A FIRST COURSE IN RATIONAL CONTINUUM MECHANICS

Volume 1

Second Edition

C. TRUESDELL

A FIRST COURSE IN

Rational Continuum Mechanics

VOLUME 1

General Concepts

Second Edition

Corrected, Revised and Augmented

This is Volume 71 in
PURE AND APPLIED MATHEMATICS

A complete list of titles in this series appears at the end of this volume.

A FIRST COURSE IN

Rational Continuum Mechanics

VOLUME 1

General Concepts

Second Edition

Corrected, Revised, and Augmented

Clifford A. Truesdell, III

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
Boston San Diego New York
London Sydney Tokyo Toronto

This book is printed on acid-free paper. ∞

ACADEMIC PRESS, INC.
1250 Sixth Avenue, San Diego, CA 92101

United Kingdom Edition published by
ACADEMIC PRESS LIMITED
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Truesdell, C. (Clifford), date.
A first course in rational continuum mechanics / Clifford Ambrose Truesdell III, — 2nd ed. corr., rev., and augmented.
p. cm. — (Pure and applied mathematics ; v. 71)
Includes bibliographical references.
Contents: v. 1. General concepts.
ISBN 0-12-701300-8 (v. 1 : alk. paper)
1. Continuum mechanics. I. Title. II. Title: Rational continuum mechanics. III. Series: Pure and applied mathematics (Academic Press) : 71.
QA3.P8 vol. 71, 1991
[QA808.2]
510 s—dc20
[531] 89-17628
CIP

Printed in the United States of America
91 92 93 94 9 8 7 6 5 4 3 2 1

To the members of
The Society for Natural Philosophy,
founded at Baltimore in 1963.

This Page Intentionally Left Blank

. . . Deduction [and] Induction . . . render the indefinite definite; Deduction explicates; Induction evaluates: that is all. Over the chasm that yawns between the ultimate goal of science and such ideas of Man's environment as . . . he managed to communicate to some fellow, we are building a cantilever bridge of induction, held together by scientific struts and ties. Yet every plank of its advance is first laid by Retroduction alone, that is to say, by the spontaneous conjectures of instinctive reason. . . .

C. S. PIERCE
Scientific Metaphysics (1908), ¶475

[While] the creative power of pure thought is at work, the outside world asserts itself again; through the real phenomena it forces new questions upon us; it opens up new fields of mathematical science; and while we try to gain these new fields of science for the realm of pure thought, we often find the answers to old unsolved problems and so at the same time best further the old theories. . . .

Besides, it is wrong to think that rigor in proof is the enemy of simplicity. Numerous examples establish the opposite, that the rigorous method is also the simpler and the easier to grasp. The pursuit of rigor compels us to discover simpler arguments; also, often it clears the path to methods susceptible of more development than were the old, less rigorous ones. . . .

While I insist upon rigor in proofs as a requirement for a perfect solution of a problem, I should like, on the other hand, to oppose the opinion that only the concepts of analysis, or even those of arithmetic alone, are susceptible of a fully rigorous treatment. This opinion, occasionally advocated by eminent men, I consider entirely mistaken. Such a one-sided interpretation of the requirement of rigor would soon lead us to ignore all concepts that derive from geometry, mechanics, and physics, to shut off the flow of new material from the outside world, and finally, indeed, as a last consequence to reject the concepts of the continuum and of the irrational number. What an important, vital nerve would be cut, were we to root 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 from the theories of natural science, the task is set for mathematics to investigate the principles underlying these ideas and establish them upon a simple and complete system of axioms in such a way that in exactness and in application to proof the new ideas shall be no whit inferior to the old arithmetical concepts.

To new concepts correspond, necessarily, new symbols. These we choose in such a way that they remind us of the phenomena which gave rise to the formation of the new concepts. . . .

If we do not succeed in solving a mathematical problem, it is often because we have failed to recognize the more general standpoint from which the problem before us appears only as a single link in a chain of related problems. . . . This way to find general methods is certainly the most practicable and the surest, for he who seeks for methods without having a definite problem in mind mainly seeks in vain.

A role still more important than generalization's in dealing with mathematical problems is played, I believe, by specialization. Perhaps in most cases where we seek in vain for the answer to a question the cause of failure lies in our having not yet or not completely solved problems simpler and easier than the one in hand. Everything depends then on finding these easier problems and effecting the solution of them by use of tools as perfect as possible and of concepts susceptible to generalization. This rule is one of the most important levers for overcoming mathematical difficulties. . . .

