in accordance with temporal categories: present, past, future. If we tried to remove all words, prefixes, and suffixes with a temporal or spatial meaning from the language we would surely all but destroy it. The system of temporal and spatial concepts of a natural language such as English constitutes a *verbal space-time structure*. It is not a very precise system, but it serves very well as a framework for the common experiences of human life.

The geometry of the ancient Greeks and the spatial and temporal concepts of the mechanics of GALILEO and NEWTON may be viewed as being refinements of the intuitive verbal space-time structure, refinements which resulted in a very precise mathematical system. This system, which I call the *classical space-time structure*, provides the basis for several very successful branches of physics, chief among them the mechanics of particle systems of NEWTON and the mechanics of rigid bodies of EULER. In Sect. 2 I shall give a brief outline of a modern version of the classical space-time structure.

Until the beginning of the 19-th century there were very few people, if there were any, who could even imagine a system that might replace classical space-time. KANT, for example, regarded the valid statements of classical geometry and mechanics as being "a priori" and "synthetic"; i.e., he considered them to be truths about reality not derived from experience and yet not mere tautologies. This judgement reflects the view that the classical space-time structure is not just an expedient framework for physical experience, but is indeed the only conceivable such framework. It is a very understandable view, because before the

invention of non-Euclidean geometry it must have been impossible to imagine how another space-time structure might be formulated.

A non-classical space-time structure cannot easily be described entirely in words, because words have connotations that imply the classical structure. The term "space-time" itself has misleading connotations when applied to a nonclassical system. The natural language keeps us in the prison of classical space-time. The only language within which a non-classical structure can be unambiguously formulated is the language of axiomatic mathematics. The methods of modern mathematics make it possible to fabricate almost at will structures that can play the role of classical space-time.

One may ask why anybody would wish to consider a nonclassical space-time at all. The reason is, of course, that the classical structure, while adequate as a basis for the concepts of ordinary experience and of the older branches of physics, is inadequate for some of the newer physical disciplines, inadequate, in particular, for relativistic and quantum physics. These disciplines require space-time structures that radically deviate from the classical one. It is not the purpose of this lecture to elaborate on these structures. Rather, I here content that classical space-time is not well suited even for some venerable branches of mechanics, and I shall develop another structure which I believe to be more appropriate for these branches. This structure, which I call *neo-classical space-time*, will be presented in Sect. 3; and in Sects. 4–6 I shall indicate how it can be used for developing classical mechanics.

Many of the concepts of classical space-time reflect the fact that we humans live on this solid earth, which in daily life is always available as a frame of reference for specifying "locations". The first blow to the classical system was dealt by COPERNICUS, who deprived the earth of its once secure place as the universal frame of reference. The concept of a "location" in interstellar space is much more problematical than that of a location on earth. NEWTON'S *absolute space*, as the set of all possible locations, has since its inception been regarded with unease by most thinkers. Nonetheless, the "Newtonian" mechanics built upon this concept of absolute space has been very successful. Actually, use of the concept of absolute space is one among several ways of accounting for the phenomenon of inertia. Classical space-time with its absolute space has been most successful in those branches of mechanics in which inertia plays the central role. Such is not the case, however, in many of the branches of the mechanics of continuous media, where inertia is often of minor importance or sometimes even altogether neglected. In these branches of mechanics absolute space is an artificial and inappropriate concept. If it is used anyway, it is necessary to compensate for its arbitrariness by introducing a requirement of invariance, called the

principle of frame-indifference or objectivity (c.f. [1], Sect. 19). The description of mechanics in terms of the neo-classical space-time, which has no room for an absolute space, shows clearly why this principle is needed.

2. The Event-World of Classical Space-Time

The *absolute space* is a set E consisting of *points (locations)* x, y, \dots . The set E is endowed with a mathematical structure defined by a *distance function* δ which associates with each pair x, y of points a number $\delta(x, y)$, the *distance from x to y* . The distance function δ is subject to certain axioms which ensure that δ is a Euclidean metric that gives E the structure of a *Euclidean space*. This Euclidean structure makes it possible to define a unique *translation space* V_E of E , which is a vector space with inner product consisting of automorphisms of E . (For details, see [2], Sect. 4.)

The *event-world* of classical space-time is the set $W = E \times R$ of all pairs (x, t) , where $x \in E$ and $t \in R$, the set of real numbers. The point x is called the *location*, the number t the *time* of the event (x, t) .

3. The Event-World of Neo-Classical Space-Time

The *event-world* of neo-classical space-time is a set W consisting of *events* e, f, \dots . The set W is endowed with a mathematical structure defined by a *time-lapse function* τ and a *distance function* δ , subject to the axioms $(T_1) - (T_4)$ and $(D_1) - (D_3)$ stated below.

(T_1) The time-lapse function τ assigns to each pair e, f of events a number $\tau(e, f)$, called the *time-lapse* between e and f .

(T_2) For any $e, f \in W$,

$$\tau(e, f) = -\tau(f, e). \quad (3.1)$$

(T_3) For any $e, f, g \in W$,

$$\tau(e, f) + \tau(f, g) = \tau(e, g). \quad (3.2)$$

(T_4) For any $e \in W$ and any $t \in R$ there is a $f \in W$ such that $\tau(e, f) = t$.

We say that e is *earlier* than, *later* than, or *simultaneous* with f according to whether $\tau(e, f) > 0$, < 0 , or $= 0$.

The set of all pairs of simultaneous events

$$S = \{(e, f) \mid \tau(e, f) = 0\} \quad (3.4)$$

is an equivalence relation on W , as can easily be seen to follow from (T_2) and (T_3) . This equivalence relation determines a partition Γ of W into classes T of simultaneous events such that

$$S = \bigcup_{T \in \Gamma} T \times T. \quad (3.5)$$

The equivalence classes T will be called *instantaneous spaces* or simply *instants*. If $e \in T$ we say that *the event e happens at the instant T* .

The value of the time-lapse $\tau(e, f)$ depends only on the instants T and S at which e and f happen. Therefore, we can define unambiguously a time-lapse

$$\bar{\tau}(T, S) = \tau(e, f) \quad \text{if } e \in T, f \in S \quad (3.6)$$

between two instants T and S .

(D_1) The distance function δ assigns to each pair $(e, f) \in S$, i.e., to each pair of simultaneous events, a number $\delta(e, f)$, called the *instantaneous distance* between e and f .

(D_2) For each instant T , the restriction δ_T of δ to $T \times T$ is a Euclidean metric on T .

To say that δ_T is a Euclidean metric means that it gives to the instant T the structure of a Euclidean space. The translation space of T will be denoted by V_T .

(D_3) For each instant T , the dimension of the translation space V_T is 3.

Physically, the values $\tau(e, f)$ of the time-lapse function τ are to be interpreted as the results of time-measurements with clocks. The axioms (T_2)–(T_4) reflect familiar experiences with such measurements. The values $\delta(e, f)$ of the distance function δ are to be interpreted as the results of distance measurements with measuring sticks. The value $\delta(e, f)$ is defined only when e and f are simultaneous because each distance measurement is made at a particular instant. The axioms (D_2) and (D_3) are the abstract of thousands of years of experience with distance measurements.

An *automorphism* α of the event-world W is a one-to-one mapping of W onto itself which preserves time-lapses and distances. Thus, an automorphism α satisfies

$$\tau(\alpha(e), \alpha(f)) = \tau(e, f) \quad \text{for all } e, f \in W \quad (3.7)$$

and

$$\delta(\alpha(e), \alpha(f)) = \delta(e, f) \quad \text{for all } (e, f) \in S. \quad (3.8)$$

If $U \subset W$ we write

$$U^\alpha = \{\alpha(e) \mid e \in U\} \quad (3.9)$$

for the set of all images under α of events in U . In this way, the mapping $T \rightarrow T^\alpha$ of T onto itself is an automorphism of T in the sense that it preserves the time-lapse function $\bar{\tau}$ defined by (3.6).

An automorphism α of W also induces isomorphisms $V_T \rightarrow V_{T^\alpha}$ between the translation spaces of the instantaneous spaces.

4. Material Universes, Motions

A *material universe*¹ is a set \mathcal{U} consisting of *bodies* B, C, \dots . It is assumed that \mathcal{U} is partially ordered by a relation $<$, and we say that C is a *part of* B if $C < B$. The body B is said to be *separate* from the body C if B and C have no part in common, i.e., if there is no $D \in \mathcal{U}$ such that both $D < B$ and $D < C$.

The material universe appropriate to a system of discrete particles consists of the collection of all subsets of a finite set, whose members represent the particles. In this case, the relation $<$ is the set-inclusion. For continuum mechanics, however, more complicated material universes must be considered.

A *motion* of the material universe \mathcal{U} is a function M which assigns to each body $B \in \mathcal{U}$ a subset $M(B)$ of W such that

$$M(B) \subset M(C) \quad \text{if } B < C \quad (4.1)$$

and

$$M(B) \cap T \neq \emptyset \quad \text{for each } T \in \Gamma. \quad (4.2)$$

The set $M(B)$ is called the *set of events experienced by* B during the motion M , or simply the *world-tube* of B .

The requirement (4.1) states that a part experiences fewer events than the whole, and (4.2) expresses the fact that bodies cannot appear out of nothing nor disappear into nothing.

An automorphism α of the event-world W induces a transformation $M \rightarrow M^\alpha$ on motions; it is defined by

$$M^\alpha(B) = (M(B))^\alpha \quad \text{for all } B \in \mathcal{U}. \quad (4.3)$$

The above is only the beginning of a kinematics based on the neo-classical space-time. A more detailed development will be presented in future publications.

5. Force Systems, Dynamical Processes

A *force system* for a material universe \mathcal{U} is a function φ which assigns to every triple (B, C, T) , where B and C are separate bodies in \mathcal{U} and $T \in \Gamma$ is an instant, a vector $\varphi(B, C, T)$ in the translation space V_T of T . The value $\varphi(B, C, T)$ is called the *force exerted by the body* C *on the body* B *at the instant* T . Force systems are subject to restrictions which will not be stated here. (They are similar to the ones given in [3], Sect. 3, and [4], Sect. 4.)

¹ The concept of material universe used here differs from the one of references [3] and [4]. In these papers, bodies are assumed to be certain subsets of a universal set, while here they need not be sets at all.

A *dynamical process* for the material universe \mathcal{U} is a pair $\Pi = (M, \varphi)$, where M is a motion of \mathcal{U} and φ a force system for \mathcal{U} , such that the fundamental laws of balance of forces and moments are satisfied. A precise statement of these laws in the present framework and a detailed treatment of dynamics will be given in future publications. The totality of all dynamical processes for \mathcal{U} will be denoted by \mathcal{D} .

An automorphism α of W induces a transformation $\varphi \rightarrow \varphi^\alpha$ on force systems; it is defined by

$$\varphi^\alpha(B, C, T^\alpha) = \alpha(\varphi(B, C, T)), \quad (5.1)$$

where the right hand side is the image under the isomorphism $V_T \rightarrow V_{T^\alpha}$ induced by α . The automorphism α induces the transformation

$$\Pi = (M, \varphi) \rightarrow \Pi^\alpha = (M^\alpha, \varphi^\alpha) \quad (5.2)$$

on the set \mathcal{D} of all dynamical processes, where M^α is defined by (4.3) and φ^α by (5.1). The fundamental laws of balance are invariant under transformations of the form (5.2), so that dynamical processes are transformed into dynamical processes.

In order to have consistency of the theory described here with the conventional approaches to classical mechanics one must include inertial forces in the force systems φ on an equal footing with other kinds of forces (cf. [3]).

6. Constitutive Classes

The nature of many problems in mechanics can roughly be described as follows: among all conceivable dynamical processes for a material universe, select the one that will actually occur. In order to make such a selection, one must know something about the particular material properties of the bodies which belong to the material universe. Conventionally such properties are described by "force laws", "stress-strain relations", or similar types of constitutive laws.

A way of making precise the concept of a material property within the present framework is that of using the notion of a constitutive class: A *constitutive class* for a pair B, C of separate bodies is a subset $\mathcal{C}(B, C)$ of the set \mathcal{D} of all dynamical processes, subject to the following requirements:

(I) If $\Pi_1 = (M_1, \varphi_1)$ and $\Pi_2 = (M_2, \varphi_2)$ are two dynamical processes such that

$$M_1(D) = M_2(D) \quad (6.1)$$

for all parts $D < B$ or $D < C$ and

$$\varphi_1(B, C, T) = \varphi_2(B, C, T) \quad (6.2)$$

for all instants $T \in I$, then $\Pi_1 \in \mathcal{C}(B, C)$ if and only if $\Pi_2 \in \mathcal{C}(B, C)$.

(II) $\mathcal{C}(B, C)$ is stable under automorphisms; i.e., if $II \in \mathcal{C}(B, C)$ and if α is an automorphism of W , then $II^\alpha \in \mathcal{C}(B, C)$, where II^α is defined by (5.2).

The first of these requirements is a principle of irrelevance, stating that the material properties of the bodies B and C concern only B and C , and not anything else in the universe. The second requirement is a principle of homogeneity for the event-world, expressing the condition that events may have no individuality beyond the one conferred to them by motions of material universes.

In conventional treatments of mechanics material properties are defined by means of *constitutive equations*. Such constitutive equations actually define classes of dynamical processes in the sense of Sect. 5 if a suitable concept of *frame of reference* is employed. One can show that these constitutive equations must satisfy the principle of material frame-indifference and the principle of local action (see [1], Sect. 23) if they are to define constitutive classes in the sense described above.

REFERENCES

- [1] TRUESDELL, C., and W. NOLL: The non-linear field theories of mechanics. In: *Handbuch der Physik* (S. FLÜGGE, ed.), vol. III/3. Berlin-Göttingen-Heidelberg: Springer 1965.
- [2] NOLL, W.: Euclidean geometry and Minkowskian chronometry. *Am. Math. Monthly* **71**, 129—144 (1964).
- [3] — La Mécanique Classique Basée Sur Un Axiome D'Objectivité, *Colloque internat. sur la méthode axiomatique dans les mécaniques classiques et nouvelles* 1959, p. 47—56. Paris 1963.
- [4] — The foundations of classical mechanics in the light of recent advances in continuum mechanics. *Proc. of the Berkeley Symposium on the Axiomatic Method*, p. 266—281, Amsterdam 1959.

Materially Uniform Simple Bodies with Inhomogeneities

WALTER NOLL

Contents

1. Introduction	1
2. Deformations	3
3. Continuous Bodies	4
4. Tangent Spaces	5
5. Simple Bodies, Material Isomorphisms, Intrinsic Isotropy Groups	6
6. Material Uniformity, Uniform References, Relative Isotropy Groups	7
7. Vector and Tensor Fields	10
8. Relative Gradients, Brackets	12
9. Affine Connections, Torsion, Curvature	15
10. Material Connections	18
11. Inhomogeneity	20
12. Relative Riemannian Structures, Contortion	22
13. Contorted Aeolotropy	24
14. Special Types of Materially Uniform Bodies	26
15. CAUCHY'S Equation of Balance	27
References	32

1. Introduction

The basic concepts of the theory of simple materials have been introduced in reference [1] (see also the exposition in [2], Chapter C III). Here I present a detailed study of the structure of bodies that consist of a uniform simple material yet are not necessarily homogeneous.

After assembly of the necessary mathematical tools in Sects. 2–4, the concept of a *simple body* is introduced in Sect. 5. This concept is more inclusive than the one described in [1] because it can be appropriate not only to mechanical material properties, but also to thermal, optical, electrical, magnetic, or any other type of material properties. A body may be simple with respect to any particular such material property or to any combination of them. The physical theory relevant to these properties need not be made explicit.

In Sect. 6 a precise definition of a *materially uniform simple body* is given. The nature of the coherence of a uniform body with respect to the local material properties under consideration can be described mathematically in terms of what I call a *material uniformity* or in terms of what I call a *uniform reference*. In general, neither of these is uniquely determined by the simple body structure, but the degree of non-uniqueness can be delimited precisely. There may or may not exist uniform references that are gradients of global configurations. If they do exist, the body is homogeneous, and the theory becomes trivial.

This paper supersedes an unpublished preliminary study written by the author in 1963. Section 34 of reference [2] is a summary of that study.

Sections 7–9 contain an exposition of the mathematical prerequisites necessary to describe the local behavior of material uniformities and uniform references that possess a degree of smoothness. In the remainder of the paper, such smoothness is always assumed. A material uniformity is then equivalent to an affine connection, which is defined in Sect. 10 and called a *material connection*. The Cartan torsion of this connection describes locally the deviation from homogeneity and is therefore called, in Sect. 11, the *inhomogeneity* of the given material uniformity.

Associated with each smooth uniform reference is also a Riemannian structure on the body, and the relation of this structure to the material connection is studied in Sect. 12. The difference between the Riemannian connection and the material connection determines what I call the *contortion* of the given uniform reference. Contortion and inhomogeneity determine one another.

Of particular interest is a special type of non-homogeneity called *contorted aeolotropy* in Section 13. It generalizes the more familiar *curvilinear aeolotropy*. In contorted aeolotropy, the deviation from homogeneity is given by a distribution of rotations on a suitable global configuration, and the contortion describes the local behavior of this distribution. The curvature of the Riemannian structure mentioned before describes locally the deviation from contorted aeolotropy.

Section 14 contains a number of results that apply when the response functions of the body have special properties, especially with respect to material symmetry.

The usual version of CAUCHY's equation of balance (*cf.* [2], (16.6)) is very useful only when applied to bodies that are homogeneous. For applications to materially uniform but inhomogeneous bodies, a new version of CAUCHY's equation, derived in Sect. 15, is much more suitable than the usual one. This new version gives rise, for example, to a definite differential equation for the theory of inhomogeneous but materially uniform elastic bodies.

Unfortunately, there is no easily accessible exposition of the coordinate-free type of modern differential geometry that is the most appropriate for the applications in this study. The monograph of LANG [3], although it explains some of the concepts used here, does not contain sufficient material and emphasizes matters not relevant in the present context. For this reason I develop in this paper all mathematical tools as they come to be needed, tailored to the requirements of the intended applications.

There is a large literature on theories of *continuous distributions of dislocations*, proposed in various forms by KONDO, NYE, BILBY, BULLOUGH, SMITH, SEEGER, KRÖNER, GÜNTHER, and others¹. Motivated by heuristic considerations, mostly concerning lattice defects in crystals, these authors lay down *a priori* certain geometric structures to describe distributions of dislocations. These geometric structures are formally of the same type as some of those occurring in the present paper. The conceptual status of the theory presented here, however, is very different. I show that once a constitutive assumption defining a materially uniform simple body is laid down, the geometric structures of the body are determined. The geometry is thus the *natural outcome*, not the first assumption, of the theory. Since the underlying constitutive assumption is very general, the real materials to which the theory can be expected to apply need be neither crystalline, nor elastic, nor solid.

¹ For details and references I refer to the expository articles [4] and [5].

2. Deformations

We shall employ the concept of *absolute physical space*², as is customary in classical physics. This space \mathcal{E} , whose elements $\mathbf{x}, \mathbf{y}, \dots$, we call *spatial points*, has the structure of a three-dimensional Euclidean point space³. The translation space of \mathcal{E} is denoted by \mathcal{V} ; it is a three-dimensional inner product space. The elements $\mathbf{u}, \mathbf{v}, \dots$, of \mathcal{V} are called *spatial vectors*. The translation which carries $\mathbf{x} \in \mathcal{E}$ to $\mathbf{y} \in \mathcal{E}$ is denoted by $\mathbf{y} - \mathbf{x} \in \mathcal{V}$, and $\mathbf{x} + \mathbf{u}$ denotes the point into which $\mathbf{x} \in \mathcal{E}$ is carried by the translation $\mathbf{u} \in \mathcal{V}$. The inner product of two spatial vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$. Of course, $\mathbf{u} \cdot \mathbf{v} \in \mathcal{R}$, where \mathcal{R} is the set of all real numbers.

The set of all linear transformations $\mathbf{L}: \mathcal{V} \rightarrow \mathcal{V}$ of \mathcal{V} into itself is denoted by \mathcal{L} . The composition of $\mathbf{L} \in \mathcal{L}$ with $\mathbf{M} \in \mathcal{L}$ is denoted by $\mathbf{ML} \in \mathcal{L}$. The identity transformation on \mathcal{V} is denoted by $\mathbf{1} \in \mathcal{L}$. The transpose of $\mathbf{L} \in \mathcal{L}$ is denoted by \mathbf{L}^T , so that $\mathbf{u} \cdot \mathbf{L}\mathbf{v} = \mathbf{L}^T \mathbf{u} \cdot \mathbf{v}$ holds for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. The trace and determinant of $\mathbf{L} \in \mathcal{L}$ are denoted by $\text{tr } \mathbf{L}$ and $\det \mathbf{L}$, respectively. The set \mathcal{L} of all linear transformations has the natural structure of a nine-dimensional algebra. It is also endowed with a natural inner product, whose values are given by $\mathbf{L} \cdot \mathbf{M} = \text{tr}(\mathbf{LM}^T)$. A transformation $\mathbf{L} \in \mathcal{L}$ is said to be *invertible* if it is a bijection (*i.e.*, one-to-one and onto). In this case, there exists an *inverse* $\mathbf{L}^{-1} \in \mathcal{L}$ so that $\mathbf{LL}^{-1} = \mathbf{L}^{-1}\mathbf{L} = \mathbf{1}$. The invertible members of \mathcal{L} form a group $\ell \subset \mathcal{L}$ under composition; it is called the *linear group* of \mathcal{V} . Important subgroups of ℓ are the *unimodular group*

$$u = \{ \mathbf{H} \in \ell \mid |\det \mathbf{H}| = 1 \}$$

and the *orthogonal group*

$$o = \{ \mathbf{Q} \in \ell \mid \mathbf{Q}\mathbf{Q}^T = \mathbf{1} \}.$$

Of course, o is a subgroup of u .

Consider a mapping $\varphi: \mathcal{G} \rightarrow \mathcal{E}'$ of an open subset $\mathcal{G} \subset \mathcal{E}$ into a point-space or vector-space \mathcal{E}' . Let \mathcal{V}' be the translation space of \mathcal{E}' ($\mathcal{V}' = \mathcal{E}'$ if \mathcal{E}' is already a vectorspace) and let $\mathcal{L}(\mathcal{V}, \mathcal{V}')$ be the space of all linear transformations of \mathcal{V} into \mathcal{V}' . We say that φ is of class C^1 if there is a continuous mapping $\nabla\varphi: \mathcal{G} \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V}')$ such that

$$\varphi(\mathbf{x} + \mathbf{u}) = \varphi(\mathbf{x}) + (\nabla\varphi(\mathbf{x}))\mathbf{u} + \sigma(\mathbf{x}, \mathbf{u}),$$

where

$$\lim_{|\mathbf{u}| \rightarrow 0} \frac{1}{|\mathbf{u}|} \sigma(\mathbf{x}, \mathbf{u}) = \mathbf{0}$$

holds for all $\mathbf{x} \in \mathcal{G}$. The mapping $\nabla\varphi$, if it exists, is uniquely determined by φ and is called the *gradient* of φ . If $\nabla\varphi$ exists and is itself of class C^1 , we say that φ is of class C^2 . The gradient of $\nabla\varphi$ is denoted by $\nabla^{(2)}\varphi$ and is called the *second gradient* of φ . Continuing in this manner, we say that φ is of class C^r , if it is of class C^{r-1} and if its $(r-1)$ st gradient $\nabla^{(r-1)}\varphi$ is of class C^1 . The gradient of $\nabla^{(r-1)}\varphi$ is denoted by $\nabla^{(r)}\varphi$. We say that φ is of class C^0 if it is merely continuous. If φ is of class C^2 , its second gradient has the symmetry property $((\nabla^{(2)}\varphi)\mathbf{u})\mathbf{v} = ((\nabla^{(2)}\varphi)\mathbf{v})\mathbf{u}$, $\mathbf{u}, \mathbf{v} \in \mathcal{V}$.

² The considerations of this paper can be adapted to the neoclassical space-time explained in [6]. When this is done, absolute space must be replaced by suitably defined "instantaneous spaces".

³ The exact meaning of this term is explained in [7], Sect. 4.

The modifier “of class C^r ” may apply, in particular, to a *scalar field*, i.e. a mapping $f: \mathcal{G} \rightarrow \mathcal{R}$, a *vector field*, i.e. a mapping $h: \mathcal{G} \rightarrow \mathcal{V}$, or a *tensor field*, i.e., a mapping $T: \mathcal{G} \rightarrow \mathcal{L}$. A one-to-one mapping $\lambda: \mathcal{G} \rightarrow \mathcal{E}$ is called a *deformation of class C^r* ($r \geq 1$) if it is not only of class C^r but if also the values of its gradient are invertible, i.e. if $\nabla \lambda(x) \in \ell$ for all $x \in \mathcal{G}$.

The members of the linear group ℓ are also called *local deformations*, so that a (global) deformation has a gradient whose values are local deformations.

3. Continuous Bodies

A physical object can often be described mathematically by the concept of a *body* \mathcal{B} , which is a set whose members X, Y, \dots , are called *material points*, and which is endowed with a structure defined by a class C of mappings $\kappa: \mathcal{B} \rightarrow \mathcal{E}$. The mappings $\kappa \in C$ are called the *configurations* of \mathcal{B} (in the space \mathcal{E}). The spatial point $\kappa(X) \in \mathcal{E}$ is called the *place* of the material point $X \in \mathcal{B}$ in the configuration κ .

We say that \mathcal{B} is a *continuous body of class C^p* ($p \geq 1$) if the class C of configurations satisfies the following axioms:

(C 1) Every $\kappa \in C$ is one-to-one and its range $\kappa(\mathcal{B})$ is an open subset of \mathcal{E} , which is called the *region occupied by \mathcal{B} in the configuration κ* .

(C 2) If $\gamma, \kappa \in C$ then the composite⁴ $\lambda = \gamma \circ \kappa: \kappa(\mathcal{B}) \rightarrow \gamma(\mathcal{B})$ is a deformation of class C^p , which is called the *deformation of \mathcal{B} from the configuration κ into the configuration γ* .

(C 3) If $\kappa \in C$ and if $\lambda: \kappa(\mathcal{B}) \rightarrow \mathcal{E}$ is a deformation of class C^p , then $\lambda \circ \kappa \in C$. The mapping $\lambda \circ \kappa$ is called the *configuration obtained from the configuration κ by the deformation λ* .

In the remainder of this paper we shall always assume that \mathcal{B} is a continuous body of class C^p , $p \geq 1$.

The axioms (C 1)–(C 3) ensure that the class C endows the body \mathcal{B} with the structure of a “ C^p -manifold modelled on \mathcal{E} ” in the sense of LANG ([3], Ch. II, §1). Topologically, it is a very simple manifold because it can be mapped out with a single “chart” (“configuration” in our terminology).

Of central importance for the present paper is the concept of a *local configuration*⁵ at a material point X . Two (global) configurations κ and γ are said to be *equivalent at X* , and we write⁶

$$\kappa \sim_X \gamma \quad \text{if} \quad \nabla(\kappa \circ \gamma^{-1})|_{\gamma(X)} = \mathbf{1}. \quad (3.1)$$

It is an immediate consequence of the chain rule for gradients that \sim_X is an equivalence relation on C . The resulting partition of C is denoted by \mathcal{C}_X , and its members K_X, G_X, \dots , i.e. the equivalence classes, are called *local configurations at X* . Instead of writing $\kappa \in K_X$ when κ is a member of the class K_X we often write

$$\nabla \kappa(X) = K_X \quad (3.2)$$

⁴ Composition of mappings other than linear mappings is denoted by \circ . The inverse of a one-to-one mapping κ is denoted by κ^{-1} .

⁵ The term “configuration gradient” was used and another meaning was assigned to the term “local configuration” in [I].

⁶ For better reading, we sometimes write $f|_x$ instead of $f(x)$ for the value of f at x .

and say that the local configuration \mathbf{K}_X is the *gradient at X* of the (global) configuration κ .

Let $\mathbf{K}_X, \mathbf{G}_X \in \mathcal{C}_X$ be two local configurations, and let $\kappa \in \mathbf{K}_X, \gamma \in \mathbf{G}_X$. It is easily seen that the local deformation $\mathcal{V}(\gamma \circ \kappa^{-1})|_{\kappa(X)} \in \ell$ depends only on \mathbf{K}_X and \mathbf{G}_X , and not on the particular choices of $\kappa \in \mathbf{K}_X$ and $\gamma \in \mathbf{G}_X$. We denote this local deformation by $\mathbf{G}_X \mathbf{K}_X^{-1}$ and call it the *local deformation from the local configuration \mathbf{K}_X into the local configuration \mathbf{G}_X* . Using the notation (3.2), we then have

$$\mathcal{V}(\gamma \circ \kappa^{-1})|_{\kappa(X)} = \mathcal{V}\gamma(X) [\mathcal{V}\kappa(X)]^{-1}. \quad (3.3)$$

If $\mathbf{K}_X \in \mathcal{C}_X$ is a local configuration and $L \in \ell$ any local deformation, we can define a new local configuration $L\mathbf{K}_X \in \mathcal{C}_X$ by

$$L\mathbf{K}_X = \{\lambda \circ \kappa \mid \mathcal{V}\lambda|_{\kappa(X)} = L, \mathcal{V}\kappa(X) = \mathbf{K}_X\}. \quad (3.4)$$

We call $L\mathbf{K}_X$ the *local configuration obtained from the local configuration \mathbf{K}_X by the local deformation L*. Clearly, we have the rules

$$(\mathbf{G}_X \mathbf{K}_X^{-1})\mathbf{K}_X = \mathbf{G}_X, \quad (L\mathbf{K}_X)\mathbf{K}_X^{-1} = L. \quad (3.5)$$

4. Tangent Spaces

Consider pairs $(\mathbf{K}_X, \mathbf{u})$, where $\mathbf{K}_X \in \mathcal{C}_X$ is a local configuration at X and $\mathbf{u} \in \mathcal{V}$ a spatial vector. We say that two such pairs $(\mathbf{K}_X, \mathbf{u})$ and $(\mathbf{G}_X, \mathbf{v})$ are *equivalent* if

$$(\mathbf{K}_X \mathbf{G}_X^{-1})\mathbf{v} = \mathbf{u}. \quad (4.1)$$

It follows from the rules (3.5) that (4.1) does indeed define an equivalence relation. The resulting equivalence classes are called *tangent vectors* u_X, v_X, \dots at X . The totality of all these tangent vectors is denoted by \mathcal{T}_X and is called the *tangent space* at $X \in \mathcal{B}$. Let $u_X \in \mathcal{T}_X$ and $\mathbf{K}_X \in \mathcal{C}_X$ be given and let $(\mathbf{G}_X, \mathbf{v})$ be any pair belonging to the class u_X . Now, if $(\mathbf{K}_X, \mathbf{u})$ is to belong to u_X then (4.1) must hold. Therefore, we see that $u_X \in \mathcal{T}_X$ and $\mathbf{K}_X \in \mathcal{C}_X$ determine a unique spatial vector $\mathbf{u} \in \mathcal{V}$ such that $(\mathbf{K}_X, \mathbf{u}) \in u_X$. We can therefore use the notation

$$\mathbf{u} = \mathbf{K}_X u_X, \quad u_X = \mathbf{K}_X^{-1} \mathbf{u} \quad \text{if } (\mathbf{K}_X, \mathbf{u}) \in u_X, \quad (4.2)$$

and we see that \mathbf{K}_X determines a one-to-one mapping of the tangent space \mathcal{T}_X onto the space \mathcal{V} of spatial vectors. The tangent space \mathcal{T}_X has the natural structure of a three-dimensional vector space, with addition defined by

$$u_X + v_X = \mathbf{K}_X^{-1}(\mathbf{u} + \mathbf{v}) \quad \text{if } u_X = \mathbf{K}_X^{-1} \mathbf{u}, \quad v_X = \mathbf{K}_X^{-1} \mathbf{v} \quad (4.3)$$

and multiplication with scalars by

$$a u_X = \mathbf{K}_X^{-1}(a \mathbf{u}) \quad \text{if } u_X = \mathbf{K}_X^{-1} \mathbf{u}, \quad a \in \mathcal{R}. \quad (4.4)$$

It is immediately seen that these definitions of $u_X + v_X$ and $a u_X$ are legitimate because the results are independent of the choice of the local configuration \mathbf{K}_X used to represent u_X and v_X in \mathcal{V} . The local configurations can be identified with the invertible linear transformations of \mathcal{T}_X onto \mathcal{V} .

Given a local configuration $\mathbf{K}_X \in \mathcal{C}_X$, we can define an inner product $\mathbf{u}_X * \mathbf{v}_X$ of $\mathbf{u}_X, \mathbf{v}_X \in \mathcal{T}_X$ by

$$\mathbf{u}_X * \mathbf{v}_X = (\mathbf{K}_X \mathbf{u}_X) \cdot (\mathbf{K}_X \mathbf{v}_X). \quad (4.5)$$

However, we obtain different inner products on \mathcal{T}_X for different choices of \mathbf{K}_X , and hence \mathcal{T}_X is *not* naturally an inner product space.

5. Simple Bodies, Material Isomorphisms, Intrinsic Isotropy Groups

To describe mathematically the physical characteristics of a body \mathcal{B} we must endow \mathcal{B} with additional structure. Some of these characteristics, such as elasticity, viscosity, heat capacity, and electrical conductivity, are local, *i. e.*, they are attached to the individual material points $X \in \mathcal{B}$ rather than to the body as a whole. Other characteristics, such as mutual gravitation and internal radiative heat transfer, involve more than one material point. We deal here only with local characteristics. The physical response of the body \mathcal{B} at a particular material point $X \in \mathcal{B}$ and a particular time will depend on the configuration κ of \mathcal{B} at that time. It may happen that only the local configuration $\nabla \kappa(X)$ at X determined by κ , and no other properties of κ , has an influence on the response. If this is the case, we say that the *material at X is simple*. We say that the whole of \mathcal{B} is simple or that \mathcal{B} is a *simple body* if the material at X is simple for all $X \in \mathcal{B}$.

We assume that a possible physical response at a material point is given mathematically by specifying an element from a set \mathbf{R} of mathematical objects. The nature of \mathbf{R} depends on the particular physical phenomena to be described. For example, in the theory of elasticity \mathbf{R} consists of all possible ‘stress tensors’, *i. e.*, of all symmetric linear transformations of \mathcal{V} into \mathcal{V} . In the mechanical theory of simple materials with fading memory, \mathbf{R} consists of “memory functionals” that relate relative deformation histories to stresses and are subject to certain smoothness requirements. In theories that include non-mechanical effects \mathbf{R} consists of functions or functionals whose independent and dependent variables have interpretations as local temperatures, energy or entropy densities, heat fluxes, electric or magnetic field strengths, polarizations, magnetizations, electric currents, *etc.* For the purpose of the present paper, no specific assumptions about the nature of \mathbf{R} need be made.

We can now make our definition of a simple body precise:

Definition 1. *Let \mathbf{R} be a set, whose elements we call **response descriptors**. A continuous body \mathcal{B} of class C^p will be called a **simple body** with respect to \mathbf{R} if it is endowed with a structure by a function \mathfrak{G} which assigns to each material point $X \in \mathcal{B}$ a mapping*

$$\mathfrak{G}_X: \mathcal{C}_X \rightarrow \mathbf{R}. \quad (5.1)$$

The value $\mathfrak{G}_X(\mathbf{G}_X)$ is the response descriptor of the material at X in any configuration γ of \mathcal{B} such that $\nabla \gamma(X) = \mathbf{G}_X$.

The mappings \mathfrak{G}_X cannot be entirely arbitrary, for they are subject to restrictions imposed by general physical principles such as the principle of frame-indifference and the principle of dissipation. These restrictions need not be made explicit here.

We would like to give now a precise meaning to the statement that the material at one point $X \in \mathcal{B}$ is the same as the material at another point $Y \in \mathcal{B}$. We cannot construe this statement to mean that \mathfrak{G}_X and \mathfrak{G}_Y are the same, for they have different domains and hence cannot be directly compared. However, we can connect the domains \mathcal{C}_X and \mathcal{C}_Y if an isomorphism $\Phi_{XY}: \mathcal{T}_Y \rightarrow \mathcal{T}_X$ of the tangent space at Y onto the tangent space at X is given. Recalling that a local configuration $\mathbf{G}_X \in \mathcal{C}_X$ can be regarded as a mapping $\mathbf{G}_X: \mathcal{T}_X \rightarrow \mathcal{V}$, we can let $\mathbf{G}_X \in \mathcal{C}_X$ correspond to the composition $\mathbf{G}_X \Phi_{XY} \in \mathcal{C}_Y$. We are thus led to the following definition:

Definition 2. An invertible linear transformation $\Phi_{XY}: \mathcal{T}_Y \rightarrow \mathcal{T}_X$ is called a *material isomorphism* from \mathcal{T}_Y onto \mathcal{T}_X if

$$\mathfrak{G}_X(\mathbf{G}_X) = \mathfrak{G}_Y(\mathbf{G}_X \Phi_{XY}) \quad (5.2)$$

holds for all $\mathbf{G}_X \in \mathcal{C}_X$.

To say that the material at X is the same as the material at Y means that there exists a material isomorphism from \mathcal{T}_X onto \mathcal{T}_Y .

It follows immediately from Definition 2 that if $\Phi_{XY}: \mathcal{T}_Y \rightarrow \mathcal{T}_X$ and $\Phi_{YZ}: \mathcal{T}_Z \rightarrow \mathcal{T}_Y$ are material isomorphisms, so is their composition $\Phi_{XY} \Phi_{YZ}: \mathcal{T}_Z \rightarrow \mathcal{T}_X$. Also, if $\Phi_{XY}: \mathcal{T}_Y \rightarrow \mathcal{T}_X$ is a material isomorphism, so is its inverse $\Phi_{XY}^{-1}: \mathcal{T}_X \rightarrow \mathcal{T}_Y$. If we denote the set of all material isomorphisms from \mathcal{T}_X onto \mathcal{T}_Y by g_{YX} , these facts can be expressed by⁷

$$g_{ZY} g_{YX} = g_{ZX}, \quad g_{YX} = g_{XY}^{-1}, \quad \text{valid if these sets are not empty.} \quad (5.3)$$

It is clear that g_{XX} , the set of all material isomorphisms of \mathcal{T}_X onto itself, is a subgroup of the linear group ℓ_X of \mathcal{T}_X , which consists of all invertible linear transformations of \mathcal{T}_X . We write

$$g_X = g_{XX} \quad (5.4)$$

and call g_X the *intrinsic isotropy group* of the material at X . For any $\Phi_{XY} \in g_{XY}$ one easily establishes the relations

$$g_{XY} = g_X \Phi_{XY} g_Y, \quad g_X = \Phi_{XY} g_Y \Phi_{XY}^{-1}. \quad (5.5)$$

It follows from (5.5)₂ that if a material isomorphism $\Phi_{XY}: \mathcal{T}_Y \rightarrow \mathcal{T}_X$ exists, *i.e.* if the material at X is the same as the material at Y , then the intrinsic isotropy groups g_X and g_Y are isomorphic.

6. Material Uniformity, Uniform References, Relative Isotropy Groups

A simple body \mathcal{B} is said to be *materially uniform* if the material at any two of its points is the same, *i.e.* if g_{XY} is never empty. From now on we assume that \mathcal{B} is a materially uniform simple body. We select a member $\Phi'(X, Y)$ from each g_{XY} and thereby define a function Φ' which assigns to each pair (X, Y) of material points of \mathcal{B} a material isomorphism from \mathcal{T}_Y onto \mathcal{T}_X . Choose $X_0 \in \mathcal{B}$ arbitrarily

⁷ If g and \mathcal{h} are sets of linear transformations of any kind such that the composition LM makes sense whenever $L \in g, M \in \mathcal{h}$, we write $g\mathcal{h} = \{LM \mid L \in g, M \in \mathcal{h}\}$. If the $L \in g$ are invertible, we write $g^{-1} = \{L^{-1} \mid L \in g\}$. Also, we write $Kg = \{KL \mid L \in g\}$ if KL makes sense for all $L \in g$.

and define Φ by

$$\Phi(X, Y) = \Phi'(X, X_0) \Phi'(Y, X_0)^{-1}. \quad (6.1)$$

It follows from (5.3) that $\Phi(X, Y) \in \mathcal{G}_{XY}$. Moreover, we have

$$\Phi(Z, Y) \Phi(Y, X) = \Phi(Z, X), \quad \Phi(X, X) = \mathbf{1}_X, \quad (6.2)$$

where $\mathbf{1}_X$ is the identity transformation of \mathcal{T}_X .

Definition 3. A function Φ which assigns to each pair (X, Y) of material points of \mathcal{B} a material isomorphism $\Phi(X, Y) \in \mathcal{G}_{XY}$ is called a **material uniformity** if (6.2) holds.

The construction (6.1) shows that the materially uniform bodies are those that admit material uniformities. It follows from (5.5) and (6.2) that any two material uniformities Φ and $\hat{\Phi}$ are related by

$$\hat{\Phi}(X, Y) = \mathfrak{B}(X) \Phi(X, Y) \mathfrak{B}(Y)^{-1}, \quad (6.3)$$

where \mathfrak{B} is a function on \mathcal{B} whose values $\mathfrak{B}(X)$ belong to the intrinsic isotropy groups \mathcal{G}_X .

Definition 4. A function \mathbf{K} on \mathcal{B} whose values $\mathbf{K}(X) \in \mathcal{C}_X$ are local configurations is called a **reference** for \mathcal{B} . If, moreover,

$$\Phi(X, Y) = \mathbf{K}(X)^{-1} \mathbf{K}(Y) \quad (6.4)$$

is a material isomorphism of \mathcal{T}_Y onto \mathcal{T}_X for any $X, Y \in \mathcal{B}$, then \mathbf{K} is called a **uniform reference** for \mathcal{B} .

Actually, (6.2) holds if Φ is defined by (6.4), so that Φ is a material uniformity if \mathbf{K} is a uniform reference. Hence, every uniform reference \mathbf{K} determines a material uniformity Φ through (6.4). Conversely, if a material uniformity Φ and a local configuration $\mathbf{K}_{X_0} \in \mathcal{C}_{X_0}$ for a particular material point $X_0 \in \mathcal{B}$ are given, then there exist a unique uniform reference \mathbf{K} such that (6.4) and $\mathbf{K}(X_0) = \mathbf{K}_{X_0}$ hold. In fact, \mathbf{K} is given by

$$\mathbf{K}(X) = \mathbf{K}_{X_0} \Phi(X_0, X). \quad (6.5)$$

Therefore, every material uniformity has representations (6.4) in terms of uniform references.

If κ is a (global) configuration, then $\nabla \kappa$, which assigns to X the local configuration $\nabla \kappa(X)$ at X , i.e., the equivalence class to which κ belongs, is a reference, called the *gradient* of the configuration κ . We say that a body is *homogeneous* if it admits a gradient as a uniform reference. Of course, not every reference is a gradient, and it may happen that none of the uniform references of a materially uniform body is a gradient.

Let \mathbf{K} be a uniform reference. Every local configuration $\mathbf{G}_X \in \mathcal{C}_X$ can be characterized by the local deformation $\mathbf{F} = \mathbf{G}_X \mathbf{K}(X)^{-1} \in \ell$ from $\mathbf{K}(X)$ into \mathbf{G}_X , so that

$$\mathbf{G}_X = \mathbf{F} \mathbf{K}(X). \quad (6.6)$$

Substituting (6.6) and (6.4) into (5.2) with the choice $\Phi_{XY} = \Phi(X, Y)$, we see that

$$\mathfrak{G}_X(FK(X)) = \mathfrak{G}_Y(FK(Y)) \quad (6.7)$$

must hold for all $F \in \ell$ and all $X, Y \in \mathcal{B}$. Conversely, if (6.7) holds for all $F \in \ell$ and all $X, Y \in \mathcal{B}$, then K is a uniform reference. This result may be formulated as follows:

Theorem 1. *A reference K for \mathcal{B} is uniform if and only if there is a function $\mathfrak{H}_K: \ell \rightarrow \mathbf{R}$ which satisfies*

$$\mathfrak{H}_K(F) = \mathfrak{G}_X(FK(X)) \quad (6.8)$$

for all $X \in \mathcal{B}$ and all $F \in \ell$.

The function \mathfrak{H}_K , which assigns to each local deformation a response descriptor, will be called the *response function* of the body *relative to the uniform reference K* .

Let K be uniform reference. If we substitute (6.4) for Φ_{XY} in (5.5), we see that

$$K(X)g_XK(X)^{-1} = K(Y)g_YK(Y)^{-1}, \quad (6.9)$$

i. e. that

$$g_K = K(X)g_XK(X)^{-1} \quad (6.10)$$

is independent of X . The group g_K is a subgroup of the linear group ℓ . We call g_K *the isotropy group of the body \mathcal{B} relative to the uniform reference K* . In view of (6.10), all the intrinsic isotropy groups g_X , $X \in \mathcal{B}$, are isomorphic to the relative isotropy group g_K . It is easily seen that g_K is given in terms of the response function \mathfrak{H}_K by

$$g_K = \{P \in \ell \mid \mathfrak{H}_K(F) = \mathfrak{H}_K(FP) \text{ for all } F \in \ell\}. \quad (6.11)$$

The relation between two uniform references and the corresponding response functions and isotropy groups is described by the following theorem:

Theorem 2. *Any two uniform references K and \hat{K} are related by*

$$\hat{K}(X) = LP(X)K(X), \quad (6.12)$$

where $L \in \ell$ and where P is a function on \mathcal{B} with values in g_K .

The isotropy groups g_K and $g_{\hat{K}}$ relative to K and \hat{K} are conjugate:

$$g_{\hat{K}} = Lg_KL^{-1}. \quad (6.13)$$

The response functions $\mathfrak{H}_{\hat{K}}$ and \mathfrak{H}_K are related by the identity

$$\mathfrak{H}_{\hat{K}}(F) = \mathfrak{H}_K(FL) \quad \text{for all } F \in \ell. \quad (6.14)$$

Proof. The two material uniformities Φ and $\hat{\Phi}$ given by

$$\Phi(X, Y) = K(X)^{-1}K(Y), \quad \hat{\Phi}(X, Y) = \hat{K}(X)^{-1}\hat{K}(Y)$$

must be related by (6.3). It follows that

$$\hat{K}(Y)\mathfrak{P}(Y)K(Y)^{-1} = \hat{K}(X)\mathfrak{P}(X)K(X)^{-1} = L \in \ell$$

is independent of $X \in \mathcal{B}$. Hence (6.12) holds with the choice

$$P(X) = K(X)\mathfrak{P}(X)^{-1}K(X)^{-1}. \quad (6.15)$$

It follows from (6.10) that $P(X) \in g_K$ for all $X \in \mathcal{B}$, which proves the first assertion of the theorem. If we write (6.10) with K replaced by \hat{K} and substitute (6.12), we obtain

$$\begin{aligned} g_{\hat{K}} &= LP(X)K(X)g_XK(X)^{-1}P(X)^{-1}L^{-1} \\ &= LP(X)g_KP(X)^{-1}L^{-1}. \end{aligned}$$

Since $P(X) \in g_K$ we have $P(X)g_KP(X)^{-1} = g_K$ and hence (6.13). The identity (6.14) is derived by writing (6.8) with K replaced by \hat{K} , then substituting (6.12) and observing (6.11). Q.E.D.

The theory of isotropy groups relative to a local reference configuration at a single material point⁸ extends without change to isotropy groups relative to a uniform reference K of a whole materially uniform body. In particular, we say that the uniform reference K is *undistorted* if g_K is comparable, with respect to inclusion, to the orthogonal group σ , i.e., if either $g_K \subset \sigma$ or $\sigma \subset g_K$. If there are uniform references K such that $g_K \supset \sigma$, we say that \mathcal{B} is a *uniform isotropic body*; if there are uniform references K such that $g_K \subset \sigma$, we say that \mathcal{B} is a *uniform solid body*. It is possible that a uniform simple body has no undistorted uniform references at all; such a body would be neither a solid nor isotropic.

7. Vector and Tensor Fields

As before, we assume that \mathcal{B} is a continuous body of class C^p , $p \geq 1$.

A mapping $\psi: \mathcal{B} \rightarrow \mathcal{E}'$ of \mathcal{B} into some point-space or vector-space \mathcal{E}' is said to be of class C^r , $0 \leq r \leq p$, if for every configuration $\kappa \in \mathcal{C}$, the mapping $\psi \circ \kappa^{-1}: \kappa(\mathcal{B}) \rightarrow \mathcal{E}'$ is of class C^r . In view of the axioms for \mathcal{B} it is clear that $\psi \circ \kappa^{-1}$ is of class C^r for every $\kappa \in \mathcal{C}$ if it is of class C^r for some $\kappa \in \mathcal{C}$. These definitions apply, in particular, to *functions (scalar fields)* on \mathcal{B} , i.e. mappings $f: \mathcal{B} \rightarrow \mathcal{R}$, to *vector fields* on \mathcal{B} , i.e. mappings $h: \mathcal{B} \rightarrow \mathcal{V}$, and to *tensor fields* on \mathcal{B} , i.e. mappings $T: \mathcal{B} \rightarrow \mathcal{L}$.

A mapping h which assigns to each material point $X \in \mathcal{B}$ a tangent vector $h(X) \in \mathcal{T}_X$ is called a *tangent vector field*. We say that such a tangent vector field h is of class C^r , $0 \leq r \leq p-1$, if the vector field $(\nabla \kappa)h$ on \mathcal{B} defined by

$$(\nabla \kappa)h|_X = (\nabla \kappa(X))h(X) \quad (7.1)$$

is of class C^r for some — and hence every — configuration $\kappa \in \mathcal{C}$.

The algebra of all linear transformations of the tangent space \mathcal{T}_X into itself will be denoted by \mathcal{I}_X . A mapping \mathfrak{T} which assigns to each material point $X \in \mathcal{B}$ a linear transformation $\mathfrak{T}(X) \in \mathcal{I}_X$ is called an *intrinsic tensor field*. We say that \mathfrak{T} is of class C^r , $0 \leq r \leq p-1$, if the tensor field $(\nabla \kappa)\mathfrak{T}(\nabla \kappa)^{-1}$ on \mathcal{B} defined by

$$(\nabla \kappa)\mathfrak{T}(\nabla \kappa)^{-1}|_X = \nabla \kappa(X)\mathfrak{T}(X)(\nabla \kappa(X))^{-1} \quad (7.2)$$

is of class C^r for some — and hence every — configuration $\kappa \in \mathcal{C}$.

We shall use the term *field on \mathcal{B}* for any mapping that assigns to every $X \in \mathcal{B}$ an element of some vector space (which may consist of linear or multilinear transformations).

⁸ This theory was initiated in [1], §§19–21. An exposition is given in [2], §§31–33.

We shall employ the following scheme of notation:

$\mathcal{F}_{\mathcal{B}}^r$ = set of all functions (scalar fields) of class C^r on \mathcal{B} .

$\mathcal{V}_{\mathcal{B}}^r$ = set of all vector fields of class C^r on \mathcal{B} .

$\mathcal{T}_{\mathcal{B}}^r$ = set of all tangent vector fields of class C^r on \mathcal{B} .

$\mathcal{L}_{\mathcal{B}}^r$ = set of all tensor fields of class C^r on \mathcal{B} .

$\mathcal{I}_{\mathcal{B}}^r$ = set of all intrinsic tensor fields of class C^r on \mathcal{B} .

The set $\mathcal{F}_{\mathcal{B}}^r$ is a commutative algebra under pointwise addition and multiplication. The sets $\mathcal{V}_{\mathcal{B}}^s$, $\mathcal{T}_{\mathcal{B}}^s$, $\mathcal{L}_{\mathcal{B}}^s$, and $\mathcal{I}_{\mathcal{B}}^s$ can be made modules with respect to any of the algebras $\mathcal{F}_{\mathcal{B}}^r$, $s \leq r \leq p-1$, by defining addition and scalar multiplication with functions in $\mathcal{F}_{\mathcal{B}}^r$ pointwise. For example, if $\mathfrak{h}, \mathfrak{k} \in \mathcal{T}_{\mathcal{B}}^r$ and $f \in \mathcal{F}_{\mathcal{B}}^r$, we define $\mathfrak{h} + \mathfrak{k} \in \mathcal{T}_{\mathcal{B}}^r$ and $f \mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^r$ by

$$(\mathfrak{h} + \mathfrak{k})|_X = \mathfrak{h}(X) + \mathfrak{k}(X), \quad (f \mathfrak{h})|_X = f(X) \mathfrak{h}(X), \quad X \in \mathcal{B}. \quad (7.3)$$

The sets $\mathcal{L}_{\mathcal{B}}^r$ and $\mathcal{I}_{\mathcal{B}}^r$ become associative (but not commutative) algebras over $\mathcal{F}_{\mathcal{B}}^r$ if multiplication is defined pointwise.

It is evident that we have $\mathcal{F}_{\mathcal{B}}^r \subset \mathcal{F}_{\mathcal{B}}^s$ if $s \leq r$ and similar inclusions for the other sets in the list given above. Actually, $\mathcal{F}_{\mathcal{B}}^r$ is a subalgebra of $\mathcal{F}_{\mathcal{B}}^s$. Also, $\mathcal{V}_{\mathcal{B}}^r$ is not only a $\mathcal{F}_{\mathcal{B}}^r$ -module, but also a submodule of $\mathcal{V}_{\mathcal{B}}^s$, regarded as a $\mathcal{F}_{\mathcal{B}}^r$ -module. Analogous observations apply to the other modules and algebras of the list above.

If $T \in \mathcal{L}_{\mathcal{B}}^s$ and $\mathbf{h} \in \mathcal{V}_{\mathcal{B}}^r$ or $\mathfrak{I} \in \mathcal{I}_{\mathcal{B}}^s$ and $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^r$, we define $T\mathbf{h}$ or $\mathfrak{I}\mathfrak{h}$ pointwise, *i. e.* by

$$T\mathbf{h}|_X = T(X)\mathbf{h}(X), \quad \mathfrak{I}\mathfrak{h}|_X = \mathfrak{I}(X)\mathfrak{h}(X). \quad (7.4)$$

When $s \leq r$, one can see that $T\mathbf{h} \in \mathcal{V}_{\mathcal{B}}^s$, $\mathfrak{I}\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^s$. It is evident from (7.4)₁ that the rules

$$T(\mathbf{h} + \mathbf{k}) = T\mathbf{h} + T\mathbf{k}, \quad T(f\mathbf{h}) = fT\mathbf{h} \quad (7.5)$$

are valid. Hence every $T \in \mathcal{L}_{\mathcal{B}}^s$ gives rise to a mapping

$$T: \mathcal{V}_{\mathcal{B}}^r \rightarrow \mathcal{V}_{\mathcal{B}}^s \quad (7.6)$$

which satisfies the rules (7.5) for $\mathbf{h}, \mathbf{k} \in \mathcal{V}_{\mathcal{B}}^r$, $f \in \mathcal{F}_{\mathcal{B}}^r$. Mappings of the type (7.6) satisfying the rules (7.5) are homomorphism with respect to the $\mathcal{F}_{\mathcal{B}}^r$ -module structures of $\mathcal{V}_{\mathcal{B}}^r$ and $\mathcal{V}_{\mathcal{B}}^s$. We also call them \mathcal{F} -linear mappings. Thus, every $T \in \mathcal{L}_{\mathcal{B}}^s$ gives rise to an \mathcal{F} -linear mapping (7.6). It is remarkable that the converse is also true, *i. e.* that every \mathcal{F} -linear mapping of the type (7.6) arises from a tensor field of class C^s on \mathcal{B} :

Proposition 1. *If $T: \mathcal{V}_{\mathcal{B}}^r \rightarrow \mathcal{V}_{\mathcal{B}}^s (s \leq r)$ is \mathcal{F} -linear, then there exists a unique tensor field $\bar{T} \in \mathcal{L}_{\mathcal{B}}^s$ such that $\bar{T}\mathbf{h} = T\mathbf{h}$ holds for all $\mathbf{h} \in \mathcal{V}_{\mathcal{B}}^r$.*

Proof. Let (e_1, e_2, e_3) be a basis of \mathcal{V} . The vectors e_i can be regarded as constant vector fields on \mathcal{B} , so that $e_i \in \mathcal{V}_{\mathcal{B}}^{p-1} \subset \mathcal{V}_{\mathcal{B}}^r$. Every $\mathbf{h} \in \mathcal{V}_{\mathcal{B}}^r$ has a unique component representation

$$\mathbf{h} = \sum_i h^i e_i, \quad h^i \in \mathcal{F}_{\mathcal{B}}^r. \quad (7.7)$$

Applying the given \mathcal{F} -linear mapping T to (7.7), we obtain

$$T\mathbf{h} = \sum_i h^i T e_i. \quad (7.8)$$

Now, if there is a tensor field \bar{T} such that $\bar{T}h = Th$ for all $h \in \mathcal{V}_{\mathcal{B}}^r$, we must have, in particular, $Te_i = \bar{T}e_i$, i. e.

$$\bar{T}(X)e_i = (Te_i)|_X \quad (7.9)$$

for all $X \in \mathcal{B}$. But since (e_1, e_2, e_3) is a basis of \mathcal{V} , we can find, for each $X \in \mathcal{B}$, exactly one $\bar{T}(X) \in \mathcal{L}$ such that (7.9) holds. Since the vector fields Te_i are of class C^s , it is easily seen that the tensor field \bar{T} obtained in this way is also of class C^s . Moreover, in view of (7.8), (7.9) and the \mathcal{F} -linearity of \bar{T} we have

$$\bar{T}h = \sum_i h^i \bar{T}e_i = \sum_i h^i Te_i = T(\sum_i h^i e_i) = Th$$

for all h . Q.E.D.