[The] conviction that every mathematical problem can be solved is a powerful incentive to us as we work. We hear within us the perpetual call: *There is the problem. Seek its solution. You can find it by pure thinking, for in mathematics there is no ignorabimus!*

HILBERT
Mathematical Problems
Archiv für Mathematik und Physik (3) **1**, 44–63, 213–237 (1901).

Contents

This Page Intentionally Left Blank

Preface to the Second Edition

Volume 2 will concern the dynamics of fluids; Volume 3, equilibrium and motion of elastic bodies. As the writing and incessant revision of the manuscripts progressed, I came to see that Volume 1 ought have provided additional background, especially in kinematics. Also as the years passed, various researches on the foundations appeared which clarified, compacted, and extended what was known in 1976. These are reflected most in the revised Sections I.6, I.9, II.1 (universes of shapes), II.11, III.1, IV.8, and IV.10. Also it seemed to me that Volume 1 wanted examples, for from the beginning the student should see that mechanics solves problems at every stage in its unfolding. To bring that fact home early, I recast Sections IV.8, IV.19, and IV.21, and I added two new sections, here Sections IV.15 and IV.18. In the text carried over from the first edition, hardly a page remains unemended.

My experience in teaching suggests that the material in Chapter I is the most difficult for a beginner. Usually I began my lectures with Chapter II and then went back to Chapter I, selectively, as material in it came to be needed. Of course, an experienced student, one who knows the applications of fluid dynamics and elasticity well, should begin at the beginning.

Acknowledgment for the Second Edition. Unfortunately I cannot now recall the names of all those who sent me corrections of the first edition and suggestions for the second. Among those who helped me most in the revisions of this volume and the yet unpublished texts of those to follow I express especial gratitude to R. D. James, E. MacMillan, C.-S. Man, A. W. Marris, W. Noll, K. R. Rajagopal, M. Scheidler, R. Segev, E. Virga, C.-C. Wang, and W. O. Williams.

"Il Palazzetto"
Baltimore
1991

This Page Intentionally Left Blank

Preface to the First Edition

The mechanics of finite systems of points and rigid bodies was given a fairly definitive form by LAGRANGE's exposition in his *Méchanique Analitique*, 1788. While that book covers only certain aspects of the rational mechanics created by LAGRANGE's great predecessors, it presents them systematically and as a branch of mathematics: "Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche, ..." The physics and the applications are omitted. He who will apply and interpret the theory, or dwell upon the intricacies and mysteries of its place among the relations between mind and external nature, is expected to learn it first. While the knowledge he thus acquires does not of itself put applications into his hands, it gives him the tools to fashion them efficiently, or at least to classify, describe, and teach the applications already known. By consistently leaving applications to the appliers, LAGRANGE set them on common ground with the theorists who sought to pursue the mathematics further: Both had been trained in the same workshop and spoke the same jargon. Even today this comradeship of infancy lingers on, provided discrete systems and rigid bodies exhaust the universe of mechanical discourse.

In 1788 the mechanics of deformable bodies, which is inherently not only subtler, more beautiful, and grander but also far closer to nature than is the rather arid special case called "analytical mechanics", had been explored only in terms of isolated examples, brilliant but untypical. Unfortunately most of these fitted into LAGRANGE's scheme; those that did not, he passed over in silence. Further brilliant examples, feigned mainly upon the framework of

NEWTON's and EULER's concepts and not easily subsumed under LAGRANGE's, were created in the next century but were studied mainly for their own sakes, separately, and did not lead to a general doctrine, despite the deep and original syntheses of stress and strain forged by CAUCHY.

A hundred years after CAUCHY died began a renascence of "classical" mechanics as a whole, taking the deformable continuum as the typical body and describing it in terms of an equally specific concept of material, which had been left nebulous and physical or metaphysical before then. This new general doctrine is now fit to be learned and used by mathematicians, experimentists, and engineers and to join the old analytical mechanics as an element of common education. Physicists should be able to understand it, should they wish to. Like geometry, it is part of mathematics.

In writing a textbook of continuum mechanics at this time I imitate the example of LAGRANGE in several ways. My book offers merely a selection from the wondrous harvest of the last few decades; leaving much else unmentioned, it bases that selection on criteria of naturalness, ease, and subsumption to a general method and conceptual frame. Thus it is a short book, designed for readers who know already that applications to further cases are numberless and possibilities for further mathematical study infinite. As LAGRANGE wrote, "On ne trouvera point de Figures dans cet Ouvrage. Les méthodes que j'y expose, ne demandent ni constructions, ni raisonnemens géométriques où méchaniques, mais seulement des opérations algébriques, assujetties à une marche regulière & uniforme." This claim is as true—or as false—of the present book as of LAGRANGE's. Of course, many proofs are easier to grasp if a figure is drawn, and both teacher and student should illumine and enrich the "regular and uniform course" by sketches. Finally—and here, perhaps, lies the greatest difference between this book and others with similar titles—it follows LAGRANGE's example in presuming that the reader commands the elementary mathematics of his own day,[1] making no attempt to offer a shadowy substitute for decent modern training in algebra and calculus or to appease the notorious reluctance of old men to learn anything new. The student may well find this book easier than his teacher does.

In three respects, however, I depart from LAGRANGE's model. First, I leave important if small pieces of the arguments, and some illustrations of them, as exercises for the reader, since my experience in teaching the new mechanics as it sprouted and grew has assured me that he who does not for himself re-create

[1] The reader is expected to know the elements of measure theory. For almost everything else needed in "pure" mathematics, more than sufficient background is given in the book by R. M. BOWEN & C.-C. WANG, designed especially for students of continuum mechanics: *Introduction to Vectors and Tensors*, 2 volumes, New York and London, Plenum Press, 1976. Some more specialized works are cited below in reference to some particular theorems, as needed.

and digest the mathematics step by step will never master this doctrine. Second, while LAGRANGE's presentation bestowed upon the subject a gloss of closure and completeness which by the passage of time has been abundantly proved specious, in this book I try to present the science of "classical" mechanics even to the beginner as what it is: a magnificent array of ordered concepts and proved theorems, some of them old, even very old, and some on the frontiers of research into great unsolved problems and not yet distilled experience of nature as human eyes see it and human hands feel it. Third, the frequent attributions of major ideas and theorems to others will make it clear that I claim little of the substance for my own. The citations of other works, however, are intended not as acknowledgments of sources but as aids to the student. Those at the ends of the chapters direct him to places where further matters closely related to the text are developed; those in the footnotes, to specific details passed over in the text such as counterexamples, direct generalizations, proofs of theorems cited from other parts of mathematics, and tangent domains of modern mechanics.

Finally, I wish to thank those who have helped me to understand mechanics and to complete and purify this book. Thus above all I thank WALTER NOLL, and after him J. L. ERICKSEN, R. A. TOUPIN, B. D. COLEMAN, M. E. GURTIN, C.-C. WANG, W. O. WILLIAMS, L. SOLOMON, T. TOKUOKA, W.-L. YIN, R. C. BATRA, and D. EUVRARD. I am indebted to Mr. BATRA also for a full set of solutions to the exercises.

"Il Palazzetto"
Baltimore
May 1, 1972

Addendum. Parts 1 through 4 of this work, expertly translated into French by D. EUVRARD from my text of 1972, were published in December, 1973, by Masson et Cie in a single volume with the title *Introduction à la Mécanique Rationnelle des Milieux Continus*. Parts 1 through 5 appeared in 1975 in Russian, Первоначалиный Курс Рациональной Механики Сплошных Сред, Moscow, Мир, translated from my text of 1973 by R. V. GOLDSHTEIN & V. M. ENTOV under the guidance of P. A. ZHILIN & A. I. LUR'E. Since that time I have been able to add some material and also to work through the text again and make numerous improvements, partly in response to criticisms and suggestions offered by readers of the French book.

A question has been raised regarding the knowledge of mechanics the student is expected to have already. A good treatise on the theory of functions of a real variable does not strictly require of its readers any previous acquaintance with the subject, even in the most elementary aspects of infinitesimal calculus, yet a student armed with no more than a naked, virgin mind is unlikely to survive the first few pages. In the same way, although this book does not call upon any previous knowledge of continuum mechanics, or even of schoolboy mechanics, it is designed for students not altogether innocent of hydrodynamics and elasticity. Much as a crude and awkward first affair may furnish knowledge that, however elementary, is indispensable to him who aspires toward Venus's ultimate refinements, a bad course—something nowadays cheaply found—will serve well enough here, too.