Proposition 1 enables us to identify the set of all \mathcal{F} -linear mappings of the type (7.6) with the set $\mathcal{L}_{\mathcal{B}}^s$ of all tensor fields of class C^s on \mathcal{B} . Similarly, we can identify the set of all \mathcal{F} -linear mappings of the type

$$\mathfrak{I}: \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^s \quad (s \leq r)$$

with the set $\mathcal{I}_{\mathcal{B}}^s$ of all intrinsic tensor fields of class C^s on \mathcal{B} . The proof of this fact follows from Proposition 1 by choosing a configuration κ of \mathcal{B} and letting \mathfrak{I} correspond to $\bar{T} = \nabla_{\kappa} \mathfrak{I} (\nabla_{\kappa})^{-1}: \mathcal{V}_{\mathcal{B}}^r \rightarrow \mathcal{V}_{\mathcal{B}}^s$. The result just stated is a special case of a general proposition referring to \mathcal{F} -multilinear mappings. For later application we state another special case:

Proposition 2. *If*

$$\mathfrak{S}: \mathcal{T}_{\mathcal{B}}^r \times \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^s \quad (\text{or } \mathcal{I}_{\mathcal{B}}^s) \quad (7.10)$$

is \mathcal{F} -bilinear (i. e. \mathcal{F} -linear in each of the two variables), then there exists a unique field $\bar{\mathfrak{S}}$ on \mathcal{B} whose values $\bar{\mathfrak{S}}(X)$ are bilinear mappings

$$\bar{\mathfrak{S}}(X): \mathcal{T}_X \times \mathcal{T}_X \rightarrow \mathcal{T}_X \quad (\text{or } \mathcal{I}_X) \quad (7.11)$$

such that

$$\bar{\mathfrak{S}}(X)(\mathfrak{h}(X), \mathfrak{f}(X)) = \mathfrak{S}(\mathfrak{h}, \mathfrak{f})|_X \quad (7.12)$$

holds for all $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^r$ and all $X \in \mathcal{B}$. The function $\bar{\mathfrak{S}}$ is of class C^s (in the obvious sense).

8. Relative Gradients, Brackets

From now on we assume that \mathcal{B} is a continuous body of class C^p with $p \geq 2$.

Let $\psi: \mathcal{B} \rightarrow \mathcal{E}'$ be a mapping of class \mathcal{C}^r , $1 \leq r \leq p$, where \mathcal{E}' is some point space or vector space. Given a configuration κ of \mathcal{B} , we can then define

$$\nabla_{\kappa} \psi: \mathcal{B} \rightarrow \mathcal{L}(\mathcal{V}, \mathcal{V}') \quad (8.1)$$

where \mathcal{V}' is the translation space of \mathcal{E}' , by $\nabla_{\kappa} \psi = \nabla(\psi \circ \kappa^{-1}) \circ \kappa$, i. e.

$$\nabla_{\kappa} \psi|_X = \nabla(\psi \circ \kappa^{-1})|_{\kappa(X)}, \quad X \in \mathcal{B}. \quad (8.2)$$

We call $\nabla_{\kappa} \psi$ the *gradient of ψ relative to the configuration κ* . It is clear that $\nabla_{\kappa} \psi$ is of class C^{r-1} .

Let $\kappa, \gamma \in \mathbf{C}$ be two configurations. Taking the gradient of $\psi \circ \kappa^{-1} = (\psi \circ \gamma^{-1}) \circ (\gamma \circ \kappa^{-1})$ and using the chain rule, we see with the help of (3.3) that the gradients of ψ rela-

tive to κ and γ are related by

$$\nabla_{\kappa}\psi|_X = \nabla_{\gamma}\psi|_X \circ [\nabla\gamma(X) (\nabla\kappa(X))^{-1}]. \quad (8.3)$$

Let \mathbf{K} and \mathbf{G} be two references for \mathcal{B} (see Definition 4). We define $\mathbf{K}\mathbf{G}^{-1}$ pointwise, *i.e.* by

$$\mathbf{K}\mathbf{G}^{-1}|_X = \mathbf{K}(X)\mathbf{G}(X)^{-1}. \quad (8.4)$$

Recalling that the configuration gradients $\nabla\kappa$ and $\nabla\gamma$ are references, we see that (8.3) can then be written as

$$\nabla_{\kappa}\psi = \nabla_{\gamma}\psi \circ (\nabla\gamma(\nabla\kappa)^{-1}). \quad (8.5)$$

We say that a reference \mathbf{K} is of class C^r , $r \leq p-1$, if for some $-$ and hence every $-$ configuration $\kappa \in \mathbf{C}$ the tensor field $(\nabla\kappa)\mathbf{K}^{-1}$ is of class C^r , *i.e.* belongs to $\mathcal{L}_{\mathcal{B}}^r$. It is clear that every gradient reference $\nabla\kappa$ is of class C^{p-1} .

Let a local configuration $\mathbf{K}_X \in \mathcal{C}_X$ be given. If κ and γ both belong to the equivalence class that defines \mathbf{K}_X , which means that $\nabla\kappa(X) = \nabla\gamma(X) = \mathbf{K}_X$, we have, by (8.3), $\nabla_{\kappa}\psi(X) = \nabla_{\gamma}\psi(X)$. Hence, $\nabla_{\kappa}\psi(X)$ depends on κ only through the equivalence class $\mathbf{K}_X \in \mathcal{C}_X$ to which κ belongs, and it is legitimate to define

$$\nabla_{\mathbf{K}_X}\psi(X) = \nabla_{\kappa}\psi(X) \quad \text{if } \kappa \in \mathbf{K}_X. \quad (8.6)$$

If \mathbf{K} is a reference, we define the *gradient of ψ relative to the reference \mathbf{K}* by

$$\nabla_{\mathbf{K}}\psi|_X = \nabla_{\mathbf{K}(X)}\psi(X), \quad X \in \mathcal{B}. \quad (8.7)$$

If \mathbf{K} and \mathbf{G} are any two references for \mathcal{B} , we see that (8.3) and the definitions (8.6) and (8.7) yield the formula

$$\nabla_{\mathbf{K}}\psi = \nabla_{\mathbf{G}}\psi \circ (\mathbf{G}\mathbf{K}^{-1}), \quad (8.8)$$

which generalizes (8.5). By writing (8.8) with $\mathbf{G} = \nabla\gamma$, where $\gamma \in \mathbf{C}$, we infer that $\nabla_{\mathbf{K}}\psi$ is of class C^{r-1} if ψ is of class C^r and \mathbf{K} of class C^{r-1} .

When the range of ψ coincides with the set \mathcal{R} of real numbers, in which case we write f instead of ψ , we can identify $\nabla_{\kappa}f$ with a vector field on \mathcal{B} . Thus, if $f \in \mathcal{F}_{\mathcal{B}}^r$ then $\nabla_{\kappa}f \in \mathcal{V}_{\mathcal{B}}^{r-1}$. The formula (8.5) becomes

$$\nabla_{\kappa}f = (\nabla\gamma(\nabla\kappa)^{-1})^T \nabla_{\gamma}f. \quad (8.9)$$

Let $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{r-1}$ and $f \in \mathcal{F}_{\mathcal{B}}^r$. The function $\mathfrak{h}(f)$ on \mathcal{B} defined by

$$\mathfrak{h}(f) = \nabla_{\kappa}f \cdot (\nabla\kappa)\mathfrak{h}, \quad (8.10)$$

where the inner product is defined pointwise, does not depend on the choice of the configuration $\kappa \in \mathbf{C}$, as is easily seen with the help of (8.9). Moreover, $\mathfrak{h}(f)$ is of class C^{r-1} . Therefore, every $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{r-1}$ gives rise to a mapping

$$\mathfrak{h}: \mathcal{F}_{\mathcal{B}}^r \rightarrow \mathcal{F}_{\mathcal{B}}^{r-1}. \quad (8.11)$$

Actually, every $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{r-1}$ can be identified with a mapping of the type (8.11), because it is easily seen from (8.10) that $\mathfrak{h}_1(f) = \mathfrak{h}_2(f)$ cannot hold for all $f \in \mathcal{F}_{\mathcal{B}}^r$ unless $\mathfrak{h}_1 = \mathfrak{h}_2$. The mapping (8.11) defined by (8.10) has the following basic property, which follows immediately from the chain rule.

Proposition 3. If $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{r-1}$, if H is a real-valued function of class C^r of any number m of real variables, and if $f_1, f_2, \dots, f_m \in \mathcal{F}_{\mathcal{B}}^r$, then

$$\mathfrak{h}(H(f_1, f_2, \dots, f_m)) = \sum_{k=1}^m H_{,k}(f_1, f_2, \dots, f_m) \mathfrak{h}(f_k), \quad (8.12)$$

where $H_{,k}$ denotes the derivative of H with respect to its k^{th} variable.

Actually, the property described in Proposition 3 characterizes the tangent vector fields of class C^{r-1} and hence could have been used for their definition, but we shall neither use nor prove this fact.

Applying (8.12) to the cases when $H(\xi_1, \xi_2) = \xi_1 + \xi_2$ and $H(\xi_1, \xi_2) = \xi_1 \xi_2$, we obtain

$$\mathfrak{h}(f+g) = \mathfrak{h}(f) + \mathfrak{h}(g), \quad \mathfrak{h}(fg) = f \mathfrak{h}(g) + g \mathfrak{h}(f). \quad (8.13)$$

Let $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{r-1}$ with $r \geq 2$. Since $\mathcal{T}_{\mathcal{B}}^{r-1} \subset \mathcal{T}_{\mathcal{B}}^{r-2}$ and hence also $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{r-2}$ we can identify \mathfrak{h} and \mathfrak{f} not only with mappings from $\mathcal{F}_{\mathcal{B}}^r$ into $\mathcal{F}_{\mathcal{B}}^{r-1}$, but also with mappings from $\mathcal{F}_{\mathcal{B}}^{r-1}$ into $\mathcal{F}_{\mathcal{B}}^{r-2}$. Therefore, we can form the compositions $\mathfrak{h} \circ \mathfrak{f}$ and $\mathfrak{f} \circ \mathfrak{h}$ as mappings from $\mathcal{F}_{\mathcal{B}}^r$ into $\mathcal{F}_{\mathcal{B}}^{r-2}$. By themselves, these compositions do not correspond to tangent vector fields, but it is remarkable that the difference

$$[\mathfrak{h}, \mathfrak{f}] = \mathfrak{h} \circ \mathfrak{f} - \mathfrak{f} \circ \mathfrak{h}: \mathcal{F}_{\mathcal{B}}^r \rightarrow \mathcal{F}_{\mathcal{B}}^{r-2}, \quad (8.14)$$

called the *bracket* of \mathfrak{h} and \mathfrak{f} , has values that belong to $\mathcal{F}_{\mathcal{B}}^{r-1} (\subset \mathcal{F}_{\mathcal{B}}^{r-2})$ and does correspond to a tangent vector field:

Proposition 4. The bracket of two tangent vector fields $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{r-1} (r \geq 2)$ can be identified with the tangent vector field of class C^{r-2} given by

$$[\mathfrak{h}, \mathfrak{f}] = (\nabla_{\kappa})^{-1} [(\nabla_{\kappa} \mathbf{k}) \mathbf{h} - (\nabla_{\kappa} \mathbf{h}) \mathbf{k}], \quad \mathbf{h} = (\nabla_{\kappa}) \mathfrak{h}, \quad \mathbf{k} = (\nabla_{\kappa}) \mathfrak{f}, \quad (8.15)$$

where κ is an arbitrary configuration.

Proof. We denote the tangent vector field of class C^{r-2} defined by the right-hand side of (8.15) by \mathfrak{b} , so that

$$(\nabla_{\kappa}) \mathfrak{b} = [(\nabla_{\kappa} \mathbf{k}) \mathbf{h} - (\nabla_{\kappa} \mathbf{h}) \mathbf{k}]. \quad (8.16)$$

Now let $f \in \mathcal{F}_{\mathcal{B}}^r$. In view of (8.10), it follows from (8.16) that

$$\mathfrak{b}(f) = \nabla_{\kappa} f \cdot [(\nabla_{\kappa} \mathbf{k}) \mathbf{h} - (\nabla_{\kappa} \mathbf{h}) \mathbf{k}] \quad (8.17)$$

and from (8.15)_{2,3} that

$$(\mathfrak{h} \circ \mathfrak{f})(f) = \mathfrak{h}(\mathfrak{f}(f)) = \nabla_{\kappa}(\nabla_{\kappa} f \cdot \mathbf{k}) \cdot \mathbf{h}. \quad (8.18)$$

The rules of ordinary differential calculus yield $\nabla_{\kappa}(\nabla_{\kappa} f \cdot \mathbf{k}) \cdot \mathbf{h} = \mathbf{h} \cdot (\nabla_{\kappa}^{(2)} f) \mathbf{k} + \nabla_{\kappa} f \cdot (\nabla_{\kappa} \mathbf{k}) \mathbf{h}$. Hence, since $\nabla_{\kappa}^{(2)} f$ is symmetric, if we write (8.18) with \mathfrak{h} and \mathfrak{f} interchanged, take the difference, and then compare with (8.17), we obtain

$$\mathfrak{b}(f) = (\mathfrak{h} \circ \mathfrak{f})(f) - (\mathfrak{f} \circ \mathfrak{h})(f) = [\mathfrak{h}, \mathfrak{f}](f)$$

i. e. the desired result $\mathfrak{b} = [\mathfrak{h}, \mathfrak{f}]$. Q.E.D.

The bracket $[\mathfrak{h}, \mathfrak{f}]$ depends linearly (but not \mathcal{F} -linearly) on \mathfrak{h} and \mathfrak{f} and satisfies for $\mathfrak{h}, \mathfrak{f}, \mathfrak{l} \in \mathcal{T}_{\mathcal{B}}^{r-1}, f \in \mathcal{F}_{\mathcal{B}}^r, 2 \leq r$, the identities

$$[\mathfrak{h}, \mathfrak{f}] = -[\mathfrak{f}, \mathfrak{h}], \tag{8.19}$$

$$[\mathfrak{h}, f \mathfrak{f}] = f[\mathfrak{h}, \mathfrak{f}] + \mathfrak{h}(f) \mathfrak{f}, \tag{8.20}$$

and for $\mathfrak{h}, \mathfrak{f}, \mathfrak{l} \in \mathcal{T}_{\mathcal{B}}^r, 2 \leq r \leq p-1$, the Jacobi-identity

$$\sum_{\text{cyclic}} [\mathfrak{h}, [\mathfrak{f}, \mathfrak{l}]] = 0, \tag{8.21}$$

where the sum is taken of all terms obtained from the one written by cyclic permutation of $\mathfrak{h}, \mathfrak{f}, \mathfrak{l}$. The identity (8.19) is obvious from (8.14), and (8.21) is the result of a trivial calculation. The identity (8.20) follows from (8.14) and (8.13).

It would have been possible to define the bracket $[\mathfrak{h}, \mathfrak{f}] \in \mathcal{T}_{\mathcal{B}}^{r-1}$ for $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^r, 1 \leq r \leq p-1$, directly by (8.15), for it is easy to see that the right-hand side of (8.15) does not depend on the choice of the configuration κ .

9. Affine Connections, Torsion, Curvature

From now on we assume that \mathcal{B} is a continuous body of class $C^p, p \geq 3$.

A mapping

$$\Gamma: \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^{r-1} \tag{9.1}$$

is called an *affine connection* of class $C^{r-1} (1 \leq r \leq p-1)$ on \mathcal{B} if

$$\Gamma(\mathfrak{h} + \mathfrak{f}) = \Gamma \mathfrak{h} + \Gamma \mathfrak{f} \tag{9.2}$$

holds for all $\mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^r$ and

$$\Gamma(f \mathfrak{h}) \mathfrak{f} = f(\Gamma \mathfrak{h}) \mathfrak{f} + \mathfrak{f}(f) \mathfrak{h} \tag{9.3}$$

holds for all $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^r, f \in \mathcal{F}_{\mathcal{B}}^r$ and all $\mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{r-1}$.

If a is a real constant, then $\mathfrak{f}(a) = 0$ by the definition (8.10). Hence (9.3) reduces to $\Gamma(a \mathfrak{h}) = a \Gamma \mathfrak{h}$ when $a \in \mathcal{R}, \mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^r$. Thus, Γ is a linear mapping, but it is never \mathcal{F} -linear. The rule (9.3) resembles one of the product rules for gradient operators.

A triple (e_1, e_2, e_3) of tangent vector fields of class $C^r, r \leq p-1$, is called a *frame* of class C^r if the values $e_i(X)$ form a basis of the tangent space \mathcal{T}_X for each $X \in \mathcal{B}$. Frames of class C^{p-1} (and hence of class $C^r, r \leq p-1$) exist. For example, if (e_1, e_2, e_3) is a basis of \mathcal{V} and κ a configuration, then $e_i = (\nabla \kappa)^{-1} e_i$ defines a frame of class C^{p-1} . Every tangent vector field $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^r$ has a component representation

$$\mathfrak{h} = \sum_i h^i e_i \tag{9.4}$$

with respect to a given frame (e_1, e_2, e_3) of class C^r such that the component functions h^i belong to $\mathcal{F}_{\mathcal{B}}^r$.

Now let Γ be a connection of class C^{r-1} . Substituting (9.4) into $(\Gamma \mathfrak{h}) e_j$ and using the rules (9.2) and (9.3), we obtain

$$(\Gamma \mathfrak{h}) e_j = \sum_i [h^i (\Gamma e_i) e_j + e_j(h^i) e_i]. \tag{9.5}$$

The components Γ_{ij}^k of the three intrinsic tensor fields Γe_i with respect to the frame (e_1, e_2, e_3) are defined by

$$(\Gamma e_i) e_j = \sum_k \Gamma_{ij}^k e_k. \quad (9.6)$$

These components Γ_{ij}^k belong to $\mathcal{F}_{\mathcal{B}}^{r-1}$ and are called the *components of the connection* Γ with respect to the frame (e_1, e_2, e_3) . If we prescribe a frame (e_1, e_2, e_3) of class C^r and 27 functions $\Gamma_{ij}^k \in \mathcal{F}_{\mathcal{B}}^{r-1}$ on \mathcal{B} arbitrarily, then (9.5) and (9.6) determine a unique affine connection of class C^{r-1} .

Let Γ be a connection of class C^{r-1} having components $\Gamma_{ij}^k \in \mathcal{F}_{\mathcal{B}}^{r-1}$ with respect to the frame (e_1, e_2, e_3) of class C^r . When $1 \leq s \leq r$, then (e_1, e_2, e_3) is also of class C^s and $\Gamma_{ij}^k \in \mathcal{F}_{\mathcal{B}}^{s-1} \supset \mathcal{F}_{\mathcal{B}}^{r-1}$. Hence, (9.5) and (9.6) define an affine connection of class C^{s-1} . Therefore, the mapping $\Gamma: \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^{r-1}$ has a unique extension to $\mathcal{T}_{\mathcal{B}}^s$ that is an affine connection of class C^{s-1} . We denote this extension by the same symbol Γ . With this convention, we can say that every affine connection of class C^{r-1} is also of class C^{s-1} when $1 \leq s \leq r \leq p-1$.

Let Γ be a connection of class C^{r-1} , and hence also of class C^{s-1} when $1 \leq s \leq r$. Using the notation

$$\Gamma_{\mathfrak{h}} \mathfrak{f} = (\Gamma \mathfrak{f}) \mathfrak{h}, \quad (9.7)$$

we can identify $\Gamma_{\mathfrak{h}}$ with a mapping

$$\Gamma_{\mathfrak{h}}: \mathcal{T}_{\mathcal{B}}^s \rightarrow \mathcal{T}_{\mathcal{B}}^{s-1} \quad (9.8)$$

for any choice of s , $1 \leq s \leq r$, and any choice of $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{s-1}$. In terms of $\Gamma_{\mathfrak{h}}$ the rule (9.3) reads

$$\Gamma_{\mathfrak{h}}(f \mathfrak{f}) = f(\Gamma_{\mathfrak{h}} \mathfrak{f}) + \mathfrak{h}(f) \mathfrak{f}. \quad (9.9)$$

Moreover, $\Gamma_{\mathfrak{h}}$ depends \mathcal{F} -linearly on \mathfrak{h} .

The *Cartan-torsion* (or simply *torsion*) of the connection Γ is the mapping

$$\mathfrak{S}: \mathcal{T}_{\mathcal{B}}^r \times \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^{r-1} \quad (9.10)$$

defined by

$$\mathfrak{S}(\mathfrak{h}, \mathfrak{f}) = \Gamma_{\mathfrak{h}} \mathfrak{f} - \Gamma_{\mathfrak{f}} \mathfrak{h} - [\mathfrak{h}, \mathfrak{f}]. \quad (9.11)$$

In view of (8.19) it is obvious that \mathfrak{S} is skew in the sense that

$$\mathfrak{S}(\mathfrak{h}, \mathfrak{f}) = -\mathfrak{S}(\mathfrak{f}, \mathfrak{h}). \quad (9.12)$$

It is an almost immediate consequence of (9.9) and the rule (8.20) that \mathfrak{S} is \mathcal{F} -bilinear. Hence, by Proposition 2 (Sect. 7), the torsion \mathfrak{S} can be identified with a field on \mathcal{B} of class C^{r-1} whose value $\mathfrak{S}(X)$ at $X \in \mathcal{B}$ is a bilinear mapping from $\mathcal{T}_X \times \mathcal{T}_X$ into \mathcal{T}_X . It follows that $\mathfrak{S}(\mathfrak{h}, \mathfrak{f})$ remains meaningful for any, even discontinuous, tangent vector fields $\mathfrak{h}, \mathfrak{f}$, and that $\mathfrak{S}(\mathfrak{h}, \mathfrak{f})|_X = \mathfrak{S}(X)(\mathfrak{h}(X), \mathfrak{f}(X))$ depends on \mathfrak{h} and \mathfrak{f} only through their values at X .

Let $\mathfrak{f}, \mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{r-1}$, $2 \leq r \leq p-1$. In view of (9.8) we can regard $\Gamma_{\mathfrak{h}}$ and $\Gamma_{\mathfrak{f}}$ as mappings from $\mathcal{T}_{\mathcal{B}}^r$ into $\mathcal{T}_{\mathcal{B}}^{r-1}$ and also as mappings from $\mathcal{T}_{\mathcal{B}}^{r-1}$ into $\mathcal{T}_{\mathcal{B}}^{r-2}$. Hence we can form the compositions $\Gamma_{\mathfrak{h}} \circ \Gamma_{\mathfrak{f}}$ and $\Gamma_{\mathfrak{f}} \circ \Gamma_{\mathfrak{h}}$ and the bracket

$$[\Gamma_{\mathfrak{h}}, \Gamma_{\mathfrak{f}}] = \Gamma_{\mathfrak{h}} \circ \Gamma_{\mathfrak{f}} - \Gamma_{\mathfrak{f}} \circ \Gamma_{\mathfrak{h}} \quad (9.13)$$

as mappings from $\mathcal{T}_{\mathcal{B}}^r$ into $\mathcal{T}_{\mathcal{B}}^{r-2}$. Since $[\mathfrak{h}, \mathfrak{k}] \in \mathcal{T}_{\mathcal{B}}^{r-1}$, we can regard $\Gamma_{[\mathfrak{h}, \mathfrak{k}]}$ as a mapping from $\mathcal{T}_{\mathcal{B}}^r (\subset \mathcal{T}_{\mathcal{B}}^{r-1})$ into $\mathcal{T}_{\mathcal{B}}^{r-2}$. Hence we can define

$$\mathfrak{R}(\mathfrak{h}, \mathfrak{k}): \mathcal{T}_{\mathcal{B}}^r \rightarrow \mathcal{T}_{\mathcal{B}}^{r-2} \tag{9.14}$$

by

$$\mathfrak{R}(\mathfrak{h}, \mathfrak{k}) = [\Gamma_{\mathfrak{h}}, \Gamma_{\mathfrak{k}}] - \Gamma_{[\mathfrak{h}, \mathfrak{k}]} \tag{9.15}$$

An easy calculation, based on the definition (9.13) and the rules (9.9) and (8.20), shows that the mapping (9.15) is \mathcal{F} -linear. Hence, by an analogue of Proposition 1 (Sect. 7) for intrinsic tensor fields, $\mathfrak{R}(\mathfrak{h}, \mathfrak{k})$ can be identified with an element of $\mathcal{T}_{\mathcal{B}}^{r-2}$ and \mathfrak{R} can be regarded as a mapping

$$\mathfrak{R}: \mathcal{T}_{\mathcal{B}}^{r-1} \times \mathcal{T}_{\mathcal{B}}^{r-1} \rightarrow \mathcal{T}_{\mathcal{B}}^{r-2}, \tag{9.16}$$

which is called the *Riemann-curvature* (or simply *curvature*) of the connection Γ . It is obvious that \mathfrak{R} is skew in the sense that

$$\mathfrak{R}(\mathfrak{h}, \mathfrak{k}) = -\mathfrak{R}(\mathfrak{k}, \mathfrak{h}). \tag{9.17}$$

A short calculation shows that the mapping (9.16) is \mathcal{F} -bilinear. Therefore, by Proposition 2 (Sect. 7), we can identify the curvature \mathfrak{R} with a field on \mathcal{B} of class C^{r-2} whose value $\mathfrak{R}(X)$ at $X \in \mathcal{B}$ is a bilinear transformation from $\mathcal{T}_X \times \mathcal{T}_X$ into \mathcal{T}_X . We have $\mathfrak{R}(\mathfrak{h}, \mathfrak{k})|_X = \mathfrak{R}(X)(\mathfrak{h}(X), \mathfrak{k}(X))$, which shows that $\mathfrak{R}(\mathfrak{h}, \mathfrak{k})$ is meaningful for any tangent vector fields $\mathfrak{h}, \mathfrak{k}$.

There is an important relation between the torsion and the curvature of an affine connection:

Proposition 5. *Let Γ be an affine connection of class C^{r-1} on \mathcal{B} , $2 \leq r \leq p-1$. The torsion \mathfrak{S} and the curvature \mathfrak{R} of Γ satisfy*

$$\sum_{\text{cyclic}} \{ \Gamma_i(\mathfrak{S}(\mathfrak{h}, \mathfrak{k})) + \mathfrak{S}(\mathfrak{k}, [\mathfrak{l}, \mathfrak{h}]) - \mathfrak{R}(\mathfrak{l}, \mathfrak{h}) \mathfrak{k} \} = 0 \tag{9.18}$$

for all $\mathfrak{h}, \mathfrak{k}, \mathfrak{l} \in \mathcal{T}_{\mathcal{B}}^r$. The sum is to be taken over all terms obtained from the one written by cyclic permutation of $\mathfrak{h}, \mathfrak{k}, \mathfrak{l}$.

The identity (9.18) is often called the *First Bianchi Identity*.

Proof. Operating with Γ_i on (9.11) gives

$$\Gamma_i(\mathfrak{S}(\mathfrak{h}, \mathfrak{k})) - (\Gamma_i \circ \Gamma_{\mathfrak{h}}) \mathfrak{k} + (\Gamma_i \circ \Gamma_{\mathfrak{k}}) \mathfrak{h} + \Gamma_i [[\mathfrak{h}, \mathfrak{k}]] = 0.$$

The cyclic sum of the left side of this equation remains unchanged if the third term is changed by one and the fourth by two cyclic permutations of $\mathfrak{h}, \mathfrak{k}, \mathfrak{l}$. Hence we have

$$\begin{aligned} \sum_{\text{cyclic}} \{ \Gamma_i(\mathfrak{S}(\mathfrak{h}, \mathfrak{k})) - (\Gamma_i \circ \Gamma_{\mathfrak{h}}) \mathfrak{k} + (\Gamma_{\mathfrak{h}} \circ \Gamma_i) \mathfrak{k} + \Gamma_i [[\mathfrak{l}, \mathfrak{h}]] \} \\ = \sum_{\text{cyclic}} \{ \Gamma_i(\mathfrak{S}(\mathfrak{h}, \mathfrak{k})) - [\Gamma_i, \Gamma_{\mathfrak{h}}] \mathfrak{k} + \Gamma_i [[\mathfrak{l}, \mathfrak{h}]] \} = 0. \end{aligned}$$

Using the definitions (9.15) and (9.11), we obtain

$$\sum_{\text{cyclic}} \{ \Gamma_i(\mathfrak{S}(\mathfrak{h}, \mathfrak{k})) - \mathfrak{R}(\mathfrak{l}, \mathfrak{h}) \mathfrak{k} + \mathfrak{S}(\mathfrak{k}, [\mathfrak{l}, \mathfrak{h}]) + [[\mathfrak{k}, [\mathfrak{l}, \mathfrak{h}]]] \} = 0.$$

In view of the Jacobi identity (8.21) the last term gives no contribution and (9.18) results. Q. E. D.

Let Γ and $\overset{*}{\Gamma}$ be two connections of class C^{r-1} on \mathcal{B} . Using the notation (9.7) and observing the rule (9.9), we see that, for each s , $1 \leq s \leq r$, and each $\mathfrak{h} \in \mathcal{T}_{\mathcal{B}}^{s-1}$, the difference

$$\mathfrak{D}_{\mathfrak{h}} = \Gamma_{\mathfrak{h}} - \overset{*}{\Gamma}_{\mathfrak{h}} : \mathcal{T}_{\mathcal{B}}^s \rightarrow \mathcal{T}_{\mathcal{B}}^{s-1} \tag{9.19}$$

is actually \mathcal{F} -linear. Hence $\mathfrak{D}_{\mathfrak{h}}$ can be identified with an intrinsic tensor field of class C^{s-1} , i. e. $\mathfrak{D}_{\mathfrak{h}} \in \mathcal{T}_{\mathcal{B}}^{s-1}$. Since $\mathfrak{D}_{\mathfrak{h}}$ depends \mathcal{F} -linearly on \mathfrak{h} , \mathfrak{D} can be regarded as an \mathcal{F} -linear mapping

$$\mathfrak{D} : \mathcal{T}_{\mathcal{B}}^{s-1} \rightarrow \mathcal{T}_{\mathcal{B}}^{s-1}, \quad 1 \leq s \leq r, \tag{9.20}$$

and hence can be identified with a field of class C^{r-1} whose values $\mathfrak{D}(X)$ are linear transformations from \mathcal{T}_X into \mathcal{T}_X . The possibility of identifying $\mathfrak{D}_{\mathfrak{h}}$ and \mathfrak{D} with fields on \mathcal{B} follows from analogues of Proposition 1 (Sect. 7).

Let \mathfrak{S} and $\overset{*}{\mathfrak{S}}$ denote the torsions and \mathfrak{R} and $\overset{*}{\mathfrak{R}}$ the curvatures of Γ and $\overset{*}{\Gamma}$, respectively. If we write the definition (9.11) of the torsion for both Γ and $\overset{*}{\Gamma}$ and take the difference, we obtain

$$\mathfrak{S}(\mathfrak{h}, \mathfrak{f}) - \overset{*}{\mathfrak{S}}(\mathfrak{h}, \mathfrak{f}) = \mathfrak{D}_{\mathfrak{h}} \mathfrak{f} - \mathfrak{D}_{\mathfrak{f}} \mathfrak{h}. \tag{9.21}$$

If we write the definition (9.15) of the curvature for $\overset{*}{\Gamma}$ and substitute $\overset{*}{\Gamma}_{\mathfrak{h}} = \Gamma_{\mathfrak{h}} - \mathfrak{D}_{\mathfrak{h}}$, we find

$$\overset{*}{\mathfrak{R}}(\mathfrak{h}, \mathfrak{f}) = \mathfrak{R}(\mathfrak{h}, \mathfrak{f}) - [\Gamma_{\mathfrak{h}}, \mathfrak{D}_{\mathfrak{f}}] - [\mathfrak{D}_{\mathfrak{h}}, \Gamma_{\mathfrak{f}}] + [\mathfrak{D}_{\mathfrak{h}}, \mathfrak{D}_{\mathfrak{f}}] + \mathfrak{D}_{[\mathfrak{h}, \mathfrak{f}]}. \tag{9.22}$$

10. Material Connections

Let Φ be a material uniformity for the simple body \mathcal{B} of class C^p . (See Definition 3, Sect. 6.) We say that a tangent vector field \mathfrak{c} is *materially constant* if

$$\mathfrak{c}(X) = \Phi(X, Y) \mathfrak{c}(Y) \tag{10.1}$$

holds for all $X, Y \in \mathcal{B}$. If $X_0 \in \mathcal{B}$ is fixed and $\mathfrak{u}_{X_0} \in \mathcal{T}_{X_0}$ is prescribed arbitrarily, then

$$\mathfrak{c}(X) = \Phi(X, X_0) \mathfrak{u}_{X_0} \tag{10.2}$$

is easily seen to define a materially constant field \mathfrak{c} such that $\mathfrak{c}(X_0) = \mathfrak{u}_{X_0}$. Moreover, every materially constant field \mathfrak{c} can be obtained in this fashion. Thus (10.2) describes a one-to-one correspondence between \mathcal{T}_{X_0} and the set \mathcal{T}_{Φ} of all materially constant vector fields. This correspondence is actually a vector-space isomorphism, showing that \mathcal{T}_{Φ} is a three-dimensional vector-space when addition and multiplication with scalars in \mathcal{T}_{Φ} are defined pointwise.

Let \mathbf{K} be a uniform reference (see Definition 4, Sect. 6) and $\mathfrak{c} \in \mathcal{T}_{\Phi}$. Then it follows from (10.1) and (6.4) that $\mathbf{K}(X) \mathfrak{c}(X) = \mathbf{K}(Y) \mathfrak{c}(Y)$ for all $X, Y \in \mathcal{B}$, i. e. that

$$\mathbf{K} \mathfrak{c} = \mathfrak{c} = \text{constant}. \tag{10.3}$$

Conversely, if $c \in \mathcal{V}$, then $c = K^{-1}c$ is easily seen to be materially constant. Thus \mathcal{T}_Φ is exactly the set of all tangent vector fields c with the property (10.3).

A material uniformity Φ is said to be of class C^r , $r \leq p-1$, if $\mathcal{T}_\Phi \subset \mathcal{T}_\mathcal{B}^r$, i.e., if all tangent vector fields materially constant with respect to Φ are of class C^r . For the remainder of this paper we lay down the following:

Smoothness assumption⁹: \mathcal{B} is a materially uniform continuous body of class C^p , $p \geq 3$ ¹⁰, which admits a material uniformity Φ of class C^{p-1} .

Let Φ be a material uniformity of class C^{p-1} and let K be a uniform reference such that (6.4) holds. Then for $c \in \mathcal{V}$, $\kappa \in C$, and $c = K^{-1}c \in \mathcal{T}_\Phi$, the vector field $((\nabla \kappa)K^{-1})c = (\nabla \kappa)c$ is of class C^{p-1} because c is of class C^{p-1} . This is possible for all $c \in \mathcal{V}$ only if $(\nabla \kappa)K$ and hence K is of class C^{p-1} . Thus, if K is a uniform reference such that

$$\Phi(X, Y) = K(X)^{-1}K(Y), \quad X, Y \in \mathcal{B}, \tag{10.4}$$

then K is of class C^{p-1} .

Theorem 3. Given a material uniformity Φ of class C^{p-1} , there is a unique affine connection Γ such that $\Gamma c = 0$ holds for all materially constant tangent vector fields $c \in \mathcal{T}_\Phi$. In terms of any uniform reference K satisfying (10.4), Γ is given by

$$\Gamma h = K^{-1} \nabla_K(K h) K, \quad h \in \mathcal{T}_\mathcal{B}^1. \tag{10.5}$$

Also, Γ is of class C^{p-2} .

Proof. To prove the uniqueness, assume that Γ and $\bar{\Gamma}$ are connections such that $\Gamma c = \bar{\Gamma} c$ for all $c \in \mathcal{T}_\Phi$. Putting $\mathcal{D}_h = \Gamma_h - \bar{\Gamma}_h$, we then have $\mathcal{D}_h c = 0$ for all $c \in \mathcal{T}_\Phi$. We have seen at the end of the previous section that \mathcal{D}_h can be identified with an intrinsic tensor field in $\mathcal{S}_\mathcal{B}^0$ when $h \in \mathcal{T}_\mathcal{B}^0$. Hence $\mathcal{D}_h(X_0)c(X_0) = 0$ for all $X_0 \in \mathcal{B}$ and all $c \in \mathcal{T}_\Phi$. Since for any prescribed $u_{X_0} \in \mathcal{T}_{X_0}$ the $c \in \mathcal{T}_\Phi$ given by (10.2) has the property $c(X_0) = u_{X_0}$, it follows that $\mathcal{D}_h(X_0)u_{X_0} = 0$ for all $u_{X_0} \in \mathcal{T}_{X_0}$, i.e., that $\mathcal{D}_h(X_0) = 0$. Since $X_0 \in \mathcal{B}$ is arbitrary, we infer that $\mathcal{D}_h = 0$, i.e. that $\Gamma = \bar{\Gamma}$.

To prove the existence of Γ we choose a uniform reference K with the property (10.4), define Γ by (10.5), and show that it has all the necessary properties. It is clear that $\Gamma c = 0$ when $c \in \mathcal{T}_\Phi$ because, by (10.3), $\nabla_K(Kc) = 0$ when $c \in \mathcal{T}_\Phi$. Since K is of class C^{p-1} it follows that Γh is of class C^{p-2} when $h \in \mathcal{T}_\mathcal{B}^{p-1}$. The validity of the rules (9.2) and (9.3) follows from the validity of the analogous rules for the relative gradient ∇_K . Hence Γ is indeed an affine connection of class C^{p-2} . Q.E.D.

Definition 5. The affine connection (of class C^{p-2}) with the property $\Gamma c = 0$ for all $c \in \mathcal{T}_\Phi$ is called the **material connection** for the material uniformity Φ (of class C^{p-1}).

Theorem 4. Material connections have zero Riemann-curvature.

Proof. Let $X \in \mathcal{B}$ and $u_X \in \mathcal{T}_X$ be given. We can determine $c \in \mathcal{T}_\Phi \subset \mathcal{T}_\mathcal{B}^{p-1}$ such that $c(X) = u_X$. If Γ is the material connection for Φ we have $\Gamma_h c = 0$ for all $h \in \mathcal{T}_\mathcal{B}^0$. Hence the definition (9.15) shows that $\mathfrak{R}(h, \mathfrak{f})c = 0$ for all $h, \mathfrak{f} \in \mathcal{T}_\mathcal{B}^{p-1}$. Since $\mathfrak{R}(h, \mathfrak{f})$

⁹ C. C. WANG [8] has recently shown that the theory given here can be extended to the case when each point has a neighborhood that admits a smooth material uniformity. This can happen even when all material uniformities for the whole body are discontinuous.

¹⁰ For all considerations not referring to curvature, $p \geq 2$ is actually sufficient.

can be identified with an intrinsic tensor field it follows that

$$\mathfrak{R}(\mathfrak{h}, \mathfrak{k}) \mathfrak{c}|_X = \mathfrak{R}(\mathfrak{h}, \mathfrak{k})|_X \mathfrak{c}(X) = \mathfrak{R}(\mathfrak{h}, \mathfrak{k})|_X \mathfrak{u}_X = 0.$$

This can be valid for all $X \in \mathcal{B}$ and all $\mathfrak{u}_X \in \mathcal{T}_X$ only if $\mathfrak{R}(\mathfrak{h}, \mathfrak{k}) = 0$. Hence, since $\mathfrak{h}, \mathfrak{k} \in \mathcal{T}_{\mathcal{B}}^{p-1}$ are arbitrary, we must have $\mathfrak{R} = 0$. Q.E.D.

11. Inhomogeneity

Let Φ be a material uniformity of class C^{p-1} , let Γ be the associated material connection (of class C^{p-2}) with torsion \mathfrak{S} , and let \mathbf{K} be a uniform reference (of class C^{p-1}) such that (10.4) holds. We can define a field \mathbf{S} of class C^{p-2} with values $\mathbf{S}(X): \mathcal{V} \rightarrow \mathcal{L}$ by the condition

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = \mathbf{K} \mathfrak{S}(\mathbf{K}^{-1} \mathbf{u}, \mathbf{K}^{-1} \mathbf{v}) \quad (11.1)$$

for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. In view of the linearity of the values $\mathbf{S}(X)$, $\mathfrak{S}(X)$, and $\mathbf{K}(X)$, (11.1) continues to hold if the fixed vectors \mathbf{u} and \mathbf{v} in (11.1) are replaced by vector fields \mathbf{h} and \mathbf{k} . The following theorem shows how \mathbf{S} and hence \mathfrak{S} can be expressed directly in terms of \mathbf{K} :

Theorem 5. *Let γ be an arbitrary configuration of \mathcal{B} and*

$$\mathbf{F} = (\nabla \gamma) \mathbf{K}^{-1} \in \mathcal{L}_{\mathcal{B}}^{p-1}. \quad (11.2)$$

Then \mathbf{S} is given by

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = \mathbf{F}^{-1} [((\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{v}) \mathbf{u} - ((\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{u}) \mathbf{v}], \quad (11.3)$$

$$\mathbf{S} \mathbf{u} = \mathbf{F}^{-1} [\nabla_{\mathbf{K}}(\mathbf{F} \mathbf{u}) - (\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{u}], \quad (11.4)$$

or

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = ((\nabla_{\gamma} \mathbf{F}^{-1}) \mathbf{h}) \mathbf{k} - ((\nabla_{\gamma} \mathbf{F}^{-1}) \mathbf{k}) \mathbf{h}, \quad (11.5)$$

where $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ and $\mathbf{h} = \mathbf{F} \mathbf{u}$, $\mathbf{k} = \mathbf{F} \mathbf{v}$.

Proof. The tangent vector fields $\mathbf{K}^{-1} \mathbf{u}$ and $\mathbf{K}^{-1} \mathbf{v}$ are materially constant and hence are annihilated by Γ . Hence, the definitions (11.1) and (9.11) of \mathbf{S} and \mathfrak{S} yield

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = -\mathbf{K} [\mathbf{K}^{-1} \mathbf{u}, \mathbf{K}^{-1} \mathbf{v}] = -\mathbf{K} [(\nabla \gamma)^{-1} \mathbf{h}, (\nabla \gamma)^{-1} \mathbf{k}]. \quad (11.6)$$

Using Proposition 4, (8.15), we find

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = \mathbf{F}^{-1} [(\nabla_{\gamma} \mathbf{h}) \mathbf{k} - (\nabla_{\gamma} \mathbf{k}) \mathbf{h}]. \quad (11.7)$$

The formula (8.8), with the choices $\psi = \mathbf{h}$ and $\mathbf{G} = \nabla \gamma$, gives

$$(\nabla_{\gamma} \mathbf{h}) \mathbf{k} = (\nabla_{\gamma} \mathbf{h}) \mathbf{F} \mathbf{v} = (\nabla_{\mathbf{K}} \mathbf{h}) \mathbf{v} = \nabla_{\mathbf{K}}(\mathbf{F} \mathbf{u}) \mathbf{v} = ((\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{v}) \mathbf{u}. \quad (11.8)$$

Substituting (11.8) and the formula obtained from (11.8) by interchanging \mathbf{h} and \mathbf{k} into (11.7), we obtain (11.3) and (11.4).

To prove (11.5) we note that $\nabla_{\gamma} \mathbf{u} = \mathbf{0}$ for constant $\mathbf{u} \in \mathcal{V}$. Using one of the product rules for gradient operators we find

$$\mathbf{0} = (\nabla_{\gamma} \mathbf{u}) \mathbf{k} = \nabla_{\gamma}(\mathbf{F}^{-1} \mathbf{h}) \mathbf{k} = ((\nabla_{\gamma} \mathbf{F}^{-1}) \mathbf{k}) \mathbf{h} + \mathbf{F}^{-1}(\nabla_{\gamma} \mathbf{h}) \mathbf{k}. \quad (11.9)$$

Of course, (11.9) remains valid if we interchange \mathbf{h} and \mathbf{k} . The formula (11.5) follows from (11.7) and (11.9). Q.E.D.

Recall that the body \mathcal{B} is homogeneous if it admits a configuration gradient $\mathbf{K} = \nabla \boldsymbol{\kappa}$ as a uniform reference. We can then choose $\boldsymbol{\gamma} = \boldsymbol{\kappa}$ in (11.2), obtaining $\mathbf{F} = \mathbf{1}$, which is constant and hence has gradient zero. Thus, Theorem 5 shows that $\mathbf{S} = \mathbf{0}$ and hence $\mathfrak{S} = 0$ for suitable uniform references if the body is homogeneous. The converse of this result is not true, but it becomes true if “homogeneous” is replaced by “locally homogeneous” in a sense we shall now make precise.

Let \mathcal{N} be an open subset of a simple body \mathcal{B} of class C^p . We can give \mathcal{N} the structure of a continuous body of class C^p by letting $\boldsymbol{\gamma} : \mathcal{N} \rightarrow \mathcal{E}$ be a configuration of \mathcal{N} if $\boldsymbol{\kappa} \circ \boldsymbol{\gamma}^{-1} : \boldsymbol{\gamma}(\mathcal{N}) \rightarrow \mathcal{E}$ is of class C^p for all configurations $\boldsymbol{\kappa} \in \mathbf{C}$ of \mathcal{B} . We denote the set of all configurations of \mathcal{N} by $\mathbf{C}_{\mathcal{N}}$. If $\boldsymbol{\kappa} \in \mathbf{C}$, then the restriction of $\boldsymbol{\kappa}$ to \mathcal{N} belongs to $\mathbf{C}_{\mathcal{N}}$. However, not all configurations $\boldsymbol{\gamma} \in \mathbf{C}_{\mathcal{N}}$ of \mathcal{N} can be obtained in this manner. Still, given any $X \in \mathcal{N}$ and any configuration $\boldsymbol{\gamma}$ of \mathcal{N} , one can easily construct a configuration $\boldsymbol{\kappa}$ of \mathcal{B} such that (3.1) holds. Therefore, an equivalence class \mathbf{K}_X which defines a local configuration at X relative to \mathcal{N} can be made to correspond to the non-empty set $\{\boldsymbol{\kappa} \in \mathbf{C} \mid \nabla(\boldsymbol{\kappa} \circ \boldsymbol{\gamma}^{-1})|_{\boldsymbol{\gamma}(X)} = \mathbf{1} \text{ for all } \boldsymbol{\gamma} \in \mathbf{K}_X\}$, which is a local configuration at X relative to \mathcal{B} . This correspondence is one-to-one and can be used to identify local configurations at X relative to \mathcal{N} with local configurations at X relative to \mathcal{B} . Using this identification, we can endow \mathcal{N} with the structure of a simple body by using the restriction to \mathcal{N} of the function \mathfrak{G} which defines the simple body structure on \mathcal{B} according to Definition 1 (Sect. 5). Thus, every open subset \mathcal{N} of \mathcal{B} has a natural structure of a simple body of class C^p , *i.e.*, every open subset \mathcal{N} of \mathcal{B} can be regarded as a simple body of class C^p . Such a subset is called a *neighborhood* of a material point if it contains that point.

A simple body \mathcal{B} is called *locally homogeneous* if every $X \in \mathcal{B}$ has a neighborhood \mathcal{N} that is homogeneous. A body can be locally homogeneous without being homogeneous, even if it is simply connected.

Definition 6. *The Cartan torsion \mathfrak{S} of the material connection Γ associated with a material uniformity Φ of class C^{p-1} is called the **inhomogeneity** of Φ .*

*The field S defined by (11.1) is called the **inhomogeneity**¹¹ relative to the reference \mathbf{K} .*

This definition finds its motivation in the result already mentioned:

Theorem 6¹². *If \mathcal{B} is homogeneous, then it admits a material uniformity of class C^{p-1} with zero inhomogeneity. If \mathcal{B} admits a material uniformity of class C^{p-1} with zero inhomogeneity, then it is locally homogeneous.*

Proof. Only the second part of the theorem remains to be proved. Assume, therefore, that Φ is a material uniformity of class C^{p-1} with zero inhomogeneity. Using the same notation as before, we then have $\mathfrak{S} = 0$ and hence $\mathbf{S} = \mathbf{0}$. Theorem 5, (11.5), shows that if $\boldsymbol{\gamma} \in \mathbf{C}$ is arbitrary and \mathbf{F} defined by (11.2), $\nabla_{\boldsymbol{\gamma}} \mathbf{F}^{-1}$ has the symmetry property

$$((\nabla_{\boldsymbol{\gamma}} \mathbf{F}^{-1}) \mathbf{u}) \mathbf{v} = ((\nabla_{\boldsymbol{\gamma}} \mathbf{F}^{-1}) \mathbf{v}) \mathbf{u} \tag{11.10}$$

¹¹ It corresponds to what is called “dislocation density” in the theory of continuous distributions of dislocations (*cf.* [4]).

¹² The theorem stated in the middle of p. 90 in reference [2] is incorrect and should be replaced by Theorem 6.

for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Let $X \in \mathcal{B}$ be given and let \mathcal{N}' be a simply connected neighborhood of X . By a classical theorem of analysis, the symmetry (11.10) implies the existence of a mapping $\lambda: \gamma(\mathcal{N}') \rightarrow \mathcal{E}$ such that $\mathbf{F}^{-1} = (\nabla \lambda) \circ \gamma$ holds in \mathcal{N}' . Moreover, $\mathbf{K}, \nabla \gamma$, and hence \mathbf{F} and \mathbf{F}^{-1} being of class C^{p-1} , λ is of class C^p . Since $\nabla \lambda = \mathbf{F}^{-1} \circ \gamma^{-1}$ is invertible, it follows by the inverse function theorem that λ is locally (but not necessarily globally) invertible, *i.e.* that X has a neighborhood $\mathcal{N} \subset \mathcal{N}'$ on which λ is invertible. The mapping $\kappa = \lambda \circ \gamma$, when restricted to \mathcal{N} , is therefore a configuration of \mathcal{N} with gradient $\nabla \kappa = (\nabla \lambda \circ \gamma) \nabla \gamma = \mathbf{F}^{-1} \nabla \gamma = \mathbf{K}$ on \mathcal{N} . Hence the uniform reference \mathbf{K} on \mathcal{N} is the gradient of the configuration κ of \mathcal{N} , *i.e.*, \mathcal{N} is homogeneous. Q. E. D.

The first Bianchi identity gives rise to the following identity for the relative inhomogeneity \mathbf{S} and its gradient $\nabla_{\mathbf{K}} \mathbf{S}$ relative to \mathbf{K} :

$$\sum [(((\nabla_{\mathbf{K}} \mathbf{S}) \mathbf{u}) \mathbf{v}) \mathbf{w}) - (\mathbf{S} \mathbf{u})(\mathbf{S} \mathbf{v}) \mathbf{w}] = 0. \quad (11.11)$$

To prove (11.11), substitute $\mathbf{l} = \mathbf{K}^{-1} \mathbf{u}$, $\mathbf{f} = \mathbf{K}^{-1} \mathbf{w}$, $\mathbf{h} = \mathbf{K}^{-1} \mathbf{v}$ into (9.18), observe that $\mathfrak{R} = 0$ (Theorem 4, Sect. 10), and make use of (11.6)₁.

12. Relative Riemannian Structures, Contortion

Let \mathbf{K} be a uniform reference of class C^{p-1} . If we choose $\mathbf{K}_X = \mathbf{K}(X)$ in (4.5), then this equation defines an inner product $*$ on each of the tangent spaces \mathcal{T}_X , $X \in \mathcal{B}$. The structure on \mathcal{B} defined by these inner products will be called the *Riemannian structure of \mathcal{B} relative to the uniform reference \mathbf{K}* . If \mathbf{h} and \mathbf{f} are tangent vector fields, we define $\mathbf{h} * \mathbf{f}$ pointwise. For such fields, (4.5) then yields

$$\mathbf{h} * \mathbf{f} = (\mathbf{K} \mathbf{h}) \cdot (\mathbf{K} \mathbf{f}). \quad (12.1)$$

It is clear that $\mathbf{h} * \mathbf{f} \in \mathcal{T}_{\mathcal{B}}^r$ if $\mathbf{h}, \mathbf{f} \in \mathcal{T}_{\mathcal{B}}^r$ for $0 \leq r \leq p-1$. This fact is expressed by saying that the Riemannian structure relative to \mathbf{K} is of class C^{p-1} .

Although the following proposition is one of the basic facts of Riemannian geometry, we shall give an independent proof:

Proposition 6. *There is a unique affine connection Γ^* of class C^{p-2} with the following properties:*

- (a) *The torsion \mathfrak{S}^* of Γ^* vanishes.*
- (b) *For any $\mathbf{h}, \mathbf{f}, \mathbf{l} \in \mathcal{T}_{\mathcal{B}}^{r-1}$ the relation*

$$\mathbf{h}(\mathbf{f} * \mathbf{l}) = \mathbf{f} * \Gamma_{\mathbf{h}}^* \mathbf{l} + \mathbf{l} * \Gamma_{\mathbf{h}}^* \mathbf{f} \quad (12.2)$$

is valid.

Proof. First we assume the existence of Γ^* . Let Γ be the material connection associated with \mathbf{K} and consider the difference

$$\mathfrak{D}_{\mathbf{h}} = \Gamma_{\mathbf{h}} - \Gamma_{\mathbf{h}}^*, \quad \mathbf{h} \in \mathcal{T}_{\mathcal{B}}^{p-1}. \quad (12.3)$$

According to the results given at the end of Sect. 9, \mathfrak{D} can be identified with a field of class C^{p-2} on \mathcal{B} whose values are linear transformations from \mathcal{T}_X into \mathcal{T}_X .

Therefore, we can define a field D on \mathcal{B} of class C^{p-2} with values $D(X): \mathcal{V} \rightarrow \mathcal{L}$ by the condition

$$D u = K \mathfrak{D}_{K^{-1} u} K^{-1} \tag{12.4}$$

for all $u \in \mathcal{V}$. Since $\overset{*}{\mathfrak{S}} = 0$, the relation (9.21) and the definitions (12.4) and (11.1) give

$$(S u) v = (D u) v - (D v) u, \quad u, v \in \mathcal{V}. \tag{12.5}$$

By the Definition 5 (Sect. 10), we have $\Gamma_b c = 0$ whenever c is materially constant. Hence, since $c = K^{-1} c$ is materially constant when $c \in \mathcal{V}$, we infer from (12.3) and (12.4) that

$$K \overset{*}{\Gamma}_b (K^{-1} c) = - (D K) c \tag{12.6}$$

when $c \in \mathcal{V}$. Now let $u, v, w \in \mathcal{V}$. If we substitute $h = K^{-1} u$, $\mathfrak{f} = K^{-1} v$, and $l = K^{-1} w$ into (12.2) and observe (12.6) and (12.1) we obtain

$$h(v \cdot w) = -v \cdot (D u) w - w \cdot (D u) v. \tag{12.7}$$

Since $v \cdot w$ is constant, we have $h(v \cdot w) = 0$ (see the definition (8.10)). Therefore (12.7) states that $D u \in \mathcal{L}$ is skew for all $u \in \mathcal{V}$:

$$D u = - (D u)^T, \quad u \in \mathcal{V}. \tag{12.8}$$

The equations (12.5) and (12.8) enable us to express D in terms of S . Indeed, if we take the inner product of (12.5) with $w \in \mathcal{V}$, subtract from the resulting equation the two equations obtained from it by cyclic permutations of u, v, w , and observe (12.8), we find

$$2 u \cdot (D w) v = w \cdot (S u) v - u \cdot (S v) w - v \cdot (S w) u. \tag{12.9}$$

Since $(S u) v = - (S v) u$, (12.9) is equivalent to

$$(D u) v = \frac{1}{2} \{ [(S u) - (S u)^T] v - (S v)^T u \}. \tag{12.10}$$

Now, since S is determined by the uniform reference K , it follows from (12.10), (12.4), and (12.3) that $\overset{*}{\Gamma}$ is uniquely determined by K .

To prove the existence of a connection $\overset{*}{\Gamma}$ with the properties (a) and (b), one can define $\overset{*}{\Gamma}$ by (12.3), (12.4), and (12.10) and verify that it has all the required properties. Q.E.D.

Definition 7. *The connection $\overset{*}{\Gamma}$ determined by the conditions (a) and (b) of Proposition 6 is called the **Riemannian connection** relative to the uniform reference K . The field D determined by (12.3) and (12.4) or (12.5) and (12.8) is called the **contortion**¹³ of K .*

The term ‘‘contortion’’ will be motivated in Section 13.

¹³ It corresponds to what is called ‘‘Cosserat structure curvature’’ or ‘‘Nye curvature’’ in the theory of continuous distributions of dislocations (cf. [4]).

In general, the curvature $\overset{*}{\mathfrak{R}}$ of the Riemannian connection is not zero. We can define a field $\overset{*}{\mathbf{R}}$ with values $\overset{*}{\mathbf{R}}(X): \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{L}$ by the condition

$$\overset{*}{\mathbf{R}}(\mathbf{u}, \mathbf{v}) = \mathbf{K} \overset{*}{\mathfrak{R}}(\mathbf{K}^{-1} \mathbf{u}, \mathbf{K}^{-1} \mathbf{v}) \mathbf{K}^{-1} \quad (12.11)$$

for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$ and put $\mathfrak{h} = \mathbf{K}^{-1} \mathbf{u}$, $\mathfrak{f} = \mathbf{K}^{-1} \mathbf{v}$, $\mathfrak{l} = \mathbf{K}^{-1} \mathbf{w}$. The fields $\mathfrak{h}, \mathfrak{f}, \mathfrak{l}$ are then materially constant and hence are annihilated by Γ . Recalling that the curvature \mathfrak{R} of Γ vanishes (Theorem 4, Sect.10), we then infer from (9.22), (10.5), (11.6)₁, (12.4), and (12.11) that

$$\begin{aligned} \overset{*}{\mathbf{R}}(\mathbf{u}, \mathbf{v}) \mathbf{w} = & -(\nabla_{\mathbf{K}}(\mathbf{D} \mathbf{v}) \mathbf{u}) \mathbf{w} + (\nabla_{\mathbf{K}}(\mathbf{D} \mathbf{u}) \mathbf{v}) \mathbf{w} \\ & + (\mathbf{D} \mathbf{u})(\mathbf{D} \mathbf{v}) \mathbf{w} - (\mathbf{D} \mathbf{v})(\mathbf{D} \mathbf{u}) \mathbf{w} - \mathbf{D}((\mathbf{S} \mathbf{u}) \mathbf{v}) \mathbf{w} \end{aligned}$$

and hence

$$\overset{*}{\mathbf{R}}(\mathbf{u}, \mathbf{v}) = ((\nabla_{\mathbf{K}} \mathbf{D}) \mathbf{v}) \mathbf{u} - ((\nabla_{\mathbf{K}} \mathbf{D}) \mathbf{u}) \mathbf{v} + (\mathbf{D} \mathbf{u})(\mathbf{D} \mathbf{v}) - (\mathbf{D} \mathbf{v})(\mathbf{D} \mathbf{u}) - \mathbf{D}((\mathbf{S} \mathbf{u}) \mathbf{v}). \quad (12.12)$$

In view of (12.5) and (12.10), equation (12.12) shows that $\overset{*}{\mathbf{R}}$ can be expressed in terms of the contortion \mathbf{D} and its gradient relative to \mathbf{K} or in terms of the inhomogeneity \mathbf{S} and its gradient relative to \mathbf{K} .

13. Contorted Aeolotropy

Definition 8. A uniform reference \mathbf{K} of class C^{p-1} is called a *state of contorted aeolotropy* if there exists a configuration κ such that the tensor field

$$\mathbf{Q} = (\nabla \kappa) \mathbf{K}^{-1} \in \mathcal{L}_{\mathcal{B}}^{p-1} \quad (13.1)$$

has orthogonal values $(\mathbf{Q}(X)) \in \mathcal{o}$ for all $X \in \mathcal{B}$.

Assume that \mathbf{K} is such a state of contorted aeolotropy. Since the inner product in \mathcal{V} is preserved under orthogonal transformations, the Riemannian structure (12.1) relative to \mathbf{K} satisfies.

$$\mathfrak{h} * \mathfrak{f} = \mathbf{Q}(\mathbf{K} \mathfrak{h}) \cdot \mathbf{Q}(\mathbf{K} \mathfrak{f}) = (\nabla \kappa) \mathfrak{h} \cdot (\nabla \kappa) \mathfrak{f} \quad (13.2)$$

for all tangent vector fields \mathfrak{h} and \mathfrak{f} . It follows from (13.2) that the Riemannian connection $\overset{*}{\Gamma}$ relative to \mathbf{K} is obtained by transporting the gradient operator ∇ from $\kappa(\mathcal{B})$ into \mathcal{B} via $(\nabla \kappa)^{-1}$, so that

$$\overset{*}{\Gamma} \mathfrak{f} = (\nabla \kappa)^{-1} \nabla_{\kappa}((\nabla \kappa) \mathfrak{f}) (\nabla \kappa), \quad \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{p-1}. \quad (13.3)$$

Indeed, if $\overset{*}{\Gamma}$ is defined by (13.3), condition (a) of Proposition 6 follows from the symmetry of the second gradient and condition (b) from the rule for the differentiation of inner products. By virtue of (13.1), (13.3) is equivalent to

$$\mathbf{K} \overset{*}{\Gamma}_{\mathfrak{h}} \mathfrak{f} = \mathbf{Q}^T \nabla_{\kappa}(\mathbf{Q} \mathbf{K} \mathfrak{f}) \mathbf{Q} \mathbf{K} \mathfrak{h}, \quad \mathfrak{h}, \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^{p-1}. \quad (13.4)$$

Now let $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. If we substitute $\mathfrak{h} = \mathbf{K}^{-1} \mathbf{u}$, $\mathfrak{f} = \mathbf{K}^{-1} \mathbf{v}$ into (13.4) and observe (12.6), we obtain

$$\mathbf{Q}^T \nabla_{\kappa}(\mathbf{Q} \mathbf{v}) \mathbf{Q} \mathbf{u} = -(\mathbf{D} \mathbf{u}) \mathbf{v},$$

which, by (13.1) and (8.8), is equivalent to

$$D\mathbf{u} = -Q^T(\nabla_{\mathbf{K}}Q)\mathbf{u}, \quad \mathbf{u} \in \mathcal{V}. \quad (13.5)$$

This equation shows that the skew transformation $-D\mathbf{u}|_{X \in \mathcal{L}}$ is the instantaneous rate of change of Q at X in the direction of \mathbf{u} , if viewed in any configuration belonging to $\mathbf{K}(X)$. In other words, D describes the local behavior of the rotation field Q , which changes the given state of contorted aeolotropy \mathbf{K} into the gradient of a global configuration κ . It is this property that the term ‘‘contortion’’ for D is meant to express.

Theorem 7. *If \mathbf{K} is a state of contorted aeolotropy, then the curvature of the Riemannian connection relative to \mathbf{K} vanishes. Conversely, if the curvature of the Riemannian connection relative to \mathbf{K} vanishes, then \mathbf{K} is locally a state of contorted aeolotropy (i.e., every point in \mathcal{B} has a neighborhood \mathcal{N} such that the restriction of \mathbf{K} to \mathcal{N} is a state of contorted aeolotropy for \mathcal{N}).*

Proof. Assume first that (13.1) holds. It follows from (13.3) that $\overset{*}{\Gamma}\mathfrak{f} = 0$ if and only if $(\nabla_{\mathbf{K}})\mathfrak{f}$ is constant. Hence we could give a simple direct proof of $\overset{*}{\mathfrak{R}} = 0$ by using the same argument as we used in the proof of Theorem 4 (Sect. 10). Another proof can be obtained on the basis of (13.5) as follows:

If γ is an arbitrary configuration and $F = (\nabla_{\gamma})\mathbf{K}^{-1}$, then (13.5) is equivalent to

$$(\nabla_{\gamma}Q)\mathbf{h} = -QD(F^{-1}\mathbf{h}), \quad \mathbf{h} \in \mathcal{V}. \quad (13.6)$$

If we take the gradient ∇_{γ} of (13.6) in the direction of $\mathbf{k} \in \mathcal{V}$, we find

$$((\nabla_{\gamma}^{(2)}Q)\mathbf{k})\mathbf{h} = -Q\{((\nabla_{\mathbf{K}}D)\mathbf{v})\mathbf{u} - (D\mathbf{v})(D\mathbf{u}) + D(((\nabla_{\gamma}F^{-1})\mathbf{k})\mathbf{h})\}, \quad (13.7)$$

where $\mathbf{u} = F^{-1}\mathbf{h}$, $\mathbf{v} = F^{-1}\mathbf{k}$. Of course, because of the linearity of the values of the fields D , $\nabla_{\mathbf{K}}D$, $\nabla_{\gamma}^{(2)}Q$, etc., (13.6) and (13.7) remain valid if \mathbf{h} and \mathbf{k} are not fixed vectors but vector fields. In particular, they remain valid when \mathbf{u} and \mathbf{v} are fixed. If we interchange \mathbf{u} and \mathbf{v} and hence \mathbf{h} and \mathbf{k} in (13.7) and subtract the resulting formula from (13.7), we obtain, after observing (11.5) and (12.12),

$$((\nabla_{\gamma}^{(2)}Q)\mathbf{k})\mathbf{h} - ((\nabla_{\gamma}^{(2)}Q)\mathbf{h})\mathbf{k} = -Q\overset{*}{\mathbf{R}}(\mathbf{u}, \mathbf{v}). \quad (13.8)$$

Thus, $\overset{*}{\mathbf{R}} = \mathbf{0}$ and hence $\overset{*}{\mathfrak{R}} = 0$ follows also from the symmetry of the second gradient $\nabla_{\gamma}^{(2)}Q$.

Assume now that \mathbf{K} is a uniform reference such that $\overset{*}{\mathbf{R}} = \mathbf{0}$. Let γ be an arbitrary configuration and put $F = (\nabla_{\gamma})\mathbf{K}^{-1}$, as before. We can then regard (13.6) as a differential equation for the determination of Q . As we have seen, $\overset{*}{\mathbf{R}} = \mathbf{0}$ is an integrability condition necessary for the existence of a solution. According to a classical theorem, $\overset{*}{\mathbf{R}} = \mathbf{0}$ is also sufficient for the existence of a solution that is valid in a simply connected neighborhood \mathcal{N}' of a given point $X_0 \in \mathcal{B}$. The solution can be chosen so that for $X_0 \in \mathcal{B}$, $Q(X_0)$ has a prescribed value, which we take to be the identity $\mathbf{1}$. Since $D\mathbf{u}$ is skew for all $\mathbf{u} \in \mathcal{V}$, it follows from (13.6) that QQ^T has gradient zero and hence must be equal to $\mathbf{1}$ everywhere in \mathcal{N}' . Hence Q has

orthogonal values. To summarize: If $\overset{*}{\mathbf{R}}=\mathbf{0}$, every point in \mathcal{B} has a neighborhood \mathcal{N}' on which we can find an orthogonal-valued tensor field \mathbf{Q} (of class C^{p-1}) such that (13.5) holds.

Assume, then, that (13.5) holds on \mathcal{N}' . Combining (12.6) with (13.5) we obtain

$$\mathbf{K} \overset{*}{\Gamma}_{\mathfrak{f}} \mathbf{c} = \mathbf{Q}^T ((\nabla_{\mathbf{K}} \mathbf{Q}) \mathbf{K} \mathfrak{f}) \mathbf{c} = \mathbf{Q}^T \nabla_{\mathbf{K}} (\mathbf{Q} \mathbf{c}) \mathbf{K} \mathfrak{f}, \quad (13.9)$$

which is valid when $\mathbf{c} = \mathbf{K}^{-1} \mathbf{c}$ is materially constant. Consider the affine connection $\bar{\Gamma}$ of class C^{p-1} defined by

$$\bar{\Gamma} \mathfrak{f} = \bar{\mathbf{K}}^{-1} \nabla_{\bar{\mathbf{K}}} (\bar{\mathbf{K}} \mathfrak{f}) \bar{\mathbf{K}}, \quad \mathfrak{f} \in \mathcal{T}_{\mathcal{B}}^1, \quad (13.10)$$

where

$$\bar{\mathbf{K}} = \mathbf{Q} \mathbf{K}. \quad (13.11)$$

It is easily seen that (13.9) is equivalent to the statement that

$$\overset{*}{\Gamma} \mathbf{c} = \bar{\Gamma} \mathbf{c} \quad (13.12)$$

holds for all materially constant tangent vector fields \mathbf{c} . Using the same argument as in the uniqueness proof of Theorem 3 (Sect. 10) we conclude that $\overset{*}{\Gamma} = \bar{\Gamma}$. Since $\overset{*}{\Gamma}$ has zero torsion, an analogue of Theorem 6 (Sect. 10) shows that every point in \mathcal{N}' must have a neighborhood \mathcal{N} such that $\bar{\mathbf{K}} = \nabla \boldsymbol{\kappa}$ for some configuration $\boldsymbol{\kappa}$ of \mathcal{N} . Hence, by (13.11), we have $\mathbf{Q} = (\nabla \boldsymbol{\kappa}) \mathbf{K}^{-1}$ on \mathcal{N} , i.e., \mathbf{K} is a state of contorted aeolotropy on \mathcal{N} . Q.E.D.

A special case of contorted aeolotropy is *curvilinear aeolotropy*. It corresponds to the case when there exists an orthogonal coordinate system on $\boldsymbol{\kappa}(\mathcal{B}) \subset \mathcal{E}$ with the following property: If $(\mathbf{e}_1(X), \mathbf{e}_2(X), \mathbf{e}_3(X))$ is the orthonormal basis which consists of the unit vectors that point in the direction of the coordinate lines at $\boldsymbol{\kappa}(X)$, then $\mathbf{Q}(X)^T \mathbf{e}_i(X)$ does not depend on $X \in \mathcal{B}$.

14. Special Types of Materially Uniform Bodies

We consider first the case when the isotropy group $\mathfrak{g}_{\mathbf{K}}$ of \mathcal{B} relative to some — and hence every — uniform reference \mathbf{K} is *discrete*. Suppose that \mathbf{K} and $\hat{\mathbf{K}}$ are two *continuous* uniform references. They must be related by (6.12), where $\mathbf{P}(X) \in \mathfrak{g}_{\mathbf{K}}$ must depend continuously on X . Since $\mathfrak{g}_{\mathbf{K}}$ is discrete, this is possible only when \mathbf{P} is constant. Thus we can absorb \mathbf{P} into \mathbf{L} , and (6.12) becomes $\hat{\mathbf{K}} = \mathbf{L} \mathbf{K}$, with $\mathbf{L} = \text{const}$. If we write (6.4) for both \mathbf{K} and $\hat{\mathbf{K}}$, we see that they correspond to the same continuous material uniformity. Since every continuous material uniformity must be of the form (6.4), where \mathbf{K} is a continuous uniform reference, we have the following result:

Theorem 8. *If the isotropy groups of a materially uniform simple body \mathcal{B} are discrete, then \mathcal{B} has at most one continuous material uniformity Φ . Any two continuous uniform references $\mathbf{K}, \hat{\mathbf{K}}$ are related by*

$$\hat{\mathbf{K}} = \mathbf{L} \mathbf{K}, \quad \mathbf{L} = \text{const} \in \ell. \quad (14.1)$$

Since material connections are only associated with differentiable uniformities Φ and not with discontinuous ones, it follows from the uniqueness assertion of

Theorem 8 that in the case when the isotropy groups are discrete, the inhomogeneity \mathfrak{S} is a characteristic of the body. If the isotropy groups are non-discrete Lie groups, however, and if there are any material uniformities of class C^{p-1} at all, there will be many, and hence also many inhomogeneities \mathfrak{S} for one and the same body. This is the case, in particular, for uniform isotropic bodies.

Next, we consider a uniform *isotropic* body \mathcal{B} with an undistorted uniform reference \mathbf{K} of class C^{p-1} , so that $\sigma \in \mathfrak{g}_{\mathbf{K}}$. If \mathbf{K} is a state of contorted aeolotropy, so that (13.1) holds, it follows from Theorem 2, (6.12), that $\nabla \kappa$ is again a uniform reference and that $\mathfrak{g}_{\mathbf{K}}$ is also the isotropy group relative to $\nabla \kappa$:

Theorem 9. *If a uniform isotropic body has an undistorted state of contorted aeolotropy, it is homogeneous.*

The conclusion of the theorem becomes false when the qualifier “undistorted” is omitted; *i.e.*, there are inhomogeneous isotropic bodies with distorted states of contorted aeolotropy.

Finally, suppose there is a natural way to single out, among all uniform references for \mathcal{B} , a particular class \mathbf{U} with the following property (*P*): All members of \mathbf{U} are of class C^{p-1} and differ from one another by a field of similarity transformations with constant ratio, so that $\mathbf{K}, \mathbf{K}' \in \mathbf{U}$ implies

$$\mathbf{K}' = a \mathbf{Q} \mathbf{K}, \tag{14.2}$$

where a is a real constant and \mathbf{Q} an orthogonal valued tensor field on \mathcal{B} . For example, if \mathcal{B} is a uniform solid body that is either isotropic or has cubic symmetry then the class \mathbf{U} of all undistorted references has the property (*P*). This follows from results proved in reference [9]. Other examples are obtained by letting \mathbf{U} be the class of all uniform references \mathbf{K} such that the corresponding response functions $\mathfrak{H}_{\mathbf{K}}$ satisfy a certain special condition such as $\mathfrak{H}_{\mathbf{K}}(\mathbf{1})=0$. Such references are often called *natural references*. The nature of the response function is often such that the class \mathbf{U} of natural references has the property (*P*).

If (14.2) holds, it follows from (12.1) and the fact that orthogonal transformations preserve inner products that the two Riemannian inner products corresponding to \mathbf{K}' and \mathbf{K} differ from one another only by the constant factor a^2 . Therefore, Proposition 6 shows that the Riemannian connections relative to \mathbf{K} and \mathbf{K}' are the same, and we have the following result:

Theorem 10. *If \mathcal{B} is a uniform simple body with a distinguished class \mathbf{U} of uniform references with the property (*P*), then the Riemannian connection $\overset{*}{\Gamma}$ and its curvature $\overset{*}{\mathfrak{R}}$ are characteristics of the body.*

The assertion of Theorem 10 applies, in particular, to uniform isotropic solid bodies, for which the curvature, $\overset{*}{\mathfrak{R}}$, defined by the class of undistorted uniform references, is an intrinsic measure of deviation from homogeneity.

15. Cauchy’s Equation of Balance

We now derive a new version of CAUCHY’S equation of balance, which expresses the fact that the forces acting on every part of a given body \mathcal{B} must add to zero. In order to do so, we first derive a lemma, Proposition 7 below.

Let \mathbf{h} be a vector field and \mathbf{T} a tensor field of class C^1 on \mathcal{B} . We define the *divergence* of these fields *relative to some reference \mathbf{K}* by

$$\operatorname{div}_{\mathbf{K}} \mathbf{h} = \operatorname{tr}(\nabla_{\mathbf{K}} \mathbf{h}) \quad \text{and} \quad (\operatorname{div}_{\mathbf{K}} \mathbf{T}) \cdot \mathbf{u} = \operatorname{div}_{\mathbf{K}}(\mathbf{T}^T \mathbf{u}), \quad \mathbf{u} \in \mathcal{V}, \quad (15.1)$$

respectively. The following product rules are valid for $f \in \mathcal{F}_{\mathcal{B}}^1$, $\mathbf{h} \in \mathcal{V}_{\mathcal{B}}^1$, $\mathbf{T} \in \mathcal{L}_{\mathcal{B}}^1$:

$$\begin{aligned} \operatorname{div}_{\mathbf{K}}(f \mathbf{h}) &= (\nabla_{\mathbf{K}} f) \cdot \mathbf{h} + f \operatorname{div}_{\mathbf{K}} \mathbf{h}, \\ \operatorname{div}_{\mathbf{K}}(\mathbf{T} \mathbf{h}) &= (\operatorname{div}_{\mathbf{K}} \mathbf{T}^T) \cdot \mathbf{h} + \operatorname{tr}(\mathbf{T} \nabla_{\mathbf{K}} \mathbf{h}). \end{aligned} \quad (15.2)$$

If $\mathbf{K} = \nabla \kappa$ is a configuration gradient, we write $\operatorname{div}_{\kappa}$ instead of $\operatorname{div}_{\mathbf{K}}$.

Suppose that a uniform reference \mathbf{K} of class C^{p-1} for \mathcal{B} and a tensor field \mathbf{T} of class C^1 on \mathcal{B} are given. For any configuration γ of \mathcal{B} we then define another tensor field \mathbf{T}_{γ} of class C^1 by

$$\mathbf{T}_{\gamma} = \frac{1}{J} \mathbf{T} \mathbf{F}^T, \quad (15.3)$$

where

$$\mathbf{F} = (\nabla \gamma) \mathbf{K}^{-1}, \quad J = |\det \mathbf{F}|. \quad (15.4)$$

Proposition 7. *If*

$$\mathbf{K}(X) = \nabla \kappa(X) \quad (15.5)$$

for some $X \in \mathcal{B}$, then the divergences at X relative to \mathbf{K} of the tensor field \mathbf{T} and of the tensor field \mathbf{T}_{κ} defined by (15.3) and (15.4) are related by

$$\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\kappa}|_X = (\operatorname{div}_{\mathbf{K}} \mathbf{T} + \mathbf{T} s)|_X, \quad (15.6)$$

where the vector field s is defined, in terms of the inhomogeneity \mathbf{S} relative to \mathbf{K} , by

$$s \cdot \mathbf{u} = \operatorname{tr}(\mathbf{S} \mathbf{u}), \quad \mathbf{u} \in \mathcal{V}. \quad (15.7)$$

Proof. We make use of (15.3) and (15.4) with γ replaced by κ . We then have

$$J(\mathbf{T}_{\kappa}^T \mathbf{u}) = \mathbf{F}(\mathbf{T}^T \mathbf{u}), \quad \mathbf{u} \in \mathcal{V}. \quad (15.8)$$

Using the rule

$$\mathbf{u} \cdot \nabla_{\mathbf{K}}(\det \mathbf{F}) = (\det \mathbf{F}) \operatorname{tr}[\mathbf{F}^{-1}((\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{u})], \quad \mathbf{u} \in \mathcal{V} \quad (15.9)$$

for the differentiation of a determinant, the product rules (15.2), and the definition (15.1)₂, we see that taking $\operatorname{div}_{\mathbf{K}}$ of (15.8) yields

$$J \operatorname{tr}[\mathbf{F}^{-1}((\nabla_{\mathbf{K}} \mathbf{F})(\mathbf{T}_{\kappa}^T \mathbf{u}))] + J(\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\kappa}) \cdot \mathbf{u} = (\operatorname{div}_{\mathbf{K}} \mathbf{F}^T) \cdot \mathbf{T}^T \mathbf{u} + \operatorname{tr}[\mathbf{F} \nabla_{\mathbf{K}}(\mathbf{T}^T \mathbf{u})]. \quad (15.10)$$

Since $\mathbf{F}(X) = \mathbf{1}$, $J(X) = 1$, and $\mathbf{T}_{\kappa}(X) = \mathbf{T}(X)$ by (15.5), (15.4), and (15.3), evaluation of (15.10) at $X \in \mathcal{B}$ gives

$$\{\operatorname{tr}[(\nabla_{\mathbf{K}} \mathbf{F})(\mathbf{T}^T \mathbf{u})] + (\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\kappa}) \cdot \mathbf{u} - (\operatorname{div}_{\mathbf{K}} \mathbf{F}^T) \cdot \mathbf{T}^T \mathbf{u} - (\operatorname{div}_{\mathbf{K}} \mathbf{T}) \cdot \mathbf{u}\}_X = 0. \quad (15.11)$$

Using the rule $\operatorname{tr}[(\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{v}] = \nabla_{\mathbf{K}}(\operatorname{tr} \mathbf{F}) \cdot \mathbf{v}$, we see that (15.11) can hold for all $\mathbf{u} \in \mathcal{V}$ only if

$$\{\mathbf{T}[\nabla_{\mathbf{K}}(\operatorname{tr} \mathbf{F}) - \operatorname{div}_{\mathbf{K}} \mathbf{F}^T] + \operatorname{div}_{\mathbf{K}} \mathbf{T}_{\kappa} - \operatorname{div}_{\mathbf{K}} \mathbf{T}\}_X = \mathbf{0}. \quad (15.12)$$

On the other hand, if we evaluate (11.4) at $X \in \mathcal{B}$ and take the trace, we obtain

$$\text{tr}(\mathbf{S}\mathbf{u})|_X = \{(\text{div}_{\mathbf{K}} \mathbf{F}^T) \cdot \mathbf{u} - \nabla_{\mathbf{K}}(\text{tr} \mathbf{F}) \cdot \mathbf{u}\}_X,$$

which, in view of (15.7), is equivalent to

$$s|_X = \{\text{div}_{\mathbf{K}} \mathbf{F}^T - \nabla_{\mathbf{K}}(\text{tr} \mathbf{F})\}_X. \quad (15.13)$$

The desired result (15.6) follows from (15.12) and (15.13). Q.E.D.

Let us assume now that the body \mathcal{B} is subject to internal contact forces and external body forces¹⁴. If the forces acting on every part of \mathcal{B} are balanced and if suitable regularity assumptions are satisfied one can prove the following results (cf. [10]):

(i) With every configuration κ of \mathcal{B} one can associate a stress tensor field \mathbf{T}_{κ} of class C^1 and a body force field \mathbf{b}_{κ} of class C^0 such that the force \mathbf{f} exerted on a part \mathcal{P} of \mathcal{B} by the combined action of a separate part \mathcal{P}' of \mathcal{B} and the external world is given by

$$\mathbf{f} = \int_{\kappa(\mathcal{P})} \mathbf{b}_{\kappa} dV + \int_{\mathcal{C}} \mathbf{T}_{\kappa} \mathbf{n} dS, \quad (15.14)$$

where \mathcal{C} is the surface of contact between \mathcal{P} and \mathcal{P}' in the configuration κ and where \mathbf{n} is the unit normal to \mathcal{C} directed away from $\kappa(\mathcal{P})$.

(ii) CAUCHY'S equation of balance

$$\text{div}_{\kappa} \mathbf{T}_{\kappa} + \mathbf{b}_{\kappa} = \mathbf{0} \quad (15.15)$$

is valid on \mathcal{B} for every configuration κ .

(iii) If κ and γ are two configurations of \mathcal{B} , then the stress fields \mathbf{T}_{κ} and \mathbf{T}_{γ} and the body force fields \mathbf{b}_{κ} and \mathbf{b}_{γ} are related by

$$\mathbf{T}_{\gamma} = \frac{1}{J} \mathbf{T}_{\kappa} \mathbf{F}^T, \quad \mathbf{b}_{\gamma} = \frac{1}{J} \mathbf{b}_{\kappa}, \quad (15.16)$$

where

$$\mathbf{F} = (\nabla \gamma)(\nabla \kappa)^{-1}, \quad J = |\det \mathbf{F}| \quad (15.17)$$

(cf. equation (43 A.3) of reference [2]).

Let \mathbf{K} be a uniform reference of class C^{p-1} for \mathcal{B} . Let a particular point $X \in \mathcal{B}$ be given. It is clear from (15.16) and (15.17) that $\mathbf{T}_{\gamma}(X) = \mathbf{T}_{\kappa}(X)$ and $\mathbf{b}_{\gamma}(X) = \mathbf{b}_{\kappa}(X)$ hold whenever both κ and γ belong to the equivalence class by which the local configuration $\mathbf{K}(X)$ is defined. Thus, we can define fields $\mathbf{T}_{\mathbf{K}}$ and $\mathbf{b}_{\mathbf{K}}$ by the condition that for each $X \in \mathcal{B}$,

$$\mathbf{T}_{\mathbf{K}}(X) = \mathbf{T}_{\kappa}(X), \quad \mathbf{b}_{\mathbf{K}}(X) = \mathbf{b}_{\kappa}(X) \quad (15.18)$$

hold whenever $\nabla \kappa(X) = \mathbf{K}(X)$. We call $\mathbf{T}_{\mathbf{K}}$ and $\mathbf{b}_{\mathbf{K}}$ the *stress tensor field* and *body force field relative to the reference \mathbf{K}* . It is clear from (15.16) and (15.17) that

$$\mathbf{T}_{\gamma} = \frac{1}{J} \mathbf{T}_{\mathbf{K}} \mathbf{F}^T, \quad \mathbf{b}_{\gamma} = \frac{1}{J} \mathbf{b}_{\mathbf{K}} \quad (15.19)$$

when \mathbf{F} and J are given by (15.4), where γ is an arbitrary configuration.

¹⁴ Inertial forces should be regarded as part of the body forces.

Since $\operatorname{div}_{\mathbf{K}} \mathbf{T}|_X = \operatorname{div}_{\mathbf{K}} \mathbf{T}|_X$ whenever $\nabla \boldsymbol{\kappa}(X) = \mathbf{K}(X)$, it follows from (15.15) that

$$[\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\mathbf{K}} + \mathbf{b}_{\mathbf{K}}]_X = \mathbf{0} \quad (15.20)$$

holds whenever $\nabla \boldsymbol{\kappa}(X) = \mathbf{K}(X)$. On the other hand, Proposition 7 applies when we choose $\mathbf{T} = \mathbf{T}_{\mathbf{K}}$. Thus, by substituting (15.6) with the choice $\mathbf{T} = \mathbf{T}_{\mathbf{K}}$ into (15.20) we obtain the following result:

Theorem 11. *The stress tensor field $\mathbf{T}_{\mathbf{K}}$ and the body force field $\mathbf{b}_{\mathbf{K}}$ relative to a uniform reference \mathbf{K} satisfy the modified equation of balance*

$$\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\mathbf{K}} + \mathbf{T}_{\mathbf{K}} s + \mathbf{b}_{\mathbf{K}} = \mathbf{0}, \quad (15.21)$$

where s is defined in terms of the inhomogeneity S relative to \mathbf{K} by (15.7).

The equation (15.21) is much more useful than (15.15) for dealing with inhomogeneous materially uniform bodies. Consider, for example, an elastic body \mathcal{B} , for which the set of response descriptors is the set $\mathcal{S} \subset \mathcal{L}$ of symmetric linear transformations. According to Theorem 1 (Sect. 6) we can associate with a given uniform reference \mathbf{K} a relative elastic response function $\mathfrak{H}_{\mathbf{K}}: \ell \rightarrow \mathcal{S}$. In order that a configuration γ be compatible with a given force system, the constitutive equation

$$\mathfrak{H}_{\mathbf{K}}(\mathbf{F}) = \mathbf{T}_{\gamma}, \quad \mathbf{F} = (\nabla \gamma) \mathbf{K}^{-1} \quad (15.22)$$

must be satisfied on \mathcal{B} , where \mathbf{T}_{γ} is the stress tensor field for γ . In view of (15.19)₁, (15.22) is equivalent to

$$\mathfrak{h}_{\mathbf{K}}(\mathbf{F}) = \mathbf{T}_{\mathbf{K}}, \quad \mathbf{F} = (\nabla \gamma) \mathbf{K}^{-1}, \quad (15.23)$$

where $\mathfrak{h}_{\mathbf{K}}: \ell \rightarrow \mathcal{L}$ is defined by

$$\mathfrak{h}_{\mathbf{K}}(\mathbf{F}) = |\det \mathbf{F}| \mathfrak{H}_{\mathbf{K}}(\mathbf{F}) \mathbf{F}^{\top -1}, \quad \mathbf{F} \in \ell. \quad (15.24)$$

Assume that $\mathfrak{h}_{\mathbf{K}}$ is of class C^1 , and denote its gradient by $\mathbf{H}_{\mathbf{K}}$. For each $\mathbf{F} \in \ell$ the value $\mathbf{H}_{\mathbf{K}}(\mathbf{F})$ is then a linear transformation from \mathcal{L} into \mathcal{L} . If we take the gradient of (15.23) relative to \mathbf{K} , the chain rule yields

$$(\nabla_{\mathbf{K}} \mathbf{T}_{\mathbf{K}}) \mathbf{u} = \mathbf{H}_{\mathbf{K}}(\mathbf{F}) [(\nabla_{\mathbf{K}} \mathbf{F}) \mathbf{u}], \quad \mathbf{u} \in \mathcal{V}. \quad (15.25)$$

It follows that

$$\operatorname{div}_{\mathbf{K}} \mathbf{T}_{\mathbf{K}} = \mathbf{A}_{\mathbf{K}}(\mathbf{F}) [(\nabla_{\mathbf{K}} \mathbf{F})], \quad (15.26)$$

where $\mathbf{A}_{\mathbf{K}}$ is that function on ℓ whose values $\mathbf{A}_{\mathbf{K}}(\mathbf{F}): \mathcal{L}(\mathcal{V}, \mathcal{L}) \rightarrow \mathcal{V}$ are determined by the property that $\mathbf{A}_{\mathbf{K}}(\mathbf{F})[\mathbf{Z}] \cdot \mathbf{w}$ is the trace of the linear transformation $\mathbf{u} \rightarrow \{\mathbf{H}_{\mathbf{K}}(\mathbf{F})[\mathbf{Z} \mathbf{u}]\}^{\top} \mathbf{w}$ for all $\mathbf{w} \in \mathcal{V}$ and all $\mathbf{Z} \in \mathcal{L}(\mathcal{V}, \mathcal{L})$. Of course, $\mathbf{A}_{\mathbf{K}}$ is determined by the response function $\mathfrak{h}_{\mathbf{K}}$. If we substitute (15.26) and (15.23) into (15.21), we obtain¹⁵

$$\mathbf{A}_{\mathbf{K}}(\mathbf{F}) [(\nabla_{\mathbf{K}} \mathbf{F})] + \mathfrak{h}_{\mathbf{K}}(\mathbf{F}) s + \mathbf{b}_{\mathbf{K}} = \mathbf{0}, \quad \mathbf{F} = (\nabla \gamma)^{-1} \mathbf{K}, \quad (15.27)$$

which is the differential equation for the determination of configurations γ possible in a materially uniform elastic body. If the body is homogeneous, we can choose

¹⁵ This result, in terms of coordinates, was announced two years ago in reference [2] as equation (44.7).

$\mathbf{K} = \nabla \kappa$. Then s vanishes and (15.27) reduces to the classical differential equation of finite elasticity.

Finally, we give another application of Proposition 7. Using the fact that \mathcal{V} is three-dimensional (which was irrelevant up to now), we choose an orientation in \mathcal{V} and consider the associated cross product \times . The curl of a vector field \mathbf{h} and a tensor field \mathbf{T} on \mathcal{B} relative to some configuration γ are defined by

$$\begin{aligned} \mathbf{u} \cdot (\nabla_\gamma \mathbf{h}) \mathbf{v} - \mathbf{v} \cdot (\nabla_\gamma \mathbf{h}) \mathbf{u} &= (\text{curl}_\gamma \mathbf{h}) \cdot (\mathbf{u} \times \mathbf{v}), \\ (\text{curl}_\gamma \mathbf{T})^T \mathbf{u} &= \text{curl}_\gamma (\mathbf{T} \mathbf{u}), \end{aligned} \quad (15.28)$$

where $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. It follows from (15.28) that

$$(\text{curl}_\gamma \mathbf{T}^T) (\mathbf{u} \times \mathbf{v}) = ((\nabla_\gamma \mathbf{T}) \mathbf{v}) \mathbf{u} - ((\nabla_\gamma \mathbf{T}) \mathbf{u}) \mathbf{v}. \quad (15.29)$$

Also, we have the rule

$$\text{div}_\gamma \text{curl}_\gamma \mathbf{T} = \mathbf{0}. \quad (15.30)$$

The inhomogeneity \mathbf{S} relative to a uniform reference \mathbf{K} has the skew symmetry $(\mathbf{S} \mathbf{u}) \mathbf{v} = -(\mathbf{S} \mathbf{v}) \mathbf{u}$, $\mathbf{u}, \mathbf{v} \in \mathcal{V}$. Therefore, \mathbf{S} determines and is determined by a tensor field \mathbf{A} on \mathcal{B} such that¹⁶

$$(\mathbf{S} \mathbf{u}) \mathbf{v} = \mathbf{A} (\mathbf{u} \times \mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in \mathcal{V}. \quad (15.31)$$

If we substitute (15.31) into (11.5) and observe the rule (15.29), we see that

$$\mathbf{A} (\mathbf{u} \times \mathbf{v}) = [\text{curl}_\gamma (\mathbf{F}^{T^{-1}})] (\mathbf{F} \mathbf{v} \times \mathbf{F} \mathbf{u}). \quad (15.32)$$

Hence, since $\mathbf{F}^T (\mathbf{F} \mathbf{v} \times \mathbf{F} \mathbf{u}) = (\det \mathbf{F}) (\mathbf{v} \times \mathbf{u})$ and since $\mathbf{u} \times \mathbf{v}$ is arbitrary, (15.32) yields

$$-\text{curl}_\gamma (\mathbf{F}^{T^{-1}}) = \frac{1}{J} \mathbf{A} \mathbf{F}^T, \quad J = \det \mathbf{F}. \quad (15.33)$$

Thus, the tensor field $-\text{curl}_\gamma (\mathbf{F}^{T^{-1}})$ is obtained from \mathbf{A} by the rules (15.3), (15.4), except that the absolute value signs are omitted in the definition of J , which does not affect the validity of Proposition 7. Since by rule (15.30) we have

$$[\text{div}_\mathbf{K} \text{curl}_\mathbf{K} (\mathbf{F}^{T^{-1}})]_X = [\text{div}_\mathbf{K} \text{curl}_\mathbf{K} (\mathbf{F}^{T^{-1}})]_X = \mathbf{0},$$

(15.6) yields

$$\text{div}_\mathbf{K} \mathbf{A} + \mathbf{A} s = \mathbf{0}, \quad (15.34)$$

where s is determined by \mathbf{A} through

$$-(\mathbf{A} - \mathbf{A}^T) \mathbf{u} = s \times \mathbf{u}, \quad \mathbf{u} \in \mathcal{V}. \quad (15.35)$$

Thus, (15.34) is a differential identity for \mathbf{A} and \mathbf{K} . One can show that it is equivalent to the Bianchi identity (11.11).

Acknowledgement. The research leading to this paper was supported by the Office of Naval Research under contract NONR-760 (30).

¹⁶ The field \mathbf{A} here corresponds to what was denoted by \mathbf{A}^T in [2], Sect. 34.

References

- [1] NOLL, W., A mathematical theory of the mechanical behavior of continuous media. Arch. Rational Mech. Anal. **2**, 197—226 (1958).
- [2] TRUESDELL, C., & W. NOLL, The Non-linear Field Theories of Mechanics. Encyclopedia of Physics, Vol. III/3. Berlin-Heidelberg-New York: Springer 1965.
- [3] LANG, S., Introduction to Differentiable Manifolds. New York-London-Sydney: Interscience 1962.
- [4] KRÖNER, E.: Allgemeine Kontinuumstheorie der Versetzungen und Eigenspannungen. Arch. Rational Mech. Anal. **4**, 273—334 (1960).
- [5] SEEGER, A.: Recent advances in the theory of defects in crystals. Physica status solidi **1**, 669—698 (1961).
- [6] NOLL, W., Space-time structures in classical mechanics. Delaware Seminar in the Foundations of Physics, pp. 28—34. New York: Springer 1967.
- [7] NOLL, W., Euclidean geometry and Minkowskian chronometry. Amer. Math. Monthly **71**, 129—144 (1964).
- [8] WANG, C.-C., On the geometric structures of simple bodies, a mathematical foundation for the theory of continuous distributions of dislocations. Arch. Rational Mech. Anal., following in this issue.
- [9] COLEMAN, B. D., & W. NOLL, Material symmetry and thermostatic inequalities in finite elastic deformations. Arch. Rational Mech. Anal. **15**, 87—111 (1964).
- [10] NOLL, W., The foundations of classical mechanics in the light of recent advances in continuum mechanics. The Axiomatic Method, with special reference to geometry and physics, pp. 266—281. Amsterdam: North Holland 1959.

Carnegie-Mellon University
Pittsburgh, Pennsylvania

(Received July 22, 1967)

A New Mathematical Theory of Simple Materials

WALTER NOLL

Dedicated to Clifford Truesdell, friend and mentor of the author

Contents

1. Introduction	1
2. Mathematical Preliminaries	5
3. Continuous Bodies, Intrinsic Configurations	8
4. Deformation Processes, Body Elements	9
5. Motions in a Frame Space	11
6. Intrinsic Stress, Work	11
7. Material Elements	12
8. Internal Constraints	14
9. Material Isomorphisms, Materially Uniform Bodies	16
10. Material Symmetry, Isotropy	18
11. Uniformity and Topology in the State Space	20
12. Relaxed States	22
13. Accessibility	24
14. Semi-Elastic and Elastic Materials	25
15. Deformation Histories for Semi-Elastic Elements	28
16. The History Functional of a Semi-Elastic Element	31
17. Monotonous Processes	35
18. Processes of Constant Reduced State, States of Monotonous Flow	37
19. Material Functionals and Functions for Incompressible Fluids	39
20. Materials of the Rate Type	44
21. Outlook	48
References	49

1. Introduction

The original theory of *simple materials* was formulated by me in 1958 in reference [N2]. I proposed this theory in an attempt to unify and clarify the confusing variety of theories of mechanical behavior of materials that had been proposed in the literature up to that time. One can perhaps say that the attempt was moderately successful, considering that this first theory of simple materials has served as a foundation for a large part of the research in continuum physics since 1958, and considering that the concepts and even the notations of [N2] are now used routinely and without reference in textbooks on continuum mechanics.

Unfortunately, the success of the first theory of simple materials obscured the fact that this theory has at least three severe defects:

- (i) In the original theory it is assumed, roughly, that the present stress is determined by the infinite past history of deformation. If one assumes that the memory of the material is limited, only a finite part of the history will have an appreciable effect on the present stress. There are materials, however, for which such an assumption is not appropriate. For such materials the original theory brings the infinite past into play. This is philosophically unacceptable, because real material bodies (perhaps even the entire universe) have not existed for an infinite time. Even if one imagines the existence of a material body for the infinite past, only a finite portion of that past is knowable.
- (ii) The original theory has failed to give an adequate conceptual framework for the mathematical description of such phenomena as *plasticity*, *yield*, and *hysteresis*.
- (iii) The definition of a *material of the rate type* in the framework of the first theory (*cf.* [N2], Sect. 24 or [TN], Sect. 36) was artificial and fraught with difficulties. BERNSTEIN ([B1] and [B2]) gave another and much more natural definition, but his definition did not fit into the framework of the first theory.

The new theory of simple materials presented here is free from the three defects just described. Only deformation processes of finite duration, not infinite histories, occur in the description of the response of a *material element*. I believe the new theory is suited very well as a basis for mathematical theories describing phenomena of plasticity, although such theories have not yet been developed on this basis. Finally, BERNSTEIN'S definition of a material of the rate type fits naturally into the new framework.

The idea of a *state of a physical system* is used, informally or formally, in almost all branches of physics. It is used formally as a primitive notion, for example, in most axiomatic descriptions of quantum mechanics. Informally, *i.e.* without a precise mathematical meaning, the term "state" has also been used in much work on continuum mechanics. In the axiomatic definition of a material element presented here, the notion of a *state of the element* is used formally as a primitive notion with a precise mathematical meaning.

An important difference between the study presented here and all previous work in continuum physics (as far as I know) is that in the present study no *a-priori* existence of *frames of reference* or *physical space* is assumed and no *a-priori coordinate systems* are used for the description of the response of material bodies. In the past literature, an inordinate effort has been expended to derive the consequences of the *principle of frame-indifference* for whatever constitutive assumptions were considered. This was the case even before the principle of frame-indifference was formulated with precision and generality in [N3]; it was the case, for example, in the pioneering work of CAUCHY and STOKES. The advantage of not using frames of reference is, of course, that the principle of frame-indifference is vacuously satisfied and need not be considered explicitly at all. An increasing number of researchers recognize (mainly since the publication of my thesis [N1] in 1955) that *a-priori* coordinate systems and hence considerations of coordinate invariance are redundant and an impediment to understanding.

The intrinsic, frame-free description of material elements has a second advantage: no *a priori* commitment need be made about the space-time structure used when the interaction of an element with an environment is considered. The description given here is compatible with classical, neo-classical (*cf.* [N6]), as well as relativistic spacetime. The imbedding of material bodies in relativistic space-time will be investigated in a future paper.

We now give a brief survey of the contents of this paper.

In Section 2 we assemble the mathematical notation and terminology used later. The basic vocabulary that mathematicians have adopted in the last few years is included in order to aid those readers whose primary interest is in mechanics and who may not yet be familiar with this vocabulary. It is assumed that the reader has a good grasp of the concepts of elementary abstract linear algebra and is comfortable with vector spaces that have no inner product.

In Section 3 we introduce, somewhat informally, the notions of *intrinsic configuration* of a continuous body and of an infinitesimal body element. This intrinsic notion is different from the extrinsic concept of "configuration" I have used in past work. I now prefer the term "placement in a frame of reference" for the extrinsic concept.

Section 4 contains the formal definition of a *body element* and a rigorous discussion of (intrinsic) *deformation processes* for a body element.

In Section 5 we describe the connection between (intrinsic) deformation processes and *motions in a frame space* introduced *a posteriori*.

In Section 6 we explain the concept of *intrinsic stress* and we show how work can be computed in terms of intrinsic stresses and deformations.

Section 7 contains the formal definition for the basic concept of the present paper, that of a *material element*. The structure of a material element is governed by six axioms, of which only the first three are stated in Section 7.

In Section 8 we discuss *internal constraints* and demonstrate that no formal modifications are necessary when such constraints are present.

Section 9 deals with *material isomorphisms*, which are defined in accordance with the notion of isomorphism for arbitrary mathematical structures. A *material* is defined as an equivalence class of material elements, the equivalence being material isomorphy. Included in Section 9 are also a few remarks intended to clarify the nature of the theory of materially uniform but inhomogeneous material bodies.

In Section 10, the *symmetry group* of a material element is defined as its group of automorphisms. We also give a precise definition for the *symmetry group of a state*. States whose symmetry group is an appropriate orthogonal group are called *isotropic states*, and *isotropic material elements* are defined as those that possess isotropic states.

In Section 11 we define natural uniform structures on sections of the state space and a natural topology on the state space. A *completeness axiom* postulates that the sections be complete uniform spaces. Readers who are not familiar with the elements of abstract general topology can skip this section and just assume that a notion of convergence for states is somehow defined.

In Section 12 we state and discuss the *relaxation axiom*, which postulates that if a material element is frozen in a definite configuration, its state will approach a *relaxed state*.

In Section 13 we say what it means that one state is *accessible* from another and we state the *accessibility axiom*, which postulates that all states must be accessible from a certain relaxed starting state.

In Section 14 an *elastic* element is defined by the condition that the states be in one-to-one correspondence with the configurations. A *semi-elastic* element is defined by the condition that the *relaxed* states be in one-to-one correspondence with the configurations. Solid and fluid materials are defined as semi-elastic materials with special symmetry. Notions of *undistorted state* and *undistorted configuration* are defined so as to apply to all semi-elastic materials (not just to solids and isotropic materials, as in the original theory).

In Section 15 we show that, for semi-elastic elements, a dense set of states can be described by deformation histories of infinite duration.

In Section 16 we establish a precise position for the original theory of simple materials. It turns out that simple materials in the original sense are equivalent to semi-elastic materials in the new sense. The proof of this equivalence is by no means simple. The history functional of a semi-elastic element, which relates infinite deformation histories to stresses, must satisfy two convergence conditions. These conditions express the fact that semi-elastic materials have limited memory, as is to be expected from the discussion about the first defect of the original theory.

The considerations of Section 17 are purely kinematical. A *monotonous process* is defined by the condition that all of its segments of equal duration are congruent. A non-trivial theorem, whose proof will be given elsewhere [N8], shows that the monotonous processes are related to what have been called “substantially stagnant motions” by COLEMAN [C] and “motions with constant stretch history” by myself [N4].

In Section 18 we explain why the dynamical analysis of monotonous processes is often not much more complicated than a statical analysis. The reason is that, if the material has a sufficiently large symmetry group and if certain other conditions are satisfied, such processes leave the reduced state unchanged. The states occurring in such processes are called *states of monotonous flow*.

Materials (as distinct from material elements) can often be characterized by numerical functionals, functions, or constants. In Section 19 such a characterization is carried out in detail for incompressible fluids.

In Section 20 we define and discuss materials of the rate type. There is no evidence, conceptual or experimental, that any real material behavior fits the theory of materials of the rate type, but this theory gives mathematically concrete examples of material elements that are not semi-elastic and hence do not fit the original theory of simple materials. (The definition of a material of the rate type given in [N2] and [TN] applies, in the language of the present paper, only to materials that are both semi-elastic and of the rate type.)

In Section 21 we discuss possible generalizations, further development, and applications of the theory presented here.

2. Mathematical Preliminaries

(a) Basic Notations. The set of real numbers is denoted by \mathbb{R} and the set of non-negative reals by \mathbb{R}^+ . If $t_1, t_2 \in \mathbb{R}$ and $t_1 \leq t_2$, the closed interval from t_1 to t_2 is defined by $[t_1, t_2] = \{t \mid t_1 \leq t \leq t_2\}$. In particular, $[t, t] = \{t\}$ is a degenerate interval that reduces to a singleton. The set of natural numbers $1, 2, 3 \dots$ is denoted by \mathbb{N} .

A mapping $f: \mathcal{D} \rightarrow \mathcal{C}$ assigns to each element x in the set \mathcal{D} a value $f(x)$ in the set \mathcal{C} . The set \mathcal{D} is called the *domain* of f and \mathcal{C} is called the *codomain* of f . One expresses this situation by saying that f maps \mathcal{D} into \mathcal{C} . The set

$$\text{Range } f = \{f(x) \mid x \in \mathcal{D}\}$$

is called the *range* (or “image”) of f . The range is a subset of the codomain. If the range coincides with the codomain, one says that the mapping is *surjective* (or “onto”) and calls it a *surjection*. A mapping f is said to be *injective* (or “one-to-one”) and is called an *injection* if $f(x_1) = f(x_2)$ is possible only when $x_1 = x_2$.

The *composition* $g \circ f: \mathcal{D} \rightarrow \mathcal{B}$ of two mappings $f: \mathcal{D} \rightarrow \mathcal{C}$ and $g: \mathcal{C} \rightarrow \mathcal{B}$ is defined by $(g \circ f)(x) = g(f(x))$ for all $x \in \mathcal{D}$. The composition $g \circ f$ is meaningful only if the codomain of f coincides with (or at least is contained in) the domain of g . The *identity* mapping $1_{\mathcal{D}}: \mathcal{D} \rightarrow \mathcal{D}$ of a set \mathcal{D} is defined by $1_{\mathcal{D}}(x) = x$ for all $x \in \mathcal{D}$. A mapping $f: \mathcal{D} \rightarrow \mathcal{C}$ is said to be *invertible* (or “bijective” or “one-to-one and onto”) if it has an *inverse* $f^{-1}: \mathcal{C} \rightarrow \mathcal{D}$ with the properties $f^{-1} \circ f = 1_{\mathcal{D}}$ and $f \circ f^{-1} = 1_{\mathcal{C}}$. A mapping is invertible if and only if it is both injective and surjective. If this is the case, the mapping is also called a *bijection* (or “one-to-one correspondence”).

If the values of a mapping f are given by a formula and if no special symbol for the mapping itself is available, one uses the notation $x \mapsto f(x)$ to denote the mapping. (Note the distinction between \mapsto and \rightarrow .) For example, $t \mapsto t^2$ denotes the “squaring” mapping from \mathbb{R} into \mathbb{R} (or into \mathbb{R}^+). Another possible notation for $x \mapsto f(x)$ is $f(\cdot)$. For example, if the domain of $g: \mathcal{D}_1 \times \mathcal{D}_2 \rightarrow \mathcal{C}$ is a Cartesian product $\mathcal{D}_1 \times \mathcal{D}_2 = \{(x_1, x_2) \mid x_1 \in \mathcal{D}_1, x_2 \in \mathcal{D}_2\}$, one can obtain, for every $x_1 \in \mathcal{D}_1$, a mapping $g(x_1, \cdot): \mathcal{D}_2 \rightarrow \mathcal{C}$.

It often happens that the codomains \mathcal{C}_1 and \mathcal{C}_2 of two mappings $f_1: \mathcal{D} \rightarrow \mathcal{C}_1$ and $f_2: \mathcal{D} \rightarrow \mathcal{C}_2$ are such that there is an algebraic operation which assigns to each $y_1 \in \mathcal{C}_1$ and each $y_2 \in \mathcal{C}_2$ a $y_1 y_2$ in a set \mathcal{C}_3 . One then defines $f_1 f_2: \mathcal{D} \rightarrow \mathcal{C}_3$ by value-wise application of the operation, *i.e.* by $(f_1 f_2)(x) = f_1(x) f_2(x)$ for all $x \in \mathcal{D}$. For example, if $f: \mathcal{D} \rightarrow \mathbb{R}$ and $g: \mathcal{D} \rightarrow \mathbb{R}$ then $f + g: \mathcal{D} \rightarrow \mathbb{R}$ and $fg: \mathcal{D} \rightarrow \mathbb{R}$ are defined by value-wise addition and multiplication, *i.e.* by $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for all $x \in \mathcal{D}$.

The terms “function”, “functional”, and “transformation” are some of many synonyms for “mapping”.

(b) Vector Spaces. We deal only with finite-dimensional real vector spaces. If \mathcal{T}_1 and \mathcal{T}_2 are such spaces, we denote by

$$\text{Lin}(\mathcal{T}_1, \mathcal{T}_2) = \{A: \mathcal{T}_1 \rightarrow \mathcal{T}_2 \mid A \text{ linear}\}$$

the set of all linear mappings from \mathcal{T}_1 into \mathcal{T}_2 . The composition of two linear mappings A_1 and A_2 , meaningful if the domain of A_2 coincides with the codomain

of A_1 , is again linear and is denoted by $A_2 A_1$. (The symbol \circ is omitted in this case.) $\text{Lin}(\mathcal{T}_1, \mathcal{T}_2)$ is again a vector space, its structure being defined by taking value-wise sums and scalar multiples. The notation

$$\text{Invlin}(\mathcal{T}_1, \mathcal{T}_2) = \{A \in \text{Lin}(\mathcal{T}_1, \mathcal{T}_2) \mid A^{-1} \text{ exists}\}$$

is used for the set of invertible linear mappings from \mathcal{T}_1 into \mathcal{T}_2 . $\text{Invlin}(\mathcal{T}_1, \mathcal{T}_2)$ is non-empty only if $\dim \mathcal{T}_1 = \dim \mathcal{T}_2$. We use the abbreviations

$$\text{Lin}(\mathcal{T}) = \text{Lin}(\mathcal{T}, \mathcal{T}), \quad \text{Invlin}(\mathcal{T}) = \text{Invlin}(\mathcal{T}, \mathcal{T}).$$

$\text{Lin}(\mathcal{T})$ is not only a vector space, but an algebra with multiplication defined by composition. $\text{Invlin}(\mathcal{T})$ is not a vector space, but it is a group under composition, and it is often called the *general linear group* of \mathcal{T} . It is the group of automorphisms of the vector space \mathcal{T} .

The *dual* \mathcal{T}^* of a vector space \mathcal{T} is defined by

$$\mathcal{T}^* = \text{Lin}(\mathcal{T}, \mathbb{R}).$$

It is a vector space of the same dimension as \mathcal{T} . The elements of \mathcal{T}^* are called *linear forms* or *covectors*. The value of $u^* \in \mathcal{T}^*$ at $v \in \mathcal{T}$ is denoted by $\langle u^*, v \rangle \in \mathbb{R}$. The dual \mathcal{T}^{**} of the dual \mathcal{T}^* is identified with the original space \mathcal{T} in such a way that the value at $u^* \in \mathcal{T}^*$ of the element of $\mathcal{T}^{**} = \text{Lin}(\mathcal{T}^*, \mathbb{R})$ corresponding to $v \in \mathcal{T}$ is $\langle u^*, v \rangle$. If $v \in \mathcal{T}$ and $u^* \in \mathcal{T}^*$, we define $v \otimes u^* \in \text{Lin}(\mathcal{T})$ by

$$(v \otimes u^*)(w) = \langle u^*, w \rangle v \quad \text{for all } w \in \mathcal{T}.$$

The dual of $\text{Lin}(\mathcal{T})$ contains a distinguished element $\text{tr} \in (\text{Lin}(\mathcal{T}))^*$, called the *trace*. It is characterized by the condition that

$$\text{tr}(v \otimes u^*) = \langle u^*, v \rangle \quad \text{for all } v \in \mathcal{T} \text{ and all } u^* \in \mathcal{T}^*.$$

Another distinguished function on $\text{Lin}(\mathcal{T})$ is the *determinant* $\det: \text{Lin}(\mathcal{T}) \rightarrow \mathbb{R}$. It is not linear but preserves products, so that $\det(A_1 A_2) = (\det A_1)(\det A_2)$. Also, one has $\det(1_{\mathcal{T}}) = 1$. We use the notations

$$\text{Lin}_0(\mathcal{T}) = \{A \in \text{Lin}(\mathcal{T}) \mid \text{tr } A = 0\}$$

and

$$\text{Unim}(\mathcal{T}) = \{A \in \text{Invlin}(\mathcal{T}) \mid (\det A)^2 = 1\}.$$

$\text{Lin}_0(\mathcal{T})$ is a subspace of $\text{Lin}(\mathcal{T})$ and $\text{Unim}(\mathcal{T})$ is a subgroup of $\text{Invlin}(\mathcal{T})$. The group $\text{Unim}(\mathcal{T})$ is called the *unimodular group* of \mathcal{T} .