Some comments on the preliminary editions in French and Russian suggest need for a reminder that this is a mathematical textbook, not a treatise or a history. In attaching names to a proposition I follow the commonest usage in the mathematical literature, proclaiming respect for those to whom I think we owe that proposition, be it in entirety, be it for discovery and proof of a pilot case, be it for clearest statement or most elegant proof; a second name never indicates rediscovery but always some major improvement, and of course it would not be feasible in any discipline so broadly cultivated as rational mechanics now is to list all the persons who have done something valuable, even if I knew of them all. . . .

I thank Mr BATRA for further suggestions and for checking the manuscript of this volume. I am deeply grateful to him and to Messrs. DAFERMOS, ERICKSEN, GURTIN, MUNCASTER, NOLL, and WILLIAMS for their generous gift of time and care in correcting the proofsheets so as to remove errors and obscurities even at the last moment. For such faults as, alas, surely remain I bear an uncommon charge, for seldom has an author had the benefit of such abundant and expert aid.

I owe a double debt of gratitude to the U.S. National Science Foundation for its continued and generous support: first, for the work of some of the great savants whose discoveries are incorporated here; second, for my own long effort to compose the essence of modern rational mechanics into an easy union with the magnificent tradition from which it sprang, so that beginners might learn both old and new together and in such a way as to see each illuminate and ennoble the other.

C.T.

December 20, 1976

Addendum, 1990. The following introductory works are sound and helpful:

D. C. LEIGH, *Nonlinear Continuum Mechanics*, New York *etc.*, McGraw-Hill, 1968.

P. CHADWICK, *Continuum Mechanics*, New York, Wiley, 1976.

C.-C. WANG, *Mathematical Principles of Mechanics and Electromagnetism, Part A: Analytical and Continuum Mechanics*, New York and London, Plenum Press, 1979.

M. E. GURTIN, *An Introduction to Continuum Mechanics*, New York *etc.*, Academic Press, 1981.

PART 1

GENERAL CONCEPTS

In the following Chapters on Abstract Dynamics we confine ourselves mainly to the general principles, and the fundamental formulas and equations of the mathematics of this extensive subject; and neither seeking nor avoiding mathematical exercitations, we enter on special problems solely with a view to possible usefulness for physical science, whether in the way of the *material* of experimental investigation, or for illustrating physical principles, or for aiding in speculations of Natural Philosophy.

THOMSON & TAIT
Treatise on Natural Philosophy
(2^{nd} ed., 1883), Section 453

This Page Intentionally Left Blank

Chapter I

Bodies, Forces, and Motions

[T]he idea of *Force* . . . is a direct object of sense . . . .

THOMSON & TAIT
Treatise on Natural Philosophy
(1867), Section 207

I have restored to the concepts of *mass* and *force* their old rights. Beyond all doubt we need these *things*, for without them, there is no mechanics. Force is more than mass times acceleration, as may be seen from the basic equation itself, which always asserts that mass times acceleration equals the *sum* of the forces. Therefore, why not use the good old *words*? The concepts themselves are not unclear; it is just that the books described them often in a very metaphysical and dark way. And what matter, if the concepts are remarkably useful—perhaps a bit riddling?—if the concepts of mechanics are deeper than many find convenient, and cannot be disposed of with a few elegant words like convention and economy of thought, abstraction and idealisation?

HAMEL
On the foundations of mechanics.
Mathematische Annalen **66** (1909): 350–397.

Space, time, and force are *a priori* forms; they can be derived only from contemplation and from general principles of research. Their common relation to each other in mechanics must be regarded as something inspired indeed by experience but in its generality fixed by convention.

HAMEL
Elementare Mechanik (1912), ¶5

[I]n the concept of force lies the chief difficulty in the whole of mechanics.

HAMEL, letter to TRUESDELL,
14 October 1952

1. Rational Mechanics

Rational Mechanics is the part of mathematics that provides and develops logical models for the enforced changes of place and shape we see everyday things suffer. It describes also much of what is observed or inferred in the laboratories where professional scientists produce experiments. For example, it is always presumed as a part of the basis for design and control of scientific apparatus which physicists regard as producing decisive experimental evidence that classical mechanics itself is only an "approximate" theory of nature. Of course, all mathematical theories of nature only approximate it.