To every $A \in \text{Lin}(\mathcal{T}_1, \mathcal{T}_2)$ one can associate a unique *adjoint* $A^* \in \text{Lin}(\mathcal{T}_2^*, \mathcal{T}_1^*)$ characterized by the condition that $\langle A^* w^*, v \rangle = \langle w^*, Av \rangle$ for all $v \in \mathcal{T}_1$ and all $w^* \in \mathcal{T}_2^*$. If $A \in \text{Lin}(\mathcal{T})$, then $A^* \in \text{Lin}(\mathcal{T}^*)$ and we have $\text{tr } A = \text{tr } A^*$, $\det A^* = \det A$.

(c) Bilinear Forms, Orthogonal Groups, Inner Product Spaces. Let \mathcal{T} be a finite-dimensional real vector space. The space $\text{Lin}(\mathcal{T}, \mathcal{T}^*)$ can be identified with the space of all *bilinear forms* on \mathcal{T} , because every such bilinear form is a mapping from $\mathcal{T} \times \mathcal{T}$ into \mathbb{R} of the type

$$(v, w) \mapsto \langle Gv, w \rangle, \quad \text{with } G \in \text{Lin}(\mathcal{T}, \mathcal{T}^*).$$

The subspace

$$\text{Sym}(\mathcal{T}, \mathcal{T}^*) = \{G \in \text{Lin}(\mathcal{T}, \mathcal{T}^*) \mid \langle Gv, w \rangle = \langle Gw, v \rangle \text{ for all } v, w \in \mathcal{T}\}$$

of $\text{Lin}(\mathcal{T}, \mathcal{T}^*)$ is then identified with the space of *symmetric bilinear forms* on \mathcal{T} . The subset

$$\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) = \{G \in \text{Sym}(\mathcal{T}, \mathcal{T}^*) \mid \langle Gv, v \rangle > 0 \text{ if } v \in \mathcal{T}, v \neq 0\}$$

of $\text{Sym}(\mathcal{T}, \mathcal{T}^*)$ is identified with the set of all *positive definite symmetric bilinear forms* on \mathcal{T} . We note that

$$\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) \subset \text{Invlin}(\mathcal{T}, \mathcal{T}^*),$$

i. e., G^{-1} exists for every $G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$. The following fact will be needed: For every $G_0 \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ we have

$$\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) = \{A^* G_0 A \mid A \in \text{Invlin} \mathcal{T}\}.$$

For any $G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ we put

$$\text{Orth}(G) = \{A \in \text{Lin}(\mathcal{T}) \mid A^* G A = G\}.$$

For every $G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$, the set $\text{Orth}(G)$ is a proper subgroup of $\text{Unim}(\mathcal{T})$. We call $\text{Orth}(G)$ the *orthogonal group of G*. The following facts concerning orthogonal groups will be needed:

(i) For all $G_1, G_2 \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ we have

$$\begin{aligned} \text{Orth}(G_1) \subset \text{Orth}(G_2) &\Leftrightarrow \text{Orth}(G_1) = \text{Orth}(G_2) \\ &\Leftrightarrow G_2 = c G_1 \quad \text{for some } c > 0. \end{aligned}$$

(ii) The groups $\text{Orth}(G)$ are maximal subgroups of $\text{Unim}(\mathcal{T})$, *i. e.*, if g is a proper subgroup of $\text{Unim}(\mathcal{T})$ such that $g \supset \text{Orth}(G)$, then $g = \text{Orth}(G)$.¹

An *inner product space* \mathcal{V} is a vector space equipped with an additional structure by singling out a particular member $I \in \text{Sym}^+(\mathcal{V}, \mathcal{V}^*)$, which is called the *inner product of* \mathcal{V} . One uses the notation

$$u \cdot v = \langle Iu, v \rangle \quad \text{for } u, v \in \mathcal{V}$$

and calls $u \cdot v$ the *inner product of u and v*. Also, since $I: \mathcal{V} \rightarrow \mathcal{V}^*$ is a bijection, it is used to identify \mathcal{V} with \mathcal{V}^* . Thus, when $u \in \mathcal{V}$ is regarded as a linear form on \mathcal{V} , its value at $v \in \mathcal{V}$ is $u \cdot v$. $\text{Lin}(\mathcal{V}^*)$ is identified with $\text{Lin}(\mathcal{V})$ and hence the adjoint of $A \in \text{Lin}(\mathcal{V})$ is identified with another element of $\text{Lin}(\mathcal{V})$. This element is denoted by A^T instead of A^* and called the *transpose of A*. The space $\text{Sym}(\mathcal{V}, \mathcal{V}^*)$ is identified with the subspace

$$\text{Sym}(\mathcal{V}) = \{A \in \text{Lin}(\mathcal{V}) \mid A = A^T\}$$

of $\text{Lin}(\mathcal{V})$. The orthogonal group of the inner product I is also called the *orthogonal group of the inner product space* \mathcal{V} and denoted by

$$\text{Orth}(\mathcal{V}) = \{Q \in \text{Lin}(\mathcal{V}) \mid Q^T Q = 1_{\mathcal{V}}\}.$$

$\text{Orth}(\mathcal{V})$ is the group of automorphisms of the inner product space \mathcal{V} .

¹ This is a rather non-trivial theorem. The simplest proof I know is the one I gave in [N5].

Let \mathcal{V} be an inner product space and \mathcal{T} a vector space. If $K \in \text{Lin}(\mathcal{T}, \mathcal{V})$, then $K^* \in \text{Lin}(\mathcal{V}^*, \mathcal{T}^*) = \text{Lin}(\mathcal{V}, \mathcal{T}^*)$ and hence $K^*K \in \text{Lin}(\mathcal{T}, \mathcal{T}^*)$. Actually, we always have $K^*K \in \text{Sym}(\mathcal{T}, \mathcal{T}^*)$ and

$$\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) = \{K^*K \mid K \in \text{InvlIn}(\mathcal{T}, \mathcal{V})\}$$

if $\dim \mathcal{T} = \dim \mathcal{V}$.

A special inner product space is the numerical space \mathbb{R}^n consisting of n -tuples of real numbers. Actually, \mathbb{R}^n is equipped with even more structure than an inner product space, because \mathbb{R}^n possesses a distinguished orthonormal basis, while for inner product spaces without additional structure any orthonormal basis is as good as any other.

3. Continuous Bodies, Intrinsic Configurations

A *continuous body* \mathcal{B} is usually defined¹ as a set endowed with a structure described by a class Φ of mappings $\kappa: \mathcal{B} \rightarrow \mathcal{E}_\kappa$ whose codomains \mathcal{E}_κ are Euclidean spaces. The elements X, Y, \dots of \mathcal{B} are called *material points*. The spaces \mathcal{E}_κ have an intuitive interpretation as *frames of reference*, and a mapping $\kappa \in \Phi$ assigns to each material point $X \in \mathcal{B}$ a *place* $\kappa(X)$ in the frame of reference \mathcal{E}_κ . We will refer to the mappings in Φ as the possible **placements**² of the body \mathcal{B} .

The class Φ of placements of \mathcal{B} is subject to a number of axioms which we will not repeat here. For our purposes it is enough to know that the axioms are such as to determine, on \mathcal{B} , a unique structure of a differentiable manifold. Therefore one can associate with each material point $X \in \mathcal{B}$ a tangent space \mathcal{T}_X , which is a finite-dimensional real vector space (see [N7], Sect. 4). Intuitively, the members of \mathcal{T}_X describe the material points in an infinitesimal neighborhood of X . Therefore, we call \mathcal{T}_X the (infinitesimal) *body element* at X of the (global) body \mathcal{B} . If $\kappa \in \Phi$ is a placement of \mathcal{B} , its gradient $\nabla \kappa(X): \mathcal{T}_X \rightarrow \mathcal{V}_\kappa$ is an invertible linear mapping from \mathcal{T}_X onto the translation space \mathcal{V}_κ of \mathcal{E}_κ (see, e.g., [N7], Sects. 3 and 4). Note that \mathcal{V}_κ is an inner product space, while \mathcal{T}_X is a vector space without inner product.

We thus see that localizations of placements of a body to one of its body elements \mathcal{T}_X are described by invertible linear mappings from \mathcal{T}_X onto inner product spaces. We therefore call such mappings *placements of the body element* \mathcal{T}_X .

We observe that one can associate with each placement $\kappa: \mathcal{B} \rightarrow \mathcal{E}_\kappa$ a distance function $\delta_\kappa: \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}^+$ defined by

$$\delta_\kappa(X, Y) = d_\kappa(\kappa(X), \kappa(Y)), \quad (3.1)$$

where d_κ denotes the Euclidean distance in the frame space \mathcal{E}_κ . Two placements $\kappa_1, \kappa_2 \in \Phi$ determine the same distance function if and only if $\kappa_2 \circ \kappa_1^{-1}$ is an isometry between a subset of \mathcal{E}_{κ_1} and a subset of \mathcal{E}_{κ_2} .

Let

$$\Delta = \{\delta_\kappa \mid \kappa \in \Phi\} \quad (3.2)$$

¹ Although the concept has been used in the literature on continuum mechanics since time immemorial, the first precise mathematical definition was given in [N3] (see also [N7], Sect. 3).

² In earlier work I used the term "configuration". I now believe "placement" is more apt. Also, I would like to reserve "configuration" for the intrinsic concept to be introduced below.

be the class of all distance functions of the form (3.1). Instead of using Φ to describe the structure on \mathcal{B} and then defining Δ by (3.2) in terms of Φ , one can also use a suitable class Δ of mappings $\delta: \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}^+$ to describe the structure of a continuous body on \mathcal{B} and then define Φ to be the class of mappings $\kappa: \mathcal{B} \rightarrow \mathcal{E}_\kappa$ for which $(X, Y) \mapsto d_\kappa(\kappa(X), \kappa(Y))$ belongs to Δ . The advantage of describing the body structure in terms of the class Δ is that such a description is completely intrinsic. The elements of Δ do not involve extrinsic frames of reference, as do the elements of Φ .

We shall refer to the distance functions in Δ as the possible (*intrinsic*) **configurations** of the body \mathcal{B} . A configuration $\delta \in \Delta$ assigns to any two material points $X, Y \in \mathcal{B}$ a number $\delta(X, Y) \in \mathbb{R}^+$, which we call the *distance between X and Y in the configuration δ* .

A configuration $\delta: \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}^+$ induces, by localization, on each of the tangent spaces \mathcal{T}_X a positive definite symmetric bilinear form $\delta_X: \mathcal{T}_X \times \mathcal{T}_X \rightarrow \mathbb{R}^+$. In precise terms, δ_X is the mixed second gradient of δ with respect to its two point variables, evaluated at (X, X) . It is natural, therefore, to call such bilinear forms (*intrinsic*) *configurations of the body element \mathcal{T}_X* . Recalling that $\text{Sym}^+(\mathcal{T}_X, \mathcal{T}_X^*)$ can be identified with the set of all positive definite symmetric bilinear forms on \mathcal{T}_X (see Section 2), we see that the configurations of \mathcal{T}_X can be regarded as elements of $\text{Sym}^+(\mathcal{T}_X, \mathcal{T}_X^*)$.

A *motion* of a body \mathcal{B} is a one-parameter family of placements of \mathcal{B} (see [N2]). By a *deformation process* for a body \mathcal{B} we mean a one-parameter family of configurations of \mathcal{B} . Motions and deformation processes of body elements are defined in the same way.

From now on we shall deal almost exclusively with body elements rather than with (global) bodies. We think of a body element as a physical system in itself, regarding the rest of the continuous body to which the element belongs merely as part of the environment.

Precise formal definitions, motivated by the considerations of this section, for the concepts of a body element and of configurations, deformation processes, placements, and motions of a body element will be given in the next two sections.

4. Deformation Processes, Body Elements

Let \mathcal{G} be any set. A function P of the type

$$P: [0, d_p] \rightarrow \mathcal{G}, \quad d_p \in \mathbb{R}^+ \tag{4.1}$$

will be called a **process with values in \mathcal{G}** . The number d_p will be called the *duration* of the process and the values

$$P^i = P(0), \quad P^f = P(d_p) \tag{4.2}$$

will be referred to as the *initial and final values* of the process.

If $G \in \mathcal{G}$ and $t \in \mathbb{R}^+$, we define a process $G_{(t)}: [0, t] \rightarrow \mathcal{G}$ by

$$G_{(t)}(r) = G \quad \text{for } r \in [0, t] \tag{4.3}$$

and call it the **freeze of duration t at G** .

Let P be a process and let $t_1, t_2 \in [0, d_P]$, with $t_1 \leq t_2$. We then define a new process $P_{[t_1, t_2]}$, of duration $t_2 - t_1$, by

$$P_{[t_1, t_2]}(t) = P(t + t_1) \quad \text{for } t \in [0, t_2 - t_1]. \quad (4.4)$$

Such a process will be called a **segment** of the given process P .

Let P_1 and P_2 be processes such that $P_1^f = P_2^i$. We then define a new process $P_1 * P_2$, of duration $d_{P_1} + d_{P_2}$, by

$$(P_1 * P_2)(t) = \begin{cases} P_1(t) & \text{if } t \in [0, d_{P_1}] \\ P_2(t - d_{P_1}) & \text{if } t \in [d_{P_1}, d_{P_1} + d_{P_2}]. \end{cases} \quad (4.5)$$

We call $P_1 * P_2$ the **continuation** of P_1 with P_2 . Continuation is an associative operation, *i. e.*, we have

$$(P_1 * P_2) * P_3 = P_1 * (P_2 * P_3) \quad (4.6)$$

if $P_1^f = P_2^i = (P_2 * P_3)^i$ and $P_2^f = (P_1 * P_2)^f = P_3^i$.

We note that processes of duration zero are not excluded. Of course, all such processes are of the form $G_{(0)}$, *i. e.* they are freezes of duration zero. It is clear that for any process P

$$P = P_{(0)}^i * P = P * P_{(0)}^f. \quad (4.7)$$

Definition 4.1. A **body element** is a triple $(\mathcal{T}, \mathcal{G}, \Pi)$, where \mathcal{T} is a finite-dimensional real vector space, \mathcal{G} is a closed and connected subset of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$, and Π is a class of processes with values in \mathcal{G} which satisfies the following conditions:

- (P1) Any freeze at any $G \in \mathcal{G}$ belongs to Π .
- (P2) If P belongs to Π , so does every segment of P .
- (P3) Π is closed under continuation, *i. e.*, if $P_1, P_2 \in \Pi$ and $P_1^f = P_2^i$, then $P_1 * P_2 \in \Pi$.
- (P4) Any two elements $G_1, G_2 \in \mathcal{G}$ can be connected by a process in Π , *i. e.*, there is at least one $P \in \Pi$ such that $G_1 = P^i$ and $G_2 = P^f$.

When we speak about a body element \mathcal{T} , we consider it understood that \mathcal{T} is endowed with a structure defined by the prescription of \mathcal{G} and Π . The elements of \mathcal{G} will be called the **configurations** of \mathcal{T} and the processes in Π the **deformation processes** for \mathcal{T} .

Examples of classes Π that satisfy (P1)–(P4) are: (a) the class of all processes with values in \mathcal{G} , (b) the class of all continuous processes with values in \mathcal{G} , (c) if \mathcal{G} is arcwise connected, the class of all processes P that are continuous and have a piecewise continuous derivative \dot{P} in the sense that $\dot{P}(t)$ exist for all but a finite number of $t \in [0, d_P]$ and \dot{P} has left and right limits for all $t \in [0, d_P]$.

Remark. It cannot happen that the class Π contains only processes that are everywhere differentiable, except in the trivial case when \mathcal{G} is a singleton. To see this, we note that (P4) implies the existence of a non-constant $P \in \Pi$. If P is everywhere differentiable, there must be a $t > 0$ such that $\dot{P}(t) \neq 0$. By (P1)–(P3), the process $P_{[0, t]} * P(t)_{(t')}$, where $t' > 0$, must also belong to Π . This process does not have a derivative at t .

5. Motions in a Frame-Space

Let $(\mathcal{T}, \mathcal{G}, \Pi)$ be a body element and let \mathcal{V} be a real inner product space having the same dimension as \mathcal{T} . An element $K \in \text{InvlIn}(\mathcal{T}, \mathcal{V})$ will be called a **placement** of the element \mathcal{T} in the *frame-space* \mathcal{V} if $K^*K \in \mathcal{G}$. The set of all placements of \mathcal{T} is denoted by

$$\mathcal{P}_{\mathcal{V}} = \{K \in \text{InvlIn}(\mathcal{T}, \mathcal{V}) \mid K^*K \in \mathcal{G}\}. \quad (5.1)$$

We call K^*K the *configuration determined by the placement* K . It is easily seen that two placements $K_1, K_2 \in \mathcal{P}_{\mathcal{V}}$ determine the same configuration if and only if $K_2 K_1^{-1} \in \text{Orth}(\mathcal{V})$ and that

$$\mathcal{G} = \{K^*K \mid K \in \mathcal{P}_{\mathcal{V}}\}. \quad (5.2)$$

A process $M: [0, d_M] \rightarrow \mathcal{P}_{\mathcal{V}}$ with values in $\mathcal{P}_{\mathcal{V}}$ will be called a **motion** of \mathcal{T} in the *frame-space* \mathcal{V} if $M^*M \in \Pi$. We call M^*M the *deformation process determined by the motion* M .

The notions of *duration, initial and final placements, freezes, segments, and continuation* apply also to motions since they apply to processes of any type. Two motions M_1 and M_2 determine the same deformation process if and only if they have the same duration and if $M_2 M_1^{-1}$ has values in $\text{Orth}(\mathcal{V})$.

It is often useful to fix a particular *reference placement* $K_R \in \mathcal{P}_{\mathcal{V}}$ and to describe the possible motions of the body-element with reference to K_R . In this way one obtains all the standard tensor functions connected with a given motion M (see [TN], Chapter CI). For example,

$$F = M K_R^{-1}: [0, d_M] \rightarrow \text{Lin}(\mathcal{V}) \quad (5.3)$$

defines the *displacement tensor*¹ of M . If $P = M^*M$ is the deformation process determined by M , then

$$C = K_R^{*-1} P K_R^{-1} = F^T F: [0, d_M] \rightarrow \text{Lin}(\mathcal{V}) \quad (5.4)$$

defines the *right Cauchy-Green tensor*. If $t \in [0, d_M]$ and if we take $M(t)$, rather than K_R , as reference placement, we obtain the *relative displacement tensor*

$$F_{(t)} = M M(t)^{-1}: [0, d_M] \rightarrow \text{Lin}(\mathcal{V}) \quad (5.5)$$

and the *relative right Cauchy-Green tensor*

$$C_{(t)} = M(t)^{* -1} G M(t)^{-1} = F_{(t)}^T F_{(t)}: [0, d_M] \rightarrow \text{Lin}(\mathcal{V}). \quad (5.6)$$

If the deformation process $P = M^*M$ is of class C^k , we can define the k^{th} *Rivlin-Ericksen tensor* by

$$A_k = M^{*-1} \overset{(k)}{P} M^{-1}: [0, d_M] \rightarrow \text{Lin}(\mathcal{V}). \quad (5.7)$$

6. Intrinsic Stress, Work

It follows from the basic principles of continuum mechanics that, roughly speaking, the contact forces exerted on a body element \mathcal{T} in a placement $K \in \mathcal{P}_{\mathcal{V}}$

¹ In [TN] the term “deformation gradient” is used. For the present context, I prefer the term I used in [N1], p. 14.

can be described by a (Cauchy-) *stress tensor* $T \in \text{Sym}(\mathcal{V})$ (see Section 16 of [TN]). We define a corresponding **intrinsic stress** $S \in \text{Sym}(\mathcal{T}^*, \mathcal{T})$ by

$$S = K^{-1} T K^{*-1}. \quad (6.1)$$

It describes the contact forces acting on the element \mathcal{T} intrinsically, without reference to any frame space.

A *mechanical process* for a body element \mathcal{T} is defined to be a pair (P, S) , where $P \in \Pi$ is a deformation process for \mathcal{T} and S is a process with values in $\text{Sym}(\mathcal{T}^*, \mathcal{T})$ and of the same duration as P , i.e. $S: [0, d_P] \rightarrow \text{Sym}(\mathcal{T}^*, \mathcal{T})$. The values of S are to be interpreted as intrinsic stresses. If P is determined by a motion M of \mathcal{T} in the frame space \mathcal{V} , then

$$T = M S M^*: [0, d_P] \rightarrow \text{Sym}(\mathcal{V}) \quad (6.2)$$

gives the Cauchy-stress corresponding to the intrinsic stress S as a function on $[0, d_P]$.

If S is continuous and P continuously differentiable, then $S\dot{P}: [0, d_P] \rightarrow \text{Lin}(\mathcal{T})$ is meaningful and one can easily show (see (5.7)) that

$$\frac{1}{2} \text{tr}(T A_1) = \frac{1}{2} \text{tr}(S\dot{P}): [0, d_P] \rightarrow \mathbb{R} \quad (6.3)$$

is the *stress power*. Its value at t measures the rate at which the contact forces exerted on the element do work, per unit volume in the configuration $P(t)$. The total work done on the element during the mechanical process (P, S) , per unit volume in the initial configuration $P^i = P(0)$, is given by

$$w = \frac{1}{2} \int_0^{d_P} \text{tr}(S\dot{P}) j, \quad (6.4)$$

where $j: [0, t_P] \rightarrow \mathbb{R}^+$ is determined by

$$j^2 = \det((P^i)^{-1} P). \quad (6.5)$$

The meaning of $j(t)$ is that of the ratio of the volume of the configuration $P(t)$ to the initial volume.

7. Material Elements

We are now ready to introduce the main concept of the present paper, the concept of a material element. A material element will be a body element endowed with additional structure that is designed to describe physical material properties. For definiteness we confine ourselves to mechanical material properties, although it is not difficult to include non-mechanical physical phenomena (see Section 21). The principal feature of our definition is that it associates with the element a space of possible physical *states*. The underlying intuitive idea is that when a material element is given in a concrete physical situation, it is given in a definite state. The state determines everything about the element: its configuration, its stress, and, most importantly, the response of the element in every possible test. We imagine that a test consists of subjecting the element to a deformation process and measuring the stress at the end.

Definition 7.1. A material element is a septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ which satisfies Axioms I–VI stated below. The nature of the objects of the septuple is as follows:

- (a) $(\mathcal{T}, \mathcal{G}, \Pi)$ is a body element as in Definition 4.1. It is called the **underlying body element** of the material element.
- (b) Σ is a set, called the **state space** of the material element.
- (c) \hat{G} is a mapping $\hat{G}: \Sigma \rightarrow \mathcal{G}$ from the state space Σ into the configuration space \mathcal{G} .
- (d) \hat{S} is a mapping $\hat{S}: \Sigma \rightarrow \mathcal{S}$ from the state space Σ into the space

$$\mathcal{S} = \text{Sym}(\mathcal{T}^*, \mathcal{T}), \quad (7.1)$$

which we call the **stress space** of the element.

- (e) $\hat{\rho}$ is a mapping $\hat{\rho}: (\Sigma \times \Pi)_{\text{fit}} \rightarrow \Sigma$, whose domain

$$(\Sigma \times \Pi)_{\text{fit}} = \{(\sigma, P) \mid \sigma \in \Sigma, P \in \Pi, P^i = \hat{G}(\sigma)\} \quad (7.2)$$

is the set of all state-process pairs such that the state “fits” the initial configuration of the process.

When we speak about a material element \mathcal{T} , we consider it understood that \mathcal{T} is endowed with a structure defined by the prescription of $\mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}$, and $\hat{\rho}$. The elements σ of Σ are the possible *states* of \mathcal{T} . The value $\hat{G}(\sigma)$ is the *configuration* and $\hat{S}(\sigma)$ is the (intrinsic) *stress* determined by the state σ . The function $\hat{\rho}$ will be called the **evolution function** of \mathcal{T} . The value $\hat{\rho}(\sigma, P)$ is the *state reached* by the element if, starting from the state σ , it is subjected to the deformation process P .

In addition to (7.1) and (7.2), we shall use the following notations throughout the rest of the paper:

- (i) The set of all deformation processes beginning at the configuration $G \in \mathcal{G}$ is denoted by

$$\Pi_G = \{P \in \Pi \mid P^i = G\}. \quad (7.3)$$

- (ii) The set of all states that fit the configuration $G \in \mathcal{G}$ is denoted by

$$\Sigma_G = \{\sigma \in \Sigma \mid \hat{G}(\sigma) = G\} \quad (7.4)$$

and called the **G -section** of Σ .

- (iii) The **response functional** \tilde{S} of the element is defined by

$$\tilde{S} = \hat{S} \circ \hat{\rho}: (\Sigma \times \Pi)_{\text{fit}} \rightarrow \mathcal{S}. \quad (7.5)$$

If $\sigma \in \Sigma$ and $P \in \Pi_{\hat{G}(\sigma)}$, then $\tilde{S}(\sigma, P)$ is the *stress produced* by the process P when the initial state was σ .

It is evident that Π is the disjoint union of the Π_G , $G \in \mathcal{G}$, that Σ is the disjoint union of the sections Σ_G , $G \in \mathcal{G}$, and that $(\Sigma \times \Pi)_{\text{fit}}$ is the disjoint union of the Cartesian products $\Sigma_G \times \Pi_G$, $G \in \mathcal{G}$.

The first two axioms flow directly from the intuitive meaning of the evolution function $\hat{\rho}$. One states that the state reached after a process must fit the final configuration. The other expresses the fact that when the element is first subjected

to a process P_1 and then to a process P_2 , it must reach the same state as when it is subjected to the continuation $P_1 * P_2$ of P_1 by P_2 .

Axiom I. For all $(\sigma, P) \in (\Sigma \times \Pi)_{\text{fit}}$,

$$\hat{\rho}(\sigma, P) \in \Sigma_{P^f}, \quad \text{i.e.,} \quad \hat{G}(\hat{\rho}(\sigma, P)) = P^f. \quad (7.6)$$

It follows from Axiom I that $\hat{G}: \Sigma \rightarrow \mathcal{G}$ must be surjective, because, by condition (P4) of Definition 4.1, for any $\sigma \in \Sigma$ and any $G \in \mathcal{G}$, there must be a $P \in \Pi$ such that $(\sigma, P) \in (\Sigma \times \Pi)_{\text{fit}}$ and $P^f = G$.

Axiom II. If $P_1, P_2 \in \Pi$, $\sigma \in \Sigma_{P_1^i}$ and $P_1^f = P_2^i$, then

$$\hat{\rho}(\sigma, P_1 * P_2) = \hat{\rho}(\hat{\rho}(\sigma, P_1), P_2). \quad (7.7)$$

An immediate consequence of (7.7) is

$$\tilde{S}(\sigma, P_1 * P_2) = \tilde{S}(\hat{\rho}(\sigma, P_1), P_2). \quad (7.8)$$

The next axiom is basic. It expresses the assumption that there must be some operational way to distinguish between states. Specifically, if two states are different but fit the same configuration, there must be some process which produces different stresses with the two states as initial states.

Axiom III. For any $G \in \mathcal{G}$, if $\sigma_1, \sigma_2 \in \Sigma_G$ and

$$\tilde{S}(\sigma_1, P) = \tilde{S}(\sigma_2, P)$$

for all $P \in \Pi_G$, then $\sigma_1 = \sigma_2$.

A simple application of Axiom III gives the following result, which states that processes of zero duration leave states unchanged.

Proposition 7.1. For all $\sigma \in \Sigma$ we have

$$\hat{\rho}(\sigma, \hat{G}(\sigma)_{(0)}) = \sigma, \quad \tilde{S}(\sigma, \hat{G}(\sigma)_{(0)}) = \hat{S}(\sigma). \quad (7.9)$$

Proof. Given any $P \in \Pi_{\hat{G}(\sigma)}$, we apply (7.8) with the choice $P_1 = \hat{G}(\sigma)_{(0)} = P_{(0)}^i$ and $P_2 = P$. In view of (4.7), we get

$$\tilde{S}(\sigma, P) = \tilde{S}(\hat{\rho}(\sigma, \hat{G}(\sigma)_{(0)}), P).$$

Since P is arbitrary, Axiom III yields the desired result (7.9)₁, and (7.9)₂ is an immediate consequence. Q.E.D.

There will be three more axioms. They all involve a certain topology on the state space Σ , a topology which we will define in Section 11. Axiom IV of Section 11 requires that the state space Σ should be in a certain sense complete. Axiom V, to be stated in Section 12, postulates the existence of "relaxed states". Axiom VI, to be stated in Section 13, requires that all states must be "accessible" in an appropriate sense.

In the next section we shall describe how our conceptual framework applies even when the material is subject to "internal constraints". Sections 9 and 10 deal with material isomorphisms and material symmetry, for which Axioms IV–VI are not needed.

8. Internal Constraints

We recall (Section 4) that in order to define the structure of a body element \mathcal{T} , one must prescribe the set \mathcal{G} of its possible configurations, and \mathcal{G} must be a closed connected subset of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$. If every neighborhood in \mathcal{G} of every $G \in \mathcal{G}$ has a non-empty interior in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$, we say that the element is *free of internal constraints*.

On the other hand, for a given configuration G , it may happen that some neighborhood of G in \mathcal{G} is contained in a proper submanifold \mathcal{C}_G of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$. If this is the case, we say that the element is subject to an *internal constraint* at G with *constraint manifold* \mathcal{C}_G . For example, if \mathcal{G} is contained in a proper submanifold \mathcal{C} of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$, then \mathcal{C} is a constraint manifold for every $G \in \mathcal{G}$.

The most important internal constraint is *incompressibility*. It is defined by the condition that any two possible configurations $G_1, G_2 \in \mathcal{G}$ must have the same volume. Although it is easy to give a precise mathematical meaning to the notion of “volume of a configuration”,¹ we shall not do so here but merely remark that the square of the ratio of the volume of G_2 to the volume of G_1 is given by $\det(G_1^{-1} G_2)$ (cf. [TN], (23.11)). We say that \mathcal{T} is an **incompressible element** if for any $G_1, G_2 \in \mathcal{G}$ we have $\det G_1^{-1} G_2 = 1$. This means that \mathcal{G} is contained in

$$\mathcal{C} = \{G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*) \mid \det(G_0^{-1} G) = 1\}, \tag{8.1}$$

where $G_0 \in \mathcal{G}$ is arbitrary. Thus, for an incompressible element, (8.1) is a constraint manifold for all $G \in \mathcal{G}$.

The severest possible constraint is *rigidity*. A rigid body element has only one possible configuration, i.e. \mathcal{G} reduces to a singleton $\mathcal{G} = \{G_0\}$. The only possible deformation processes in Π are the freezes at G_0 .

Another internal constraint is *inextensibility* in some direction. We say that \mathcal{T} is *inextensible* in the direction of $v \in \mathcal{T}$ if for any $G_1, G_2 \in \mathcal{G}$ we have $\langle G_1 v, v \rangle = \langle G_2 v, v \rangle$. This means that \mathcal{G} is contained in

$$\mathcal{C} = \{G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*) \mid \langle G v, v \rangle = c\}, \tag{8.2}$$

where $c > 0$. For an inextensible element, (8.2) is a constraint manifold for all $G \in \mathcal{G}$.

The definition of a material element given in the previous sections applies whether or not internal constraints are present. However, if \mathcal{T} is subject to internal constraints, the values of $\hat{S}: \Sigma \rightarrow \mathcal{S}$ must be reinterpreted as being *intrinsic extra stresses* (cf. [TN], p. 71) rather than intrinsic stresses. The formulas (6.3) and (6.4) remain valid if \hat{S} is so reinterpreted. The indeterminacy of the extra stress must be removed by normalization. Such normalization is somewhat arbitrary, but the following condition seems to be the simplest. We note that if \mathcal{C} is a submanifold of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$, then the tangent spaces of \mathcal{C} are subspaces of $\text{Sym}(\mathcal{T}, \mathcal{T}^*)$. Also, if $S \in \text{Sym}(\mathcal{T}^*, \mathcal{T})$ and $G \in \mathcal{G}$, then $GSG \in \text{Sym}(\mathcal{T}, \mathcal{T}^*)$.

Normalization Condition. *If \mathcal{T} is subject to an internal constraint at $G \in \mathcal{G}$ with constraint manifold \mathcal{C}_G , then, for every $\sigma \in \Sigma_G$, $G\hat{S}(\sigma)G$ must belong to the tangent space of \mathcal{C}_G at G .*

¹ As an element of the space $\overset{n}{\wedge} \mathcal{T}^*$ of n -covectors, with $n = \dim \mathcal{T}$.

One can show that the tangent space of the manifold (8.1) at $G \in \mathcal{G}$ consists of all $D \in \text{Sym}(\mathcal{T}, \mathcal{T}^*)$ such that $\text{tr}(G^{-1}D) = 0$. Hence, for incompressible body elements, the Normalization Condition reduces to

$$\text{tr}(\hat{S}(\sigma)G) = 0 \quad \text{for all } G \in \mathcal{G} \text{ and all } \sigma \in \Sigma_G. \tag{8.3}$$

For a rigid material element, the constraint manifold is the zero-dimensional manifold $\mathcal{G} = \mathcal{G} = \{G_0\}$. Its tangent-space at G_0 is the zero-dimensional space whose only member is the zero element of $\text{Sym}(\mathcal{T}, \mathcal{T}^*)$. Hence the Normalization Condition gives $\hat{S}(\sigma) = 0$ for all $\sigma \in \Sigma$. Using Axiom III, one can see immediately that the state space Σ must also reduce to a singleton $\Sigma = \{\sigma_0\}$, and the entire material element structure becomes trivial.

9. Material Isomorphisms, Materially Uniform Bodies

Isomorphisms between body elements (Definition 4.1) and between material elements (Definition 7.1) will be defined in accordance with the general notion of isomorphisms for arbitrary mathematical structures.

We note that the set $\text{Invl}(\mathcal{T}_1, \mathcal{T}_2)$ is the set of vector-space isomorphisms from \mathcal{T}_1 onto \mathcal{T}_2 . Each $A \in \text{Invl}(\mathcal{T}_1, \mathcal{T}_2)$ induces an isomorphism $G \mapsto A^*GA$ from $\text{Sym}(\mathcal{T}_2, \mathcal{T}_2^*)$ onto $\text{Sym}(\mathcal{T}_1, \mathcal{T}_1^*)$ and an isomorphism $S \mapsto ASA^*$ from $\text{Sym}(\mathcal{T}_1^*, \mathcal{T}_1)$ onto $\text{Sym}(\mathcal{T}_2^*, \mathcal{T}_2)$.

Definition 9.1. Let $(\mathcal{T}_1, \mathcal{G}_1, \Pi_1)$ and $(\mathcal{T}_2, \mathcal{G}_2, \Pi_2)$ be body elements. A mapping $A: \mathcal{T}_1 \rightarrow \mathcal{T}_2$ is called an isomorphism between the body elements \mathcal{T}_1 and \mathcal{T}_2 if (i) $A \in \text{Invl}(\mathcal{T}_1, \mathcal{T}_2)$, (ii) $A^*\mathcal{G}_2A = \mathcal{G}_1$, and (iii) $A^*\Pi_2A = \Pi_1$.

Definition 9.2. Let $(\mathcal{T}_1, \mathcal{G}_1, \Pi_1, \Sigma_1, \hat{G}_1, \hat{S}_1, \hat{\rho}_1)$ and $(\mathcal{T}_2, \mathcal{G}_2, \Pi_2, \Sigma_2, \hat{G}_2, \hat{S}_2, \hat{\rho}_2)$ be material elements. A mapping $A: \mathcal{T}_1 \rightarrow \mathcal{T}_2$ is called a **material isomorphism** between the material elements \mathcal{T}_1 and \mathcal{T}_2 if A is an isomorphism between the underlying body elements of \mathcal{T}_1 and \mathcal{T}_2 and if there exists a bijection

$$\iota_A: \Sigma_1 \rightarrow \Sigma_2 \tag{9.1}$$

with the following properties:

(a) for all $\sigma \in \Sigma_1$,

$$\hat{G}_1(\sigma) = A^* \hat{G}_2(\iota_A(\sigma)) A, \tag{9.2}$$

(b) for all $\sigma \in \Sigma_1$,

$$\hat{S}_2(\iota_A(\sigma)) = A \hat{S}_1(\sigma) A^*, \tag{9.3}$$

(c) for all $\sigma \in \Sigma_1$ and all $P \in \Pi_2$ such that $\hat{G}_1(\sigma) = A^* P^i A$,

$$\iota_A(\hat{\rho}_1(\sigma, A^* P A)) = \hat{\rho}_2^*(\iota_A(\sigma), P). \tag{9.4}$$

Proposition 9.1. The bijection ι_A of (9.1)–(9.4) is uniquely determined by the material isomorphism A .

Proof. Application of (9.3) and (9.4) shows that

$$A \hat{S}_1(\hat{\rho}_1(\sigma, A^* P A)) A^* = \hat{S}_2(\hat{\rho}_2(\iota_A(\sigma), P)) \tag{9.5}$$

for all $\sigma \in \Sigma_1$ and all $P \in \Pi_2$ with $\hat{G}_1(\sigma) = A^* P^i A$. Hence, if ι_A and ι'_A are both bijections satisfying (9.1)–(9.4), we have

$$\hat{S}_2(\hat{\rho}_2(\iota_A(\sigma), P)) = \hat{S}_2(\hat{\rho}_2(\iota'_A(\sigma), P))$$

for all $P \in \Pi_2$ with $P^i = A^{*-1} \hat{G}_1(\sigma) A^{-1} = \hat{G}_2(\iota_A(\sigma)) = \hat{G}_2(\iota'_A(\sigma))$. It follows from Axiom III that $\iota_A(\sigma) = \iota'_A(\sigma)$. Since $\sigma \in \Sigma_1$ is arbitrary, we obtain $\iota_A = \iota'_A$. Q.E.D.

If A is a material isomorphism, so is its inverse A^{-1} , and the corresponding bijections ι_A and $\iota_{A^{-1}}$ are related by

$$\iota_{A^{-1}} = (\iota_A)^{-1}. \tag{9.6}$$

If A_1 and A_2 are material isomorphisms and if their composition $A_2 A_1$ makes sense, it is again a material isomorphism. The bijections ι_{A_1} , ι_{A_2} , and $\iota_{A_2 A_1}$ are related by

$$\iota_{A_2 A_1} = \iota_{A_2} \circ \iota_{A_1}. \tag{9.7}$$

If two material elements \mathcal{T}_1 and \mathcal{T}_2 are materially isomorphic we also say that they *consist of the same material*. We give a precise meaning to the term “material” by saying that a **material** is an equivalence class of material elements, the equivalence being material isomorphy.

Suppose that we have defined, on one and the same vector space \mathcal{T} , two material element structures in such a way that the identity $1_{\mathcal{T}}$ is a material isomorphism between them. By Definition 9.1, the sets \mathcal{G} and Π must be the same for both structures. Although the state spaces need not be the same *a priori*, we shall always employ the bijection $\iota_{1_{\mathcal{T}}}$ to identify one with the other. The functions \hat{G} , \hat{S} , and $\hat{\rho}$ then become the same for both structures, *i.e.* the two structures become the same. Thus, we do not distinguish between material element structures on \mathcal{T} that are isomorphic via the identity mapping of \mathcal{T} .

Remark. In Section 3 we have indicated how a continuous body \mathcal{B} can be defined and how a body element \mathcal{T}_X can be associated with each of the material points $X \in \mathcal{B}$. We say that \mathcal{B} has the structure of a **material body** if each of its body elements \mathcal{T}_X has been endowed with the structure of a material element in accordance with Definition 7.1. We say that a material body is **materially uniform** if any two of its material elements consist of the same material. (This definition is consistent with the one I gave in [TN], p. 59 and [N7], p. 7.)

Let \mathcal{B} be a materially uniform material body, and let \mathcal{A} be the set of all material isomorphisms between material elements of \mathcal{B} . It is easily seen that this set \mathcal{A} has the following properties:

- (I1) Every member of \mathcal{A} is a vector-space isomorphism between two (not necessarily distinct) tangent spaces of \mathcal{B} .
- (I2) If $A \in \mathcal{A}$, then $A^{-1} \in \mathcal{A}$.
- (I3) If $A_1, A_2 \in \mathcal{A}$ and $\text{Domain } A_2 = \text{Range } A_1$, then $A_2 \circ A_1 \in \mathcal{A}$.
- (I4) For any $X, Y \in \mathcal{B}$, there is at least one $A \in \mathcal{A}$ such that $A: \mathcal{T}_X \rightarrow \mathcal{T}_Y$.

The set \mathcal{A} endows the body-manifold \mathcal{B} with a certain “geometrical” structure. This structure is closely related to what, in the literature on abstract differential geometry, is called a *G*-structure (see [S], Chap. VII). The theory of

inhomogeneous but materially uniform bodies as developed by me in [N7] and C.-C. WANG in [W2] is essentially nothing more than a theory of the geometrical structure on a manifold \mathcal{B} defined by a set \mathcal{A} with the properties (I1)–(I4). For example, a *material connection* Γ on \mathcal{B} is defined to be an affine connection on \mathcal{B} such that every parallel transport with respect to Γ belongs to \mathcal{A} . (This concept of material connection is the one introduced by C.-C. WANG in [W2], p. 66. It is a more inclusive concept than the original concept introduced by me in [TN], p. 90 and [N7], p. 19.)

10. Material Symmetry, Isotropy

Let \mathcal{T} be a material element. A material isomorphism (see Definition 9.2) between \mathcal{T} and \mathcal{T} itself is called a *material automorphism* or a **symmetry** of \mathcal{T} . The set

$$\mathcal{g} = \text{Aut}(\mathcal{T}) \subset \text{InvlIn}(\mathcal{T}) \quad (10.1)$$

of all material automorphisms is obviously a subgroup of the general linear group $\text{InvlIn}(\mathcal{T})$. We call this group \mathcal{g} the **symmetry group** of material element \mathcal{T} .¹

The following result is an immediate corollary of Prop. 9.1 and (9.6), (9.7):

Proposition 10.1. *There is a unique homomorphism*

$$\iota: \mathcal{g} \rightarrow \text{Perm}(\Sigma) \quad (10.2)$$

from the symmetry group \mathcal{g} into the group $\text{Perm}(\Sigma)$ of all permutations of the state space Σ such that for every $A \in \mathcal{g}$ the value ι_A of ι at A satisfies the following conditions:

(i) for all $\sigma \in \Sigma$

$$\hat{G}(\sigma) = A^* \hat{G}(\iota_A(\sigma)) A, \quad (10.3)$$

and

$$\hat{S}(\iota_A(\sigma)) = A \hat{S}(\sigma) A^*, \quad (10.4)$$

(ii) for all $\sigma \in \Sigma$ and all $P \in \Pi$ such that $\hat{G}(\sigma) = A^* P^i A$,

$$\iota_A(\hat{\rho}(\sigma, A^* P A)) = \hat{\rho}(\iota_A(\sigma), P). \quad (10.5)$$

This proposition asserts that the symmetry group \mathcal{g} acts as a transformation group on the state space Σ . Therefore, we can define the *orbit* Ω_σ of $\sigma \in \Sigma$ under the action of \mathcal{g} by

$$\Omega_\sigma = \{\iota_A(\sigma) \mid A \in \mathcal{g}\}. \quad (10.6)$$

The orbits under the action of \mathcal{g} form a partition of the state space Σ . We call these orbits the **reduced states** of the given material element. Every state determines a unique reduced state. Two states that determine the same reduced state are physically equivalent in an intuitively obvious sense.

Let $\sigma \in \Sigma$ be any state. The set

$$\mathcal{g}_\sigma = \{A \in \mathcal{g} \mid \iota_A(\sigma) = \sigma\} \quad (10.7)$$

¹ In [N2], Sect. 19, and in later work (notably [TN]) I used the term “isotropy group”. Some of my colleagues convinced me that “symmetry group” is preferable.

of all symmetries that leave the state σ invariant will be called the **symmetry group of the state** σ . Of course, \mathcal{g}_σ is a subgroup of \mathcal{g} .

Proposition 10.2. *The symmetry group \mathcal{g}_σ of a state σ is a subgroup of the orthogonal group of $\widehat{G}(\sigma)$ and hence*

$$\mathcal{g}_\sigma \subset \mathcal{g} \cap \text{Orth}(\widehat{G}(\sigma)). \quad (10.8)$$

Proof. It follows from (10.3) and (10.7) that $\widehat{G}(\sigma) = A^* \widehat{G}(\sigma) A$ for all $A \in \mathcal{g}_\sigma$. This means (see Section 2) that $A \in \text{Orth}(\widehat{G}(\sigma))$. Q.E.D.

We say that a state σ is **isotropic** if

$$\mathcal{g}_\sigma = \text{Orth}(\widehat{G}(\sigma)). \quad (10.9)$$

We say that \mathcal{T} is an *isotropic material element* if it has isotropic states.

Proposition 10.3. *The symmetry group \mathcal{g} of an isotropic material element \mathcal{T} is such that either $\mathcal{g} = \mathcal{g}_\sigma$ if and only if σ is isotropic or else $\mathcal{g} \supset \text{Unim}(\mathcal{T})$.*

Proof. By the definition (10.9), σ is isotropic if and only if $\mathcal{g}_\sigma = \text{Orth}(\widehat{G}(\sigma))$. Since $\mathcal{g} \supset \mathcal{g}_\sigma$ we can apply the theorem on the maximality of an orthogonal group in the unimodular group (see Section 2) and conclude that we must have either $\mathcal{g} = \text{Orth} \widehat{G}(\sigma) = \mathcal{g}_\sigma$ or $\mathcal{g} \supset \text{Unim}(\mathcal{T})$. Since these are mutually exclusive possibilities, the assertion follows. Q.E.D.

Proposition 10.4. *The stress $\widehat{S}(\sigma)$ determined by a state σ is such that*

$$A \widehat{S}(\sigma) \widehat{G}(\sigma) = \widehat{S}(\sigma) \widehat{G}(\sigma) A \quad \text{for all } A \in \mathcal{g}_\sigma. \quad (10.10)$$

In particular, if σ is an isotropic state, then

$$\widehat{S}(\sigma) = -p \widehat{G}(\sigma)^{-1}, \quad (10.11)$$

where $p \in \mathbb{R}$.

Proof. (10.10) is an immediate consequence of (10.4), (10.7), and (10.3). If σ is isotropic, *i.e.* $\mathcal{g}_\sigma = \text{Orth}(\widehat{G}(\sigma))$, then (10.10) states that $\widehat{S}(\sigma) \widehat{G}(\sigma)$ commutes with all $A \in \text{Orth}(\widehat{G}(\sigma))$. It is well known that this can happen only if $\widehat{S}(\sigma) \widehat{G}(\sigma)$ is a multiple of the identity $1_{\mathcal{T}}$, *i.e.* if (10.11) holds. Q.E.D.

The number p of (10.11) is called the *pressure* corresponding to the isotropic state σ . For incompressible material elements, it follows from Prop. 10.4 and the normalization (8.3) that the extra-stress is zero for all isotropic states.

We recall that two states σ, σ' determine the same reduced state if one is in the orbit of the other under the action of \mathcal{g} , *i.e.* if $\sigma' = \iota_A(\sigma)$ for some $A \in \mathcal{g}$. The following proposition shows that if two states determine the same reduced state, then the symmetry groups of the two states are conjugate in \mathcal{g} and hence isomorphic.

Proposition 10.5. *If $\sigma \in \Sigma$ and $A \in \mathcal{g}$, then*

$$\mathcal{g}_{\iota_A(\sigma)} = A \mathcal{g}_\sigma A^{-1}. \quad (10.12)$$

Proof. In view of (9.6), (9.7), and the definition (10.7), we have $B \in \mathcal{g}_\sigma$ if and only if

$$\iota_{ABA^{-1}}(\iota_A(\sigma)) = (\iota_A \circ \iota_B \circ \iota_{A^{-1}} \circ \iota_A)(\sigma) = \iota_A(\iota_B(\sigma)) = \iota_A(\sigma),$$

which is valid if and only if $ABA^{-1} \in \mathcal{g}_{\iota_A(\sigma)}$. Q.E.D.

We say that a state σ is a **state of maximal symmetry** if g_σ is maximal in the class of all symmetry groups of states, *i.e.* if $g_\sigma \subset g_\tau$ is possible for $\tau \in \Sigma$ only if $g_\sigma = g_\tau$.

Proposition 10.6. *For an isotropic material element, a state is of maximal symmetry if and only if it is isotropic.*

Proof. In view of Prop. 10.3 it is trivial that isotropic states are always of maximal symmetry.

On the other hand, let σ be a state of maximal symmetry. By Prop. 10.3, we have two possibilities for g :

(a) $g = g_\tau$, with τ isotropic. In this case, (10.8) becomes

$$g_\sigma \subset g_\tau \cap \text{Orth}(\hat{G}(\sigma)) \subset g_\tau.$$

Since g_σ is maximal, the inclusions must reduce to equalities, which is possible only when $g_\sigma = g_\tau = g$. By Prop. 10.3, σ must be isotropic.

(b) $g \supset \text{Unim}(\mathcal{T})$. Let τ be an isotropic state. We can always find a $A \in \text{Unim}(\mathcal{T})$ such that $\hat{G}(\sigma) = cA^* \hat{G}(\tau) A$, with $c > 0$ (see Section 2). Since $A \in g$, we have, by (10.3), $\hat{G}(\sigma) = c \hat{G}(\tau')$, where $\tau' = l_A(\tau)$ is another isotropic state (see Prop. 10.5). It follows that $g_\tau = \text{Orth}(\hat{G}(\tau')) = \text{Orth}(\hat{G}(\sigma))$. Using (10.8), we obtain

$$g_\sigma \subset \text{Orth}(\hat{G}(\sigma)) = \text{Orth}(\hat{G}(\tau')) = g_\tau.$$

Since g_σ is maximal, the inclusion must reduce to equality, so that $g_\sigma = \text{Orth}(\hat{G}(\sigma))$. This means that σ is isotropic. Q.E.D.

It is clear from Prop. 10.5 that if a state σ is of maximal symmetry or is isotropic, so are all states that determine the same reduced state as σ . Thus, it is meaningful to speak about *reduced states of maximal symmetry* and *isotropic reduced states*.

The conditions (10.3)–(10.5) are obviously satisfied for $A = 1_{\mathcal{T}}$ and $A = -1_{\mathcal{T}}$ if we put

$$l_{1_{\mathcal{T}}} = l_{(-1_{\mathcal{T}})} = 1_{\Sigma}. \quad (10.13)$$

Hence the symmetry groups g_σ , $\sigma \in \Sigma$, and g always contain the two-element group $\{1_{\mathcal{T}}, -1_{\mathcal{T}}\}$.

11. Uniformity and Topology in the State Space

We now introduce natural uniformities on the sections of the state space of a material element and a natural topology on the whole state space. For the explanation of the concepts occurring in the following definition we refer to [BK], Chapt. II, §2, No. 3, and Chapt. I, §2, No. 4, or to [FT], Chapt. 24. Recall that the stress space $\mathcal{S} = \text{Sym}(\mathcal{T}^*, \mathcal{T})$ is a finite-dimensional real vector space and hence has a natural uniformity.

Definition 11.1. *Let \mathcal{T} be a material element with response functional \tilde{S} (see (7.5)). For every $G \in \mathcal{G}$, we call **natural uniformity** of the G -section Σ_G (see (7.4)) the coarsest uniformity on Σ_G which renders the mappings*

$$\tilde{S}(\cdot, P): \Sigma_G \rightarrow \mathcal{S}$$

uniformly continuous for all $P \in \Pi_G$ ((see 7.3)).¹ We call **natural topology** of the state space $\Sigma = \bigcup \{ \Sigma_G \mid G \in \mathcal{G} \}$ the sum of the topologies on the sections Σ_G induced by the natural uniformities of these sections.

The following result, a special case of [BK], Prop. 4 of §2, Chap. II, characterizes the natural uniformity of Σ_G .

Proposition 11.1. *A mapping $\varphi: \mathcal{U} \rightarrow \Sigma_G$ from a uniform space into the G -section Σ_G is uniformly continuous if and only if the mappings*

$$\tilde{S}(\varphi(\cdot), P): \mathcal{U} \rightarrow \mathcal{S}$$

are uniformly continuous for all $P \in \Pi_G$.

This proposition is used to prove the following result.

Proposition 11.2. *For all $P \in \Pi$, the mapping*

$$\hat{\rho}(\cdot, P): \Sigma_{P_i} \rightarrow \Sigma_{P_f}$$

is uniformly continuous.

Proof. It follows from (7.8) that

$$\tilde{S}(\hat{\rho}(\sigma, P), P') = \tilde{S}(\sigma, P * P')$$

for all $P' \in \Pi_{P_f}$. Since $\tilde{S}(\cdot, P * P'): \Sigma_{P_i} \rightarrow \mathcal{S}$ is uniformly continuous by definition of the natural uniformity on Σ_{P_i} , it follows that $\tilde{S}(\hat{\rho}(\cdot, P), P'): \Sigma_{P_i} \rightarrow \mathcal{S}$ is uniformly continuous for all $P' \in \Pi_{P_f}$. Application of Prop. 11.1 with the choice $\mathcal{U} = \Sigma_{P_i}$, $\varphi = \hat{\rho}(\cdot, P)$ yields the assertion. Q.E.D.

The following result, a special case of [BK] Prop. 10 of §8, Chap. I, characterizes the natural topology of Σ . (For the explanation of convergence of nets we refer to [K], Chapt. 2.)

Proposition 11.3. *A net $(i \mapsto \sigma_i)$ in Σ converges to $\sigma \in \Sigma$ if and only if (i) σ_i belongs eventually to some fixed section Σ_G , (ii) $\sigma \in \Sigma_G$, and (iii)*

$$\lim_i \tilde{S}(\sigma_i, P) = \tilde{S}(\sigma, P) \tag{11.1}$$

for all $P \in \Pi_G$.

Proposition 11.4. *The state space Σ , with its natural topology, is a Hausdorff space, i.e., nets in Σ cannot have more than one limit.*

Proof. If σ_1 and σ_2 are both limits of the net $(i \mapsto \sigma_i)$, we must have, by Prop. 11.3, $\hat{G}(\sigma_1) = \hat{G}(\sigma_2) = G$ and $\tilde{S}(\sigma_1, P) = \tilde{S}(\sigma_2, P)$ for all $P \in \Pi_G$. By Axiom III, Sect. 7, it follows that $\sigma_1 = \sigma_2$. Q.E.D.

The natural topology on Σ is such that all sections Σ_G , $G \in \mathcal{G}$, are both open and closed in Σ . Hence $\hat{G}: \Sigma \rightarrow \mathcal{G}$ is trivially continuous, because it is constant on each section. Also, $\hat{S}: \Sigma \rightarrow \mathcal{S}$ is continuous, because by (7.9)₂ and Definition 11.1, the restriction of \hat{S} to Σ_G is the same as $\tilde{S}(\cdot, G_{(0)})$ and hence (uniformly) continuous.

¹ This uniformity is also called the "initial uniformity" on Σ_G associated with the mappings $\tilde{S}(\cdot, P)$, $P \in \Pi_G$.

If \mathcal{T}_1 and \mathcal{T}_2 are material elements with state spaces Σ_1 and Σ_2 and $A: \mathcal{T}_1 \rightarrow \mathcal{T}_2$ a material isomorphism between \mathcal{T}_1 and \mathcal{T}_2 , then the bijection $\iota_A: \Sigma_1 \rightarrow \Sigma_2$ induced by A (see Prop. 9.1) is a homeomorphism when Σ_1 and Σ_2 are endowed with their natural topologies. This is but a special instance of the general fact that isomorphisms preserve all structures induced from the given structure. In particular, if A is a symmetry of an element \mathcal{T} , then ι_A is an auto-homeomorphism of the state space Σ of \mathcal{T} .

The sections Σ_G , with their natural uniformity described by Definition 11.1, may or may not be complete spaces. If they are not, we may form the completions $\bar{\Sigma}_G$ of the Σ_G (cf. [K], p. 196 or [BK] Chapt. II, §3) and define the *completed states space* $\bar{\Sigma}$ by $\bar{\Sigma} = \bigcup \{ \bar{\Sigma}_G \mid G \in \mathcal{G} \}$. Then Σ is dense in $\bar{\Sigma}$ and, in view of the continuity of \hat{G} and \hat{S} , we can extend the domains of these functions from Σ to $\bar{\Sigma}$, preserving continuity. Also, in view of Prop. 11.2, one can extend the domain of \hat{p} from $(\Sigma \times \Pi)_{\text{fit}}$ to $(\bar{\Sigma} \times \Pi)_{\text{fit}}$, preserving uniform continuity of $\hat{p}(\cdot, P)$ for all $P \in \Pi$. After this has been done, one obtains what is technically a new structure having $(\mathcal{T}, \mathcal{G}, \Pi)$ as the underlying body element but $\bar{\Sigma}$ instead of Σ as the state space. However, the distinction between the original structure and the new one with completed state space is physically insignificant, and artificial distinctions are avoided by means of the following postulate:

Axiom IV. *The sections Σ_G , $G \in \mathcal{G}$, with the natural uniformity described by Definition 11.1, are complete spaces.*

12. Relaxed States

In Axiom V, below, the topology of Definition 11.1 on the state space Σ of the element \mathcal{T} plays an essential role. The axiom has the following interpretation: If the element is initially in the state σ and then frozen in the corresponding configuration, its state will approach, in the limit of infinite time, a “relaxed” state.

Axiom V. *For all $\sigma \in \Sigma$, the limit*

$$\lim_{t \rightarrow \infty} \hat{p}(\sigma, \hat{G}(\sigma)_{(t)}) = \hat{\lambda}(\sigma) \tag{12.1}$$

exists (it is unique by Prop. 11.4).

The mapping $\hat{\lambda}: \Sigma \rightarrow \Sigma$ determined by (12.1) will be called the *state relaxation mapping*, and the members of its range

$$\Sigma_{\text{rel}} = \hat{\lambda}(\Sigma) \tag{12.2}$$

will be called **relaxed states**.¹

It is clear from (7.6) that

$$\hat{G}(\hat{\lambda}(\sigma)) = \hat{G}(\sigma). \tag{12.3}$$

Hence, since \hat{G} is surjective, the restriction $\hat{G}|_{\Sigma_{\text{rel}}}$ of \hat{G} to Σ_{rel} is also surjective. In other terms, $\Sigma_G \cap \Sigma_{\text{rel}}$ is non-empty for all $G \in \mathcal{G}$.

¹ This concept of a relaxed state is similar to what I have called, in a very special and now obsolete context, a “body at ease” in [N1], p. 26.

If we apply Prop. 11.3 to the net $t \mapsto \hat{\rho}(\sigma, \hat{G}(\sigma)_{(t)})$ and observe (7.8), we see that (12.1) is equivalent to

$$\lim_{t \rightarrow \infty} \tilde{S}(\sigma, G_{(t)} * P) = \tilde{S}(\hat{\lambda}(\sigma), P) \quad \text{for all } P \in \Pi_G, \quad (12.4)$$

where $G = \hat{G}(\sigma)$.

We now give several characterizations of relaxed states.

Proposition 12.1. *Let $\lambda \in \Sigma$ and $G = \hat{G}(\lambda)$. The following four conditions are equivalent:*

- (i) $\lambda \in \Sigma_{\text{rel}}$,
- (ii) for all $t \in \mathbb{R}^+$, $\hat{\rho}(\lambda, G_{(t)}) = \lambda$,
- (iii) for some $t_0 > 0$, $\hat{\rho}(\lambda, G_{(t_0)}) = \lambda$,
- (iv) $\hat{\lambda}(\lambda) = \lambda$.

Proof. First we note that, in view of (4.3) and (4.5),

$$G_{(t+r)} = G_{(r)} * G_{(t)} \quad (12.5)$$

holds for any $r, t \in \mathbb{R}^+$. For every $\sigma \in \Sigma_G$, it follows from (12.5) and (7.7) that

$$\hat{\rho}(\sigma, G_{(t+r)}) = \hat{\rho}(\hat{\rho}(\sigma, G_{(r)}), G_{(t)}). \quad (12.6)$$

We now prove (i) \Rightarrow (ii): $\lambda \in \Sigma_{\text{rel}}$ means that $\lambda = \hat{\lambda}(\sigma)$ for some $\sigma \in \Sigma$. Since (12.3) shows that $\sigma \in \Sigma_G$ we see that (12.6) is valid. Taking the limit $r \rightarrow \infty$ in (12.6) and observing Prop. 11.2 as well as Axiom V, we obtain

$$\lambda = \hat{\lambda}(\sigma) = \hat{\rho}(\hat{\lambda}(\sigma), G_{(t)}) = \hat{\rho}(\lambda, G_{(t)}),$$

i. e., condition (ii) is satisfied.

The implication (ii) \Rightarrow (iii) is trivial.

To prove (iii) \Rightarrow (iv), we assume that (iii) holds and apply (12.6) with the choice $\sigma = \lambda$, $r = t_0$, $t = (n-1)t_0$, where $n \in \mathbb{N}$:

$$\hat{\rho}(\lambda, G_{(nt_0)}) = \hat{\rho}(\hat{\rho}(\lambda, G_{(t_0)}), G_{(n-1)t_0}) = \hat{\rho}(\lambda, G_{(n-1)t_0}).$$

Since $\hat{\rho}(\lambda, G_{(t_0)}) = \lambda$ by (7.9), we see that induction over n yields $\lambda = \hat{\rho}(\lambda, G_{(nt_0)})$ for all $n \in \mathbb{N}$. Taking the limit $n \rightarrow \infty$ and using Axiom V, we get $\lambda = \hat{\lambda}(\lambda)$.

The implication (iv) \Rightarrow (i) is trivial. Q.E.D.

Prop. 12.1 has the following interpretation: Suppose a material element is initially in a state λ and then frozen in the corresponding configuration. The state will remain unchanged if and only if λ is relaxed. The material cannot return to its initial state λ in a finite time or in the limit of infinite time unless λ is relaxed.

Proposition 12.2. *Relaxation cannot decrease the symmetry group of a state, i. e., for all $\sigma \in \Sigma$, we have*

$$\mathcal{g}_\sigma \subset \mathcal{g}_{\hat{\lambda}(\sigma)}. \quad (12.7)$$

Proof. Let $G = \hat{G}(\sigma) = \hat{G}(\hat{\lambda}(\sigma))$ (see (12.3)). If $A \in \mathcal{g}_\sigma$, i. e. $\iota_A(\sigma) = \sigma$, we have, by (10.3), $G = A * G A$ and hence, for all $t \in \mathbb{R}^+$, $G_{(t)} = A * G_{(t)} A$. Writing (10.5) with

$P = G_{(t)}$, we get

$$\iota_A(\hat{\rho}(\sigma, G_{(t)})) = \hat{\rho}(\sigma, G_{(t)}).$$

Taking the limit $t \rightarrow \infty$ and observing Axiom V, we obtain $\iota_A(\hat{\lambda}(\sigma)) = \hat{\lambda}(\sigma)$ and hence $A \in \mathcal{G}_{\hat{\lambda}(\sigma)}$. Q.E.D.

The following result is a corollary to Prop. 12.2.

Proposition 12.3. *If σ is an isotropic state or a state of maximal symmetry, so is $\hat{\lambda}(\sigma)$.*

If λ is a relaxed state and A a material isomorphism, it is clear that $\iota_A(\lambda)$ is again a relaxed state. This applies, in particular, when $A \in \mathcal{G}$ is a material symmetry. It follows that if λ is relaxed, so are all states that determine the same reduced state as λ . Thus it is meaningful to speak about *relaxed reduced states*.

13. Accessibility

Let σ be a state of a material element with state space Σ . The set of all states that can be reached from σ is $\{\hat{\rho}(\sigma, P) \mid P \in \Pi_{\hat{G}(\sigma)}\}$. The closure of this set, in the natural topology of Σ described in Definition 11.1, will be denoted by

$$\Sigma_\sigma = \text{cl}\{\hat{\rho}(\sigma, P) \mid P \in \Pi_{\hat{G}(\sigma)}\}. \quad (13.1)$$

If $\tau \in \Sigma_\sigma$, we say that the state τ is **accessible** from the state σ . Thus, τ is accessible from σ if every neighborhood of τ contains states that can actually be reached from σ . We note that, by (7.9), we have $\sigma = \hat{\rho}(\sigma, \hat{G}(\sigma)_{(0)}) \in \Sigma_\sigma$, showing that every state can be reached from itself.

The following proposition states that accessibility is a transitive relation, *i. e.*, if ν is accessible from τ and τ is accessible from σ , then ν is also accessible from σ .

Proposition 13.1. *If $\tau \in \Sigma_\sigma$, then $\Sigma_\tau \subset \Sigma_\sigma$.*

Proof. First we recall that for any $\nu \in \Sigma$, $\Sigma_{\hat{G}(\nu)}$ is open in Σ and hence a neighborhood of ν . Thus, the neighborhoods of ν that are contained in $\Sigma_{\hat{G}(\nu)}$ form a neighborhood basis for ν .

Now let $\nu \in \Sigma_\tau$ and let \mathcal{T}_ν be an arbitrary neighborhood of ν with $\mathcal{T}_\nu \subset \Sigma_{\hat{G}(\nu)}$. By the definition of Σ_τ , there is a process $P_1 \in \Pi_{\hat{G}(\tau)}$ such that $\hat{\rho}(\tau, P_1) \in \mathcal{T}_\nu$. By (7.6) we have $P_1^f = \hat{G}(\nu)$. The continuity of $\hat{\rho}(\cdot, P_1): \Sigma_{\hat{G}(\tau)} \rightarrow \Sigma_{\hat{G}(\nu)}$ (see Prop. 11.2) insures that there is a neighborhood $\mathcal{T}_\tau \subset \Sigma_{\hat{G}(\tau)}$ of τ such that $\hat{\rho}(\tau', P_1) \in \mathcal{T}_\nu$ for all $\tau' \in \mathcal{T}_\tau$. The condition $\tau \in \Sigma_\sigma$ implies the existence of a process $P_2 \in \Pi_{\hat{G}(\sigma)}$ such that $\tau' = \hat{\rho}(\sigma, P_2) \in \mathcal{T}_\tau$. Hence we have $\hat{\rho}(\hat{\rho}(\sigma, P_2), P_1) \in \mathcal{T}_\nu$, *i. e.*, by (7.7),

$$\hat{\rho}(\sigma, P_2 * P_1) \in \mathcal{T}_\nu.$$

Thus we have found a state in \mathcal{T}_ν , which can be reached from σ . Since the neighborhood \mathcal{T}_ν was arbitrary, it follows that $\nu \in \Sigma_\sigma$. Q.E.D.

The next result is a direct consequence of (12.1) and Prop. 13.1:

Proposition 13.2. *For every $\sigma \in \Sigma$, if $\tau \in \Sigma_\sigma$, then $\hat{\lambda}(\tau) \in \Sigma_\sigma$, *i. e.*, if τ is accessible from σ , so is $\hat{\lambda}(\tau)$. In particular, $\hat{\lambda}(\sigma) \in \Sigma_\sigma$.*

We are now ready to state our final axiom. It postulates that every state must be accessible from some relaxed starting state λ_0 . The axiom thus limits the possible size of the state space Σ . If an axiom of this type were not imposed, there could exist two states which had no connection whatever. Physically, these states would have to be regarded as states of distinct materials.

Axiom VI. *There is at least one $\lambda_0 \in \Sigma_{\text{rel}}$ such that $\Sigma_{\lambda_0} = \Sigma$.*

If the septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ satisfies Axioms I–V but not Axiom VI, one can use the following procedure to define material element structures on \mathcal{T} : We select $\lambda_0 \in \Sigma_{\text{rel}}$ arbitrarily and put $\Sigma' = \Sigma_{\lambda_0}$. Let G' and S' be the restrictions of \hat{G} and \hat{S} to Σ' , and let $\hat{\rho}'$ be the restriction of $\hat{\rho}$ to $(\Sigma' \times \Pi_G)_{\text{fit}}$. It is not hard to verify that the septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma', \hat{G}', \hat{S}', \hat{\rho}')$ thus obtained again satisfies Axioms I–V. It obviously also satisfies Axiom VI and hence defines a structure of a material element on \mathcal{T} . This structure may depend, of course, on the particular choice of λ_0 , and hence the original septuple may give rise to not one, but many material element structures on \mathcal{T} .

The procedure just described will be illustrated by an example in Section 20.

14. Semi-Elastic and Elastic Materials

Definition 14.1. *A material element \mathcal{T} is said to be **semi-elastic** if the restriction $\hat{G}|_{\Sigma_{\text{rel}}}$ of its configuration function \hat{G} to the set Σ_{rel} of relaxed states is injective. The material element is said to be **elastic** if it is semi-elastic and if, in addition, all of its states are relaxed, so that $\Sigma = \Sigma_{\text{rel}}$.*

We have seen in Section 12 that $\hat{G}|_{\Sigma_{\text{rel}}}$ is always surjective. Hence, semi-elastic materials are characterized by the condition that $\hat{G}|_{\Sigma_{\text{rel}}}$ has an inverse

$$\bar{\lambda}: \mathcal{G} \rightarrow \Sigma_{\text{rel}}. \quad (14.1)$$

The value $\bar{\lambda}(G)$ is the relaxed state uniquely determined by the configuration G . Observing (12.3), we see that $\bar{\lambda}$ satisfies

$$\hat{\lambda}(\sigma) = \bar{\lambda}(\hat{G}(\sigma)) \quad (14.2)$$

for all $\sigma \in \Sigma$. Hence (12.1) reduces to

$$\lim_{t \rightarrow \infty} \hat{\rho}(\sigma, G_{(t)}) = \bar{\lambda}(G) \quad \text{if } \hat{G}(\sigma) = G. \quad (14.3)$$

For semi-elastic elements with a known symmetry group, one can determine the symmetry groups of its relaxed states:

Proposition 14.1. *Let \mathcal{T} be a semi-elastic element with symmetry group g . For all $\lambda \in \Sigma_{\text{rel}}$ we then have*

$$g_\lambda = g \cap \text{Orth}(\hat{G}(\lambda)), \quad (14.4)$$

where g_λ , the symmetry group of λ , is defined by (10.7).

Proof. In view of (10.8) we need only prove that $A \in g \cap \text{Orth}(\hat{G}(\lambda))$ implies $\iota_A(\lambda) = \lambda$. Now, $A \in \text{Orth}(\hat{G}(\lambda))$ means that $A^* \hat{G}(\lambda) A = \hat{G}(\lambda)$. On the other hand,

by (10.3), $A \in \mathcal{G}$ implies that $\hat{G}(\lambda) = A^* \hat{G}(\iota_A(\lambda)) A$. Thus we have $\hat{G}(\lambda) = \hat{G}(\iota_A(\lambda))$. Since λ is relaxed, so is $\iota_A(\lambda)$, and the injectivity of $\hat{G}|_{\Sigma_{\text{rel}}}$ yields $\lambda = \iota_A(\lambda)$. Q.E.D.

We say that a state of a semi-elastic element is **undistorted** if it is relaxed and of maximal symmetry. We say that the configuration $G \in \mathcal{G}$ is *undistorted* if the corresponding relaxed state $\bar{\lambda}(G)$ is undistorted.¹ Undistorted states always exist, for if σ is any state of maximal symmetry, so is the relaxed state $\bar{\lambda}(\sigma)$, as is evident from Prop. 12.2. It follows from Prop. 10.6 that for an isotropic semi-elastic element, a state is undistorted if and only if it is both relaxed and isotropic.

We call a semi-elastic element a **solid element**² if its symmetry group is contained in the orthogonal group of some configuration.

Proposition 14.2. *In a solid element a state λ is undistorted if and only if it is relaxed and satisfies $\mathcal{G}_\lambda = \mathcal{G}$.*

Proof. We need only show that for any $\lambda \in \Sigma_{\text{rel}}$, if λ is of maximal symmetry then $\mathcal{G}_\lambda = \mathcal{G}$. Since the element is solid, we have $\mathcal{G} \subset \text{Orth}(G)$ for some $G \in \mathcal{G}$. Let $\lambda_0 = \bar{\lambda}(G)$, so that $G = \hat{G}(\lambda_0)$. By (14.4) we have $\mathcal{G}_{\lambda_0} = \mathcal{G} \cap \text{Orth}(G) = \mathcal{G}$ and hence $\mathcal{G}_\lambda = \mathcal{G} \cap \text{Orth}(\hat{G}(\lambda)) \subset \mathcal{G} = \mathcal{G}_{\lambda_0}$. Since λ is of maximal symmetry, it follows that $\mathcal{G}_\lambda = \mathcal{G}_{\lambda_0} = \mathcal{G}$. Q.E.D.

Proposition 14.3. *If the symmetry group of a semi-elastic element is $\mathcal{G} = \text{Orth}(G)$ for some $G \in \mathcal{G}$, then the element is isotropic (and solid). Moreover, a state σ is isotropic if and only if $\mathcal{G}_\sigma = \mathcal{G}$. If σ is isotropic, then $\hat{G}(\sigma) = cG$ for some $c > 0$.*

Proof. We must prove that isotropic states exist. Since undistorted states exist and since the element is solid if $\mathcal{G} = \text{Orth}(G)$, it follows from Prop. 14.2 that there are states σ such that $\mathcal{G}_\sigma = \mathcal{G} = \text{Orth}(G)$. Now, in view of (10.8), $\mathcal{G}_\sigma = \mathcal{G} = \text{Orth}(G)$ implies $\text{Orth}(G) \subset \text{Orth}(G) \cap \text{Orth}(\hat{G}(\sigma))$, and hence $\text{Orth}(G) \subset \text{Orth}(\hat{G}(\sigma))$. This is possible (see Section 2) only when $\text{Orth}(\hat{G}(\sigma)) = \text{Orth}(G) = \mathcal{G} = \mathcal{G}_\sigma$, which shows that σ is isotropic.

If, conversely σ is isotropic, we have $\text{Orth}(\hat{G}(\sigma)) = \mathcal{G}_\sigma \subset \mathcal{G} = \text{Orth}(G)$. Again, this is possible only if $\mathcal{G}_\sigma = \mathcal{G}$ and $\hat{G}(\sigma) = cG$ with some $c > 0$ (see Section 2). Q.E.D.

We call a semi-elastic element a **fluid element**³ if its symmetry group \mathcal{G} contains the unimodular group $\text{Unim}(\mathcal{S})$. It follows from Prop. 14.1 that all relaxed states of a fluid element are isotropic. The converse is also valid, except when the set \mathcal{G} of configurations is *degenerate* in the sense that all $G \in \mathcal{G}$ are scalar multiples of one of them. Degeneracy is a severe constraint and of little if any physical interest.

Proposition 14.4. *If all relaxed states of a semi-elastic element are isotropic and if \mathcal{G} is non-degenerate, then the element is fluid.*

Proof. To say that all relaxed states are isotropic means that $\mathcal{G}_\lambda = \text{Orth}(\hat{G}(\lambda))$ for all $\lambda \in \Sigma_{\text{rel}}$. Hence, since $\hat{G}|_{\Sigma_{\text{rel}}}$ is surjective, it follows that $\mathcal{G} \supset \text{Orth}(G)$ for all

¹ In [TN], Sects. 31 and 33 and in earlier work ([N2], [CN1], [CN2]) I have used a related notion of “undistorted configuration”, applicable only to isotropic materials and to solids. The concept introduced here applies also to materials that are neither one nor the other.

² See [N2], Sect. 20 or [TN], Sect. 33.

³ See [N2], Sect. 21 or [TN], Sect. 32.

$G \in \mathcal{G}$. Since \mathcal{G} is non-degenerate, the groups $\text{Orth}(G)$ are not all the same (see Section 2). Therefore, we must have $g \neq \text{Orth}(G)$ for some $G \in \mathcal{G}$. By the theorem on the maximality of an orthogonal group in the unimodular group it follows that $g \supset \text{Unim}(\mathcal{T})$. Q.E.D.

For a fluid element, all configurations are undistorted and, by Prop. 10.4, the stress in all relaxed states reduces to a pressure. For an isotropic solid element, by Prop. 14.3, all undistorted configurations are scalar multiples of any one of them. An incompressible isotropic solid element has exactly one undistorted configuration.

Remark 1. For material elements that are not semi-elastic, the conclusions of Props. 14.1–4 need no longer be valid. It is conceivable, for example, that the symmetry group is an orthogonal group or even the unimodular group and yet the element is not isotropic, because there are no isotropic states.

For semi-elastic materials, Axiom VI of Section 13 yields the following strong conclusion.

Proposition 14.5. *In a semi-elastic element, every state is accessible from every other state, i.e. $\Sigma_\sigma = \Sigma$ for all $\sigma \in \Sigma$.*

Proof. Let $\tau \in \Sigma_\sigma$. By Prop. 13.2 we then also have $\hat{\lambda}(\tau) \in \Sigma_\sigma$ and hence, by Prop. 13.1, $\Sigma_{\hat{\lambda}(\tau)} \subset \Sigma_\sigma$.

Now choose a process $P \in \Pi$ such that $P^i = \hat{G}(\hat{\lambda}(\tau))$ and $P^f = \hat{G}(\lambda_0)$, where $\lambda_0 \in \Sigma_{\text{rel}}$ satisfies $\Sigma_{\lambda_0} = \Sigma$ (see Axiom VI). The existence of such a process is assured by the condition (P4) of Definition 4.1. We have

$$\hat{\rho}(\hat{\lambda}(\tau), P) \in \Sigma_{\hat{\lambda}(\tau)}, \quad P^f = \hat{G}(\hat{\rho}(\hat{\lambda}(\tau), P)) = \hat{G}(\lambda_0). \quad (14.5)$$

Using (12.3), we obtain from (14.5)₂ that $\hat{G}(\hat{\lambda}(\hat{\rho}(\hat{\lambda}(\tau), P))) = \hat{G}(\lambda_0)$. The injectivity of $\hat{G}|_{\Sigma_{\text{rel}}}$ yields $\lambda_0 = \hat{\lambda}(\hat{\rho}(\hat{\lambda}(\tau), P))$. Hence, by (14.5)₁ and Prop. 13.2, we conclude that $\lambda_0 \in \Sigma_{\hat{\lambda}(\tau)} \subset \Sigma_\sigma$. Using Prop. 13.1 again, we find $\Sigma = \Sigma_{\lambda_0} \subset \Sigma_\sigma$, i.e. $\Sigma_\sigma = \Sigma$. Q.E.D.

Remark 2. Prop. 14.5 would remain valid if Axiom VI were replaced by the weaker requirement that every state σ be accessible from some relaxed state λ_σ which may depend on σ .

The results of this section are valid, *a fortiori*, for elastic material elements. For such elements, $\hat{G}: \Sigma \rightarrow \mathcal{G}$ is a bijection, and the sections $\Sigma_G = \{\hat{\lambda}(G)\}$ are singletons. The natural topology on Σ is then the discrete topology. One can use $\hat{G}: \Sigma \rightarrow \mathcal{G}$ and its inverse $\hat{\lambda}: \mathcal{G} \rightarrow \Sigma$ to identify the state space Σ with the set \mathcal{G} of configurations. If this is done one must remember, however, that \mathcal{G} is equipped with the discrete topology when considered as the state space.

In order to define the structure of an elastic element on a given body element $(\mathcal{T}, \mathcal{G}, \Pi)$, it is sufficient to prescribe a *stress function* $\hat{S}: \mathcal{G} \rightarrow \mathcal{S} = \text{Sym}(\mathcal{T}^*, \mathcal{T})$. For the state space one can then take $\Sigma = \mathcal{G}$, for \hat{G} the identity mapping of \mathcal{G} , and for $\hat{\rho}$ the trivial mapping given by $\hat{\rho}(P^i, P) = P^f$.

15. Deformation Histories for Semi-Elastic Elements

As in Section 4, let \mathcal{G} be any set. We observe the notation and terminology of Section 4. A function H of the type

$$H: \mathbb{R}^+ \rightarrow \mathcal{G} \quad (15.1)$$

will be called a **history** with values in \mathcal{G} . Its value

$$H(0) = H^f \quad (15.2)$$

will be called the *final value* of H .

If $G \in \mathcal{G}$ is given, we can define a history $G_{(\infty)}: \mathbb{R}^+ \rightarrow \mathcal{G}$ by

$$G_{(\infty)}(s) = G \quad \text{for } s \in \mathbb{R}^+; \quad (15.3)$$

we call this history the *constant history* at G .

For any $s \in \mathbb{R}^+$, we define the s -section $H^{(s)}: \mathbb{R}^+ \rightarrow \mathcal{G}$ of a history H by

$$H^{(s)}(r) = H(s+r) \quad \text{for } r \in \mathbb{R}^+. \quad (15.4)$$

It is again a history. The s' -section of the s -section of H is the same as the $(s+s')$ -section of H , *i.e.*,

$$H^{(s)}(s') = H^{(s+s')} \quad \text{for } s, s' \in \mathbb{R}^+. \quad (15.5)$$

Of course, we have $H^{(0)} = H$.

If H is a history and $s \in \mathbb{R}^+$, we can define a process $H_{[s]}$, of duration s , by

$$H_{[s]}(t) = H(s-t) \quad \text{for } t \in [0, s]. \quad (15.6)$$

We call $H_{[s]}$ the s -segment of H . We have

$$(H_{[s]})^i = H(s), \quad (H_{[s]})^f = H^f = H(0). \quad (15.7)$$

The s -segment $(H^{(r)})_{[s]}$ of the r -section $H^{(r)}$ of H is a segment of the segment $H_{[s+r]}$, namely

$$(H^{(r)})_{[s]} = (H_{[s+r]})_{[0, s]}. \quad (15.8)$$

If H is a history and P a process, of duration d_P , such that $H^f = P^i$, we define the **continuation** $H * P: \mathbb{R}^+ \rightarrow \mathcal{G}$ of H with P by

$$(H * P)(s) = \begin{cases} H(s - d_P) & \text{if } s \geq d_P \\ P(d_P - s) & \text{if } s \in [0, d_P]. \end{cases} \quad (15.9)$$

We have, for all $s \in \mathbb{R}^+$,

$$H = H^{(s)} * H_{[s]}. \quad (15.10)$$

The continuation is associative, *i.e.*

$$(H * P_1) * P_2 = H * (P_1 * P_2) \quad \text{if } H^f = P_1^i, \quad P_1^f = P_2^i.$$

The s -section of a continuation $H * P$ is given by

$$(H * P)^{(s)} = \begin{cases} H^{(s-d_P)} & \text{if } s \geq d_P \\ H * P_{[0, d_P-s]} & \text{if } s \in [0, d_P]. \end{cases} \quad (15.11)$$

For the s -segment of $H * P$ we have

$$(H * P)_{[s]} = \begin{cases} P_{[d_P - s, d_P]} & \text{if } s \in [0, d_P] \\ H_{[s - d_P]} * P & \text{if } s \geq d_P. \end{cases} \quad (15.12)$$

Definition 15.1. Let $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \widehat{G}, \widehat{S}, \widehat{\rho})$ be a semi-elastic material element, so that $\widehat{G}|_{\Sigma_{rel}}$ is invertible with inverse $\bar{\lambda}: \mathcal{G} \rightarrow \Sigma_{rel}$. A history $H: \mathbb{R}^+ \rightarrow \mathcal{G}$ is called a **deformation history** for the element \mathcal{T} if

- (i) for all $s \in \mathbb{R}^+$, the segment $H_{[s]}$ belongs to Π (in view of (15.8) and (P2) of Definition 4.1, we then also have $(H^{(r)})_{[s]} \in \Pi$ for all $s, r \in \mathbb{R}^+$), and
- (ii) the limit

$$\lim_{s \rightarrow \infty} \widehat{\rho}(\bar{\lambda}(H^{(r)}(s)), (H^{(r)})_{[s]}) \in \Sigma_{H^{(r)}}$$

exists for all $r \in \mathbb{R}^+$.

The set of all deformation histories for \mathcal{T} will be denoted by Φ and, in view of condition (ii), we can define $\bar{\rho}: \Phi \rightarrow \Sigma$ by

$$\bar{\rho}(H) = \lim_{s \rightarrow \infty} \widehat{\rho}(\bar{\lambda}(H(s)), H_{[s]}). \quad (15.13)$$

We use the notation

$$\Phi_G = \{H \in \Phi \mid H^f = G\} \quad (15.14)$$

for the set of all deformation histories for \mathcal{T} that end at the configuration $G \in \mathcal{G}$.

We note that if H belongs to Φ , so do all of its sections $H^{(s)}$, $s \in \mathbb{R}^+$. Also, in view of (12.1), all constant histories (15.3) belong to Φ , and by Prop. 12.1 (iv), we have

$$\bar{\rho}(G_{(\infty)}) = \bar{\lambda}(G) \quad (15.15)$$

for all $G \in \mathcal{G}$. The next observation is somewhat less trivial:

Proposition 15.1. If $H \in \Phi$ and $P \in \Pi_{H^f}$, then $H * P \in \Phi$ and we have

$$\bar{\rho}(H * P) = \widehat{\rho}(\bar{\rho}(H), P). \quad (15.16)$$

Proof. By (15.9) and (15.12) we have, for $s \geq d_P$,

$$\widehat{\rho}(\bar{\lambda}((H * P)(s)), (H * P)_{[s]}) = \widehat{\rho}(\bar{\lambda}(H(s - d_P)), H_{[s - d_P]} * P).$$

Application of (7.7) yields

$$\widehat{\rho}(\bar{\lambda}((H * P)(s)), (H * P)_{[s]}) = \widehat{\rho}(\widehat{\rho}(\bar{\lambda}(H(s - d_P)), H_{[s - d_P]}), P).$$

Taking the limit $s \rightarrow \infty$ and using the continuity of $\widehat{\rho}(\cdot, P)$ (see Prop. 11.2), we see that (15.13) gives the desired result (15.16). Q.E.D.

Observing (15.10) and (15.15), we obtain the following special cases of (15.16):

$$\bar{\rho}(H) = \widehat{\rho}(\bar{\rho}(H^{(s)}), H_{[s]}), \quad s \in \mathbb{R}^+, H \in \Phi, \quad (15.17)$$

$$\bar{\rho}(P_{(\infty)}^i * P) = \widehat{\rho}(\bar{\lambda}(P^i), P), \quad P \in \Pi. \quad (15.18)$$

Proposition 15.2. *The function $\bar{\rho}: \Phi \rightarrow \Sigma$ satisfies*

$$\lim_{s \rightarrow \infty} \bar{\rho}(H(s)_{(\infty)} * H_{[s]}) = \bar{\rho}(H) \quad \text{for all } H \in \Phi \quad (15.19)$$

and

$$\lim_{t \rightarrow \infty} \bar{\rho}(H * H_{(t)}^f) = \bar{\rho}(H_{(\infty)}^f) \quad \text{for all } H \in \Phi. \quad (15.20)$$

Proof. If we write (15.18) with the choice $P = H_{[s]}$ and substitute the result into (15.13), we obtain (15.19).

To prove (15.20), we apply (15.16) with the choice $P = H_{(t)}^f$ and take the limit $t \rightarrow \infty$. Observing (14.3), we obtain

$$\lim_{t \rightarrow \infty} \bar{\rho}(H * H_{(t)}^f) = \lim_{t \rightarrow \infty} \hat{\rho}(\bar{\rho}(H), H_{(t)}^f) = \bar{\lambda}(H^f).$$

In view of (15.15) this is the desired result. Q.E.D.

The next result states that arbitrarily close to any state there is a state that can be described by a deformation history.

Proposition 15.3. *The range of $\bar{\rho}$ is dense in the state space Σ , i. e.,*

$$\Sigma = \text{cl } \Sigma_{\text{hist}}, \quad \text{where } \Sigma_{\text{hist}} = \text{Range } \bar{\rho}. \quad (15.21)$$

Proof. Let λ be any relaxed state. It follows from the definition of Σ_λ (cf. (13.1)) and from (15.18) that

$$\Sigma_\lambda = \text{cl } \{ \bar{\rho}(P_{(\infty)}^i * P) \mid P \in \Pi_{\hat{G}(\lambda)} \}.$$

Since this set is the closure of the range of the restriction of $\bar{\rho}$ to histories of the form $P_{(\infty)}^i * P$ with $P^i = \hat{G}(\lambda)$, it follows *a fortiori* that $\Sigma_\lambda = \text{cl } \Sigma_{\text{hist}}$. By Prop. 14.5, this is the desired result. Q.E.D.

For semi-elastic elements, it is often sufficient to confine one's attention to the space Σ_{hist} of states that correspond to deformation histories. The remaining states, which are limits of states determined by histories, are of little physical interest.

Finally, we describe how $\bar{\rho}(H)$ changes under a symmetry.

Proposition 15.4. *If $A \in \mathcal{g}$ is a symmetry of a semi-elastic element, then $A^* \Phi A = \Phi$ and*

$$\bar{\rho}(A^* H A) = \iota_{A^{-1}}(\bar{\rho}(H)) \quad \text{for all } H \in \Phi. \quad (15.22)$$

Proof. Let $A \in \mathcal{g}$. It follows from (10.3) and the definition of $\bar{\lambda}$ as the inverse $\hat{G}|_{\Sigma_{\text{rel}}}$ that

$$\bar{\lambda}(A^* G A) = \iota_{A^{-1}}(\bar{\lambda}(G)) \quad (15.23)$$

for all $G \in \mathcal{g}$. Now let $H \in \Phi$. Using (15.23) with $G = H(s)$ and then (10.5) with the choice $P = H_{[s]}$ and $\sigma = \iota_{A^{-1}}(\bar{\lambda}(H(s)))$, we obtain

$$\begin{aligned} \hat{\rho}(\bar{\lambda}((A^* H A)(s)), (A^* H A)_{[s]}) &= \hat{\rho}(\iota_{A^{-1}}(\bar{\lambda}(H(s))), A^* H_{[s]} A) \\ &= \iota_{A^{-1}}(\hat{\rho}(\bar{\lambda}(H(s)), H_{[s]})). \end{aligned}$$

Recalling that $\iota_{A^{-1}} = \iota_{A^{-1}}$ (cf. (9.6)) is a homeomorphism and remembering the definition (15.13) of $\bar{\rho}$, we see, by taking the limit $s \rightarrow \infty$, that $A^* H A \in \Phi$ and that (15.22) holds. Q.E.D.

16. The History Functional of a Semi-Elastic Element

Let \mathcal{T} be a semi-elastic element, Φ the set of its deformation histories, and $\bar{\rho}: \Phi \rightarrow \Sigma$ the function defined by (15.13). We define the **history functional** \bar{S} of the element by

$$\bar{S} = \hat{S} \circ \bar{\rho}: \Phi \rightarrow \mathcal{S}, \quad (16.1)$$

where $\mathcal{S} = \text{Sym}(\mathcal{T}^*, \mathcal{T})$ is the stress space. If we apply \hat{S} to (15.16), we see that \bar{S} is related to the response functional $\tilde{S} = \hat{S} \circ \hat{\rho}$ by

$$\tilde{S}(\bar{\rho}(H), P) = \bar{S}(H * P) \quad \text{for } H \in \Phi, P \in \Pi_{H^f}. \quad (16.2)$$

A knowledge of the history functional is sufficient for deciding whether two deformation histories determine the same state. Indeed, it follows from Axiom III, Section 7, and (16.2) that

$$\bar{\rho}(H_1) = \bar{\rho}(H_2) \Leftrightarrow \left\{ \begin{array}{l} H_1^f = H_2^f \text{ and } \bar{S}(H_1 * P) = \bar{S}(H_2 * P) \\ \text{for all } P \in \Pi \text{ with } P^i = H_1^f = H_2^f \end{array} \right\}. \quad (16.3)$$

We shall show in Theorem 16.1, below, that a knowledge of the set Φ of deformation histories and of the history functional $\bar{S}: \Phi \rightarrow \mathcal{S}$ characterizes the semi-elastic element structure on \mathcal{T} . We first give a complete list of conditions that are satisfied by Φ . They are easy consequences of the results of the previous section.

- (H1) Every $H \in \Phi$ is a history $H: \mathbb{R}^+ \rightarrow \mathcal{G}$ with values in a closed and connected subset \mathcal{G} of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$.
- (H2) For all $G \in \mathcal{G}$ we have $G_{(\infty)} \in \Phi$.
- (H3) If $H \in \Phi$ and $s \in \mathbb{R}^+$, then $H^{(s)} \in \Phi$.
- (H4) If $H, K \in \Phi$ and $s \in \mathbb{R}^+$ with $K^f = H(s)$, then $K * H_{[s]} \in \Phi$.
- (H5) For any $G_1, G_2 \in \mathcal{G}$ there is a $H \in \Phi$ such that $G_1, G_2 \in \text{Range } H$.

The set Π of all deformation processes for \mathcal{T} is determined by Φ as follows:

$$\Pi = \{H_{[s]} \mid H \in \Phi, s \in \mathbb{R}^+\}. \quad (16.4)$$

Next, we show that the history functional \bar{S} has the following two convergence properties:

- (F1) For every $H \in \Phi$ we have

$$\lim_{s \rightarrow \infty} \bar{S}(H(s)_{(\infty)} * H_{[s]}) = \bar{S}(H). \quad (16.5)$$

- (F2) If $H \in \Phi$ and $P \in \Pi_{H^f}$, so that $H^f = P^i$, then

$$\lim_{t \rightarrow \infty} \bar{S}(H * H_{(t)}^f * P) = \bar{S}(H_{(\infty)}^f * P). \quad (16.6)$$

Property (F1) is an immediate consequence of (15.19) and the continuity of \hat{S} . If we apply the convergence criterion for states given by Prop. 11.3 to (15.20) and observe (16.2), we see that (F2) must be valid. Applying the same convergence criterion to the condition (ii) of Definition 15.1 and observing (15.15) and (15.16),

we see that the class Φ must satisfy the following *saturation condition* with respect to \bar{S} :

(S) *If H is a history with values in \mathcal{G} such that $H_{[s]} \in \Pi$ for every $s \in \mathbb{R}^+$ and such that*

$$\lim_{s \rightarrow \infty} \bar{S}((H^{(r)}(s))_{(\infty)} * (H^{(r)})_{[s]} * P)$$

exists whenever $r \in \mathbb{R}^+$ and $P^i = H(r)$, then H belongs to Φ .

Theorem 16.1. *Let \mathcal{T} be a finite-dimensional real vector space, and let Φ be a class of histories with the properties (H1)–(H5). Let Π be defined by (16.4). Finally, let $\bar{S}: \Phi \rightarrow \mathcal{S}$ be a functional with the properties (F1), (F2) and assume that Φ satisfies the saturation condition (S) with respect to \bar{S} . Then there is a unique semi-elastic element structure on \mathcal{T} for which Φ is the class of deformation histories and \bar{S} the history functional.*

Proof. First we assume the existence of a semi-elastic element structure $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ whose set of deformation histories is Φ and whose history functional is \bar{S} , and we prove the uniqueness of this structure. It is clear that Π must be the set defined by (16.4) in terms of Φ and that \mathcal{G} must be the set mentioned in (H1). Hence the underlying body element $(\mathcal{T}, \mathcal{G}, \Pi)$ is determined by Φ .

Guided by (16.3), we now introduce an equivalence relation \sim on Φ by

$$H_1 \sim H_2 \Leftrightarrow \left\{ \begin{array}{l} H_1^f = H_2^f \text{ and } \bar{S}(H_1 * P) = \bar{S}(H_2 * P) \\ \text{for all } P \in \Pi \text{ with } P^i = H_1^f = H_2^f \end{array} \right\}. \quad (16.7)$$

Let Σ' be the quotient of Φ with respect to this equivalence relation. In other words, the elements of Σ' are the equivalence classes determined by \sim . We use the notation $\bar{\rho}': \Phi \rightarrow \Sigma'$ for the quotient mapping, which assigns to each $H \in \Phi$ the equivalence class $\bar{\rho}'(H) \in \Sigma'$ to which H belongs. It follows from (16.3) that there is a bijection $\iota: \Sigma' \rightarrow \Sigma_{\text{hist}} = \text{Range } \bar{\rho}$ such that $\bar{\rho} = \iota \circ \bar{\rho}'$. The bijection ι associates with every equivalence class $\sigma' \in \Sigma'$ the common value of $\bar{\rho}$ at all members $H \in \sigma'$. We now define $\hat{G}': \Sigma' \rightarrow \mathcal{G}$ by $\hat{G}' = \hat{G} \circ \iota$, $\hat{S}': \Sigma' \rightarrow \mathcal{S}$ by $\hat{S}' = \hat{S} \circ \iota$, and $\hat{\rho}': (\Sigma' \times \Pi)_{\text{fit}} \rightarrow \Sigma'$ by $\hat{\rho}'(\cdot, \cdot) = \hat{\rho}(\iota(\cdot), \cdot)$. We then have

$$\hat{G}'(\sigma') = H^f, \quad \hat{S}'(\sigma') = \bar{S}(H) \quad \text{if } H \in \sigma' \quad (16.8)$$

and

$$\hat{\rho}'(\sigma', P) = \bar{\rho}'(H * P) \quad \text{if } H \in \sigma', P^i = H^f. \quad (16.9)$$

It is clear from (16.8) and (16.9) that \hat{G}' , \hat{S}' , and $\hat{\rho}'$ are completely determined by Φ and \bar{S} .

We have obtained a structure $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma', \hat{G}', \hat{S}', \hat{\rho}')$ which is determined by Φ and \bar{S} alone. It is easily seen that this structure satisfies Axioms I–III. Hence we can apply Definition 11.1 to it and endow each section $\Sigma'_G = \{\sigma' \in \Sigma' \mid \hat{G}'(\sigma') = G\}$ with a natural uniformity and $\Sigma' = \bigcup \{\Sigma'_G \mid G \in \mathcal{G}\}$ with a natural topology. The functions \hat{G}' , \hat{S}' , and $\hat{\rho}'(\cdot, P)$ are continuous with respect to this topology. Moreover, it is easily seen that the bijection $\iota: \Sigma' \rightarrow \Sigma_{\text{hist}}$ is a homeomorphism when Σ_{hist} is endowed with the topology induced by the natural topology of Σ .

Next, we consider the completions $\bar{\Sigma}'_G$ and $\Sigma^* = U\{\bar{\Sigma}'_G \mid G \in \mathcal{G}\}$ as described at the end of Section 11, and we extend the functions \hat{G}' , \hat{S}' , and $\hat{\rho}'$ so as to obtain

a structure $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$ which satisfies not only Axioms I–III but also Axiom IV. Since Axiom IV is satisfied for $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$, we can also extend the homeomorphism $\iota: \Sigma' \rightarrow \Sigma_{\text{hist}}$ to a homeomorphism $\iota^*: \Sigma^* \rightarrow \text{cl } \Sigma_{\text{hist}} = \Sigma$ (see (15.21)). The relations $\hat{G}' = \hat{G} \circ \iota$, $\hat{S}' = \hat{S} \circ \iota$, and $\iota(\hat{\rho}'(\cdot, \cdot)) = \hat{\rho}(\iota(\cdot), \cdot)$ remain satisfied after \hat{G}' , \hat{S}' , $\hat{\rho}'$, and ι have been extended to \hat{G}^* , \hat{S}^* , $\hat{\rho}^*$, and ι^* . It follows that the identity $1_{\mathcal{T}}$ is a material isomorphism between the material element $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ and the structure $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$ determined by Φ and \bar{S} , the bijection $\iota_{1_{\mathcal{T}}}$ being the same as ι^* . Since we have agreed, under these circumstances (see Section 9), to employ ι^* to identify Σ with Σ^* and not to distinguish between the two structures, we have the following conclusion: *If there is a semi-elastic element structure on \mathcal{T} whose set of deformation histories is Φ and whose history functional is \bar{S} , it must be given by $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$.*

We now prove the existence of a semi-elastic element structure on \mathcal{T} whose set of deformation histories is Φ and whose history functional is \bar{S} . If \mathcal{G} is the set mentioned in (H 1) and if Π is defined by (16.4), it is easily inferred from (H 1)–(H 5) that Π satisfies the conditions (P 1)–(P 4) of Definition 4.1 and hence that $(\mathcal{T}, \mathcal{G}, \Pi)$ endows \mathcal{T} with the structure of a body element.

Consider the quotient set Σ' of Φ with respect to the equivalence relation defined by (16.7) and the quotient mapping $\bar{\rho}': \Phi \rightarrow \Sigma'$. It follows from (16.7) that $H_1 \sim H_2$ implies $H_1^f = H_2^f$, $\bar{S}(H_1) = \bar{S}(H_2)$, and $H_1 * P \sim H_2 * P$ when $P^i = H_1^f = H_2^f$. Thus, it is meaningful to *define* \hat{G}' , \hat{S}' by (16.8) and $\hat{\rho}'$ by (16.9). We introduce a corresponding response functional $\hat{S}' = \bar{S}' \circ \hat{\rho}'$, so that

$$\tilde{S}'(\sigma', P) = \bar{S}(H * P) \quad \text{if } H \in \sigma', \quad P^i = H^f. \quad (16.10)$$

It is easily verified that the septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma', \hat{G}', \hat{S}', \hat{\rho}')$ thus constructed satisfies Axioms I–III. Observing (16.10), it follows from Condition (F 2) for \bar{S} that \tilde{S}' satisfies (12.4) when $\hat{\lambda}': \Sigma' \rightarrow \Sigma'_{\text{rel}}$ is defined by

$$\hat{\lambda}'(\sigma') = \bar{\rho}'(H_{(\infty)}^f) \quad \text{if } H \in \sigma'. \quad (16.11)$$

Therefore Axiom V is also satisfied, because it is equivalent to (12.4). The set Σ'_{rel} consists of the equivalence classes that contain constant histories. The restriction of \hat{G}' to Σ'_{rel} has an inverse $\bar{\lambda}': \mathcal{G} \rightarrow \Sigma'_{\text{rel}}$ given by

$$\bar{\lambda}'(G) = \bar{\rho}'(G_{(\infty)}), \quad (16.12)$$

so that $\bar{\lambda}'(G)$ is the equivalence class determined by the constant history $G_{(\infty)}$. Thus, the defining condition of Definition 14.1 for a semi-elastic element is satisfied.

Next we construct the completion $\Sigma^* = \bigcup \{ \bar{\Sigma}'_G \mid G \in \mathcal{G} \}$ and the extensions \hat{G}^* , \hat{S}^* , and $\hat{\rho}^*$ as described in the uniqueness proof above. The septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$ thus constructed satisfies not only Axioms I, II, III, and V, but also Axiom IV. Note that the completion introduces no new relaxed states, so that $\Sigma^*_{\text{rel}} = \Sigma'_{\text{rel}}$. To show that the accessibility axiom, Axiom VI, is satisfied, let $G \in \mathcal{G}$ and $H \in \Phi$ be given. For each $s \in \mathbb{R}^+$ we select a process $P_s \in \Pi$ such that $P_s^i = G$, $P_s^f = H(s)$. This is possible because Π satisfies (P 4) of Definition 4.1. Using the assumption that \bar{S} satisfies both (F 1) and (F 2), we easily infer that

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \bar{S}(G_{(\infty)} * P_s * H(s)_{(t)} * H_{[s]} * P') = \bar{S}(H * P') \quad (16.13)$$

holds for all $P' \in \Pi_{H'}$. Observing (16.9), (16.10), (16.11), and using the convergence criterion of Prop. 11.3, we see that (16.13) is equivalent to

$$\lim_{s \rightarrow \infty} \lim_{t \rightarrow \infty} \hat{\rho}'(\bar{\lambda}'(G), P_s * H(s)_{(t)} * H_{[s]}) = \bar{\rho}'(H)$$

in the natural topology of $\Sigma' = \text{Range } \bar{\rho}'$. It follows that $\bar{\rho}'(H)$ is accessible from $\bar{\lambda}'(G)$. Since H was arbitrary, we conclude that $\Sigma' \subset \Sigma_{\bar{\lambda}'(G)}^*$, and since Σ' is dense in Σ^* , we find $\Sigma^* = \Sigma_{\bar{\lambda}'(G)}^*$, i.e. that Axiom VI is satisfied. Therefore $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$ is a semi-elastic material element structure.

It remains to be shown that Φ is the set of deformation histories and \bar{S} the history functional of the semi-elastic element just constructed. It follows from (H2), (H3), (H4), and (F1), using also (16.10) and (16.12), that

$$\lim_{s \rightarrow \infty} \bar{S}'(\bar{\lambda}'(H^{(r)}(s)), (H^{(r)})_{[s]} * P) = \bar{S}(H^{(r)} * P) \tag{16.14}$$

for all $H \in \Phi$, all $r \in \mathbb{R}^+$, and $P \in \Pi_{H^{(r)}}$. From the convergence criterion of Prop. 11.3 we see that condition (ii) as well as condition (i) of Definition 15.1 is satisfied for all $H \in \Phi$ and hence that Φ is contained in the set of deformation histories for the constructed element. The fact that Φ is actually equal to this set of deformation histories is a consequence of the saturation condition (S). It is not hard to verify, finally, that \bar{S} is the history functional of the constructed element. Q.E.D.

Remark. If Φ and $\bar{S}: \Phi \rightarrow \mathcal{S}$ are such that (H1)–(H5) and (F1), (F2), but not necessarily the saturation condition (S), are satisfied, the construction of the semi-elastic element structure $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma^*, \hat{G}^*, \hat{S}^*, \hat{\rho}^*)$ described in the proof above can still be carried out. However, the set Φ^* of deformation histories for the element so constructed will then be larger than the original set Φ . To obtain Φ^* from Φ , one must saturate Φ by joining to it all histories for which the limits of condition (S) exist.

Theorem 16.1 states, in essence, that a semi-elastic material element structure on \mathcal{T} is the same as a structure defined by a triple $(\mathcal{T}, \Phi, \bar{S})$ subject to the conditions (H1)–(H5), (F1), (F2), and (S) as axioms. It follows that material isomorphisms between semi-elastic elements can be described as isomorphisms between structures of the type $(\mathcal{T}, \Phi, \bar{S})$. More precisely, we have the following corollary to Theorem 16.1.

Proposition 16.2. *Let \mathcal{T}_1 and \mathcal{T}_2 be semi-elastic elements whose sets of deformation histories are Φ_1 and Φ_2 and whose history functionals are \bar{S}_1 and \bar{S}_2 , respectively. A mapping $A \in \text{Invlin}(\mathcal{T}_1, \mathcal{T}_2)$ is a material isomorphism between \mathcal{T}_1 and \mathcal{T}_2 if and only if $\Phi_1 = A^* \Phi_2 A$ and*

$$\bar{S}_1(A^* H A) = A^{-1} \bar{S}_2(H) A^{*-1} \quad \text{for all } H \in \Phi_2. \tag{16.15}$$

Applying this proposition to material automorphism, we obtain the following corollary.

Proposition 16.3. *Let \mathcal{T} be a semi-elastic element whose set of deformation histories is Φ and whose history functional is \bar{S} . A mapping $A \in \text{Invlin}(\mathcal{T})$ belongs to the symmetry group g of \mathcal{T} if and only if $\Phi = A^* \Phi A$ and*

$$\bar{S}(A^* H A) = A^{-1} \bar{S}(H) A^{*-1} \quad \text{for all } H \in \Phi. \tag{16.16}$$

17. Monotonous Processes

We deal with processes and histories as described in Sections 4 and 15. Let \mathcal{T}_1 and \mathcal{T}_2 be two real vector spaces of equal finite dimension. We say that a process P_1 with values in $\text{Sym}^+(\mathcal{T}_1, \mathcal{T}_1^*)$ is *congruent* to a process P_2 with values in $\text{Sym}^+(\mathcal{T}_2, \mathcal{T}_2^*)$ if P_1 and P_2 have the same duration and if there is an $A \in \text{InvlIn}(\mathcal{T}_1, \mathcal{T}_2)$ such that

$$P_1 = A^* P_2 A. \tag{17.1}$$

For example, if $(\mathcal{T}_1, \mathcal{G}_1, \Pi_1)$ and $(\mathcal{T}_2, \mathcal{G}_2, \Pi_2)$ are body elements (cf. Definition 4.1) and if A is an isomorphism between these elements (cf. Definition 9.1), then every process in Π_1 is of the form $A^* P_2 A$ with $P_2 \in \Pi_2$ and hence congruent to a process in Π_2 . It is clear that congruence is an equivalence relation.

Definition 17.1. We say that a process P with values in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ is **monotonous** if P is continuous and if any two segments of P of equal duration are congruent. We say that a history H with values in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ is **monotonous** if all of its segments $H_{[s]}$, $s \in \mathbb{R}^+$, are monotonous.

Proposition 17.1. A process P with values in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ is monotonous if and only if there is a function $A: [0, d_p] \rightarrow \text{InvlIn}(\mathcal{T})$ such that

$$P(r+t) = A(r)^* P(t) A(r) \quad \text{for } r \in [0, d_p], t \in [0, d_p - r]. \tag{17.2}$$

Proof. If P is monotonous, then $P_{[0, d_p - r]}$ and $P_{[r, d_p]}$ are congruent, and (17.2) expresses this congruence. If (17.2) holds, then $P_{[r, r+d]}$ and $P_{[0, d]}$ are congruent if $r \leq d_p - d$. Hence all segments of P of duration d are congruent to $P_{[0, d]}$ and hence to one another, i.e., P is monotonous. Q.E.D.

Every freeze is obviously monotonous. A complete characterization of all monotonous processes is given in the following theorem.

Theorem 17.1. A process P with values in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ is monotonous if and only if there is an element $E \in \text{Lin}(\mathcal{T})$ such that¹

$$P(t) = \exp(tE^*) P^i \exp(tE) \quad \text{for } t \in [0, d_p]. \tag{17.3}$$

Moreover, if P is monotonous, then (17.3) holds for exactly one E in the subalgebra of $\text{Lin}(\mathcal{T})$ that is generated by the subset $\{P(t_1)^{-1} P(t_2) \mid t_1, t_2 \in [0, d_p]\}$ of $\text{Lin}(\mathcal{T})$.

The proof of this rather non-trivial theorem will be given in a separate paper [N8]. If it were known in advance that both P and A of (17.2) are differentiable, one could prove the assertion by the method used in the proof of Theorem 1 of [N4]. The hard part of the proof of Theorem 17.1 is to show that *a priori* continuity and differentiability assumptions on A are unnecessary.

If (17.3) holds, we call P the **monotonous process of exponent E , initial configuration P^i , and duration d_p** . It is clear that every segment of a monotonous process of exponent E is again a monotonous process of exponent E .

¹ If $E \in \text{Lin}(\mathcal{T})$, then $\exp E \in \text{InvlIn}(\mathcal{T})$ is defined by $\exp E = \sum_{n=1}^{\infty} \frac{1}{n!} E^n$ or as the value $A(1)$ at $t=1$ of the solution of the initial value problem $\dot{A}(t) = EA(t)$, $A(0) = 1$.

In general, the exponent is not uniquely determined by the monotonous process P . For example, if $P = G_{(d)}$ is a freeze, then (17.3) holds not only when $E = 0$, but also when E is skew-symmetric relative to G , *i.e.* when $E^*G = -GE$. In this case, the subalgebra mentioned in Theorem 17.1 consists of the scalar multiples of the identity of \mathcal{T} , and $E = 0$ is the only exponent in that subalgebra for which (17.3) holds.

Proposition 17.2. *A history H with values in $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ is monotonous if and only if it is of the form*

$$H(s) = \exp(-sE^*)H^f \exp(-sE), \quad s \in \mathbb{R}^+, \quad (17.4)$$

with $E \in \text{Lin}(\mathcal{T})$. If (17.4) holds, then every segment of H is a monotonous process of exponent E .

Proof. If (17.4) holds, it is easily seen (*cf.* (15.6) and (15.7)) that

$$H_{[s]}(t) = \exp(tE^*)H(s)\exp(tE), \quad t \in [0, s]$$

and hence that $H_{[s]}$ is monotonous with exponent E for each $s \in \mathbb{R}^+$, which means that H is monotonous.

Now assume that H is monotonous, and choose $r \in \mathbb{R}^+$, $r > 0$. Since $H_{[r]}$ is then monotonous, we can apply Theorem 17.1 to $H_{[r]}$ obtaining

$$H(r-t) = H_{[r]}(t) = \exp(tE^*)H(r)\exp(tE) \quad \text{for } t \in [0, r]. \quad (17.5)$$

The choice $t = r$ gives $H_{[r]}^f = H^f = \exp(rE^*)H(r)\exp(rE)$. Solving this equation for $H(r)$, substituting the result into (17.5) and then putting $s = r - t$, we derive

$$H(s) = \exp(-sE^*)H^f \exp(-sE) \quad \text{for } s \in [0, r]. \quad (17.6)$$

We must show that (17.6) holds not only for $s \in [0, r]$ but also when $s > r$. Now, if $s > r$, we select $r' > s$ and observe that an equation of the form (17.6), but possibly with an exponent E' different from E , holds for all $s' \in [0, r']$. It follows that H is analytic on $[0, r']$. By the principle of analytic continuation, (17.6) must also hold for s . Q.E.D.