The things mechanics represents by mathematical constructs include animals and plants, mountains and the atmosphere, oceans and the subterraneous riches, the whole orb which is the seat of our life and experience, heavenly objects both old and new, and the elements out of which these things seem to be composed: earth, water, air, and fire. As its name suggests, mechanics represents also the contrivances of man's artifice: fountains and engines and vehicles, bridges and fabrics, instruments of music and warfare, sewers and rockets. All these things mechanics models, but models crudely. Like any other branch of mathematics, it abstracts and evolves the common features of what it represents, setting aside most of the detail. As is necessary in any science which aims not merely to describe but also to predict, it seeks to select and correlate the simple out of the manifold and insuperable complexity of nature. Simplicity, while it does not ensure success in a branch of mechanics, is necessary there. A complicated theory in mechanics, although it may be socially or sociably useful at a particular time and place, does not enlighten and hence does not endure. Finally, since our experience grows with time and in proportion to our ingenuity, while the progress of mathematics enables us to manage easily and neatly mathematical ideas and operations of greater and greater scope, mechanics cannot be a closed science but must contain or at least be provided with means of improving or refining the models it presently possesses and also of constructing new ones.

Mechanics does not study natural things directly. Instead, it considers *bodies*, which are mathematical concepts designed to abstract some common features of many natural things. One such feature is the *mass* assigned to each body. Always, a natural body is at any one *instant* found to occupy a set of *places*; that set is the *shape* of that body at that instant. The theory of places, which is called *geometry*, was created long ago and thus lies ready to hand for application in mechanics. The change of shape undergone by a body from one instant to another is called the *motion* of that body, and description of motion, or *kinematics*, is the second part of the foundation of mechanics. Third, motions of bodies are conceived as resulting from or at least being invariably accompanied by the action of *forces*. Thus, mechanics provides a *mathematical model*, or, better, an infinite class of models, for certain aspects of nature.

In the words of NEWTON,

> ... Rational Mechanics will be the science of motions resulting from any forces whatsoever and of the forces required to produce any motions, accurately proposed and generated.

Mechanics rests upon three substructures: a universe of bodies, a geometry with its kinematics, and a theory of forces. These substructures provide the concepts mechanics is to connect. Relations among places, the shapes of bodies, forces, and times are of two kinds: the *general* ones, common to all bodies in an assigned universe, appropriate to a branch of mechanics, and the *particular* ones, which within a given branch distinguish one class of such bodies from another. The general relations are of two kinds: *statics*, which compares putative equilibria; and *dynamics*, which describes *motions*. The particular relations are called *constitutive*. They define *materials*, which are mathematical idealizations of the materials encountered in nature. Typical branches concern mass-points, three-dimensional continua, plates, shells, membranes, rods, jets, and strings. Typical constitutive classes are the bodies called rigid or solid or fluid, isotropic or anisotropic.

The chapter now begun presents mechanics of a fairly general kind, rendered concrete by illustrations from the theories of continua and of mass-points. From Chapter II onward we shall treat only the mechanics of continua occupying three-dimensional shapes.

I cannot develop all of mechanics from explicit axioms.[1] So as to reach the level at which we may formulate and study constitutive relations, we shall pass lightly over the foundations of general mechanics.

While the presentation is lacunary and informal, it is abstract. The reader who is content to take bodies, the event world, frames of reference, motions, and forces for granted may skip this chapter and pass to the next one, which begins the formal treatment of continuum mechanics along traditional lines. The traditional approach to mechanics is in no way incorrect, but it fails to satisfy modern standards of criticism and explicitness. Therefore, some parts of the foundations of mechanics heretofore left in the penumbrae of intuition and metaphysics I shall here present in an explicit, compact mathematical style, notably the theories of substantial universes of bodies (Sections I.2 and I.3), systems of forces (Section I.5), and the universe of shapes of continua (Section

[1]The sixth of the problems HILBERT set for the twentieth century to solve was to formulate an axiomatic structure for physics, and especially for mechanics. Apart from a noteworthy attempt of HAMEL in 1909, this problem was given scarcely any serious attention until it was taken up by NOLL in 1957. The content of Chapter I of this book derives essentially from the work of NOLL and those who have accepted, applied, and extended his ideas.