If (17.4) holds, we say that H is the *monotonous history of exponent E and final configuration H^f* . If H is a monotonous history of exponent E , so are all its sections $H^{(s)}$ (see (15.4)) and, since $\exp(r+s)E = (\exp sE)(\exp rE)$, we have

$$H^{(s)} = \exp(-sE^*)H \exp(-sE). \quad (17.7)$$

We recall the definitions and terminology of Section 5 and consider a body element $(\mathcal{T}, \mathcal{G}, \Pi)$ as well as a frame-space \mathcal{V} . We say that a motion $M: [0, d_M] \rightarrow \mathcal{P}_{\mathcal{V}}$ is *monotonous* if the deformation process $M^*M: [0, d_M] \rightarrow \mathcal{G}$ determined by it is monotonous.

Proposition 17.3. *Let M be a motion of \mathcal{T} in \mathcal{V} and $F_{(0)} = MM_i^{-1}$ its displacement tensor relative to the initial placement $M_i = M(0)$. The motion M is monotonous if and only if $F_{(0)}$ is of the form¹*

¹ Comparison of this result with Theorem I of [N4] shows that the class of monotonous motions is identical to the class of motions introduced by COLEMAN [C] under the name of "substantially stagnant motions", which I later called "motions with constant stretch history" in [N4] and [TN], Sect. 109.

$$F_{(0)}(t) = Q(t) \exp(tB), \quad (17.8)$$

where $B \in \text{Lin}(\mathcal{V})$ and $Q: [0, d_M] \rightarrow \text{Orth}(\mathcal{V})$, $Q(0) = 1_{\mathcal{V}}$.

Proof. By Theorem 17.1, M^*M and hence M is monotonous if and only if

$$M^*(t)M(t) = \exp(tE^*)M^i M^i \exp(tE). \quad (17.9)$$

If (17.9) holds, then M and $t \mapsto M^i \exp(tE)$ determine the same process. Hence, as was remarked in Section 5,

$$M(t)(M^i \exp(tE))^{-1} = Q(t) \quad (17.10)$$

belongs to $\text{Orth}(\mathcal{V})$. If we put $B = M_i E M_i^{-1}$ so that $M_i \exp(tE) M_i^{-1} = \exp(tB)$, we see that (17.10) implies (17.8). Conversely, if (17.8) holds, it is easily seen that (17.9) is valid with $E = M_i^{-1} B M_i$, and hence that M is monotonous. Q.E.D.

An important special type of monotonous deformation process is one whose exponent E satisfies $E^2 = 0$ but $E \neq 0$. We call such a process a **simple shearing**. Since $\exp(tE) = 1 + tE$ when $E^2 = 0$, (17.3) shows that a simple shearing P has the form

$$P(t) = P^i + t(E^* P^i + P^i E) + t^2 E^* P^i E, \quad t \in [0, d_P]. \quad (17.11)$$

A motion which determines a simple shearing is called a *viscometric motion* (cf. [CMN]).

18. Processes of Constant Reduced State, States of Monotonous Flow

Definition 18.1. Let \mathcal{T} be a material element with symmetry group \mathcal{g} . We say that a deformation process $P \in \Pi$ is of **constant reduced state** relative to a state $\sigma \in \Sigma_{P^i}$ if the states $\hat{\rho}(\sigma, P_{[0, t]})$, $t \in [0, d_P]$, all belong to the reduced state Ω_σ determined by σ .

To say that P is of constant reduced state relative to $\sigma \in \Sigma_{P^i}$ means (cf. (10.6)) that there is a function $A: [0, d_P] \rightarrow \mathcal{g}$ such that

$$\hat{\rho}(\sigma, P_{[0, t]}) = \iota_{A(t)}(\sigma), \quad t \in [0, d_P]. \quad (18.1)$$

Since $\hat{\rho}(\sigma, P_{[0, 0]}) = \sigma$ by Prop. 7.1, we may assume that $A(0) = 1_{\mathcal{g}}$.

It is clear from Prop. 12.1 (ii) that all freezes $G_{(t)}$ are of constant reduced state (actually of constant state) relative to any relaxed state λ with $\hat{G}(\lambda) = G$. If the symmetry group \mathcal{g} is large enough, one can expect to find non-constant processes of constant reduced state among the monotonous processes in Π , as is shown in the proposition below. We say that $\sigma \in \Sigma$ is a **state of monotonous flow** with *exponent* E if $\exp(sE) \in \mathcal{g}$ for all $s \in \mathbb{R}^+$ and if every monotonous process $P \in \Pi$ of exponent E and initial configuration $P^i = \hat{G}(\sigma)$ satisfies

$$\hat{\rho}(\sigma, P_{[0, t]}) = \iota_{\exp(-tE)}(\sigma) \quad (18.2)$$

and hence is of constant reduced state relative to σ . The states of monotonous flow with exponent 0 are identical to the relaxed states (cf. Prop. 12.1 (ii)). If \mathcal{g}

is a Lie group, then the condition that $\exp(sE) \in \mathcal{G}$ for all $s \in \mathbb{R}^+$ means that E belongs to the Lie algebra of \mathcal{G} .

Proposition 18.1. *Let H be a monotonous history, with values in \mathcal{G} and of exponent E , such that*

- (i) *the segments $H_{[s]}$, $s \in \mathbb{R}^+$, belong to Π ,*
- (ii) *for all $s \in \mathbb{R}^+$, $\exp(sE)$ belongs to the symmetry group \mathcal{G} ,*
- (iii) *for some $\sigma_0 \in \Sigma_{H^t}$, the limit*

$$\lim_{s \rightarrow \infty} \hat{\rho}(\iota_{\exp(sE)}(\sigma_0), H_{[s]}) = \sigma \quad (18.3)$$

exists.

Then σ is a state of monotonous flow with exponent E .

Proof. We use the abbreviation

$$A(t) = \exp(-tE), \quad t \in \mathbb{R}. \quad (18.4)$$

It follows from (17.4) and (17.5) that

$$A(r)^* H_{[r]}(r) A(r) = A(t)^* H^f A(t), \quad r \in [0, t]. \quad (18.5)$$

On the other hand, if P is a monotonous process of exponent E and initial configuration H^f , it follows from (17.3) that $A(r)^* P_{[0, r]}(r) A(r) = H^f$, $r \in [0, t]$, $t \in [0, d_P]$. Substituting this expression for H^f into (18.5) and using the fact that $A(r)$ and $A(t)$ commute, we obtain

$$H_{[t]} = A(t)^* P_{[0, t]} A(t), \quad t \in [0, d_P]. \quad (18.6)$$

Since $H_{[t]} \in \Pi$ and $A(t)^{-1} \in \mathcal{G}$, we infer that $P_{[0, t]}$ belongs to Π . Moreover, by (15.10) and (17.7), (18.6) gives

$$A(t)^* (H * P_{[0, t]}) A(t) = H \quad (18.7)$$

and hence

$$A(t)^* (H_{[s]} * P_{[0, t]}) A(t) = H_{[t+s]}. \quad (18.8)$$

By (10.5), it follows from (18.8) that

$$\iota_{A(t)} \hat{\rho}(\iota_{A(t)}^{-1}(\iota_{A(-s)}(\sigma_0)), H_{[t+s]}) = \hat{\rho}(\iota_{A(-s)}(\sigma_0), H_{[s]} * P_{[0, t]}).$$

If we apply (7.7) and observe $\iota_{A(-(t+s))} = \iota_{A(t)}^{-1} \circ \iota_{A(-s)}$, we obtain

$$\iota_{A(t)}(\hat{\rho}(\iota_{A(-(t+s))}(\sigma_0), H_{[t+s]})) = \hat{\rho}(\hat{\rho}(\iota_{A(-s)}(\sigma_0), H_{[s]}), P_{[0, t]}).$$

Taking the limit $s \rightarrow \infty$ and observing (18.3), we get

$$\iota_{A(t)}(\sigma) = \hat{\rho}(\sigma, P_{[0, t]}),$$

which shows that (18.2) holds and hence that σ is a state of monotonous flow. Q.E.D.

The following result is a corollary to Prop. 18.1.

Proposition 18.2. *Let \mathcal{T} be a semi-elastic element and Φ the set of deformation histories for \mathcal{T} (cf. Definition 15.1). If H belongs to Φ and is monotonous of ex-*

ponent E such that $\exp(sE) \in \mathcal{G}$ for all $s \in \mathbb{R}^+$, then $\bar{\rho}(H)$ (see (15.13)) is a state of monotonous flow with exponent E .

Proof. H satisfies the condition (i) of Prop. 18.1 because H is in Φ , the condition (ii) by hypothesis, and the condition (iii) with $\sigma = \bar{\rho}(H)$ and $\sigma_0 = \bar{\lambda}(H^f)$ because $\iota_{\exp(sE)}(\bar{\lambda}(H^f)) = \bar{\lambda}(H(s))$. Q.E.D.

For a process of constant reduced state it is possible to determine the stress at time t from a knowledge of the initial stress and the symmetry-valued function A of (18.1) alone. Indeed, it follows from (18.1) and (10.4) that the stress $\tilde{S}(\sigma, P_{[0, t]})$ at time t is given by

$$\tilde{S}(\sigma, P_{[0, t]}) = A(t) \hat{S}(\sigma) A(t)^*. \quad (18.9)$$

Remark. It is (18.9) that often makes the dynamical analysis of processes of constant reduced state as simple as a statical analysis. For example, (18.9) is at the root of the dynamical analysis of viscometric flows of incompressible fluids (cf. [CMN]) as well as of the incompressible “subfluids” of WANG [W1].

For semi-elastic elements, it is possible to give a lower bound on the size of the symmetry group of a state of monotonous flow of the form $\bar{\rho}(H)$.

Proposition 18.3. *Let H be a monotonous history of exponent E that belongs to the set Φ of deformation histories of a semi-elastic element \mathcal{T} . Let*

$$\text{Comm}(E) = \{B \in \text{Invlin}(\mathcal{T}) \mid BE = EB\} \quad (18.10)$$

be the group of all invertible transformations that commute with E . The symmetry group $\mathcal{G}_{\bar{\rho}(H)}$ then satisfies

$$\mathcal{G}_{\bar{\rho}(H)} \supset \mathcal{G} \cap \text{Orth}(H^f) \cap \text{Comm}(E). \quad (18.11)$$

Proof. If $B \in \text{Orth}(H^f) \cap \text{Comm}(E)$, it is clear from (17.4) that $B^{*-1}HB^{-1} = H$. If also $B \in \mathcal{G}$, it follows from Prop. 15.4 that $\bar{\rho}(H) = \bar{\rho}(B^{*-1}HB^{-1}) = \iota_B(\bar{\rho}(H))$, i.e. that $B \in \mathcal{G}_{\bar{\rho}(H)}$. Q.E.D.

The Propositions 18.2 and 18.3 apply, in particular, to fluid elements, for which $\mathcal{G} \supset \text{Unim}(\mathcal{T})$ (see Section 14). Since $\exp(sE) \in \text{Unim}(\mathcal{T})$ if and only if $\text{tr } E = 0$, we see that for fluids all monotonous deformation histories H in Φ with traceless exponent define a state $\bar{\rho}(H)$ of monotonous flow. The symmetry group of this state satisfies

$$\mathcal{G}_{\bar{\rho}(H)} \supset \text{Orth}(H^f) \cap \text{Comm}(E). \quad (18.12)$$

By Prop. 10.4 the stress $\bar{S}(H) = \hat{S}(\bar{\rho}(H))$ is such that $\bar{S}(H)H^f$ commutes with all transformations in $\text{Orth}(H^f) \cap \text{Comm}(E)$.

19. Material Functionals and Functions for Incompressible Fluids

We recall (Section 9) that a material is defined to be an equivalence class of material elements, the equivalence being material isomorphy. It is often possible to characterize such an equivalence class by means of numerical functionals. We call such functionals *material functionals* for the material in question. Numerical functions and numbers derived from a material functional are called *material*

functions and material constants. For simplicity we consider only the case when the material is an incompressible fluid.

Proposition 19.1. *The set \mathcal{G} of configurations of an incompressible fluid element \mathcal{T} has the form*

$$\begin{aligned}\mathcal{G} &= \{G \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*) \mid \det G_0^{-1} G = 1\} \\ &= \{A^* G_0 A \mid A \in \text{Unim } \mathcal{T}\},\end{aligned}\tag{19.1}$$

where $G_0 \in \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$. The symmetry group g of \mathcal{T} is $g = \text{Unim}(\mathcal{T})$.

Proof. Let $G_0 \in \mathcal{G}$. By definition of a fluid element (Section 14) we have $\text{Unim}(\mathcal{T}) \subset g$. Hence, since $A^* \mathcal{G} A = \mathcal{G}$ for all $A \in g$ (see Definition 9.1 (ii)), \mathcal{G} must contain the set $\mathcal{G}_1 = \{A^* G_0 A \mid A \in \text{Unim}(\mathcal{T})\}$. On the other hand, since the element is incompressible, \mathcal{G} must be contained in the set \mathcal{C} given by (8.1). Now, since $\det(G_0^{-1}(A^* G_0 A)) = (\det A)^2 = 1$ holds if and only if $A \in \text{Unim}(\mathcal{T})$ and since $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) = \{A^* G_0 A \mid A \in \text{InvlIn}(\mathcal{T})\}$ (see Section 2), it follows that $\mathcal{G}_1 = \mathcal{C}$ and hence $\mathcal{G} = \mathcal{G}_1 = \mathcal{C}$, which is (19.1). Also, $A^* \mathcal{G} A = \mathcal{G}$ can hold only if $A \in \text{Unim}(\mathcal{T})$, which implies $g = \text{Unim}(\mathcal{T})$. Q.E.D.

We assume now that \mathcal{G} is given by (19.1) and that a frame space \mathcal{V} has been selected (see Section 5). The set $\mathcal{P}_{\mathcal{V}}$ of all placements corresponding to \mathcal{G} , as defined by (5.1), is easily seen to have the form

$$\mathcal{P}_{\mathcal{V}} = \{K_0 A \mid A \in \text{Unim}(\mathcal{T})\},\tag{19.2}$$

where $K_0 \in \text{Lin}(\mathcal{T}, \mathcal{V})$ is such that $K_0^* K_0 \in \mathcal{G}$.

Given any set Φ of histories, with values in \mathcal{G} and such that $A^* \Phi A = \Phi$ for all $A \in \text{Unim}(\mathcal{T})$, we define

$$\Phi_{\mathcal{V}} = \{K^*{}^{-1} H K^{-1} \mid H \in \Phi, K \in \mathcal{P}_{\mathcal{V}}, K^* K = H(0)\}.\tag{19.3}$$

This set $\Phi_{\mathcal{V}}$ consists of histories with values in $\text{Sym}^+(\mathcal{V}) \cap \text{Unim}(\mathcal{V})$ and all $C \in \Phi_{\mathcal{V}}$ satisfy $C(0) = 1_{\mathcal{V}}$. Moreover, $\Phi_{\mathcal{V}}$ satisfies $Q^T \Phi_{\mathcal{V}} Q = \Phi_{\mathcal{V}}$ for all $Q \in \text{Orth}(\mathcal{V})$, because if $K \in \mathcal{P}_{\mathcal{V}}$ and $K^* K = H(0)$, then also $QK \in \mathcal{P}_{\mathcal{V}}$ and $(QK)^*(QK) = H(0)$ for all $Q \in \text{Orth}(\mathcal{V})$.

Conversely, suppose that $\Phi_{\mathcal{V}}$ is a set of histories with values in $\text{Sym}^+(\mathcal{V}) \cap \text{Unim}(\mathcal{V})$ and such that $C(0) = 1_{\mathcal{V}}$ for all $C \in \Phi_{\mathcal{V}}$ and $Q^T \Phi_{\mathcal{V}} Q = \Phi_{\mathcal{V}}$ for all $Q \in \text{Orth}(\mathcal{V})$. Then

$$\Phi = \{K^* C K \mid C \in \Phi_{\mathcal{V}}, K \in \mathcal{P}_{\mathcal{V}}\}\tag{19.4}$$

defines a set of histories with values in \mathcal{G} and such that $A^* \Phi A = \Phi$ for all $A \in \text{Unim}(\mathcal{T})$.

We say that a mapping $\bar{T}: \Phi_{\mathcal{V}} \rightarrow \text{Sym}(\mathcal{V})$ is an *isotropic functional* if

$$\bar{T}(Q^T C Q) = Q^T \bar{T}(C) Q \quad \text{for all } C \in \Phi_{\mathcal{V}} \text{ and all } Q \in \text{Orth}(\mathcal{V}).\tag{19.5}$$

The following lemma is the key for the construction of material functionals.

Lemma 19.2. *Let Φ be a set of histories with values in \mathcal{G} such that $A^* \Phi A = \Phi$ for all $A \in \text{Unim}(\mathcal{T})$, and let $\Phi_{\mathcal{V}}$ be defined by (19.3). If $\bar{S}: \Phi \rightarrow \text{Sym}(\mathcal{T}^*, \mathcal{T})$ satisfies*

$$\bar{S}(A^* H A) = A^{-1} \bar{S}(H) A^{*-1} \quad \text{for all } H \in \Phi \text{ and all } A \in \text{Unim}(\mathcal{T}),\tag{19.6}$$

then there is a unique isotropic functional $\bar{T}: \Phi_{\mathcal{V}} \rightarrow \text{Sym}(\mathcal{V})$ such that

$$\bar{T}(C) = K\bar{S}(K^*CK)K^* \quad \text{for all } C \in \Phi_{\mathcal{V}} \text{ and all } K \in \mathcal{P}_{\mathcal{V}}. \quad (19.7)$$

Conversely, let $\Phi_{\mathcal{V}}$ be a set of histories with values in $\text{Sym}^+(\mathcal{V}) \cap \text{Unim}(\mathcal{V})$ such that $C(0) = 1_{\mathcal{V}}$ for all $C \in \Phi_{\mathcal{V}}$ and $Q^T\Phi_{\mathcal{V}}Q = \Phi_{\mathcal{V}}$ for all $Q \in \text{Orth}(\mathcal{V})$, and let Φ be defined by (19.4). If $\bar{T}: \Phi_{\mathcal{V}} \rightarrow \text{Sym}(\mathcal{V})$ is an isotropic functional, then there is a unique functional $\bar{S}: \Phi \rightarrow \text{Sym}(\mathcal{T}^*, \mathcal{T})$ which satisfies (19.6) and (19.7).

Proof. Assume Φ and \bar{S} are given and that (19.6) holds. Select $K_0 \in \mathcal{P}_{\mathcal{V}}$ and define \bar{T} by

$$\bar{T}(C) = K_0\bar{S}(K_0^*CK_0)K_0^*. \quad (19.8)$$

Now let $K \in \mathcal{P}_{\mathcal{V}}$ be arbitrary. By (19.2), there is a $A \in \text{Unim}(\mathcal{T})$ such that $K = K_0A$. Using (19.6) and (19.8), we get

$$K\bar{S}(K^*CK)K^* = K_0A\bar{S}(A^*K_0^*CK_0A)A^*K_0^* = \bar{T}(C),$$

i. e., (19.7) is valid. Since $QK_0 \in \mathcal{P}_{\mathcal{V}}$ whenever $Q \in \text{Orth}(\mathcal{V})$, we can substitute QK_0 for K in (19.7). Using (19.8) again, we obtain

$$\bar{T}(C) = QK_0\bar{S}(K_0^*Q^T C Q K_0)K_0^*Q^T = Q\bar{T}(Q^T C Q)Q^T.$$

Hence \bar{T} satisfies (19.5) and hence is an isotropic functional.

Assume now, conversely, that $\Phi_{\mathcal{V}}$ and an isotropic functional $\bar{T}: \Phi_{\mathcal{V}} \rightarrow \text{Sym}(\mathcal{V})$ are given. Let $H \in \Phi$, with Φ defined by (19.4). If $K_1, K_2 \in \mathcal{P}_{\mathcal{V}}$ satisfy $K_1^*K_1 = K_2^*K_2 = H(0)$, we have $K_1^{*-1}HK_1^{-1}, K_2^{*-1}HK_2^{-1} \in \Phi_{\mathcal{V}}$ and there is a $Q \in \text{Orth}(\mathcal{V})$ such that $K_2 = QK_1$. Since \bar{T} is isotropic, we obtain

$$\begin{aligned} K_2^{-1}\bar{T}(K_2^{*-1}HK_2^{-1})K_2^{*-1} &= K_1^{-1}Q^T\bar{T}(QK_1^{*-1}HK_1^{-1}Q^T)QK_1^{*-1} \\ &= K_1^{-1}\bar{T}(K_1^{*-1}HK_1^{-1})K_1^{*-1}. \end{aligned}$$

Therefore it is meaningful to define \bar{S} by

$$\bar{S}(H) = K^{-1}\bar{T}(K^{*-1}HK^{-1})K^{*-1} \quad \text{if } K^*K = H(0), \quad (19.9)$$

because the right side depends only on H , not on the particular choice of K . It is clear that (19.7) holds when \bar{S} is defined by (19.9). For any $A \in \text{Unim}(\mathcal{T})$ we have $(KA)^*(KA) = A^*K^*KA = A^*H(0)A$. Hence (19.9) remains valid when H is replaced by A^*HA and K by KA . It follows that \bar{S} also satisfies (19.6). Q.E.D.

We assume now that \mathcal{T} is an incompressible fluid element, that Φ is the set of all deformation histories for \mathcal{T} (see Definition 15.1), and that \bar{S} is the history functional of \mathcal{T} (see Section 16). It follows from Prop. 16.3 and Prop. 19.1 that \bar{S} satisfies (19.6) and hence that Lemma 19.2 can be applied to it. Thus, given any frame space \mathcal{V} , we can associate with the element \mathcal{T} a set $\Phi_{\mathcal{V}}$ of histories with values in $\text{Sym}^+(\mathcal{V}) \cap \text{Unim}(\mathcal{V})$ and an isotropic functional \bar{T} with domain $\Phi_{\mathcal{V}}$. Since \bar{S} satisfies the normalization condition (8.3) for incompressible elements, it easily follows from (19.7) that the values of \bar{T} must belong to the space

$$\text{Sym}_0(\mathcal{V}) = \{T \in \text{Sym}(\mathcal{V}) \mid \text{tr } T = 0\} \quad (19.10)$$

of traceless symmetric tensors.

If we take for \mathcal{V} the numerical space \mathbb{R}^n , we denote the resulting isotropic functional by \mathfrak{I} and call it the **material functional** of the element. The domain $\Phi_{\mathbb{R}^n}$ of \mathfrak{I} consists of histories whose values are unimodular and positive-definite symmetric matrices. The values of \mathfrak{I} are symmetric matrices with zero trace. By (19.7), the material functional \mathfrak{I} is related to the history functional $\bar{\mathfrak{S}}$ in such a way that

$$\mathfrak{I}(C) = K\bar{\mathfrak{S}}(K^*CK)K^* \quad \text{for all } C \in \Phi_{\mathbb{R}^n} \quad (19.11)$$

and all placements K of \mathcal{T} into the numerical space \mathbb{R}^n .

Proposition 19.3. *Two incompressible fluid elements are materially isomorphic if and only if they have the same material functional.*

Proof. Recall that for semi-elastic elements \mathcal{T} , the structure is determined by the triple $(\mathcal{T}, \Phi, \bar{\mathfrak{S}})$ (see Section 16). Now, if \mathcal{T}_1 and \mathcal{T}_2 are two incompressible fluid elements with the same material functional \mathfrak{I} and if K_1 and K_2 are arbitrary placements of \mathcal{T}_1 and \mathcal{T}_2 into \mathbb{R}^n , then it follows from (19.11) and Prop. 16.2 that $K_2^{-1}K_1$ is a material isomorphism. Conversely, if $A \in \text{Invlin}(\mathcal{T}_1, \mathcal{T}_2)$ is a material isomorphism and K_1 a placement of \mathcal{T}_1 into \mathbb{R}^n , then $K_2 = AK_1$ is a placement of \mathcal{T}_2 and Prop. 16.2 shows that (19.11) defines the same material functional for both \mathcal{T}_1 and \mathcal{T}_2 . Q.E.D.

Using the properties (H1)–(H5) of Φ listed in Section 16, one can easily show that the domain $\Phi_{\mathbb{R}^n}$ of a material functional of an incompressible fluid must have the following properties:

(D1) *Every $C \in \Phi_{\mathbb{R}^n}$ is a history with values in $\text{Sym}^+(\mathbb{R}^n) \cap \text{Unim}(\mathbb{R}^n)$.*

(D2) *For every $C \in \Phi_{\mathbb{R}^n}$, $C(0) = 1_{\mathbb{R}^n}$ and $(1_{\mathbb{R}^n})_{(\infty)} \in \Phi_{\mathbb{R}^n}$.*

(D3) *If $C \in \Phi_{\mathbb{R}^n}$, $s \in \mathbb{R}^+$ and $U \in \text{Unim}(\mathbb{R}^n)$ such that $U^T U = C(s)$, then*

$$U^{T-1} C(s) U^{-1} \in \Phi_{\mathbb{R}^n}.$$

(D4) *If $C, D \in \Phi_{\mathbb{R}^n}$, $s \in \mathbb{R}^+$ and $U \in \text{Unim}(\mathbb{R}^n)$ such that $U^T U = C(s)$, then*

$$U^T D U * C_{[s]} \in \Phi_{\mathbb{R}^n}.$$

(D5) *For all $C_0 \in \text{Sym}^+(\mathbb{R}^n) \cap \text{Unim}(\mathbb{R}^n)$ there is a $C \in \Phi_{\mathbb{R}^n}$ such that $C_0 \in \text{Range}(C)$.*

The properties (F1), (F2), of Section 16, for the history functionals imply that the material functional $\mathfrak{I}: \Phi_{\mathbb{R}^n} \rightarrow \text{Sym}_0(\mathbb{R}^n)$, in addition to being isotropic, must have the following convergence properties.

(M1) *For all $C \in \Phi_{\mathbb{R}^n}$ we have*

$$\lim_{s \rightarrow \infty} \mathfrak{I}(C(s)_{(\infty)} * C_{[s]}) = \mathfrak{I}(C). \quad (19.12)$$

(M2) *For all $C, D \in \Phi_{\mathbb{R}^n}$, $s \in \mathbb{R}^n$ and $U \in \text{Unim}(\mathbb{R}^n)$ such that $U^T U = C(s)$ we have*

$$\lim_{t \rightarrow \infty} \mathfrak{I}(U^t D U * C(s)_{(t)} * C_{[s]}) = \mathfrak{I}(C(s)_{(\infty)} * C_{[s]}). \quad (19.13)$$

Prop. 19.3 states, in essence, that every incompressible fluid is characterized by an isotropic material functional \mathfrak{I} whose domain $\Phi_{\mathbb{R}^n}$ satisfies (D1)–(D5)

and which has the convergence properties (M1) and (M2). Conversely, every functional $\mathfrak{I}: \Phi_{\mathbb{R}^n} \rightarrow \text{Sym}_0(\mathbb{R}^n)$ of this type determines an incompressible fluid material. Indeed, if we select a vector space \mathcal{F} of dimension n and any $G_0 \in \text{Sym}^+(\mathcal{F}, \mathcal{F}^*)$, we can define \mathcal{G} by (19.1), then $\mathcal{P}_{\mathbb{R}^n}$ by (19.2), then Φ by (19.4), and finally \bar{S} by (19.9). The triple $(\mathcal{F}, \Phi, \bar{S})$ then satisfies (H1)–(H5) and (F1), (F2) of Section 16. As remarked there, one can then construct a structure of a semi-elastic material element on \mathcal{F} whose set Φ^* of deformation histories is obtained by saturation of Φ with respect to \bar{S} . The element is incompressible and fluid, and its material functional is obtained from the given \mathfrak{I} by extension from $\Phi_{\mathbb{R}^n}$ to an appropriate saturation of $\Phi_{\mathbb{R}^n}$.

Let \mathfrak{I} be a material functional of an incompressible fluid and assume that the domain of \mathfrak{I} contains all histories of the form $s \mapsto \exp(-sB^T) \exp(-sB)$ with B in the space $\text{Lin}_0(\mathbb{R}^n)$ of all traceless matrices. We then define $t: \text{Lin}_0(\mathbb{R}^n) \rightarrow \text{Sym}_0(\mathbb{R}^n)$ by

$$t(B) = \mathfrak{I}(s \mapsto \exp(-sB^T) (\exp -sB)). \quad (19.14)$$

Since \mathfrak{I} is an isotropic functional, t is an isotropic function in the sense that

$$t(Q^T B Q) = Q t(B) Q^T \quad \text{for all } B \in \text{Lin}_0(\mathbb{R}^n) \text{ and all } Q \in \text{Orth}(\mathbb{R}^n). \quad (19.15)$$

The following result shows that the material function t derived from \mathfrak{I} by (19.14) describes the behavior of the incompressible fluid determined by \mathfrak{I} in all monotonous flow states (see Section 18).

Proposition 19.4. *Let \mathcal{F} be any incompressible fluid element with material functional \mathfrak{I} . Then for every $G \in \mathcal{G}$ and every $E \in \text{Lin}_0(\mathcal{F})$ there is a unique state σ of monotonous flow with exponent E and such that $\hat{G}(\sigma) = G$. The intrinsic stress $\hat{S}(\sigma)$ satisfies*

$$\hat{S}(\sigma) = K^{-1} t(KEK^{-1}) K^{*-1} \quad (19.16)$$

for all $K \in \mathcal{P}_{\mathbb{R}^n}$ such that $K^* K = G$.

Proof. Let $G \in \mathcal{G}$ and $E \in \text{Lin}_0(\mathcal{F})$ be given. Let H be the monotonous history with exponent E and final value $H^f = G$. This history is given by (17.4). We now choose $K \in \mathcal{P}_{\mathbb{R}^n}$ such that $K^* K = G$. It is easily seen that we then have

$$H = K^* C K, \quad \text{where } C = (s \mapsto \exp(-sB^T) \exp(-sB)) \text{ with } B = KEK^{-1}. \quad (19.17)$$

Since $C \in \Phi_{\mathbb{R}^n}$ by assumption, we infer from (19.4) that $H \in \Phi$. As we have seen at the end of Section 18, it follows that $\sigma = \bar{\rho}(H)$ is the state of monotonous flow with exponent E and configuration $\hat{G}(\sigma) = H^f = G$. The relation (19.16) is an immediate consequence of (19.11), (19.14), and (19.17). Q.E.D.

The material function t determines, in particular, the *viscometric functions* τ, σ_1, σ_2 . These are given by

$$\tau(\kappa) = t_{12}(\kappa B_0), \quad \sigma_1(\kappa) = (t_{11} - t_{33})(\kappa B_0), \quad \sigma_2(\kappa) = (t_{22} - t_{33})(\kappa B_0), \quad (19.20)$$

for $\kappa \in \mathbb{R}$. Here t_{ij} , $i, j = 1, 2, 3$ denote the components of t and B_0 is the matrix $B_0 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. As is well known (see, e.g., [N4], Sect. 4), the viscometric functions describe the behavior of the fluid for all simple shearings (see end of Section 17).

20. Materials of the Rate Type

We consider a body element $(\mathcal{T}, \mathcal{G}, \Pi)$ such that Π is the class of piecewise continuously differentiable processes in the sense described in example (c) after Definition 4.1. We use the notation

$$\dot{\mathcal{G}}_G = \{\dot{P}(t) \mid P \in \Pi, t \in [0, d_P], \dot{P}(t) \text{ exists}, P(t) = G\} \quad (20.1)$$

for the class of all possible derivatives of processes at a time at which they have the value G .

We assume that with each $G \in \mathcal{G}$, there is associated a closed subset \mathcal{S}_G of the stress space $\mathcal{S} = \text{Sym}^+(\mathcal{T}^*, \mathcal{T})$ and that we are given a function

$$\mathfrak{H}: \mathcal{D} \rightarrow \mathcal{S}$$

whose domain \mathcal{D} consists of triples (G, \dot{G}, S) as follows:

$$\mathcal{D} = \{(G, \dot{G}, S) \mid G \in \mathcal{G}, \dot{G} \in \dot{\mathcal{G}}_G, S \in \mathcal{S}_G\}. \quad (20.2)$$

We assume that \mathfrak{H} satisfies the following conditions:

- (i) If P is continuously differentiable and $S^i \in \mathcal{S}_{P^i}$, the initial value problem

$$\dot{S}(t) = \mathfrak{H}(P(t), \dot{P}(t), S(t)), \quad S(0) = S^i \quad (20.3)$$

has a unique solution $S: [0, d_P] \rightarrow \mathcal{S}$. We denote the dependence of the final value $S(d_P) = S^f$ on P and on the initial value S^i by

$$S^f = \mathfrak{S}(P, S^i), \quad (20.4)$$

so that $t \mapsto \mathfrak{S}(P_{[0, t]}, S^i)$ is the solution of (20.3).

- (ii) The final value $S^f = \mathfrak{S}(P, S^i)$ belongs to \mathcal{S}_{P^f} and the dependence of the final value on the initial value is uniformly continuous, i.e. the function

$$\mathfrak{S}(P, \cdot): \mathcal{S}_{P^i} \rightarrow \mathcal{S}_{P^f} \quad (20.5)$$

is uniformly continuous for each continuously differentiable $P \in \Pi$.

- (iii) If $G \in \mathcal{G}$ and $S \in \mathcal{S}_G$, then

$$\lim_{t \rightarrow \infty} \mathfrak{S}(G_{(t)}, S) = \hat{\mathfrak{S}}(G, S) \quad (20.6)$$

exists.

If \mathfrak{H} is a reasonably regular function, then one can apply general theorems on ordinary differential equations to verify Conditions (i) and (ii). Condition (iii) says that the solutions of $\dot{S}(t) = \mathfrak{H}(G, 0, S(t))$ have limits as $t \rightarrow \infty$ for every (constant) $G \in \mathcal{G}$. This condition is closely related to one I have described in [N1], pp. 52–53.

We now attempt to use the function \mathfrak{S} for endowing \mathcal{T} with the structure of a material element. For the state space Σ we take

$$\Sigma = \{(G, S) \mid G \in \mathcal{G}, S \in \mathcal{S}_G\}. \quad (20.7)$$

Thus, the states are configuration-stress pairs. The functions \hat{G} and \hat{S} merely assign to each pair its first and second member, so that

$$\hat{G}(G, S) = G, \quad \hat{S}(G, S) = S. \quad (20.8)$$

In order to define the evolution function $\hat{\rho}$, we first note that there is a natural way to extend the solution functional \mathfrak{S} of (20.3) in such a way that $\mathfrak{S}(P, S^i)$ becomes meaningful for *all* $P \in \Pi$ with $S^i \in \mathcal{S}_{P^i}$, not just those that have a continuous derivative for all $t \in [0, d_P]$. The general member $P \in \Pi$ is of the form

$$P = P_1 * P_2 * \cdots * P_m,$$

where the P_k are continuously differentiable. We then define $\mathfrak{S}(P, S^i)$ inductively by

$$\mathfrak{S}(P_1 * \cdots * P_{m-1} * P_m, S^i) = \mathfrak{S}(P_m, \mathfrak{S}(P_1 * \cdots * P_{m-1}, S^i)). \quad (20.9)$$

With \mathfrak{S} thus extended, it makes sense to define $\hat{\rho}$ by

$$\hat{\rho}((G, S), P) = (P^f, \mathfrak{S}(P, S)), \quad P \in \Pi_G, S \in \mathcal{S}_G. \quad (20.10)$$

The response functional $\tilde{S} = \hat{S} \circ \hat{\rho}$ satisfies

$$\tilde{S}((G, S), P) = \mathfrak{S}(P, S) \quad \text{if } P^i = G \quad (20.11)$$

and hence is essentially the same as the solution functional of (20.3).

It is very easy to verify that the septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ thus constructed satisfies Axioms I–III. We note that for all $G \in \mathcal{G}$, there is a natural one-to-one correspondence $S \leftrightarrow (G, S)$ between \mathcal{S}_G and the G -section Σ_G of the state space. It is not hard to prove, using Condition (ii) and (20.11), that the natural uniformity on Σ_G (see Definition 11.1) is identical to the uniformity on Σ_G induced by the natural uniformity of $\mathcal{S}_G \subset \mathcal{S}$. Since \mathcal{S}_G was assumed to be closed in \mathcal{S} and since \mathcal{S} is complete, it follows that \mathcal{S}_G and hence Σ_G is complete. Thus, Axiom IV, Section 11, is satisfied. The validity of the relaxation axiom, Axiom V, Section 12, is now an immediate consequence of Condition (iii), and it follows from (20.6) that the relaxation mapping $\hat{\lambda}$ is given by

$$\hat{\lambda}(G, S) = (G, \hat{\mathfrak{S}}(G, S)). \quad (20.12)$$

We have proved that $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma, \hat{G}, \hat{S}, \hat{\rho})$ satisfies Axioms I–V. The accessibility axiom (Axiom VI, Section 13), however, will *not* be satisfied unless the sets \mathcal{S}_G of stresses compatible with a given configuration are properly selected. Indeed Axiom VI is equivalent to the requirement that the \mathcal{S}_G be of the form

$$\mathcal{S}_G = \text{cl} \{ \mathfrak{S}(P, S_0) \mid P^i = G_0, P^f = G \}, \quad (20.13)$$

where $G_0 \in \mathcal{G}$ and $S_0 \in \mathcal{S}_{G_0}$ such that $\hat{\mathfrak{S}}(G_0, S_0) = S_0$.

If Axiom VI is not satisfied, one can apply the procedure described at the end of Section 13 to construct material element structures on \mathcal{T} . In the present

case, this procedure goes as follows: Select $G_0 \in \mathcal{G}$ and $S_1 \in \mathcal{S}_{G_1}$ arbitrarily and compute $S_0 = \hat{\mathfrak{C}}(G_0, S_1)$ using (20.7). Define \mathcal{S}'_G to be the right side of (20.13) and then determine $(\Sigma', \hat{G}', \hat{S}', \hat{\rho}')$ in terms of the \mathcal{S}'_G in the same way as $(\Sigma, \hat{G}, \hat{S}, \hat{\rho})$ were determined in terms of the \mathcal{S}_G . The septuple $(\mathcal{T}, \mathcal{G}, \Pi, \Sigma', \hat{G}', \hat{S}', \hat{\rho}')$ then satisfies all axioms and hence defines the structure of a material element on \mathcal{T} .

A material element obtained from a function \mathfrak{H} satisfying Conditions (i)–(iii) in the manner just described is called a **material element of the rate type**. As we have seen, one and the same function \mathfrak{H} may define more than one material element structure on \mathcal{T} .

We now consider the case when \mathcal{S}_G is a subspace (rather than merely a subset) of $\text{Sym}(\mathcal{T}, \mathcal{T}^*)$ and when the given function \mathfrak{H} is linear in its second variable, *i. e.* if

$$\mathfrak{H}(G, \cdot, S) \in \text{Lin}(\mathcal{S}_G, \mathcal{S}) \quad (20.14)$$

for every $G \in \mathcal{G}$ and $S \in \mathcal{S}_G$. A material element whose structure is defined by such a function will be called a **hypo-elastic element**¹. Since (20.14) implies that $\mathfrak{H}(G, 0, S) = 0$, it follows that the solutions of $\dot{S}(t) = \mathfrak{H}(G, 0, S(t))$ are constants and hence that Condition (iii) is automatically satisfied with

$$\hat{\mathfrak{C}}(G, S) = S \quad (20.15)$$

for all $G \in \mathcal{G}$ and $S \in \mathcal{S}_G$. In view of (20.12) it follows that *for a hypo-elastic element all states are relaxed*. Thus, a hypo-elastic element cannot be semi-elastic unless it is elastic. If the element is elastic, then the sets \mathcal{S}_G must be singletons $\mathcal{S}_G = \{\hat{S}(G)\}$. It can be proved that the elastic elements which are also hypo-elastic are those whose stress function \hat{S} satisfies certain differentiability and invertibility conditions².

When (20.14) holds, then the solution functional \mathfrak{S} of (20.4) is easily seen to have the property

$$S^f = \mathfrak{S}(P, S^i) \Leftrightarrow S^i = \mathfrak{S}(P_{\text{rev}}, S^f), \quad (20.16)$$

where the *reversed process* P_{rev} is defined by $P_{\text{rev}}(t) = P(d_P - t)$, $t \in [0, d_P]$. It follows from (20.16), as is easy to show, that accessibility (see Section 13) becomes an equivalence relation on Σ when Σ is defined by (20.7). Even when \mathfrak{H} is a very simple function there can be more than one equivalence class, as is shown by an example due to BERNSTEIN ([B1]; see also [TN], Sect. 100). Axiom V is satisfied only after the \mathcal{S}_G are replaced by subsets \mathcal{S}'_G which consist of all those $S \in \mathcal{S}_G$ for which (G, S) belongs to a particular equivalence class in Σ . The state space Σ' defined by (20.7) in terms of the \mathcal{S}'_G then is the same as this particular equivalence class. Thus, each of the equivalence classes in Σ defines on \mathcal{T} the structure of a hypo-elastic element³.

Remark 1. In the past literature (*cf.* [N2], Sect. 24 or [TN], Sect. 36) materials of the rate type have been defined by differential equations involving Cauchy-stresses and motions in a frame space rather than intrinsic stresses and deformation

¹ This notion of hypo-elasticity is more inclusive than the original one of TRUESDELL [T].

² The method of proof is the same as the one used, in a special situation, in Sect. 15b of [N1].

³ This procedure was first used by BERNSTEIN [B2] to define a concept of "material" in hypo-elasticity.

processes. It is instructive to consider the following special non-intrinsic constitutive equation of the rate type:

$$T = \mathfrak{G}(T, A_1). \tag{20.17}$$

Here, T is the Cauchy-stress, $\overset{*}{T}$ any invariant time flux of T , and A_1 the first Rivlin-Ericksen tensor. Apart from a possible dependence on density, (20.17) is the constitutive equation for what I called “hygrosteric” materials in [N1]. If \mathfrak{G} is linear in its second variable, then (20.17) is the constitutive equation of hypo-elasticity in the original sense of TRUESDELL ([T]; see also [TN], Chapter DIV).

It will be convenient to choose for $\overset{*}{T}$ the time flux given by

$$\overset{*}{T} = \dot{T} - LT - TL^T, \tag{20.18}$$

where $L(t) = \dot{F}_{(t)}(t)$, with $F_{(t)}$ defined by (5.5). Using (5.7), (6.2), (5.5), and (20.18), we can express A_1 , T , and $\overset{*}{T}$ in terms of the motion M , the intrinsic stress S , and the process $P = M^* M$:

$$A_1 = M^{*-1} \dot{P} M^{-1}, \quad T = MSM^*, \quad \overset{*}{T} = M \dot{S} M^*. \tag{20.19}$$

Now, (20.17) can serve to define material elements of the rate type in the intrinsic sense only if, after substitution of (20.19), equation (20.17) becomes a special case of (20.3)₁. This is the case if and only if there is a function \mathfrak{H} of the type described at the start of this section such that

$$\mathfrak{H}(G, \dot{G}, S) = K^{-1} \mathfrak{G}(KSK^*, K^{*-1} \dot{G} K^{-1}) K^{*-1} \tag{20.20}$$

holds for all (G, \dot{G}, S) in the domain of \mathfrak{H} and all placements K such that $K^* K = G$. Since $K^* K = K_1^* K_1 = G$ is valid if and only if $K_1 = QK$ for some $Q \in \text{Orth}(\mathcal{V})$, it follows that (20.20) can be valid for all K with $K^* K = G$ only if \mathfrak{G} is an isotropic function, *i. e.* if

$$\mathfrak{G}(QTQ^T, QA_1Q^T) = Q \mathfrak{G}(T, A_1) Q^T \tag{20.21}$$

for all $Q \in \text{Orth}(\mathcal{V})$ and all (T, A_1) in the domain of \mathfrak{G} . We have thus recovered the well known result ([N1], Sect. 7) that (20.17) cannot define a material unless \mathfrak{G} is isotropic.

If (20.20) holds, then \mathfrak{H} satisfies

$$\mathfrak{H}(A^* GA, A^* \dot{G} A, A^{-1} SA^{*-1}) = A^{-1} \mathfrak{H}(G, \dot{G}, S) A^{*-1} \tag{20.22}$$

for all (G, \dot{G}, S) in the domain \mathcal{D} of \mathfrak{H} and all A in the set

$$\mathcal{g} = \{A \in \text{Invlin}(\mathcal{T}) \mid (G, \dot{G}, S) \in \mathcal{D} \Rightarrow (A^* GA, A^* \dot{G} A, A^{-1} SA^{*-1}) \in \mathcal{D}\}. \tag{20.23}$$

Indeed, if we replace K by KA in (20.20) so that $G = K^* K$ becomes $(KA)^*(KA) = A^* K^* KA = A^* GA$, if we replace also \dot{G} by $A^* \dot{G} A$ and S by $A^{-1} SA^{*-1}$, we see that (20.22) must hold. Now, we recall that in order to define a material element structure by means of \mathfrak{H} , one may have to replace the sets \mathcal{S}_G by suitable subsets. In view of (20.2), this means that \mathcal{D} has to be replaced by a suitable subset. After

this has been done, (20.23) can easily be shown to be the symmetry group of the element.

One and the same constitutive equation (20.17) may determine several material element structures on \mathcal{T} , and these structures need not all have the same symmetry group.

Remark II. More general materials of the rate type are obtained if the initial value problem (20.3) is replaced by an initial value problem for a differential equation of higher order. No new insight is gained by considering such materials.

21. Outlook

We give here a list of possible generalizations, further developments, and applications of the ideas presented in this paper.

(a) Abstract theory of systems with memory. The basic features of the theory presented here can be used to define a very general concept of *physical systems with memory*. In the definition of such a system, one uses as primitive notions a set \mathcal{G} of *abstract configurations* and an *action space* \mathcal{S} , whose elements are assumed to describe the possible actions of the system on the environment. A system with memory is then defined to be a septuple $(\mathcal{G}, \Pi, \Sigma, \hat{G}, \mathcal{S}, \hat{S}, \hat{\rho})$ whose entries $\Pi, \Sigma, \hat{G}, \hat{S}$, and $\hat{\rho}$ are as described in Definitions 4.1 and 7.1. The only difference is that \mathcal{G} and \mathcal{S} are now abstract sets which need not be defined concretely in terms of a vector space \mathcal{T} .

Axioms I–VI remain in force. It must be assumed that the action space \mathcal{S} is endowed *a priori* with a uniform topology. This is necessary in order that Definition 11.1 and hence Axioms IV–VI remain meaningful.

A material element in the sense of Definition 7.1 is a very special kind of physical system with memory.

Note Added in Proof. The structure of an abstract system with memory described here has many features in common with the structure used in “mathematical systems theory”. (See, e.g., WILLEMS, J. C.: Dissipative dynamical systems I. Arch. Rational Mech. Anal. 45, 322–351 (1972), and the literature cited there.)

(b) Thermodynamic theory of material elements. To include thermal and energetic phenomena, one should modify Definition 4.1 by assuming that \mathcal{G} is a closed and connected subset of $\text{Sym}^+(\mathcal{T}, \mathcal{T}^*) \times \mathbb{R}^{++} \times \mathcal{T}^*$, where \mathbb{R}^{++} denotes the set of all positive reals. The members of \mathcal{G} are triples $G = (C, \theta, g)$, where C is interpreted as a configuration, θ as a temperature of the material element, and g as a temperature gradient across the element.

In Definition 7.1, Equation (7.1) should be replaced by

$$\mathcal{S} = \text{Sym}(\mathcal{T}^*, \mathcal{T}) \times \mathbb{R} \times \mathcal{T} \times \wedge^n \mathcal{T}^*, \quad (21.1)$$

where $\wedge^n \mathcal{T}^*$, the n^{th} exterior power of \mathcal{T}^* ($n = \dim \mathcal{T}$), is the one-dimensional space of possible volumes of the element. The members of \mathcal{S} are quadruples (S, η, h, ψ) , where S is interpreted as an intrinsic stress, η as an entropy per unit volume, h as an intrinsic heat flux, and ψ as a free energy.

Let $P = (\bar{C}, \bar{\theta}, \bar{g}) : [0, d_P] \rightarrow \mathcal{G}$ be a process (i.e. $P \in \Pi$), let $\sigma \in \Sigma$ be a state such that $P^i = \hat{G}(\sigma)$ and let

$$(\bar{S}, \bar{\eta}, \bar{h}, \bar{\psi}) : [0, d_P] \rightarrow \mathcal{S} \tag{21.2}$$

be defined with the help of the response functional \tilde{S} (see (7.5)) by

$$(\bar{S}(t), \bar{\eta}(t), \bar{h}(t), \bar{\psi}(t)) = \tilde{S}(\sigma, P_{[0, t]}), \quad t \in [0, d_P]. \tag{21.3}$$

One then defines the *rate of entropy production* $\bar{\gamma} : [0, d_P] \rightarrow \hat{\Lambda} \mathcal{T}^*$ corresponding to σ and P by

$$\bar{\theta} \bar{\gamma} = v_{\bar{c}} \left(\frac{1}{2} \text{tr}(\bar{S} \dot{\bar{C}}) - \bar{\eta} \dot{\bar{\theta}} + \frac{1}{\bar{\theta}} \langle \bar{g}, \dot{\bar{h}} \rangle \right) - \dot{\bar{\psi}}, \tag{21.4}$$

where $v_{\bar{c}}(t) = v_{\bar{c}(t)}$ is the volume in $\hat{\Lambda} \mathcal{T}^*$ corresponding to the configuration $\bar{C}(t)$.

In a detailed development of a thermodynamic theory of material elements, the Definition (21.4) should play a fundamental role.

I believe, also, that the framework just described can serve to simplify and clarify the axiomatic approach to thermodynamics developed by W. A. DAY [D].

(c) Theories of plastic behavior. D. R. OWEN ([O1], [O2]) has developed, within the framework of the old theory of simple materials, concepts and results intended to describe in mathematical terms what physicists usually call plastic behavior. I believe that the new theory presented here provides a more natural setting for such concepts and that it can be used to simplify and clarify OWEN's ideas.

We note that the theory of states of monotonous flow given in Section 18 applies not only to semi-elastic materials, but also to materials that exhibit plastic behavior, provided the materials possess sufficient symmetry. I expect, therefore, that one can develop concrete descriptions of plastic monotonous flows, in particular viscometric flows.

(d) Theories of fading memory. In the framework of the old concept of simple material, several theories of fading memory are known (see [TN], Chapter CV; also [CM] and the literature cited there). Analogous theories for the new concept given here remain to be created. Such theories should involve continuity assumptions imposed upon the response functional

$$\tilde{S}(\sigma, \cdot) : \Pi_G \rightarrow \mathcal{S} \quad (G = \hat{G}(\sigma)). \tag{21.5}$$

Suitable topologies on Π_G , which are needed to make continuity meaningful, can be defined in terms of a natural metric on the configuration space \mathcal{G} . I shall demonstrate the existence of such a natural metric in a future paper.

For semi-elastic materials, it should not be too difficult to translate the known theories of fading memory into the intrinsic setting of this paper. It would be interesting if one could also create theories of fading memory for non-semi-elastic materials.

Acknowledgement. The research leading to this paper was supported by the National Science Foundation under Grant GP-28430. I am grateful to Dr. F. BLOOM for help in proofreading.

References

- [B1] BERNSTEIN, B., Hypo-elasticity and elasticity. Arch. Rational Mech. Anal. **6**, 89–104 (1960).
- [B2] BERNSTEIN, B., Remarks on materials of the rate type and the principle of determinism. Trans. Soc. Rheol. **5**, 35–40 (1961).
- [BK] BOURBAKI, N., Topologie Générale, Chapitres 1 et 2, quatrième édition. Paris: Hermann 1965.
- [C] COLEMAN, B. D., Kinematical concepts with applications in the mechanics and thermodynamics of incompressible viscoelastic fluids. Arch. Rational Mech. Anal. **9**, 273–300 (1962).
- [CM] COLEMAN, B. D., & V. J. MIZEL, On the general theory of fading memory. Arch. Rational Mech. Anal. **29**, 18–31 (1968).
- [CMN] COLEMAN, B. D., H. MARKOVITZ, & W. NOLL, Viscometric Flows of Non-Newtonian Fluids. New York: Springer 1966.
- [CN1] COLEMAN, B. D., & W. NOLL, On the thermostatics of continuous media. Arch. Rational Mech. Anal. **4**, 97–128 (1959).
- [CN2] COLEMAN, B. D., & W. NOLL, Material symmetry and thermostatic inequalities in finite elastic deformations. Arch. Rational Mech. Anal. **15**, 87–111 (1964).
- [D] DAY, W. A., Thermodynamics based on a work axiom. Arch. Rational Mech. Anal. **31**, 1–34 (1968).
- [FT] FAIRCHILD, W. W., & C. IONESCU-TULCEA, Topology. Philadelphia-London-Toronto: W. B. Saunders 1971.
- [K] KELLEY, J. L., General Topology. Princeton: Van Nostrand 1955.
- [N1] NOLL, W., On the continuity of the solid and fluid states. J. Rational Mech. Anal. **4**, 3–81 (1955).
- [N2] NOLL, W., A mathematical theory of the mechanical behavior of continuous media. Arch. Rational Mech. Anal. **2**, 197–226 (1958).
- [N3] NOLL, W., The foundations of classical mechanics in the light of recent advances in continuum mechanics. Proc. of the Berkeley Symposium on the Axiomatic Method, p. 266–281, Amsterdam 1959.
- [N4] NOLL, W., Motions with constant stretch history. Arch. Rational Mech. Anal. **11**, 97–105 (1962).
- [N5] NOLL, W., Proof of the maximality of the orthogonal group in the unimodular group. Arch. Rational Mech. Anal. **18**, 100–102 (1965).
- [N6] NOLL, W., Space-time structures in classical mechanics. Delaware Seminar in the Foundations of Physics, p. 28–34. Berlin-Heidelberg-New York: Springer 1967.
- [N7] NOLL, W., Materially uniform simple bodies with inhomogeneities. Arch. Rational Mech. Anal. **27**, 1–32 (1967).
- [N8] NOLL, W., The representation of monotonous processes by exponentials. To appear.
- [O1] OWEN, D. R., Thermodynamics of materials with elastic range. Arch. Rational Mech. Anal. **32**, 91–112 (1968).
- [O2] OWEN, D. R., A mechanical theory of materials with elastic range. Arch. Rational Mech. Anal. **37**, 85–110 (1970).
- [T] TRUESDELL, C., Hypo-elasticity. J. Rational Mech. Anal. **4**, 83–133 (1955).
- [TN] TRUESDELL, C., & W. NOLL, The Non-linear Field Theories of Mechanics. Encyclopedia of Physics, Vol. III/3. Berlin-Heidelberg-New York: Springer 1965.
- [S] STERNBERG, S., Lectures on Differential Geometry. Englewood Cliffs: Prentice Hall 1964.
- [W1] WANG, C. C., A general theory of subfluids. Arch. Rational Mech. Anal. **20**, 1–40 (1965).
- [W2] WANG, C. C., On the geometric structures of simple bodies, a mathematical foundation for the theory of continuous distributions of dislocations. Arch. Rational Mech. Anal. **27**, 33–94 (1967).

Department of Mathematics
Carnegie-Mellon University
Pittsburgh

(Received May 12, 1972)

*Lectures on the Foundations
of Continuum Mechanics and Thermodynamics*

WALTER NOLL

Expository Memoir, invited by the Editors

Lectures on the Foundations of Continuum Mechanics and Thermodynamics

WALTER NOLL

Expository Memoir, invited by the Editors

Contents

1. Introduction	62
I. Kinematics of Continuous Bodies	
2. Euclidean spaces	63
3. Neo-classical space-time	65
4. Kinematical processes, frames, motions	66
5. Differential calculus in Euclidean spaces	69
6. Continuous bodies	70
7. Configurations, deformations	72
II. Thermomechanics of Continuous Bodies	
8. Interactions	75
9. Equilibrium of forces, mechanical processes	79
10. Thermodynamic processes, entropy production	82
11. Constitutive assumptions, the dissipation principle	84
Appendix: The concept of a material universe	87
References	92

1. Introduction

This is an outline of a series of lectures I delivered at the Technion in Haifa, Israel, in the summer of 1972. This outline gives my view of what the basic concepts of modern continuum mechanics and thermodynamics are and how they should be presented to graduate students of mathematics and theoretical mechanics. This view evolved gradually over the past eight years. A first version was given in a series of lectures I delivered at the Summer Session in Bressanone, Italy, in 1965 (reference [1]). The material was reworked several times for an introductory graduate course repeatedly given at Carnegie-Mellon University and for lecture series given at the University of Karlsruhe, Germany, in 1968 and in Jabłonna, Poland, in 1970. This paper is the latest version, and it renders reference [1] obsolete.

The basic mathematical notations and tools used in the present paper, mainly from linear algebra, are described in Section 2 of reference [2].

I. Kinematics of Continuous Bodies

2. Euclidean Spaces

A *metric* on a set \mathcal{E} is a function

$$d: \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}^+$$

which associates with each pair (x, y) of points in \mathcal{E} a number $d(x, y)$, called the *distance* from x to y . Euclidean geometry can be based entirely on an appropriate metric.

The *group of isometries* \mathcal{I} for the metric d is the set of all invertible mappings of \mathcal{E} onto itself that leave distances unchanged:

$$\mathcal{I} = \{ \alpha: \mathcal{E} \rightarrow \mathcal{E} \mid \alpha \text{ invertible, } d(x, y) = d(\alpha(x), \alpha(y)) \text{ for all } x, y \in \mathcal{E} \}.$$

It is clear that \mathcal{I} is a group under composition.

If the metric d is to describe Euclidean geometry it must have certain properties. It is intuitively clear that translations are isometries, and we characterize Euclidean metrics in terms of the translations as follows:

It may happen that \mathcal{I} contains a subgroup \mathcal{V} with the following properties:

- (E1) \mathcal{V} is commutative, *i.e.*, $u \circ v = v \circ u$ for all $u, v \in \mathcal{V}$.
- (E2) \mathcal{V} is transitive, *i.e.*, for any $x, y \in \mathcal{E}$ there is a $u \in \mathcal{V}$ such that $u(x) = y$.
- (E3) \mathcal{V} acts freely, *i.e.*, if $v(x) = x$ for some $x \in \mathcal{E}$, then $v = 1_{\mathcal{E}}$.
- (E4) There is a scalar multiplication $\mathbb{R} \times \mathcal{V} \rightarrow \mathcal{V}$ which makes \mathcal{V} a vector space when composition in \mathcal{V} is taken as addition, and there is an inner product I on the vector space \mathcal{V} such that

$$d(x, y)^2 = u \cdot u = \langle Iu, u \rangle \quad \text{when } u(x) = y.$$

The properties (E1)–(E4) are intuitively plausible when \mathcal{V} is the group of translations.

Uniqueness Theorem. *There is at most one subgroup \mathcal{V} of \mathcal{I} with the properties (E1)–(E4). Moreover, if \mathcal{V} is such a subgroup, the scalar multiplication and the inner product on \mathcal{V} required for (E4) are uniquely determined.*

The proof of this theorem will not be given; it is not trivial but not too difficult (see, *e.g.*, reference [3]).

Definition. The function $d: \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}^+$ is called a *Euclidean metric* on \mathcal{E} if \mathcal{I} has a subgroup \mathcal{V} with the properties (E1)–(E4). If a Euclidean metric d is prescribed, we say that d endows \mathcal{E} with the structure of a *Euclidean space*.

The inner product space \mathcal{V} determined by d is then called the *translation space* of \mathcal{E} .

Let \mathcal{E} be a Euclidean space with translation space \mathcal{V} . The following notations are useful:

- (i) if $u, v \in \mathcal{V}$, write $u + v$ for $v \circ u$,
- (ii) write $0 \in \mathcal{V}$ for $1_{\mathcal{E}}$,
- (iii) if $u \in \mathcal{V}$, write $-u$ for u^{-1} .

(iv) if $x \in \mathcal{E}$ and $u \in \mathcal{V}$, write $x+u$ for $u(x)$.

(v) if $x, y \in \mathcal{E}$, write $u=y-x$ for the unique $u \in \mathcal{V}$ for which $u(x)=y$.

If these notations are used, the rules of the ordinary algebra of addition and subtraction are valid.

Let \mathcal{E}_1 and \mathcal{E}_2 be Euclidean spaces, with metrics d_1 and d_2 and with translation spaces \mathcal{V}_1 and \mathcal{V}_2 . An invertible mapping $\alpha: \mathcal{E}_1 \rightarrow \mathcal{E}_2$ is called an *isomorphism* from \mathcal{E}_1 onto \mathcal{E}_2 if it preserves distance, *i. e.* if

$$d_1(x, y) = d_2(\alpha(x), \alpha(y))$$

for all $x, y \in \mathcal{E}_1$. An isomorphism from \mathcal{V}_1 onto \mathcal{V}_2 is a linear mapping which preserves the inner product. The following result describes the isomorphisms of Euclidean spaces:

Representation Theorem. *Every isomorphism $\alpha: \mathcal{E}_1 \rightarrow \mathcal{E}_2$ from one Euclidean space onto another induces an isomorphism $\bar{\alpha}: \mathcal{V}_1 \rightarrow \mathcal{V}_2$ between the corresponding translation spaces in such a way that*

$$\alpha(x) - \alpha(q) = \bar{\alpha}(x - q) \quad (1)$$

holds for all $x, q \in \mathcal{E}_1$.

Proof. We define $\mathcal{V}'_2 = \{\alpha \circ u \circ \alpha^{-1} \mid u \in \mathcal{V}_1\}$. It is easily seen that \mathcal{V}'_2 is a subgroup of the group \mathcal{S}_2 of isometries of \mathcal{E}_2 . It is also easily verified that \mathcal{V}'_2 satisfies the conditions (E1)–(E3). We define a scalar multiplication on \mathcal{V}'_2 by putting $a(\alpha \circ u \circ \alpha^{-1}) = \alpha \circ (au) \circ \alpha^{-1}$ for $a \in \mathbb{R}$, $u \in \mathcal{V}_1$ and an inner product on \mathcal{V}'_2 by putting

$$(\alpha \circ u \circ \alpha^{-1}) \cdot (\alpha \circ v \circ \alpha^{-1}) = u \cdot v \quad \text{when } u, v \in \mathcal{V}_1.$$

It then becomes clear that (E4) is also satisfied by \mathcal{V}'_2 . By the uniqueness theorem, it follows that $\mathcal{V}_2 = \mathcal{V}'_2$. Moreover, the mapping $\bar{\alpha}: \mathcal{V}_1 \rightarrow \mathcal{V}_2$ given by

$$\bar{\alpha}(u) = \alpha \circ u \circ \alpha^{-1}, \quad u \in \mathcal{V}_1 \quad (2)$$

is an isomorphism from \mathcal{V}_1 onto \mathcal{V}_2 . Equation (2) is equivalent to

$$\bar{\alpha}(u)(\alpha(q)) = \alpha(u(q)) \quad \text{for all } q \in \mathcal{E}_1, u \in \mathcal{V}_1.$$

If we put $u = x - q$ and use the notations (i)–(v) above, we get the desired result. Q. E. D.

The case when $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E}$ gives rise to the following

Corollary. *To every isometry $\alpha \in \mathcal{S}$ of a Euclidean space \mathcal{E} with translation space \mathcal{V} corresponds a unique $Q_\alpha \in \text{Orth}(\mathcal{V})$ such that*

$$\alpha(x) - \alpha(q) = Q_\alpha(x - q) \quad (3)$$

for all $x, q \in \mathcal{E}$.

The dimension of a Euclidean space is defined to be the dimension of its translation space. Finite-dimensional Euclidean spaces are isomorphic if and only if they have the same dimension.

3. Neo-Classical Spacetime

A basic primitive notion of any physical theory is an *event-world* \mathcal{W} . It is a set whose elements e, f, \dots are called *events*. Intuitively, events are actual or possible “atoms of experience” and the event world is a mathematical idealization of experience.

The classical time concept can be described by a *time-lapse function*

$$\hat{i}: \mathcal{W} \times \mathcal{W} \rightarrow \mathbb{R},$$

whose value $\hat{i}(e, f)$ gives the time-lapse between the events e and f , as measured by a stop watch. We agree to count $\hat{i}(e, f)$ as positive when e is earlier than f , negative when e is later than f . Our intuitive idea of time requires that \hat{i} have the following properties:

$$(T1) \quad \hat{i}(e, f) = -\hat{i}(f, e) \text{ for all } e, f \in \mathcal{W}.$$

$$(T2) \quad \hat{i}(e, f) + \hat{i}(f, g) = \hat{i}(e, g) \text{ for all } e, f, g \in \mathcal{W}.$$

$$(T3) \quad \text{For every } e \in \mathcal{W} \text{ and every } t \in \mathbb{R} \text{ there is a } f \in \mathcal{W} \text{ such that } \hat{i}(e, f) = t.$$

We say that e is *earlier* than, *later* than, or *simultaneous* with f according to whether $\hat{i}(e, f) > 0$, < 0 , or $= 0$. Simultaneity is easily seen to be an equivalence relation on \mathcal{W} . Its graph is the set

$$\mathcal{S} = \{(e, f) \in \mathcal{W} \times \mathcal{W} \mid \hat{i}(e, f) = 0\} \quad (1)$$

of all pairs of simultaneous events. The simultaneity relation determines a partition Γ of \mathcal{W} into classes τ of simultaneous events so that

$$\mathcal{W} = \bigcup \{\tau \mid \tau \in \Gamma\}, \quad \mathcal{S} = \bigcup \{\tau \times \tau \mid \tau \in \Gamma\}, \quad (2)$$

the unions being disjoint.

The classes $\tau \in \Gamma$ are called *instants*. If $e \in \tau$ we say that *the event e happens at the instant τ* .

The time-lapse $\hat{i}(e, f)$ depends only on the instants τ and σ at which e and f happen. Therefore one can define a *time-lapse function*

$$\bar{i}: \Gamma \times \Gamma \rightarrow \mathbb{R}$$

between instants by

$$\bar{i}(\tau, \sigma) = \hat{i}(e, f) \quad \text{if } e \in \tau, f \in \sigma.$$

The absolute value $|\bar{i}|$ can easily be shown to be a Euclidean metric on Γ . The corresponding translation space of Γ can be identified with the reals \mathbb{R} . If the notations of Section 2 are used, we then have

$$\tau - \sigma = \bar{i}(\sigma, \tau), \quad \sigma + t = \tau \quad \text{if } t = \bar{i}(\sigma, \tau)$$

for $\tau, \sigma \in \Gamma$, $t \in \mathbb{R}$.

Distances between events, as measured by measuring rods, are meaningful only when the events are simultaneous, because a measurement has to take place at a particular instant. Thus, distance measurements are described by a function

$$\hat{d}: \mathcal{S} \rightarrow \mathbb{R}^+,$$

where \mathcal{S} is the set (1) of all simultaneous pairs of events. Experience with distance measurements shows that \hat{d} should have the following basic properties:

- (D1) For each instant $\tau \in \Gamma$, the restriction $d_\tau: \tau \times \tau \rightarrow \mathbb{R}^+$ of \hat{d} to $\tau \times \tau \subset \Gamma$ is a Euclidean metric and hence endows τ with the structure of a Euclidean space.
- (D2) For each instant $\tau \in \Gamma$ the dimension of the translation space \mathcal{V}_τ of τ is 3.

Definition. We say that a set \mathcal{W} is a *neo-classical event-world* if it is endowed with the structure defined by a prescribed time-lapse function \hat{t} with the properties (T1)–(T3) and a prescribed distance function \hat{d} with the properties (D1), (D2).

The internal symmetry of a neo-classical event-world \mathcal{W} is described by its group of automorphisms. An automorphism α of \mathcal{W} is an invertible mapping of \mathcal{W} onto itself which preserves time lapses and distances, i. e. satisfies

$$\begin{aligned} t(\alpha(e), \alpha(f)) &= \hat{t}(e, f) & \text{for all } e, f \in \mathcal{W}, \\ \hat{d}(\alpha(e), \alpha(f)) &= \hat{d}(e, f) & \text{for all } (e, f) \in \mathcal{S}. \end{aligned}$$

The following result is an easy consequence of the Representation Theorem of Section 2.

Theorem. *To every automorphism α of a neo-classical event world \mathcal{W} corresponds a unique number $t_\alpha \in \mathbb{R}$ and to each instant $\tau \in \Gamma$ a unique isomorphism $\bar{\alpha}_\tau: \mathcal{V}_\tau \rightarrow \mathcal{V}_{\tau+t_\alpha}$ such that $\alpha(e) \in \tau + t_\alpha$ if $e \in \tau$ and*

$$\alpha(e) - \alpha(f) = \bar{\alpha}_\tau(e - f) \quad \text{if } e, f \in \tau.$$

4. Kinematical Processes, Frames, Motions

A *material system* \mathcal{B} is a set of material points X, Y, \dots . These are idealizations of material objects or markings in or on material objects.

From now on we assume a neo-classical event-world \mathcal{W} given once and for all. By a *time-interval* \mathcal{T} we mean a connected subset of the set Γ of all instants, connected in the sense that if $\tau \in \mathcal{T}$ and $\tau + t \in \mathcal{T}$, then also $\tau + s \in \mathcal{T}$ for all $s \in [0, t]$.

A mapping

$$\chi: \mathcal{B} \times \mathcal{T}_x \rightarrow \mathcal{W},$$

where \mathcal{T}_x is a time-interval, will be called a **kinematical process** if

$$\chi(X, \tau) \in \tau \quad \text{for all } X \in \mathcal{B} \text{ and all } \tau \in \mathcal{T}_x.$$

The event $\chi(X, \tau)$ is interpreted to be the event experienced by the material point X at the instant τ .

The set

$$\{\chi(X, \tau) \mid \tau \in \mathcal{T}_x\}$$

of all events experienced by X is called the *worldline of X* in the kinematical process χ .

Let $f: \mathcal{T} \rightarrow \mathcal{E}$ be a function on a time-interval \mathcal{T} with values in some Euclidean space \mathcal{E} . The limit

$$\dot{f}(\tau) = \lim_{s \rightarrow 0} \frac{f(\tau+s) - f(\tau)}{s} = \frac{d}{d\tau} f(\tau),$$

if it exists, is called the *derivative* of f at $\tau \in \mathcal{T}$. The value $\dot{f}(\tau)$ belongs to the translation space \mathcal{V} of \mathcal{E} . If $\dot{f}(\tau)$ exists for all $\tau \in \mathcal{T}$ we say that f is differentiable and call $\dot{f}: \mathcal{T} \rightarrow \mathcal{V}$ its derivative. We say that f is of class C^1 if \dot{f} exists and is continuous.

The n^{th} derivative $f^{(n)}$ of f is defined inductively by

$$f^{(0)} = f, \quad f^{(1)} = \dot{f}, \quad f^{(2)} = \ddot{f} = (\dot{f})', \quad f^{(n)} = (f^{(n-1)})'.$$

We say that f is of class C^n if $f^{(n)}$ exists and is continuous.

A material system \mathcal{B} is assumed to be endowed with a structure defined by the prescription of a non-empty class \mathfrak{R} of admissible kinematical processes. The members of \mathfrak{R} are assumed to be such that for any two material points $X, Y \in \mathcal{B}$, the function

$$\tau \mapsto \hat{d}(\chi(X, \tau), \chi(Y, \tau))$$

is of class C^2 .

The derivative $\frac{d}{d\tau} \hat{d}(\chi(X, \tau), \chi(Y, \tau))$ is called the *relative speed* of X and Y at τ and the second derivative $\frac{d^2}{d\tau^2} \hat{d}(\chi(X, \tau), \chi(Y, \tau))$ the *relative scalar acceleration* of X and Y at τ . A kinematical process χ is called a *rigid process* if $\hat{d}(\chi(X, \tau), \chi(Y, \tau))$ is independent of τ for all $X, Y \in \mathcal{B}$, i.e., if the relative speed of any two material points is always zero.

A material system \mathcal{E} is called a *rigid system* if its class \mathfrak{R} of admissible kinematical processes is defined in terms of a distance function $d: \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}$ by $\mathfrak{R} = \{\beta \mid \beta: \mathcal{E} \times \mathcal{T}_\beta \rightarrow \mathcal{W} \text{ is a kinematical process and } d(x, y) = \hat{d}(\beta(x, \tau), \beta(y, \tau)) \text{ for all } x, y \in \mathcal{E} \text{ and all } \tau \in \mathcal{T}_\beta\}$. If $\beta \in \mathfrak{R}$ and $\tau \in \mathcal{T}_\beta$, we define

$$\beta_\tau: \mathcal{E} \rightarrow \tau \quad \text{by} \quad \beta_\tau(x) = \beta(x, \tau). \quad (1)$$

If β_τ is invertible for some $\beta \in \mathfrak{R}$ and some $\tau \in \Gamma$ then β_τ is an isomorphism and d must be a Euclidean metric on \mathcal{E} . In this case, \mathcal{E} is called a **frame of reference** (or simply *frame*). The processes in \mathfrak{R} will then be called *reference processes*.

We now assume that a frame \mathcal{E} with Euclidean metric d and class \mathfrak{R} of reference processes is given. Let \mathcal{B} be a material system, unrelated to \mathcal{E} , and let \mathfrak{R} be the class of admissible processes of \mathcal{B} .

We assume that if $\chi \in \mathfrak{R}$ and if α is an automorphism of \mathcal{W} , then also $\chi^\alpha \in \mathfrak{R}$, where χ^α is defined by $\chi^\alpha(X, \tau) = \alpha(\chi(X, \tau - t_\alpha))$. For every $\chi \in \mathfrak{R}$ and every $\tau \in \mathcal{T}_\chi$, we define

$$\chi_\tau: \mathcal{B} \rightarrow \tau \quad \text{by} \quad \chi_\tau(X) = \chi(X, \tau). \quad (2)$$

If a reference process $\beta \in \mathfrak{R}$ with $\mathcal{T}_\beta = \mathcal{T}_\beta$ is given, we define the **motion** $\mu: \mathcal{B} \times \mathcal{T}_\beta \rightarrow \mathcal{E}$ of \mathcal{B} relative to the frame \mathcal{E} , determined by $\chi \in \mathfrak{R}$, by

$$\mu(X, \tau) = \mu_\tau(X) \quad \text{with} \quad \mu_\tau = \beta_\tau^{-1} \circ \chi_\tau, \quad \tau \in \mathcal{T}_\beta.$$

Let \mathfrak{M} be the class of all motions obtained in this way, i. e.

$$\mathfrak{M} = \{ \mu \mid \mu_\tau = \beta_\tau \circ \chi_\tau, \chi \in \mathfrak{K}, \tau \in \mathcal{T}_\chi \}.$$

It follows easily from the invariance of \mathfrak{K} under the automorphisms of \mathcal{W} that \mathfrak{M} is independent of the choice of $\beta \in \mathfrak{R}$. For every choice $\beta \in \mathfrak{R}$ with $\mathcal{T}_\beta = \mathcal{T}_\chi$ the relations

$$\mu_\tau = \beta_\tau^{-1} \circ \chi_\tau, \quad \chi_\tau = \beta_\tau \circ \mu_\tau \quad (3)$$

define a one-to-one correspondence between the class \mathfrak{K} of kinematical processes of \mathcal{B} and the class \mathfrak{M} of motions of \mathcal{B} relative to the frame \mathcal{E} . This correspondence depends, of course, on the choice of the reference process $\beta \in \mathfrak{R}$.

Let a kinematical process $\chi \in \mathfrak{K}$ be given. If $\beta, \beta^* \in \mathfrak{R}$ are two reference processes (with $\mathcal{T}_\chi = \mathcal{T}_\beta = \mathcal{T}_{\beta^*}$) we obtain two motions $\mu, \mu^* \in \mathfrak{M}$ corresponding to χ by

$$\mu_\tau = \beta_\tau^{-1} \circ \chi_\tau, \quad \mu_\tau^* = \beta_\tau^{*-1} \circ \chi_\tau, \quad \tau \in \mathcal{T}_\chi. \quad (4)$$

It follows from (4) that these motions μ and μ^* are related by

$$\mu_\tau^* = \alpha_\tau \circ \mu_\tau, \quad \alpha_\tau: \mathcal{E} \rightarrow \mathcal{E}, \quad (5)$$

where $\alpha_\tau = \beta_\tau^{*-1} \circ \beta_\tau$, for each $\tau \in \mathcal{T}_\chi$, is an isometry of the frame \mathcal{E} . Using the representation (3) of Section 2, we infer from (5) that

$$\mu^* = c + Q(\mu - q), \quad (6)$$

where $q \in \mathcal{E}$ can be chosen arbitrarily, $c: \mathcal{T}_\chi \rightarrow \mathcal{E}$ is defined by $c(\tau) = \alpha_\tau(q)$, and $Q: \mathcal{T}_\chi \rightarrow \text{Orth}(\mathcal{V})$ is defined by $Q(\tau) = Q_{\alpha_\tau}$.

If a motion μ of \mathcal{B} relative to \mathcal{E} is of class C^2 , we can define a corresponding velocity field $\dot{\mu}: \mathcal{B} \times \mathcal{T}_\mu \rightarrow \mathcal{V}$ and a corresponding acceleration field $\ddot{\mu}: \mathcal{B} \times \mathcal{T}_\mu \rightarrow \mathcal{V}$ by

$$\dot{\mu}(X, \tau) = \frac{d}{d\tau} \mu(X, \tau), \quad \ddot{\mu}(X, \tau) = \frac{d^2}{d\tau^2} \mu(X, \tau).$$

Assume that the motion μ , corresponding to a given kinematical process χ relative to a reference process β , is of class C^2 . Let β^* be another reference process such that the functions c and Q of (6) determined by β^* are also of class C^2 . Then the motion μ^* corresponding to χ and β^* , given by (6), is also of class C^2 . The velocity and acceleration fields of μ and μ^* are easily seen to be related by

$$\dot{\mu}^* - Q\dot{\mu} = \dot{c} + A(\mu^* - c), \quad (7)$$

$$\ddot{\mu}^* - Q\ddot{\mu} = \ddot{c} + 2A(\dot{\mu}^* - \dot{c}) + (\dot{A} - A^2)(\mu^* - c), \quad (8)$$

where

$$A: \mathcal{T}_\chi \rightarrow \text{Skew}(\mathcal{V}) \quad \text{is defined by } A = \dot{Q}Q^T$$

$$(\text{Skew}(\mathcal{V})) = \{ A \in \text{Lin}(\mathcal{V}) \mid A = -A^T \}.$$

It is possible to associate with a kinematical process χ and a reference process β an instantaneous velocity field v by $v_\tau = \beta_\tau \circ \dot{\mu}_\tau: \mathcal{B} \rightarrow \mathcal{V}$, where $\mu_\tau = \beta_\tau^{-1} \circ \chi_\tau$. If the right side of (7) were zero, one could prove that v_τ is independent of the choice of the reference process. Since the right side of (7) is not zero, it follows that there is no way of associating a velocity field with the kinematical process itself,

independent of any choice of reference process. We express this fact by saying that the velocity is *frame-dependent*. Of course, the acceleration is also frame-dependent. Later, we will see that there are certain tensor fields derived from the velocity that are *frame-indifferent*, *i.e.* independent of the choice of reference process.

5. Differential Calculus in Euclidean Spaces

Let $\phi: \mathcal{D} \rightarrow \mathcal{D}'$ be a mapping from an open subset \mathcal{D} of a Euclidean space \mathcal{E} into an open subset \mathcal{D}' of a Euclidean space \mathcal{E}' . Let \mathcal{V} and \mathcal{V}' be the translation spaces of \mathcal{E} and \mathcal{E}' , respectively. We say that ϕ is *differentiable at* $x \in \mathcal{D}$ if there is a $L \in \text{Lin}(\mathcal{V}, \mathcal{V}')$ such that

$$\phi(x + \mathbf{u}) - \phi(x) = L\mathbf{u} + o(\mathbf{u})$$

for \mathbf{u} in a neighborhood of $\mathbf{0} \in \mathcal{V}$. The notation $o(\mathbf{u})$ indicates that $\mathbf{u} \mapsto o(\mathbf{u})$ is some function of order small o . This means that for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|o(\mathbf{u})| < \varepsilon |\mathbf{u}| \quad \text{when} \quad |\mathbf{u}| < \delta.$$

Roughly, $o(\mathbf{u})$ can be neglected compared to $L\mathbf{u}$ for small \mathbf{u} . The linear mapping L is called the *gradient of ϕ at x* and is denoted by $\nabla\phi(x)$. The gradient $\nabla\phi(x)$ is a linear approximation to the mapping $\mathbf{u} \mapsto \phi(x + \mathbf{u}) - \phi(x)$. A mapping $\phi: \mathcal{D} \rightarrow \mathcal{D}'$ is said to be of class C^1 if ϕ is differentiable at every $x \in \mathcal{D}$ and if the gradient function

$$\nabla\phi: \mathcal{D} \rightarrow \text{Lin}(\mathcal{V}, \mathcal{V}')$$

is continuous. It is possible to define on $\text{Lin}(\mathcal{V}, \mathcal{V}')$ a natural structure of an inner product space in the following way: If $L, M \in \text{Lin}(\mathcal{V}, \mathcal{V}')$ then

$$M^* \in \text{Lin}(\mathcal{V}'^*, \mathcal{V}^*) = \text{Lin}(\mathcal{V}', \mathcal{V})$$

and hence $LM^* \in \text{Lin}(\mathcal{V})$. If we put $L \cdot M = \text{tr}(LM^*)$, then $(L, M) \mapsto L \cdot M$ is easily proved to be an inner product on $\text{Lin}(\mathcal{V}, \mathcal{V}')$.

Since every inner product space can be regarded as a Euclidean space, it is meaningful to discuss the possibility that $\nabla\phi$ is of class C^1 . If this is the case, we say that ϕ is of class C^2 . In a similar way, one can define what is meant by saying that ϕ is of class C^n , $n=1, 2, \dots$.

The most important theorem of the differential calculus is the *chain rule*: Let $\phi: \mathcal{D} \rightarrow \mathcal{D}'$ and $\psi: \mathcal{D}' \rightarrow \mathcal{D}''$ be mappings of class C^1 , where \mathcal{D} , \mathcal{D}' , and \mathcal{D}'' are open subsets of Euclidean spaces \mathcal{E} , \mathcal{E}' , and \mathcal{E}'' . The composition $\psi \circ \phi: \mathcal{D} \rightarrow \mathcal{D}''$ is then again of class C^1 and its gradient is given by

$$\nabla(\psi \circ \phi)(x) = \nabla\psi(\phi(x)) \nabla\phi(x), \quad x \in \mathcal{D}.$$

Moreover, if ϕ and ψ are both of class C^n , $n > 1$, so is $\psi \circ \phi$.

Another important theorem is the *product rule*:

Let $\phi: \mathcal{D} \rightarrow \mathcal{V}_1$, $\psi: \mathcal{D} \rightarrow \mathcal{V}_2$ be mappings of class C^1 , where \mathcal{D} is an open subset of a Euclidean space and where \mathcal{V}_1 and \mathcal{V}_2 are inner product spaces.

Let $B: \mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \mathcal{V}_3$ be a bilinear function. Then

$$\eta = B(\phi(\cdot), \psi(\cdot)): \mathcal{D} \rightarrow \mathcal{V}_3$$

is again of class C^1 and its gradient is given by

$$(\nabla \eta(x)) \mathbf{u} = B((\nabla \phi(x)) \mathbf{u}, \psi(x)) + B(\phi(x), (\nabla \psi(x)) \mathbf{u}).$$

Finally, we have the following *symmetry theorem*:

If $\phi: \mathcal{D} \rightarrow \mathcal{D}'$ is of class C^2 , then $\nabla \nabla \phi: \mathcal{D} \rightarrow \text{Lin}(\mathcal{V}, \text{Lin}(\mathcal{V}, \mathcal{V}'))$ has symmetric values in the sense that

$$(((\nabla \nabla \phi)(x)) \mathbf{u}) \mathbf{v} = (((\nabla \nabla \phi)(x)) \mathbf{v}) \mathbf{u}$$

for all $x \in \mathcal{D}$ and all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$.

6. Continuous Bodies

We describe here precise mathematical structures intended to idealize the vague physical notion of continuous body. Such structures depend on the choice of a class \mathbf{D} of mappings called **displacements**. This class is assumed to satisfy the following requirements:

- (i) the members of \mathbf{D} are invertible mappings whose domains and codomains are open subsets of Euclidean spaces.
- (ii) Every mapping in \mathbf{D} is of class C^1 .
- (iii) If $\lambda \in \mathbf{D}$ then $\lambda^{-1} \in \mathbf{D}$.
- (iv) If $\lambda, \mu \in \mathbf{D}$ and $\text{Range}(\lambda) = \text{Domain}(\mu)$ then $\mu \circ \lambda \in \mathbf{D}$.
- (v) Isometric (*i.e.* distance-preserving) bijections between open subsets of Euclidean spaces belong to \mathbf{D} .

The following classes \mathbf{D} are important examples of classes of displacements:

- (a) $\mathbf{D} = \mathbf{D}_n$ = class of all restrictions to open sets of C^n -diffeomorphisms between Euclidean spaces ($n \geq 1$).
- (b) $\mathbf{D} = \mathbf{D}_n^{\text{iso}}$ = class of all isochoric (*i.e.* volume-preserving) mappings belonging to \mathbf{D}_n .
- (c) $\mathbf{D} = \mathbf{D}^{\text{is}}$ = class of all isometric bijections between open sets of Euclidean spaces.

Definition. We say that a material system \mathcal{B} is a **continuous body** of type \mathbf{D} if it is endowed with a structure defined by a non-empty class \mathbf{P} of mappings subject to the following conditions

- (B1) The members of \mathbf{P} are invertible mappings from \mathcal{B} onto open subsets of Euclidean spaces.
- (B2) If $\kappa, \gamma \in \mathbf{P}$, then $\kappa \circ \gamma^{-1} \in \mathbf{D}$.
- (B3) If $\kappa \in \mathbf{P}$, $\lambda \in \mathbf{D}$ and $\text{Range}(\kappa) = \text{Domain}(\lambda)$, then $\lambda \circ \kappa \in \mathbf{P}$.

The members of \mathbf{P} are called the **placements** of \mathcal{B} . The range \mathcal{R}_κ of a $\kappa \in \mathbf{P}$ is called the *region occupied* by \mathcal{B} in the placement κ .

The Euclidean space \mathcal{E}_κ of which the range \mathcal{R}_κ of $\kappa \in \mathbf{P}$ is an open subset is called the *range space* of κ . Its translation space is denoted by \mathcal{V}_κ .

We say that \mathcal{B} is a *body of class C^n* if $D \subset D_n$, an *incompressible body* of class C^n if $D \subset D_n^{\text{iso}}$, and a *rigid body* if $D = D^{\text{rig}}$.

Let \mathcal{F} be some Euclidean space with translation space \mathcal{U} and let \mathcal{B} be a body of class C^n . We say that the mapping

$$f: \mathcal{B} \rightarrow \mathcal{F}$$

is of class C^r ($0 \leq r \leq n$) if $f \circ \kappa^{-1}: \mathcal{R}_\kappa \rightarrow \mathcal{F}$ is of class C^r for *some* $\kappa \in \mathbf{P}$. It follows from (B2) and the chain rule that $f \circ \kappa^{-1}$ is then of class C^r for *every* $\kappa \in \mathbf{P}$. If f is of class C^1 and if $\kappa \in \mathbf{P}$, we define

$$\nabla_\kappa f: \mathcal{B} \rightarrow \text{Lin}(\mathcal{V}_\kappa, \mathcal{U})$$

by

$$\nabla_\kappa f(X) = \nabla(f \circ \kappa^{-1})(\kappa(X)), \quad X \in \mathcal{B} \quad (1)$$

and call it the *gradient of f in the placement κ* .

If κ and γ are two placements, it follows from the chain rule that

$$\nabla_\gamma f = (\nabla_\kappa f)(\nabla_\gamma \kappa). \quad (2)$$

We say that $\nabla_\gamma \kappa$ is the *gradient of the displacement $\lambda = \gamma \circ \kappa^{-1}$ from κ to γ* .

The intuitive notion of an ‘‘infinitesimal element’’ of a continuous body can be made precise in the following way.

Let $X \in \mathcal{B}$ and let γ be a placement of \mathcal{B} . If \mathbf{u} varies in a neighborhood of zero in the translation space \mathcal{V}_γ of the range space of γ , then $\gamma(X) + \mathbf{u}$ varies in a neighborhood of $\gamma(X)$ in the region occupied by the body in the configuration γ , and hence $\gamma^{-1}(\gamma(X) + \mathbf{u})$ varies in a neighborhood of $X \in \mathcal{B}$. If κ is another placement then

$$\gamma^{-1}(\gamma(X) + \mathbf{u}) = \kappa^{-1}(\kappa(X) + \mathbf{v})$$

describes the same point in the neighborhood of X if

$$\lambda(\gamma(X) + \mathbf{u}) = \kappa(X) + \mathbf{v}, \quad \lambda = \kappa \circ \gamma^{-1}. \quad (3)$$

Approximating the displacement λ by its gradient in the neighborhood of $\gamma(X)$, we see that (3) can be written

$$(\nabla_\gamma \kappa(X))\mathbf{u} + o(\mathbf{u}) = \mathbf{v}.$$

Thus, the two vectors $\mathbf{u} \in \mathcal{V}_\gamma$ and $\mathbf{v} \in \mathcal{V}_\kappa$ describe approximately the same material point if $(\nabla_\gamma \kappa(X))\mathbf{u} = \mathbf{v}$.

Definition. We say that the pairs (γ, \mathbf{u}) and (κ, \mathbf{v}) , where $\kappa, \gamma \in \mathbf{P}$, $\mathbf{u} \in \mathcal{V}_\gamma$, $\mathbf{v} \in \mathcal{V}_\kappa$, are *equivalent at $X \in \mathcal{B}$* and write $(\gamma, \mathbf{u}) \sim_X (\kappa, \mathbf{v})$ if

$$(\nabla_\gamma \kappa(X))\mathbf{u} = \mathbf{v}. \quad (4)$$

The equivalence classes determined by the equivalence relation thus defined are called the *tangent vectors* at X . The set of all tangent vectors at X is called the **tangent space** at X and is denoted by \mathcal{T}_X .

The tangent vectors at X describe, approximately, the points in a neighborhood of X independently of any placement. Thus, the notion of tangent space is a mathematization of the intuitive notion of an "infinitesimal element about X ".

Let $f: \mathcal{B} \rightarrow \mathcal{F}$ be a mapping of class C^1 from \mathcal{B} into some Euclidean space \mathcal{F} with translation space \mathcal{U} . It follows from (2) and (4) that $(\nabla_{\kappa} f(X))v = (\nabla_{\gamma} f(X))u$ if $(\gamma, u) \sim_X (\kappa, v)$. Hence we can define a mapping

$$\nabla f(X): \mathcal{T}_X \rightarrow \mathcal{U} \quad (5)$$

by

$$(\nabla f(X))t = (\nabla_{\kappa} f(X))v \quad \text{if } (\kappa, v) \in t \in \mathcal{T}_X. \quad (6)$$

The following theorem justifies the term "tangent vectors" for the elements of \mathcal{T}_X .

Theorem. *The tangent space \mathcal{T}_X has a unique natural vector space structure such that for all $f: \mathcal{B} \rightarrow \mathcal{F}$ of class C^1 , the mapping (5) defined by (6) is linear.*

Proof. If $\kappa \in P$, then $K = \nabla \kappa(X): \mathcal{T}_X \rightarrow \mathcal{V}_{\kappa}$ is easily seen to be an invertible mapping. The vector-space structure on \mathcal{T}_X is obtained by transporting the vector-space structure of \mathcal{V}_{κ} to \mathcal{T}_X by means of K^{-1} . It is easily seen that the structure on \mathcal{T}_X thus obtained is independent of the choice of $\kappa \in P$. Q. E. D.

Remark. Given $\kappa \in P$, one can also transport the inner product of \mathcal{V}_{κ} to \mathcal{T}_X by means of K^{-1} . But one gets different inner products on \mathcal{T}_X for different choices of $\kappa \in P$. Hence \mathcal{T}_X has no intrinsic natural inner product.

7. Configurations, Deformations

Let \mathcal{B} be a continuous body. With every placement $\kappa \in P$ we can associate a distance function

$$d_{\kappa}: \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}^+$$

by putting

$$d_{\kappa}(X, Y) = d(\kappa(X), \kappa(Y)) = |\kappa(Y) - \kappa(X)|. \quad (1)$$

(If u belongs to some inner product space, we define the magnitude $|u|$ of u by $|u| = \sqrt{u \cdot u}$.)

The distance functions obtained in this manner will be called **configurations** of \mathcal{B} , so that

$$C = \{d_{\kappa} \mid \kappa \in P\}$$

is the set of all configurations of \mathcal{B} . Two placements $\kappa, \gamma \in P$ determine the same configuration $d_{\kappa} = d_{\gamma}$, if and only if the displacement $\lambda = \gamma \circ \kappa^{-1}$ from κ to γ is an isometric bijection.

It is possible to recover the class P of placements of \mathcal{B} from a knowledge of the set C of configurations of \mathcal{B} , because P is the class of all invertible mappings $\kappa: \mathcal{B} \rightarrow \mathcal{R}_{\kappa}$ onto open subsets \mathcal{R}_{κ} of Euclidean spaces \mathcal{E}_{κ} such that $d(X, Y) = |\kappa(Y) - \kappa(X)|$ holds for some $d \in C$.

Let $d \in C$ be a configuration of \mathcal{B} and let $X \in \mathcal{B}$ be a material point. The definition (6) of Section 6 can be applied to $d^2(\cdot, Y): \mathcal{B} \rightarrow \mathbb{R}$, where $Y \in \mathcal{B}$. We denote the gradient at $X \in \mathcal{B}$ of this function by

$$(\nabla d^2(\cdot, Y))(X) = \nabla_1 d^2(X, Y) \in \text{Lin}(\mathcal{T}_X, \mathbb{R}) = \mathcal{T}_X^*.$$

Since $Y \in \mathcal{B}$ is arbitrary we thus obtain a function $\nabla_1 d^2(X, \cdot): \mathcal{B} \rightarrow \mathcal{T}_X^*$. The definition (6) of Section 6 can be applied again, with Y instead of X , and we obtain

$$\nabla_2 \nabla_1 d^2(X, Y) = \nabla(\nabla_1 d^2(X, \cdot))(Y) \in \text{Lin}(\mathcal{T}_Y, \mathcal{T}_X^*).$$

Using the fact that $d^2(X, Y) = d^2(Y, X)$ and the symmetry theorem of Section 5, one can easily prove that

$$G_d = \nabla_2 \nabla_1 d^2(X, X) \in \text{Sym}^+(\mathcal{T}_X, \mathcal{T}_X^*). \quad (2)$$

We say that G_d is the *configuration of the infinitesimal element \mathcal{T}_X induced by the configuration d of the body \mathcal{B}* . The set

$$\mathbf{G} = \{G_d \mid d \in \mathbf{C}\} \subset \text{Sym}^+(\mathcal{T}_X, \mathcal{T}_X^*)$$

is then the set of all configurations of \mathcal{T}_X .

There are several other ways to characterize the induced configurations G_d of \mathcal{T}_X :

(i) If $\mathbf{t} \in \mathcal{T}_X$ and $(\kappa, \mathbf{u}) \in \mathbf{t}$, then

$$\langle G_d \mathbf{t}, \mathbf{t} \rangle = \frac{d^2}{ds^2} d^2(X, \kappa^{-1}(\kappa(X) + s\mathbf{u}))|_{s=0}.$$

(ii) If $\kappa \in \mathbf{P}$ is a placement and $K = \nabla \kappa(X) \in \text{Invlin}(\mathcal{T}_X, \mathcal{V}_\kappa)$, then

$$G_{d_\kappa} = K^* K; \quad (3)$$

hence G_{d_κ} is just the “inner product” on \mathcal{T}_X obtained by transporting the inner product of \mathcal{V}_κ to \mathcal{T}_X via K^{-1} .

A **deformation process** of a body \mathcal{B} is a function of the type

$$p: [0, r] \rightarrow \mathbf{C}, \quad r \in \mathbb{R}^+.$$

The value of p at $t \in [0, r]$ is denoted by p_t . The number r is called the *duration* of the process. With every kinematical process

$$\chi: \mathcal{B} \times \mathcal{T}_\chi \rightarrow \mathcal{W} \quad \text{of } \mathcal{B}, \quad \text{where } \mathcal{T}_\chi = [\tau_0, \tau_0 + r]$$

is a closed time-interval, we can associate a deformation process $p: [0, r] \rightarrow \mathbf{C}$ by

$$p_t(X, Y) = \hat{d}(\chi(X, \tau_0 + t), \chi(Y, \tau_0 + t)), \quad t \in [0, r]. \quad (4)$$

Similarly, with every motion $\mu: \mathcal{B} \times \mathcal{T}_\mu \rightarrow \mathcal{E}$ of \mathcal{B} relative to some frame \mathcal{E} , where $\mathcal{T}_\mu = [\tau_0, \tau_0 + r]$, we can associate a deformation process $p: [0, r] \rightarrow \mathbf{C}$ by

$$p_t(X, Y) = d(\mu(X, \tau_0 + t), \mu(Y, \tau_0 + t)), \quad t \in [0, r]. \quad (5)$$

Theorem A. *Two kinematical processes χ and χ^* of \mathcal{B} determine the same deformation process if and only if there is an automorphism α of the event-world \mathcal{W} such that $\chi^* = \chi^\alpha$. (Recall that $\chi^\alpha(X, \tau) = \alpha(\chi(X, \tau - t_\alpha))$.)*

Theorem B. *Two motions μ and μ^* of \mathcal{B} relative to \mathcal{E} determine the same deformation process if and only if there is a kinematical process χ and two reference processes β*

and β^* such that μ and μ^* are the motions corresponding to χ, β and χ, β^* , respectively. This is the case if and only if μ and μ^* are connected by a relation of the form (6) of Section 4.

A continuous body \mathcal{B} can be endowed with the structure of a material system in the sense of Section 4 in the following manner: A class \mathfrak{D} of deformation processes of \mathcal{B} is prescribed. The class \mathfrak{R} of admissible kinematical processes then consist of all those that determine a deformation process in \mathfrak{D} .

In view of Theorem A, the condition that $\chi \in \mathfrak{R} \Rightarrow \chi^\alpha \in \mathfrak{R}$ for every automorphism α of \mathcal{W} is automatically satisfied. In view of Theorem B, the class \mathfrak{M} of admissible motions of \mathcal{B} relative to the frame \mathcal{E} consists of all those that determine a deformation process in \mathfrak{D} .

If μ is a motion of \mathcal{B} , then $\mu_\tau: \mathcal{B} \rightarrow \mathcal{E}$ is called the *placement of \mathcal{B} in \mathcal{E} at time $\tau \in \mathcal{T}_\mu$* . The gradient

$$M(\tau) = \nabla \mu_\tau(X) \in \text{Lin}(\mathcal{T}_X, \mathcal{V})$$

is called the *placement of the body element \mathcal{T}_X in \mathcal{V} at time τ* . The function $M: \mathcal{T}_\mu \rightarrow \text{Lin}(\mathcal{T}_X, \mathcal{V})$ is called the *motion of the body element \mathcal{T}_X* induced by the motion μ of \mathcal{B} .

If $p: [0, r] \rightarrow \mathcal{C}$ is a deformation process of \mathcal{B} , then

$$P: [0, r] \rightarrow \mathcal{G}, \quad P(t) = G_{p_t}$$

is called the *deformation process of \mathcal{T}_X induced by the deformation process p of \mathcal{B}* .

It is often useful to consider body elements in isolation, as if they were disconnected from the continuous body from which they are obtained.

Definition. A **body element** is a 3-dimensional vector space \mathcal{T} endowed with structure defined by the prescription of a set $\mathcal{G} \subset \text{Sym}^+(\mathcal{T}, \mathcal{T}^*)$ of *configurations* and a set Π of *deformation processes* $P: [0, r] \rightarrow \mathcal{G}, r \in \mathbb{R}^+$.

A *placement* of the element \mathcal{T} in a frame space \mathcal{V} (inner product space) is defined to be an element $K \in \text{InvLin}(\mathcal{T}, \mathcal{V})$ such that $K^* K \in \mathcal{G}$. A *motion* of \mathcal{T} in the frame space \mathcal{V} is a function $M: [0, r] \rightarrow \text{InvLin}(\mathcal{T}, \mathcal{V})$ such that $M^* M \in \Pi$. It is often useful to fix a particular *reference placement* K_R and describe the possible motions in relation to K_R . The following functions from $[0, r]$ into $\text{Lin}(\mathcal{V})$ are obtained in this fashion:

(i) *Displacement tensor:*

$$F = MK_R^{-1}.$$

(ii) *Right Cauchy-Green tensor:*

$$C = F^T F = K_R^*{}^{-1} P K_R^{-1}, \quad P = M^* M.$$

In the case when $K_R = M(t)$ for some $t \in [0, r]$, we get:

(iii) *Relative displacement tensor:*

$$F_{(t)} = MM(t)^{-1}.$$

(iv) *Relative right Cauchy-Green tensor*:

$$C_{(t)} = F_{(t)}^T F_{(t)} = M(t)^{* -1} P M(t)^{-1}.$$

(v) k^{th} *Rivlin-Ericksen tensor*:

$$A_k = M^{* -1} P M^{-1}, \quad A_k(t) = C_{(t)}^{(k)}(t).$$

If L is defined by $L(t) = \dot{F}_{(t)}(t)$, we have $\frac{1}{2}A = D = \frac{1}{2}(L + L^T)$, which is called the *stretching rate* (or *rate of deformation*). The *spin* (or *vorticity*) W is defined by $W = \frac{1}{2}(L - L^T)$.

The functions C , $C_{(t)}$, A_k , and D are frame-indifferent, F , $F_{(t)}$, L , and W are frame-dependent.

II. Thermomechanics of Continuous Bodies

8. Interactions

In this section we deal with what we call a **material universe** Ω , whose members \mathcal{A} , \mathcal{B} , \mathcal{C} , ... we call *material objects*. We assume that Ω is endowed with a structure defined by a relation \prec and we read $\mathcal{A} \prec \mathcal{B}$ as “ \mathcal{A} is a part of \mathcal{B} ”.

We assume that Ω and \prec satisfies the 6 axioms stated below, which are all intuitively plausible.

(M1) $\mathcal{A} = \mathcal{B}$ if and only if $\mathcal{A} \prec \mathcal{B}$ and $\mathcal{B} \prec \mathcal{A}$.

(M2) $\mathcal{A} \prec \mathcal{B}$ and $\mathcal{B} \prec \mathcal{C}$ implies $\mathcal{A} \prec \mathcal{C}$.

These two axioms state that \prec is a partial order on Ω . Given $\mathcal{A}, \mathcal{B} \in \Omega$, there can be at most one $\mathcal{C} \in \Omega$ such that $\mathcal{C} \prec \mathcal{A}, \mathcal{B}$ and $\mathcal{C}' \prec \mathcal{A}, \mathcal{B} \Rightarrow \mathcal{C}' \prec \mathcal{C}$. If such a \mathcal{C} exists, we call it the *greatest common part* of \mathcal{A} and \mathcal{B} and write

$$\mathcal{C} = \mathcal{A} \wedge \mathcal{B}.$$

Given $\mathcal{A}, \mathcal{B} \in \Omega$, there can be at most one $\mathcal{D} \in \Omega$ such that $\mathcal{A}, \mathcal{B} \prec \mathcal{D}$ and $\mathcal{A}, \mathcal{B} \prec \mathcal{D}' \Rightarrow \mathcal{D} \prec \mathcal{D}'$. If such a \mathcal{D} exists, we call it the *least envelope* of \mathcal{A} and \mathcal{B} and we write

$$\mathcal{D} = \mathcal{A} \vee \mathcal{B}.$$

(M3) There are two elements $\emptyset, \infty \in \Omega$, called the *material nothing* and the *material all*, such that $\emptyset \prec \mathcal{A} \prec \infty$ for all $\mathcal{A} \in \Omega$.

If $\mathcal{A} \wedge \mathcal{B} = \emptyset$ we say that \mathcal{A} and \mathcal{B} are *separate*.

(M4) For each $\mathcal{A} \in \Omega$ there is exactly one $\mathcal{A}^e \in \Omega$, called the *exterior* of \mathcal{A} , such that $\mathcal{A} \wedge \mathcal{A}^e = \emptyset$ and $\mathcal{A} \vee \mathcal{A}^e = \infty$.

(M5) If $\mathcal{A} \wedge \mathcal{B}^e = \emptyset$, then $\mathcal{A} \prec \mathcal{B}$.

(M6) For all $\mathcal{A}, \mathcal{B} \in \Omega$, $\mathcal{A} \wedge \mathcal{B}$ exists.

One can prove, from (M1)–(M6), that the operations \wedge and \vee define on Ω the structure of a Boolean algebra. The proof (taken from reference [1]) is given in the Appendix.

We use the notation

$$(\Omega \times \Omega)_{\text{sep}} = \{(\mathcal{A}, \mathcal{B}) \in \Omega \times \Omega \mid \mathcal{A} \wedge \mathcal{B} = \emptyset\}$$

for the set of all separate pairs of bodies, and the notation

$$\Omega_{\mathcal{B}} = \{\mathcal{P} \in \Omega \mid \mathcal{P} \prec \mathcal{B}\}$$

for the set of all parts of \mathcal{B} .

A mapping $F: \Omega' \rightarrow \mathcal{U}$ from some subset Ω' of Ω into some vector space \mathcal{U} is said to be *additive* if

$$F(\mathcal{A} \vee \mathcal{B}) = F(\mathcal{A}) + F(\mathcal{B}) \quad (1)$$

whenever $\mathcal{A}, \mathcal{B}, \mathcal{A} \vee \mathcal{B} \in \Omega'$ and $\mathcal{A} \wedge \mathcal{B} = \emptyset$.

Definition. A mapping

$$I: (\Omega \times \Omega)_{\text{sep}} \rightarrow \mathcal{U},$$

where \mathcal{U} is some vector space, is called an **interaction** if $I(\cdot, \mathcal{A}^e): \Omega_{\mathcal{A}} \rightarrow \mathcal{U}$ and $I(\mathcal{A}^e, \cdot): \Omega_{\mathcal{A}} \rightarrow \mathcal{U}$ are additive for every $\mathcal{A} \in \Omega$.

Lemma. If I is an interaction and $\mathcal{A} \wedge \mathcal{B} = \emptyset$ then

$$I(\mathcal{A}, \mathcal{B}) + I(\mathcal{B}, \mathcal{A}) = I(\mathcal{A}, \mathcal{A}^e) + I(\mathcal{B}, \mathcal{B}^e) - I(\mathcal{A} \vee \mathcal{B}, (\mathcal{A} \vee \mathcal{B})^e). \quad (2)$$

Proof. Let $\mathcal{C} = \mathcal{A} \vee \mathcal{B}$. Then $\mathcal{C}^e \wedge \mathcal{A} = \emptyset$, $\mathcal{C}^e \wedge \mathcal{B} = \emptyset$, and

$$\mathcal{A}^e = \mathcal{C}^e \vee \mathcal{B}, \quad \mathcal{B}^e = \mathcal{C}^e \vee \mathcal{A}.$$

The additivity of $I(\mathcal{A}, \cdot)$ and $I(\mathcal{B}, \cdot)$ gives

$$I(\mathcal{A}, \mathcal{A}^e) + I(\mathcal{B}, \mathcal{B}^e) = I(\mathcal{A}, \mathcal{C}^e) + I(\mathcal{A}, \mathcal{B}) + I(\mathcal{B}, \mathcal{C}^e) + I(\mathcal{B}, \mathcal{A}).$$

Using the additivity of $I(\cdot, \mathcal{C}^e)$, we get the desired result (2). Q.E.D.

The following theorem is an immediate consequence of the Lemma.

Theorem A. An interaction I satisfies the “law of action and reaction”

$$I(\mathcal{A}, \mathcal{B}) = -I(\mathcal{B}, \mathcal{A}) \quad (3)$$

for all $\mathcal{A}, \mathcal{B} \in \Omega$ with $\mathcal{A} \wedge \mathcal{B} = \emptyset$ if and only if the mapping $\mathcal{A} \mapsto I(\mathcal{A}, \mathcal{A}^e)$ is additive.

An interaction I is said to be **balanced** if

$$I(\mathcal{A}, \mathcal{A}^e) = 0 \quad \text{for all } \mathcal{A} \in \Omega. \quad (4)$$

Theorem A implies that the law (3) of action and reaction holds for balanced interactions.

We now assume that a continuous body \mathcal{B} of class C^1 is given and that for some (and hence every) placement κ of \mathcal{B} the region \mathcal{R}_κ occupied by \mathcal{B} is bounded and has a piecewise C^1 boundary $\partial \mathcal{R}_\kappa$. One can then complete \mathcal{B} by joining to \mathcal{B} a boundary $\partial \mathcal{B}$. The placements of \mathcal{B} can be extended to the completion

$$\bar{\mathcal{B}} = \mathcal{B} \cup \partial \mathcal{B}$$

of \mathcal{B} in such a way that for every $\kappa \in \mathbf{P}$, κ is continuous in $\overline{\mathcal{B}}$ and

$$\kappa(\partial\mathcal{B}) = \partial\mathcal{R}_\kappa.$$

For every subset \mathcal{P} of \mathcal{B} , we say that the boundary $\partial\mathcal{P} \subset \overline{\mathcal{B}}$ is piecewise C^1 if $\kappa(\partial\mathcal{P})$ is piecewise C^1 for some (and hence every) placement $\kappa \in \mathbf{P}$.

We assume, further, that we have singled out a certain class $\Omega_{\mathcal{B}}$ of open subsets of \mathcal{B} such that

(i) Every $\mathcal{P} \in \Omega_{\mathcal{B}}$ has a piecewise C^1 boundary.

(ii) If $\mathcal{P} \in \Omega_{\mathcal{B}}$, then

$$\mathcal{P}^b = \text{int}(\mathcal{B} \setminus \mathcal{P}) \in \Omega_{\mathcal{B}}. \quad (5)$$

(iii) If $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$, then

$$\mathcal{P} \wedge \mathcal{Q} = \mathcal{P} \cap \mathcal{Q} \in \Omega_{\mathcal{B}}. \quad (6)$$

(iv) If $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$, then

$$\mathcal{P} \vee \mathcal{Q} = \text{int cl}(\mathcal{P} \cup \mathcal{Q}) \in \Omega_{\mathcal{B}}. \quad (7)$$

(v) The class $\Omega_{\mathcal{B}}$ is as large as possible (in a sense not made precise here).

We now construct a material universe Ω from a given continuous body \mathcal{B} and a prescribed class $\Omega_{\mathcal{B}}$ of *parts of* \mathcal{B} as explained above, and from a given Euclidean space \mathcal{E} . The class Ω consists of $\Omega_{\mathcal{B}}$ and objects of the form $\mathcal{P} \vee \mathcal{E}$ with $\mathcal{P} \in \Omega_{\mathcal{B}}$. The “is a part of” relation \prec is defined in such a way that $\mathcal{P} \prec \mathcal{Q}$ means $\mathcal{P} \subset \mathcal{Q}$ if $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$ and

$$\mathcal{E} = \mathcal{B}^e. \quad (8)$$

In words, the Euclidean space \mathcal{E} is regarded as the exterior of the body, or as the *external world*. For the parts of \mathcal{B} , the greatest common part and the least envelope are given by (6) and (7), respectively. If $\mathcal{P} \in \Omega_{\mathcal{B}}$, we have

$$\mathcal{P}^e = \mathcal{P}^b \vee \mathcal{E}, \quad (9)$$

where \mathcal{P}^b is given by (5).

Let I be an interaction in the material universe Ω just constructed with values in a normed space \mathcal{U} . The restriction

$$I^{\text{int}}: (\Omega_{\mathcal{B}} \times \Omega_{\mathcal{B}})_{\text{sep}} \rightarrow \mathcal{U}$$

of I to separate pairs of parts of \mathcal{B} is called the **internal interaction** determined by I .

The additive function $I^{\text{ext}}: \Omega_{\mathcal{B}} \rightarrow \mathcal{U}$ defined by $I^{\text{ext}}(\mathcal{P}) = I(\mathcal{P}, \mathcal{E})$ is called the **external action** determined by I . It is clear that for any $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$ with $\mathcal{P} \subset \mathcal{Q}$, we have

$$I(\mathcal{P}, \mathcal{Q}^e) = I^{\text{int}}(\mathcal{P}, \mathcal{Q}^b) + I^{\text{ext}}(\mathcal{P}). \quad (10)$$

The following two assumptions are made about most of the interactions that occur in continuum physics:

(I1) The internal interaction I^{int} is a *contact interaction* and it is *area-continuous*. More precisely, for some (and hence every) placement κ of \mathcal{B} , there is a $k > 0$ such that

$$|I^{\text{int}}(\mathcal{P}, \mathcal{Q})| < k \text{Area}_\kappa(\partial\mathcal{P} \cap \partial\mathcal{Q})$$

holds for all $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$ with $\mathcal{P} \wedge \mathcal{Q} = \emptyset$.

- (I2) The external action I^{ext} can be by *contact and at a distance* and it is *area-volume-continuous*. More precisely, for some (and hence every) placement κ of \mathcal{B} , there are $k_1, k_2 > 0$ such that

$$|I^{\text{ext}}(\mathcal{P})| < k_1 \text{Vol}_\kappa(\mathcal{P}) + k_2 \text{Area}_\kappa(\partial\mathcal{P} \cap \partial\mathcal{B}).$$

for all $\mathcal{P} \in \Omega$.

These assumptions have the following consequence:

Theorem B. *Let I be interaction that satisfies the conditions (I1) and (I2) and let κ be a placement of \mathcal{B} . One can then find an integrable function*

$$f_\kappa: \mathcal{B} \rightarrow \mathcal{U}$$

and, for each oriented surface $\mathcal{S} \subset \overline{\mathcal{B}}$ belonging to a suitable class, an integrable function

$$\bar{j}_{\mathcal{S}, \kappa}: \mathcal{S} \rightarrow \mathcal{U}$$

such that for any $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$ with $\mathcal{P} \subset \mathcal{Q}$, we have

$$I(\mathcal{P}, \mathcal{Q}^e) = \int_{\mathcal{P}} f_\kappa d(\text{Vol}_\kappa) + \int_{\mathcal{S}} \bar{j}_{\mathcal{S}, \kappa} d(\text{Area}_\kappa), \quad (11)$$

where \mathcal{S} is the surface of contact between \mathcal{P} and \mathcal{Q} , oriented in such a way that \mathcal{P} and \mathcal{Q} are on the interior side of \mathcal{S} . The functions f_κ and $\bar{j}_{\mathcal{S}, \kappa}$ are uniformly bounded.

The proof of this theorem is not difficult but too technical to be presented here. The essential parts of this proof are contained in Section III and the Appendix of reference [4]. The following theorem is non-trivial and surprising. The original proof is contained in reference [5].

Theorem C. *Let I be an interaction that satisfies the conditions (I1) and (I2) so that the conclusions of Theorem B are valid. Assume that $\mathcal{P} \mapsto I(\mathcal{P}, \mathcal{P}^e)$ is volume-continuous on $\Omega_{\mathcal{B}}$ (i.e., for some, and hence every, placement κ of \mathcal{B} , there is a $k > 0$ such that $|I(\mathcal{P}, \mathcal{P}^e)| \leq k \text{Vol}_\kappa(\mathcal{P})$ holds for all $\mathcal{P} \in \Omega_{\mathcal{B}}$).*

Then there is a function

$$\bar{j}_\kappa: \overline{\mathcal{B}} \times \mathcal{V}_\kappa^1 \rightarrow \mathcal{U},$$

where $\mathcal{V}_\kappa^1 = \{\mathbf{n} \in \mathcal{V}_\kappa \mid |\mathbf{n}| = 1\}$, such that

$$\bar{j}_{\mathcal{S}, \kappa}(X) = \bar{j}_\kappa(X, \mathbf{n}) \quad (12)$$

whenever \mathbf{n} is the exterior unit normal to \mathcal{S} at $X \in \mathcal{S}$ in the placement κ .

The following theorem is a classical non-trivial result that was proved, in essence, by CAUCHY in 1823. A modern proof, which I have used in courses and lectures since 1958 but never published, is reproduced in Section VIII of reference [4].

Theorem D. *If the hypotheses of Theorem C are valid and if the function*

$$\bar{j}_\kappa(\cdot, \mathbf{n}): \overline{\mathcal{B}} \rightarrow \mathcal{U}$$

of Theorem C is continuous for every $\mathbf{n} \in \mathcal{V}_\kappa^1$, then there is a continuous function

$$j_\kappa: \bar{\mathcal{B}} \rightarrow \text{Lin}(\mathcal{V}_\kappa, \mathcal{U})$$

such that

$$\bar{j}_\kappa(X, \mathbf{n}) = j_\kappa(X) \mathbf{n}. \quad (13)$$

It is unfortunate that nobody has been able, so far, to prove the conclusion of Theorem D on the basis of the hypotheses of Theorem C alone, without the *ad hoc* continuity assumption of Theorem D. A step towards filling this hole in the foundations of continuum thermomechanics has been made by GURTIN, MIZEL & WILLIAMS in 1967 (reference [6]), who proved the existence of a measurable function j_κ under the assumption that $\bar{j}_\kappa(\cdot, \mathbf{n})$ is measurable for each \mathbf{n} .

9. Equilibrium of Forces, Mechanical Processes

We assume in this section that \mathcal{B} is a continuous body of class C^1 , that $\Omega_{\mathcal{B}}$ is a class of parts of \mathcal{B} , that $\mathcal{E} = \mathcal{B}^e$ is a Euclidean space representing the external world, and that Ω is the material universe constructed from $\Omega_{\mathcal{B}}$ and \mathcal{E} as explained in the previous section.

An interaction \mathbf{f} in Ω with values in the translation space \mathcal{V} of \mathcal{E} will be called a **force system** for \mathcal{B} . We consider only force systems that satisfy the conditions (I1) and (I2) of the previous section. It is then possible to associate, with each part $\mathcal{P} \in \Omega_{\mathcal{B}}$ a bounded vector-valued Borel measure $\mathbf{f}_{\mathcal{P}}$ on the Borel subsets of \mathcal{P} such that

$$\mathbf{f}_{\mathcal{P}}(\mathcal{C}) = \int_{\mathcal{C}} \mathbf{b}_\kappa \, d(\text{Vol}_\kappa) + \int_{\mathcal{C} \cap \partial\mathcal{P}} \mathbf{t}_{\partial\mathcal{P}, \kappa} \, d(\text{Area}_\kappa), \quad (1)$$

holds for all placements κ of \mathcal{B} in \mathcal{E} , where \mathbf{b}_κ and $\mathbf{t}_{\partial\mathcal{P}, \kappa}$ correspond to \mathbf{f}_κ and $\bar{\mathbf{f}}_{\mathcal{S}, \kappa}$, with $\mathcal{S} = \partial\mathcal{P}$, of Theorem B of the previous section. This measure describes, intuitively, the forces exerted on \mathcal{P} and its parts by the exterior of \mathcal{P} . It is meaningful to integrate continuous functions on \mathcal{P} relative to the measure $\mathbf{f}_{\mathcal{P}}$. For example, if $\mathbf{v}: \mathcal{P} \rightarrow \mathcal{V}$ is a continuous function, we have

$$\int_{\mathcal{P}} \mathbf{v} \cdot d\mathbf{f}_{\mathcal{P}} = \int_{\mathcal{P}} (\mathbf{b}_\kappa \cdot \mathbf{v}) \, d(\text{Vol}_\kappa) + \int_{\partial\mathcal{P}} (\mathbf{t}_{\partial\mathcal{P}, \kappa} \cdot \mathbf{v}) \, d(\text{Area}_\kappa). \quad (2)$$

A vector field $\mathbf{k}: \mathcal{E} \rightarrow \mathcal{V}$ is called an **infinitesimal isometry** of \mathcal{E} if the solution $\xi(s) = \alpha(s, x)$ of the initial value problem $\dot{\xi} = \mathbf{k}(\xi)$, $\xi(0) = x$, determines an isometry $x \mapsto \alpha(s, x)$ for every $s \in \mathbb{R}$. The Representation Theorem for isometries (Corollary at the end of Section 2) is easily seen to imply the following result.

Representation Theorem for Infinitesimal Isometries. *To every infinitesimal isometry \mathbf{k} of \mathcal{E} corresponds a unique $A_{\mathbf{k}} \in \text{Skew}(\mathcal{V})$ such that*

$$\mathbf{k}(x) = \mathbf{k}(q) + A_{\mathbf{k}}(x - q) \quad (3)$$

holds for all $x, q \in \mathcal{E}$.

Definition. We say that a force system \mathbf{f} for \mathcal{B} is **in equilibrium** in the placement κ of \mathcal{B} in \mathcal{E} if for every part \mathcal{P} of \mathcal{B} and every infinitesimal isometry \mathbf{k} of \mathcal{B}

$$\int_{\mathcal{P}} (\mathbf{k} \circ \kappa) \cdot d\mathbf{f}_{\mathcal{P}} = 0 \quad (4)$$

holds.

The right hand side of (4) can be interpreted as the rate at which the forces do work in a “virtual” rigid motion (*i.e.* a motion obtained from a rigid kinematical process; see Section 4).

The following result is obtained by applying the Representation Theorem above. The details of the proof are given in reference [7], Section 7.

Theorem A. *A force system f for \mathcal{B} is in equilibrium in the placement κ of \mathcal{B} in \mathcal{E} if and only if the following two balance laws are valid:*

I. *Balance of Forces. For every part \mathcal{P} of \mathcal{B}*

$$\text{holds.} \quad f(\mathcal{P}, \mathcal{P}^e) = \mathbf{0} \quad (5)$$

II. *Balance of Moments. For every part \mathcal{P} of \mathcal{B} and every $q \in \mathcal{E}$ the tensor*

$$K_\kappa(\mathcal{P}, q) = \left(\int_{\mathcal{P}} (\mathbf{r} \circ \kappa) \otimes d\mathbf{f}_{\mathcal{P}} \right)^T \in \text{Lin}(\mathcal{V}) \quad (6)$$

is symmetric, where $\mathbf{r}: \mathcal{E} \rightarrow \mathcal{V}$ is defined by $\mathbf{r}(x) = x - q$.

We note that the condition I does not involve the placement κ at all. Hence, if the forces are not balanced, the force system cannot be in equilibrium in *any* placement.

Since we have assumed that the force system f satisfies the conditions (I1) and (I2) of the previous section, we can apply Theorems C and D of the previous Section to a balanced force system. If it were not for the *ad hoc* continuity condition of Theorem D, we could conclude that there exists, for every placement κ of \mathcal{B} in \mathcal{E} a field

$$T_\kappa: \bar{\mathcal{B}} \rightarrow \text{Lin}(\mathcal{V})$$

such that for any $\mathcal{P} \in \Omega_{\mathcal{B}}$ and any Borel subset \mathcal{C} of $\bar{\mathcal{P}}$ we have

$$f_{\mathcal{P}}(\mathcal{C}) = \int_{\mathcal{C}} \mathbf{b}_\kappa d(\text{Vol}_\kappa) + \int_{\mathcal{C} \cap \partial\mathcal{P}} T_\kappa \mathbf{n}_\kappa d(\text{Area}_\kappa), \quad (7)$$

where $\mathbf{n}_\kappa: \partial\mathcal{P} \rightarrow \mathcal{V}$ is the unit normal to $\kappa(\partial\mathcal{P})$, directed towards the exterior of \mathcal{P} .

The field T_κ , if it exists, is called the **stress** determined by f in the placement κ .

For an arbitrary force system f , the tensor $K_\kappa(\mathcal{P}, q)$ defined by (6) is called the **astatic load** on \mathcal{P} relative to q in κ . Its skew-symmetric part is called the **moment** relative to q of the forces acting on \mathcal{P} in the placement κ . A proof of the following theorem is given in Section 5 of reference [8].

Theorem B. *Assume that f satisfies the balance of force condition I of Theorem A and that the stress T_κ exists for some and hence all placements κ of \mathcal{B} . Then the astatic load $K_\kappa(\mathcal{P}, q) = K_\kappa(\mathcal{P})$ is independent of q and given by*

$$K_\kappa(\mathcal{P}) = \int_{\mathcal{P}} T_\kappa d(\text{Vol}_\kappa). \quad (8)$$

Since the left side of (8) cannot be symmetric for *every* part \mathcal{P} of \mathcal{B} unless T_κ has symmetric values, we obtain the following consequence of Theorems A and B:

Theorem C. *A force system f that satisfies the hypotheses of Theorem B is in equilibrium in the placement κ of \mathcal{B} in \mathcal{E} if and only if the stress T_κ has symmetric values.*

Assume now that f is a force system that has a representation of the form (7) for some, and hence all, placements κ . Condition I, expressing balance of forces, is then equivalent to the condition that

$$\int_{\mathcal{P}} \mathbf{b}_\kappa d(\text{Vol}_\kappa) + \int_{\partial\mathcal{P}} T_\kappa \mathbf{n}_\kappa d(\text{Area}_\kappa) = \mathbf{0} \quad (10)$$

hold for all parts \mathcal{P} of \mathcal{B} . If T_κ is of class C^1 , this condition is equivalent to *Cauchy's Law*:

$$\text{div}_\kappa T_\kappa + \mathbf{b}_\kappa = \mathbf{0}. \quad (11)$$

If Cauchy's Law (11) holds for one placement κ , it holds automatically for all placements.

Definition. A **mechanical process** is a triple (χ, β, f) with the following properties:

- (a) $\chi: \mathcal{B} \times \mathcal{T} \rightarrow \mathcal{W}$ is a kinematical process,
- (b) $\beta: \mathcal{E} \times \mathcal{T} \rightarrow \mathcal{W}$ is a reference process,
- (c) For each $\tau \in \mathcal{T}$, f_τ is a force system in Ω that satisfies the conditions (I1) and (I2) of the previous section.
- (d) For each $\tau \in \mathcal{T}$ and each $\mathcal{P} \in \Omega$, the rate of work

$$P(\mathcal{P}, \tau) = \int_{\mathcal{P}} \dot{\mu}_\tau^* \cdot df_\tau, \quad (12)$$

of the forces f_τ in a motion μ^* determined by χ and a reference process β^* is independent of the reference process β^* .

It follows from Equation (7) of Section 4 that the Condition (d) is equivalent to the following:

- (d') For each $\tau \in \mathcal{T}$ the force system f_τ is in equilibrium in the placement μ_τ , where μ is the motion determined by χ and β .

If the continuity condition of Theorem D of the previous section could be eliminated, one would have the following result: *A mechanical process can be characterized by a triple (χ, β, T) , where $\chi: \mathcal{B} \times \mathcal{T} \rightarrow \mathcal{W}$ is a kinematical process of \mathcal{B} , $\beta: \mathcal{E} \times \mathcal{T} \rightarrow \mathcal{W}$ a reference process of the external world \mathcal{E} , and $T: \mathcal{B} \times \mathcal{T} \rightarrow \text{Sym}(\mathcal{V})$ a symmetric stress tensor field.*

The force-systems f_τ , $\tau \in \mathcal{T}$, of the process (χ, β, T) are determined as follows: If $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$, $\mathcal{P} \wedge \mathcal{Q} = \emptyset$, then

$$\mathbf{f}_\tau(\mathcal{P}, \mathcal{Q}) = \int_{\partial\mathcal{P} \cap \partial\mathcal{Q}} T_\tau \mathbf{n}_\tau d(\text{Area}_\tau), \quad (13)$$

where \mathbf{n}_τ is the exterior unit normal in the placement $\mu_\tau = \beta_\tau^{-1} \circ \chi_\tau$ and Area_τ denotes the surface area in that placement. For each $\tau \in \mathcal{T}$ there is a unique body force field \mathbf{b}_τ such that (10) holds with κ replaced by μ_τ (or simply τ). If T_τ is of class C^1 , then \mathbf{b}_τ is actually given by $\mathbf{b}_\tau = -\text{div}_{\mu_\tau} T_\tau$. We have

$$\mathbf{f}_\tau(\mathcal{P}, \mathcal{E}) = \int_{\partial\mathcal{B} \cap \partial\mathcal{P}} T_\tau \mathbf{n}_\tau d(\text{Area}_\tau) + \int_{\mathcal{P}} \mathbf{b}_\tau d(\text{Vol}_\tau). \quad (14)$$

If we put $\mathbf{f}_\tau(\mathcal{E}, \mathcal{P}) = -\mathbf{f}_\tau(\mathcal{P}, \mathcal{E})$ when $\mathcal{P} \in \Omega_{\mathcal{B}}$, then \mathbf{f}_τ is completely determined by (13) and (14).

10. Thermodynamic Processes, Entropy Production

As in the last section we assume that \mathcal{B} is a continuous body of class C^1 , and that Ω is a material universe constructed from \mathcal{B} and an external world \mathcal{E} .

Definition. A **thermodynamic process** is a septuple $(\chi, \beta, \mathbf{f}, E, H, S, \theta)$ whose entries have the following nature:

- (α) $(\chi, \beta, \mathbf{f})$ is a mechanical process on a time interval \mathcal{T} and the motion μ determined by χ and β is of class C^1 with velocity $\mathbf{v} = \dot{\mu}$.
- (β) E is a function of the type $E: \Omega_{\mathcal{B}} \times \mathcal{T} \rightarrow \mathbb{R}$, whose value $E(\mathcal{P}, \tau) = E_{\tau}(\mathcal{P})$ is called the *energy content* of \mathcal{P} at time $\tau \in \mathcal{T}$. It is assumed that E_{τ} is additive and that E_{τ} and \dot{E}_{τ} (defined by $\dot{E}_{\tau}(\mathcal{P}) = \dot{E}(\mathcal{P}, \tau)$) are volume continuous.
- (γ) H is a function of the type $H: (\Omega \times \Omega)_{\text{sep}} \times \mathcal{T} \rightarrow \mathbb{R}$, whose value $H(\mathcal{P}, \mathcal{A}, \tau) = H_{\tau}(\mathcal{P}, \mathcal{A})$ is called the *rate of heat transfer* from \mathcal{A} to \mathcal{P} at time $\tau \in \mathcal{T}$. It is assumed that H_{τ} is an interaction that satisfies the conditions (I1) and (I2) of Sect. 8.
- (δ) S is a function of the type $S: \Omega_{\mathcal{B}} \times \mathcal{T} \rightarrow \mathbb{R}$, whose value $S(\mathcal{P}, \tau) = S_{\tau}(\mathcal{P})$ is called the *entropy content* of \mathcal{P} at time $\tau \in \mathcal{T}$. It is assumed that S_{τ} and \dot{S}_{τ} are volume continuous.
- (ε) θ is a function of the type $\theta: \bar{\mathcal{B}} \times \mathcal{T} \rightarrow \mathbb{R}^{++}$ whose value $\theta(\chi, \tau)$ is called the *temperature* at $\chi \in \bar{\mathcal{B}}$ at time $\tau \in \mathcal{T}$.

The septuple is subject to the following axiom, called the **law of energy balance**: For all parts \mathcal{P} of \mathcal{B} and all $\tau \in \mathcal{T}$, we have

$$\dot{E}_{\tau}(\mathcal{P}) = H_{\tau}(\mathcal{P}, \mathcal{P}^e) + P_{\tau}(\mathcal{P}), \quad (1)$$

where

$$P_{\tau}(\mathcal{P}) = \int_{\mathcal{P}} \mathbf{v}_{\tau} \cdot d\mathbf{f}_{\tau, \mathcal{P}}, \quad \mathbf{v}_{\tau}(X) = \dot{\mu}(X, \tau), \quad (2)$$

is the rate of work of the forces acting on \mathcal{P} in the motion μ determined by χ and β .

Recall that by Condition (d) of the previous section, P is independent of the reference process β and depends on μ only through χ . Using the results of the previous section, one can express P in terms of the stress as follows:

Theorem A. Let κ be a placement of \mathcal{B} in \mathcal{E} . If the force systems \mathbf{f}_{τ} are described by a stress $T_{\kappa}: \bar{\mathcal{B}} \times \mathcal{T} \rightarrow \text{Lin}(\mathcal{V})$, then the rate of working is given by

$$P(\mathcal{P}) = \int_{\mathcal{P}} \text{tr}(T_{\kappa} \nabla_{\kappa} \mathbf{v}) d(\text{Vol}_{\kappa}). \quad (3)$$

From now on we omit the subscript τ and consider dependence on time as understood.

It follows from the assumed additivity and volume continuity of E_{τ} , \dot{E}_{τ} , S_{τ} and \dot{S}_{τ} that for every placement κ there exists an *energy density* $\varepsilon_{\kappa}: \mathcal{B} \times \mathcal{T} \rightarrow \mathbb{R}$ and an *entropy density* $\eta_{\kappa}: \mathcal{B} \times \mathcal{T} \rightarrow \mathbb{R}$ such that

$$E(\mathcal{P}) = \int_{\mathcal{P}} \varepsilon_{\kappa} d(\text{Vol}_{\kappa}), \quad \dot{E}(\mathcal{P}) = \int_{\mathcal{P}} \dot{\varepsilon}_{\kappa} d(\text{Vol}_{\kappa}) \quad (4)$$

and

$$S(\mathcal{P}) = \int_{\mathcal{P}} \eta_{\kappa} d(\text{Vol}_{\kappa}), \quad \dot{S}(\mathcal{P}) = \int_{\mathcal{P}} \dot{\eta}_{\kappa} d(\text{Vol}_{\kappa}). \quad (5)$$

If the force systems can be described by a stress field, then Theorem A shows that the rate of working P is additive and volume continuous. Since \dot{E} is additive and volume continuous by assumption, it follows from the law (1) of energy balance that $\mathcal{P} \mapsto H(\mathcal{P}, \mathcal{P}^e)$ is additive and volume continuous. The hypotheses of Theorem C of Section 8 are therefore satisfied for the interaction H . If it were not for the continuity condition of Theorem D of Section 9 we could conclude that with each $\mathcal{P} \in \Omega_{\mathcal{B}}$ one can associate a bounded Borel measure $H_{\mathcal{P}}$ on \mathcal{P} of the form

$$H_{\mathcal{P}}(\mathcal{C}) = \int_{\mathcal{C}} r_{\kappa} d(\text{Vol}_{\kappa}) + \int_{\mathcal{C} \cap \partial \mathcal{P}} \mathbf{q}_{\kappa} \cdot \mathbf{n}_{\kappa} d(\text{Area}_{\kappa}) \quad (6)$$

where κ is an arbitrary placement of \mathcal{B} .

We have $H_{\mathcal{Q}}(\mathcal{P}) = H(\mathcal{P}, \mathcal{Q}^e)$ whenever $\mathcal{P}, \mathcal{Q} \in \Omega_{\mathcal{B}}$, $\mathcal{P} \subset \mathcal{Q}$. The function $r_{\kappa}: \mathcal{B} \times \mathcal{T} \rightarrow \mathbb{R}$ is called the *heat supply* and $\mathbf{q}_{\kappa}: \mathcal{B} \times \mathcal{T} \rightarrow \mathcal{V}$ the *heat flux* determined by H in the placement κ .

If we substitute (3), (4) and (6), with $\mathcal{C} = \mathcal{P}$, into the energy balance equation (1), we obtain

$$\int_{\mathcal{P}} (\dot{\varepsilon}_{\kappa} - \text{tr}(T_{\kappa} \nabla_{\kappa} \mathbf{v}) - r_{\kappa}) d(\text{Vol}_{\kappa}) - \int_{\partial \mathcal{P}} \mathbf{q}_{\kappa} \cdot \mathbf{n}_{\kappa} d(\text{Area}_{\kappa}) = 0. \quad (7)$$

If \mathbf{q}_{κ} is of class C^1 , then (7) is valid for all parts \mathcal{P} of \mathcal{B} if and only if the *local energy balance equation*

$$\dot{\varepsilon}_{\kappa} = \text{tr}(T_{\kappa} \nabla_{\kappa} \mathbf{v}) + \text{div}_{\kappa} \mathbf{q}_{\kappa} + r_{\kappa} \quad (8)$$

holds.

One can derive a result for thermodynamical process analogous to the one for mechanical processes stated at the end of the previous section: If continuity conditions could be ignored, every thermodynamical process could be characterized by a septuple $(\chi, \beta, T, \varepsilon, \mathbf{q}, \eta, \theta)$, where $T, \varepsilon, \mathbf{q}, \eta, \theta$ are functions on $\mathcal{B} \times \mathcal{T}$ representing the stress, energy density, heat flux, entropy density, and temperature corresponding to the actual placements μ_{τ} of the motion determined by χ and β . The local energy balance equation (8) then determines the heat supply $r(X, \tau) = r_{\mu_{\tau}}(X, \tau)$.

We assume now that a thermodynamical process is given. Using the measures $H_{\mathcal{Q}}$ defined by (6), we define the rate of *entropy transfer* from \mathcal{Q}^e to $\mathcal{P} \in \Omega_{\mathcal{B}}$, $\mathcal{P} \subset \mathcal{Q}$, by

$$M(\mathcal{P}, \mathcal{Q}^e) = \int_{\mathcal{P}} \frac{1}{\theta} dH_{\mathcal{Q}} = \int_{\mathcal{P}} \frac{1}{\theta} r_{\kappa} d(\text{Vol}_{\kappa}) + \int_{\partial \mathcal{P} \cap \partial \mathcal{Q}^e} \frac{1}{\theta} \mathbf{q}_{\kappa} \cdot \mathbf{n}_{\kappa} d(\text{Area}_{\kappa}). \quad (9)$$

Definition. The *rate of entropy production* $\Gamma: \Omega_{\mathcal{B}} \times \mathcal{T} \rightarrow \mathbb{R}$ of a thermodynamic process is defined by

$$\Gamma(\mathcal{P}, \tau) = \dot{S}(\mathcal{P}, \tau) - M(\mathcal{P}, \mathcal{P}^e, \tau). \quad (10)$$

If the heat flux \mathbf{q}_{κ} and the temperature θ are of class C^1 , one can use the divergence theorem in (9). After substitution of (5)₂ and (9) into (10) one then obtains

$$\Gamma(\mathcal{P}) = \int_{\mathcal{P}} \gamma_{\kappa} d(\text{Vol}_{\kappa}), \quad (11)$$

where $\gamma_{\kappa}: \mathcal{B} \times \mathcal{T} \rightarrow \mathbb{R}$, the *local rate of entropy production*, is given by

$$\gamma_{\kappa} = \dot{\eta}_{\kappa} - \frac{1}{\theta} r_{\kappa} - \text{div}_{\kappa} \left(\frac{1}{\theta} \mathbf{q}_{\kappa} \right). \quad (12)$$

Using (8), we can eliminate r_κ from (12) and obtain

$$\theta \gamma_\kappa = \theta \dot{\eta}_\kappa - \dot{\varepsilon}_\kappa + \text{tr}(T_\kappa \nabla_\kappa \mathbf{v}) + \frac{1}{\theta} \mathbf{q}_\kappa \cdot \nabla_\kappa \theta. \quad (13)$$

In view of (11) we see that $\Gamma(\mathcal{P}) \geq 0$ can hold for all parts \mathcal{P} of \mathcal{B} if and only if $\gamma_\kappa \geq 0$. Hence (13) gives the following result:

Theorem B. *If the thermodynamic process is such that for some, and hence every, placement κ of \mathcal{B} the stress T_κ and heat flux \mathbf{q}_κ exist and if \mathbf{q}_κ and the temperature θ are of class C^1 , then the rate of entropy production Γ has non-negative values if and only if*

$$\theta \dot{\eta}_\kappa - \dot{\varepsilon}_\kappa + \text{tr}(T_\kappa \nabla_\kappa \mathbf{v}) + \frac{1}{\theta} \mathbf{q}_\kappa \cdot \nabla_\kappa \theta \geq 0 \quad (14)$$

holds for some, and hence every, placement κ .

Remark. GURTIN & WILLIAMS in 1967 (reference [4]) have developed a foundation of thermodynamics in which the rate of entropy transfer M is regarded as one of the primitive data in the definition of a thermodynamical process. Using certain physically well motivated axioms, they then derive from M and H not one but two temperatures, a conductive temperature and a radiative temperature. They then proceed to show that under many, but not all, circumstances, the two temperatures are the same.

11. Constitutive Assumptions, the Dissipation Principle

Constitutive assumptions are mathematical statements that describe particular physical circumstances which govern the behavior of particular bodies and their environments. A constitutive assumption is a relation involving the seven functions $\chi, \beta, \mathbf{f}, E, H, S, \theta$ that define a thermodynamic process. Processes that satisfy this relation are said to be *admissible* for the constitutive assumption in question.

External constitutive assumptions describe the influence of the environment \mathcal{E} on the body \mathcal{B} . We describe those assumptions that are very often made in continuum physics:

- (a) There are two additive and volume-continuous functions $m_i, m_g: \Omega_{\mathcal{B}} \rightarrow \mathbb{R}^+$ and a scalar field $\phi: \mathcal{E} \times \mathcal{T} \rightarrow \mathbb{R}$ of class C^1 such that

$$\mathbf{f}^{\text{ext}}(\mathcal{P}) = - \int_{\mathcal{P}} \mathbf{a} \, d m_i + \int_{\mathcal{P}} \nabla \phi \circ \mu \, d m_g \quad \text{when } \bar{\mathcal{P}} \subset \mathcal{B} \quad (1)$$

holds for all admissible processes. Here \mathbf{f}^{ext} is the external force defined by \mathbf{f} , i. e. $\mathbf{f}^{\text{ext}}(\mathcal{P}) = \mathbf{f}(\mathcal{P}, \mathcal{E})$, and $\mathbf{a} = \ddot{\mu}$ is the acceleration for the motion μ defined by χ and β . m_i is called the *inertial mass* and m_g the *gravitational mass*. The first term on the right side of (1) is called the *inertial force*, the second is called the *gravitational force* acting on \mathcal{P} . The field ϕ is called the *gravitational potential* of the environment \mathcal{E} .

It is an experimental fact that the ratio $m_i(\mathcal{P})/m_g(\mathcal{P})$ is not only independent of $\mathcal{P} \in \Omega_{\mathcal{B}}$, but actually the same for all bodies ever encountered in nature. By suitable choice of units, one therefore can take $m_i = m_g = m$, replace (1) by

$$\mathbf{f}^{\text{ext}}(\mathcal{P}) = \int_{\mathcal{P}} (-\mathbf{a} + \nabla \phi \circ \mu) \, d m \quad \text{when } \bar{\mathcal{P}} \subset \mathcal{B}, \quad (2)$$

and call the function $m: \Omega_{\mathcal{B}} \rightarrow \mathbb{R}^+$ the *mass*. Since m is assumed to be volume continuous, we can associate with every placement κ of \mathcal{B} a *mass density* $\rho_\kappa: \mathcal{B} \rightarrow \mathbb{R}^+$ such that

$$m(\mathcal{P}) = \int_{\mathcal{P}} \rho_\kappa \, d(\text{Vol}_\kappa). \quad (3)$$

The condition (2) can be expressed in terms of ρ_κ and the body force density \mathbf{b}_κ by

$$\mathbf{b}_\kappa = \rho_\kappa (-\mathbf{a} + \nabla \phi \circ \mu). \quad (4)$$

- (b) If $\bar{\mathcal{P}} \not\subset \mathcal{B}$, *i.e.* if $\partial\mathcal{P} \cap \partial\mathcal{B} \neq \emptyset$, a surface integral over $\partial\mathcal{P} \cap \partial\mathcal{B}$ is added to the right side of (2). This integral represents the *surface traction* acting on \mathcal{B} . A variety of assumptions on this surface traction represent a variety of physical circumstances.
- (c) It is usually assumed that there is no heat transfer at a distance from the environment, *i.e.* that

$$H^{\text{ext}}(\mathcal{P}) = H(\mathcal{P}, \mathcal{E}) = 0 \quad \text{when } \bar{\mathcal{P}} \subset \mathcal{B}. \quad (5)$$

In terms of the heat supply, (5) means that

$$r_\kappa = 0 \quad (6)$$

for some, and hence all, placements κ of \mathcal{B} .

- (d) If $\partial\mathcal{P} \cap \partial\mathcal{B} \neq \emptyset$, then $H^{\text{ext}}(\mathcal{P})$ is set equal to a surface integral over $\partial\mathcal{P} \cap \partial\mathcal{B}$. This integral represents the heat supplied to \mathcal{B} by contact with the environment. A variety of assumptions on this integral represent a variety of physical conditions.

Internal constitutive assumptions describe the internal physical constitution of the body \mathcal{B} . They are relations that involve only the internal portion $(\chi, \bar{\beta}f^{\text{int}}, E, H^{\text{int}}, S, \theta)$ of the thermodynamic process $(\chi, \beta, \mathbf{f}, E, H, S, \theta)$ considered. The kinematical process β of the environment \mathcal{E} should not occur and only the internal parts $\bar{\beta}f^{\text{int}}$ and H^{int} of the interactions \mathbf{f} and H should be involved. It is $\bar{\beta}_\tau f_\tau^{\text{int}}$ rather than f_τ^{int} that describes the internal forces at time τ intrinsically, without reference to the environment \mathcal{E} .

A fundamental restriction on the possible internal constitutive assumptions must be imposed:

Dissipation Principle. *For all members of the class of thermodynamic processes that are admissible for an internal constitutive assumption, the rate of entropy production must have non-negative values.*

A body \mathcal{B} is said to consist of a **simple material** if its physical behaviour is completely determined by the physical properties of its infinitesimal elements. In order to make this idea more precise, we introduce the concept of a thermodynamic process of a body element as follows:

Let $(\chi, \beta, \mathbf{f}, E, H, S, \Theta)$ be a thermodynamic process, let μ be the motion determined by χ and β , and let $X \in \mathcal{B}$ be a material point. Assume that the process is defined on the closed time interval $\mathcal{T} = [\tau_0, \tau_0 + r]$. The function $M: [0, r] \rightarrow$

Invlin $(\mathcal{T}_X, \mathcal{V})$ defined by

$$M(t) = \nabla \mu_{\tau+t}(X), \quad t \in [0, r] \quad (7)$$

describes the motion of the body element \mathcal{T}_X (see Sect. 7). The deformation process P of \mathcal{T}_X induced by the deformation process p induced by the kinematical process χ is

$$P = M^* M: [0, r] \rightarrow \mathbf{G} \subset \text{Sym}^+(\mathcal{T}_X, \mathcal{T}_X^*). \quad (8)$$

We define the *intrinsic stress*

$$S: [0, r] \rightarrow \text{Sym}(\mathcal{T}_X^*, \mathcal{T}_X)$$

in terms of the stress T_{μ_τ} (assumed to exist) by

$$S(t) = M(t)^{-1} T_{\mu_{\tau_0+t}}(X, \tau_0 + t) M(t)^{* -1}. \quad (9)$$

Also, we define the *intrinsic heat flux*

$$\mathbf{q}: [0, r] \rightarrow \mathcal{T}_X$$

in terms of the heat fluxes \mathbf{q}_{μ_τ} (assumed to exist) by

$$\mathbf{q}(t) = M(t)^{-1} \mathbf{q}_{\mu_{\tau_0+t}}(X, \tau_0 + t). \quad (10)$$

The *intrinsic temperature gradient* $\mathbf{g}: [0, r] \rightarrow \mathcal{T}_X^*$ is defined by

$$\mathbf{g}(t) = \nabla \theta(X, \tau_0 + t) = (\nabla_{\mu_{\tau_0+t}} \theta(X, \tau_0 + t)) M(t). \quad (11)$$

It is useful to choose a *reference configuration* $G_R \in \mathbf{G}$ of \mathcal{T}_X and to introduce an energy density $\varepsilon_R: [0, r] \rightarrow \mathbb{R}$ and an entropy density $\eta_R: [0, r] \rightarrow \mathbb{R}$ by

$$\varepsilon_R(t) = \varepsilon_\kappa(X, \tau_0 + t), \quad \eta_R(t) = \eta_\kappa(X, \tau_0 + t), \quad (12)$$

where $\kappa \in \mathbf{P}$ is such that $G_R = \nabla \kappa(X)^* \nabla \kappa(X)$. It is easily seen that ε_R and η_R depend on κ only through G_R .

The septuple $(P, S, \varepsilon_R, \mathbf{q}, \eta_R, \theta, \mathbf{g})$ of functions on $[0, r]$ with values in \mathbf{G} , $\text{Sym}(\mathcal{T}_X^*, \mathcal{T}_X)$, \mathbb{R} , \mathcal{T}_X , \mathbb{R} , \mathbb{R}^{++} , and \mathcal{T}_X^* , respectively, is called the *thermodynamic process of the element* \mathcal{T}_X induced by the thermodynamic process $(\chi, \beta, \mathbf{f}, E, H, S, \theta)$ of \mathcal{B} . It is easily seen that the induced process depends only on the internal portion $(\chi, \beta \mathbf{f}^{\text{int}}, E, H^{\text{int}}, S, \theta)$ of the given process.

We say that \mathcal{B} is a *simple body* if the internal constitutive assumptions for \mathcal{B} are relations involving only the thermodynamical processes of the body elements.

We introduce a *local volume* $j_R: [0, r] \rightarrow \mathbb{R}^{++}$ by

$$j_R(t)^2 = \det(G_R^{-1} P(t)), \quad t \in [0, r], \quad (13)$$

and a *local entropy production* $\gamma_R: [0, r] \rightarrow \mathbb{R}$ by

$$\gamma_R(t) = \gamma_\kappa(X, \tau_0 + t), \quad t \in [0, r], \quad (14)$$

when $G_R = \nabla \kappa(X)^* \nabla \kappa(X)$. It follows from Equation (13) of the previous section that

$$\theta \gamma_R = \theta \dot{\eta}_R - \dot{\varepsilon}_R + j_R \left\{ \frac{1}{2} \text{tr}(S \dot{P}) + \frac{1}{\theta} \langle \mathbf{g}, \mathbf{q} \rangle \right\}, \quad (15)$$

which shows that γ_R depends only on the thermodynamic process of the element \mathcal{T}_X under consideration. Theorem B of the previous Section implies the following result:

Theorem A. *The dissipation principle is satisfied for a material element \mathcal{T} if and only if for every admissible thermodynamical process $(P, S, \varepsilon_R, q, \eta_R, \theta, g)$ of \mathcal{T} , the local entropy production γ_R defined by (15) has non-negative values.*

We now show by an example how Theorem A can be applied. A *Navier-Stokes fluid with Fourier heat conduction* is described by 6 material functions $\bar{p}, \bar{\lambda}, \bar{\mu}, \bar{\varepsilon}_R, \bar{\theta}$, and $\bar{\kappa}$ from $\mathbb{R}^{++} \times \mathbb{R}$ into \mathbb{R} . The constitutive relations are the following:

(i) Navier-Stokes relation:

$$PSP = (-p + \lambda \operatorname{tr}(P^{-1} \dot{P}))P + \mu \dot{P}, \tag{16}$$

where

$$p = \bar{p}(j_R, \eta_R), \quad \lambda = \bar{\lambda}(j_R, \eta_R), \quad \mu = \bar{\mu}(j_R, \eta_R). \tag{17}$$

(ii) Equations of state:

$$\varepsilon_R = \bar{\varepsilon}_R(j_R, \eta_R), \quad \theta = \bar{\theta}(j_R, \eta_R). \tag{18}$$

(iii) Fourier's law of heat conduction:

$$Pq = \kappa g, \quad \text{where } \kappa = \bar{\kappa}(j_R, \eta_R). \tag{19}$$

The following result is a consequence of Theorem A, but the proof, given in reference [9], is not trivial.

Theorem B. *The constitutive relations (16)–(19) for a Navier-Stokes fluid with Fourier heat conduction satisfy the dissipation principle if and only if the following conditions are satisfied:*

- (a) $\bar{\theta} = \bar{\varepsilon}_{R,2}$ (temperature relation),
- (b) $\bar{p} = -\bar{\varepsilon}_{R,1}$ (pressure relation),
- (c) $\bar{\mu} \geq 0, \bar{\lambda} + \frac{2}{3}\bar{\mu} \geq 0$ (viscosity inequalities),
- (d) $\bar{\kappa} \geq 0$ (heat conductivity inequality).

Remark. The formulation of the dissipation principle given here is due to COLEMAN & NOLL (1963) (reference [9]), who also gave the first non-trivial application, of which Theorem B is a special case. Since then, the principle has been applied by many authors to a variety of constitutive relations, often with surprising results.

Appendix: The Concept of a Material Universe

A *material universe* Ω is a set endowed with a structure defined by a relation \prec , subject to the axioms (B1)–(B6) below. The elements A, B, C, \dots are called *material objects* or *bodies*.

- (B1) *Certain pairs (A, B) of bodies are related by \prec ; we write $A \prec B$ if this is the case and say that A is a part of B .*
- (B2) *$A = B$ if and only if both $A \prec B$ and $B \prec A$.*
- (B3) *If $A \prec B$ and $B \prec C$, then $A \prec C$.*

The axioms (B1)–(B3) state that \prec is a *partial ordering* of Ω .

If $B \prec A$ and $C \prec A$ we call A an *envelope* of $\{B, C\}$ and write $\{B, C\} \prec A$. If $\{B, C\} \prec D$ and if $\{B, C\} \prec A$ implies $D \prec A$ we say that D is the *least envelope* of $\{B, C\}$ and we write $D = B \vee C$, so that

$$\{B, C\} \prec A \quad \text{implies} \quad \{B, C\} \prec B \vee C \prec A. \quad (1)$$

A least envelope may or may not exist, but if it does, it is uniquely determined by B and C . If $A \prec B$ and $A \prec C$ we say that A is a *common part* of $\{B, C\}$ and write $A \prec \{B, C\}$. If $D \prec \{B, C\}$ and if $A \prec \{B, C\}$ implies $A \prec D$ we say that D is the *greatest common part* of $\{B, C\}$ and write $D = B \wedge C$, so that

$$A \prec \{B, C\} \quad \text{implies} \quad A \prec B \wedge C \prec \{B, C\}. \quad (2)$$

If there is a greatest common part it is unique. Envelopes, the greatest envelope, common parts, and the greatest common part of an arbitrary collection of bodies are defined in a similar manner. Let $\{A_i | i \in I\}$ be such a collection, its members A_i marked with indices i taken from an index set I . If the least envelope of the collection exists, we denote it by $\bigvee_{i \in I} A_i$; if the greatest common part exists, we denote it by $\bigwedge_{i \in I} A_i$.

Two bodies A and B are said to be *separate* if they have no common part. It will be convenient to adjoin to the material universe Ω two *improper bodies*, the *null-body* \emptyset and the *universal body* ∞ . We extend the relation \prec to the extended universe

$$\Omega' = \Omega \cup \{\emptyset, \infty\} \quad (3)$$

by putting

$$\emptyset \prec A \prec \infty \quad \text{for all } A \in \Omega'. \quad (4)$$

Of course, the relation \prec , when extended by (4), remains a partial ordering. We have

$$A \wedge B = \emptyset, \quad A, B \in \Omega \quad (5)$$

if and only if A and B are separate.

The following rules are easily established:

I. $A \prec B$ if and only if $A \wedge B = A$ or $A \vee B = B$.

II. If $A \prec B$ and if $A \wedge C$, $B \wedge C$ exist, then

$$A \wedge C \prec B \wedge C. \quad (6)$$

III. If $A \prec B$ and if $A \vee C$, $B \vee C$ exist, then

$$A \vee C \prec B \vee C. \quad (7)$$

IV. If $A \prec B$ and if $B \wedge C = \emptyset$, then $A \wedge C = \emptyset$.

V. If $A \wedge B$ and $B \wedge C$ exist, then

$$(A \wedge B) \wedge C = A \wedge (B \wedge C) = A \wedge B \wedge C, \quad (8)$$

provided either $(A \wedge B) \wedge C$ or $A \wedge (B \wedge C)$ exist.

It is possible that $A \wedge B \wedge C$ exists but $A \wedge B$ or $B \wedge C$ do not.

We are now ready to state the next axiom:

(B4) *For each body $A \in \Omega$ there is exactly one body $A^e \in \Omega$ with the property that $\{A, A^e\}$ have neither a common part nor an envelope; i. e.,*

$$A \wedge A^e = \emptyset, \quad A \vee A^e = \infty. \quad (9)$$

The body A^e is called *the exterior* of the body A . If we put $\emptyset^e = \infty$, $\infty^e = \emptyset$, then (9) is valid for all $A \in \Omega'$. It follows directly from the definition of A^e that

$$(A^e)^e = A \quad (10)$$

holds for every $A \in \Omega'$. Also,

$$A \prec B \text{ implies } A \wedge B^e = \emptyset, \quad (11)$$

which is a consequence of (9)₁ and rule IV. The next axiom postulates that the converse of (11) is valid:

(B5) *If A and the exterior of B have no common part, then A is a part of B .*

It follows from (B5) and (11) that

$$A \prec B \text{ if and only if } A \wedge B^e = \emptyset. \quad (12)$$

By (12) and (10) we have $A \prec B$ if and only if $B^e \wedge (A^e)^e = \emptyset$. Hence, using (12) again, we find that

$$A \prec B \text{ if and only if } B^e \prec A^e. \quad (13)$$

The following propositions are corollaries of (13): Let $\{A_i | i \in I\}$ be a collection of bodies. If $\bigvee_{i \in I} A_i$ exists, so does $\bigwedge_{i \in I} (A_i^e)$ and

$$\bigwedge_{i \in I} (A_i^e) = (\bigvee_{i \in I} A_i)^e. \quad (14)$$

If $\bigwedge_{i \in I} A_i$ exists, so does $\bigvee_{i \in I} (A_i^e)$ and

$$\bigvee_{i \in I} (A_i^e) = (\bigwedge_{i \in I} A_i)^e. \quad (15)$$

The following result is basic to the theory of bodies:

Lemma. *Assume that A_1, A_2 , and B are bodies and that $A_1 \wedge B, A_2 \wedge B$, and $A_1 \vee A_2$ exist. Then we have*

$$\{A_1 \wedge B, A_2 \wedge B\} \prec \{B, A_1 \vee A_2\}; \quad (16)$$

furthermore, if

$$\{A_1 \wedge B, A_2 \wedge B\} \prec C \quad (17)$$

and

$$D \prec \{B, A_1 \vee A_2\}, \quad (18)$$

then

$$D \prec C. \quad (19)$$

Proof. It follows from the definitions (1) and (2) that $A_i \wedge B \prec B$, and $A_i \wedge B \prec A_i \prec A_1 \vee A_2$, $i = 1, 2$, which immediately gives (16).

Assume now that (17) and (18) hold. It follows from (17) and (11) that

$$\{A_i \wedge B\} \wedge C^e = \emptyset, \quad i = 1, 2. \quad (20)$$

Suppose that E is a common part of $\{D, C^e\}$, *i.e.*

$$E \prec \{D, C^e\}. \quad (21)$$

By (18) we have $E \prec D \prec B$, and hence

$$E \prec \{B, C^e\}. \quad (22)$$

Now, if G_i is a common part of $\{E, A_i\}$,

$$G_i \prec \{E, A_i\}, \quad (23)$$

then $G_i \prec E \prec B$ by (22) and hence $G_i \prec \{B, A_i\}$, which is equivalent to

$$G_i \prec A_i \wedge B, \quad i = 1, 2. \quad (24)$$

Using rule IV, we infer from (24) and (20) that $G_i \wedge C^e = \emptyset$. But by (23) and (22) we also have $G_i \prec E \prec C^e$ and hence, by rule IV, $G_i = G_i \wedge G_i \prec G_i \wedge C^e = \emptyset$, *i.e.* $G_i = \emptyset$. We have shown that (23) implies $G_i = \emptyset$, which means that

$$E \wedge A_i = \emptyset, \quad i = 1, 2. \quad (25)$$

Applying (10) and (12), we infer from (25) that $A_i \prec E^e$, $i = 1, 2$, which is equivalent to $A_1 \vee A_2 \prec E^e$, or, by (13), to

$$E \prec (A_1 \vee A_2)^e. \quad (26)$$

On the other hand, it follows from (21) and (18) that

$$E \prec D \prec A_1 \vee A_2. \quad (27)$$

The relations (26) and (27) can both hold only if $E = \emptyset$. We have shown that (21) implies $E = \emptyset$, which means that $D \wedge C^e = \emptyset$. But by (12) this is equivalent to (19). Q.E.D.

The following four propositions are easy consequences of the Lemma:

(A) *If $A_1 \wedge B$, $A_2 \wedge B$, and $A_1 \vee A_2$ exist, then*

$$(A_1 \vee A_2) \wedge B = (A_1 \wedge B) \vee (A_2 \wedge B), \quad (28)$$

provided either the left hand side or the right hand side exists.

(B) *If $A_1 \vee B$, $A_2 \vee B$ and $A_1 \wedge A_2$ exist, then*

$$(A_1 \wedge A_2) \vee B = (A_1 \vee B) \wedge (A_2 \vee B), \quad (29)$$

provided either the left hand side or the right hand side exists.

(C) *If $A \wedge B$ and $A^e \wedge B$ exist, then*

$$B = (A \wedge B) \vee (A^e \wedge B). \quad (30)$$

(D) *We have*

$$A = B \vee C, \quad B \wedge C = \emptyset \quad (31)$$

if and only if

$$B \prec A, \quad C = A \wedge B^e. \quad (32)$$

Proposition (A) is an immediate corollary of the Lemma; (B) results from applying (A) to A_1^e , A_2^e , and B^e and using (14) and (15); (C) corresponds to the special case $A_1 = A$, $A_2 = A^e$ of (A). To prove (D), assume that (31) holds. Then $B \prec A$ follows from (31)₁. Using (12) and (10) we infer from (31)₂ that $C \prec B^e$, i.e. that $C \wedge B^e = C$. Using (31)₁ and (A), we obtain

$$A \wedge B^e = (B \vee C) \wedge B^e = (B \wedge B^e) \vee (C \wedge B^e) = \emptyset \vee C = C,$$

which proves (32). Assume, conversely, that (32) holds. We then have $A \wedge B = B$ and hence, by (C),

$$A = (B \wedge A) \vee (B^e \wedge A) = B \vee C.$$

Also, $\emptyset = \emptyset \wedge A = (B \wedge B^e) \wedge A = B \wedge (B^e \wedge A) = B \wedge C$, which proves (31).

We now state the final axiom:

(B6) *If $A, B \in \Omega$ are not separate, then the greatest common part $A \wedge B \in \Omega$ exists.*

It follows from (B6) that in the extended universe defined by (3), $A \wedge B$ exists for all $A, B \in \Omega'$. By (15), the least envelope $A \vee B = (A^e \wedge B^e)^e \in \Omega'$ exists always also. Equations (28) and (29) state that the operations \wedge and \vee are distributive with respect to one another. This fact, together with axiom (B4), can be expressed by saying that Ω' is a *complemented distributive lattice*.

Suppose that a body B is given. A finite collection $\mathfrak{P} = \{P_i | i=1, \dots, n\}$ is called a *partition of B* if

$$B = \bigvee_{i=1}^n P_i \quad \text{and} \quad P_i \wedge P_j = 0 \quad \text{if} \quad i \neq j. \quad (33)$$

We say that the partition $\mathfrak{P}' = \{P'_i | i=1, \dots, n'\}$ is a *refinement* of \mathfrak{P} if for each $P'_i \in \mathfrak{P}'$ there is a $P_j \in \mathfrak{P}$ such that $P'_i \prec P_j$.

Theorem. *Any two partitions $\mathfrak{P} = \{P_i | i=1, \dots, n\}$, $\mathfrak{Q} = \{Q_j | j=1, \dots, m\}$ have a common refinement; i.e., there is a partition which is a refinement of both \mathfrak{P} and \mathfrak{Q} .*

Proof. The required refinement consists of the nm parts $P_i \wedge Q_j$, $i=1, \dots, n$, $j=1, \dots, m$. The fact that these parts do indeed form a refinement of both \mathfrak{P} and \mathfrak{Q} follows easily from the distributive laws (28), (29), and from $P_i \wedge Q_j \prec \{P_i, Q_j\}$. Q.E.D.

The axioms (B1)–(B6) and the theorems derived from them reflect our common sense experiences with physical bodies. It is necessary here to make the mathematical description of bodies independent of any imbedding in “space”, simply because there is no such thing as “space” in our development.

The following purely mathematical special examples illustrate the concept of a material universe:

(α) Let Ω' consist in all subsets of an arbitrary set X and let \prec represent set inclusion. We have $A \wedge B = A \cap B$ (intersection), $A \vee B = A \cup B$ (union) and $A^e = A^c$

(complement). The Newtonian mechanics of particle systems can be based on such a universe when X is a finite set.

(β) Let Ω' consist in all closures of open sets in an arbitrary topological space and let \prec represent set inclusion. In this case the greatest common part of two bodies A, B is *not* always their intersection, but rather $A \wedge B = \overline{\bar{A} \cap \bar{B}}$ (a superimposed \circ denotes the interior and a superimposed bar denotes the closure). The least envelope of a collection $\{A_i | i \in I\}$ is given by $\bigvee_{i \in I} A_i = \overline{\bigcup_{i \in I} \bar{A}_i}$, *i.e.* by the closure of the union of the interiors of the sets in the collection. This least envelope is equal to the union $\bigcup_{i \in I} A_i$ if the collection is finite, but not necessarily if it is infinite. The exterior of a body is $A^e = \bar{A}^c$, *i.e.* the closure of the complement.

(γ) Let Ω consist in all finite unions of closed polyhedra and closed exteriors of polyhedra in a Euclidean space. Ω satisfies the axioms of a material universe when \prec represents set inclusion. Greatest common parts, least envelopes, and exteriors are given by the same formulas as in example (β).

(δ) Let Ω consist in all closed regions with piecewise smooth boundaries, in a Euclidean space. If \prec represents set inclusion, then the axioms (B1)–(B5) are satisfied, but (B6) is not. Since this example is of importance in conventional continuum mechanics, it may be desirable to develop the mechanics of material universes without postulating (B6). However, this leads to considerable technical difficulties, and the issue has not yet been resolved.

Acknowledgement. The research leading to this paper was supported by grants from the National Science Foundation.

References

1. NOLL, W., The foundations of mechanics. Non-Linear Continuum Theories, C.I.M.E. Lectures, 1965, pp. 159–200, Rome (1966)
2. NOLL, W., A new mathematical theory of simple materials. *Archive for Rational Mech. Anal.* **48**, 1–50 (1972)
3. NOLL, W., Euclidean geometry and Minkowskian chronometry. *American Mathematical Monthly* **71**, 129–144 (1964)
4. GURTIN, M. E., & W. O. WILLIAMS, An axiomatic foundation for continuum thermodynamics. *Archive for Rational Mech. Anal.* **26**, 83–117 (1967)
5. NOLL, W., The foundations of classical mechanics in the light of recent advances in continuum mechanics. *Proceedings of the Berkeley Symposium on the Axiomatic Method*, 226–281, Amsterdam, 1959
6. GURTIN, M. E., V. J. MIZEL & W. O. WILLIAMS, A note on Cauchy's stress theorem. *J. Math. Anal. Appl.* **22**, 398–401 (1968)
7. NOLL, W., La Mécanique Classique, Basée sur un Axiome d'Objectivité. *Colloque International sur la Méthode Axiomatique dans les Mécaniques Classiques et Nouvelles* (1959), pp. 47–56, Paris, 1963
8. COLEMAN, B. D., & W. NOLL, On the thermostatics of continuous media. *Archive for Rational Mech. Anal.* **4**, 97–128 (1959)
9. COLEMAN, B. D., & W. NOLL, The thermodynamics of elastic materials with heat conduction and viscosity. *Archive for Rational Mech. Anal.* **13**, 167–178 (1963)

Department of Mathematics
Carnegie-Mellon University
Pittsburgh

(Received June 7, 1973)

Springer Tracts in Natural Philosophy

Edited by B.D. Coleman
Co-Editors: S.S. Antman, R. Aris,
L. Collatz, J.L. Ericksen, P. Germain,
M.E. Gurtin, C. Truesdell
Subscribers to "Zentralblatt für
Mathematik" are entitled to a 10 per
cent preferential price.

Volume 1

R.M. Gundersen
Linearized Analysis of One-Dimen-
sional Magnetohydrodynamic Flows
10 figures. X, 119 pages. 1964
Cloth DM 22,-; US \$9.00
ISBN 3-540-03216-9

From the reviews: "... This monograph is devoted to the discussion of weakly non-isotropic quasi-one-dimensional flows of an ideal, inviscid, perfectly conducting compressible fluid subject to a transverse magnetic field. The analysis is based on characteristic perturbations developed by the author. By starting with some standard solution of the quasi-linear equations of magneto-hydrodynamics, he linearizes them about this solution and studies the resulting perturbation problem. The method is applied to many interesting physical problems, starting with the propagation of free and piston driven shock waves down a uniform tube. The next application is to uniform steady flow perturbed by the local addition of heat. Thereafter a chapter is devoted to simple waves and another to the growth and decay of shock waves. In all these problems, the magnetic field is applied normal to the flow field. The final chapter deals with the extension of the perturbation technique to flows with an oblique magnetic field. Although the book is devoted to a narrow aspect of magnetohydro-dynamics, it is a valuable addition to the literature."

Mathematical Reviews

Volume 5

B.D. Coleman, H. Markovitz, W. Noll
Visometric Flows of Non-Newtonian
Fluids. Theory and Experiment
37 figures. XII, 130 pages. 1966
Cloth DM 22,-; US \$9.00
ISBN 3-540-03672-5

From the reviews: "... The present book is an attempt on the part of the authors (two of them trained as chemists and one as a mathematician) to convey the theoretical and experi-mental significance of the authors' original papers in terms understandable to the perceptive, if mathematically unsophisticated, reader. It is a remark-ably trim volume. In approximately 60 pages the reader is given precise definitions of the Coleman-Noll vocabulary, which includes simple fluid, viscometric flow, motion, and deformation. Without a paragraph that fails to introduce a key notion, the authors proceed to put these definitions to work, showing how a general approach can be used to discuss Poiseuille flow, Couette flow, torsional flow, and helical flow. The theoretical concepts are introduced with a compactness which the accomplished mechanicist will admire and which the novice will, at times, deplore. Nevertheless, the basic ideas are all here. The second half of the book is devoted to experimental means of exploiting the theory. This is a valuable contribution. References are abundant, and many of the subtle and not so subtle experimental pitfalls are dis-cussed ..."

American Scientist

Prices are subject to change without notice



Springer-Verlag
Berlin
Heidelberg
New York