II.1). In the theory of contact forces and body forces (Sections III.1 and III.3) modern, precise treatments appear alongside the corresponding classical ones.

Natural things usually are endowed with a *hotness* or *hotnesses*, which are rendered numerical through assignment of scales of *temperature*. Natural things may also absorb and emit *heat*. *Rational Thermomechanics*[1] is a group of mathematical theories that interrelate motion, force, hotness, and heat through a general structure, thus providing a framework more general than Rational Mechanics in that it allows models for a greater class of natural things. On the other hand, in thermomechanics the theory of constitutive relations is in part less general, for that theory delivers restrictions that, if pulled back into mechanics, would narrow its scope.

This textbook stops short of thermomechanics.

2. Universes of Bodies

Most collections of bodies $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}, \ldots, \mathscr{X}$ conform with the mathematical structure of a *Boolean lattice* or *complemented distributive lattice*, often called a *Boolean algebra*.[2] The student who is familiar with this structure may pass directly to the next section. Here, following NOLL, we shall simply list in order of their immediacy the properties common to all bodies in most theories within mechanics and prove some theorems concerning them.

The set $\mathbf{\Omega}$ of all bodies of some particular kind is called a *universe*. At the very beginning in any branch of mechanics, a universe is specified, though in older work the reader was expected to infer the particular universe from the context. If the body $\mathscr{B}$ is a *part* of the body $\mathscr{C}$, we write $\mathscr{B} \prec \mathscr{C}$. The relation $\prec$ gives $\mathbf{\Omega}$ the structure of a *partially ordered set*, defined by familiar axioms:

Axiom B1. $\mathscr{B} \prec \mathscr{B}$.

Axiom B2. $(\mathscr{X} \prec \mathscr{B})\ \&\ (\mathscr{B} \prec \mathscr{X}) \quad \Rightarrow \quad \mathscr{X} = \mathscr{B}$.

Axiom B3. $(\mathscr{B} \prec \mathscr{C})\ \&\ (\mathscr{C} \prec \mathscr{D}) \quad \Rightarrow \quad \mathscr{B} \prec \mathscr{D}$.

[1]Rational Thermomechanics is a field still under active discussion and development. Among the books that deal with it are

C. TRUESDELL, *Rational Thermodynamics*, N.Y. *etc.*, McGraw-Hill, 1968; second edition, with a historical introit and appendices by several authors, N.Y. *etc.*, Springer-Verlag, 1984.

W. A. DAY, *The Thermodynamics of Simple Materials with Fading Memory*, N.Y. *etc.*, Springer-Verlag, 1972.

D. R. OWEN, *A First Course in the Mathematical Foundations of Thermodynamics*, New York *etc.*, Springer-Verlag, 1984.

I. MÜLLER, *Thermodynamics*, Boston *etc.*, Pitman Publishing, 1985.

J. SERRIN (editor), *New Perspectives in Thermodynamics*, Berlin *etc.*, Springer-Verlag, 1986.

[2]For the general theory see the book by R. SIKORSKY, *Boolean Algebras*, 3rd edition, Berlin *etc.*, Springer, 1969.

That is, $\mathscr{B}$ is a part of itself; $\mathscr{B}$ is not a part of any other of its parts; and if $\mathscr{B}$ is a part of $\mathscr{C}$, while $\mathscr{C}$ is a part of $\mathscr{D}$, then $\mathscr{B}$ is a part of $\mathscr{D}$. Another wording of Axiom B2 is, a body is the greatest of its parts.

To picture the relations among bodies, it may help to consider $\mathbf{\Omega}$ as being the collection of all open sets in the Euclidean plane and to take $\prec$ as being the sign of inclusion, $\subset$, so that the suggestive sketches often called "Venn diagrams" are easy to draw. This illustration is only one of many. Others, including the universes commonly presumed in mechanics, will be presented in the next section.

When the bodies $\mathscr{B}$ and $\mathscr{C}$ are given, neither need be a part of the other, but often they are both parts of a third one, $\mathscr{D}$. Such a $\mathscr{D}$ is called an *envelope* of $\mathscr{B}$ and $\mathscr{C}$. If, further, there is a body $\mathscr{A}$ that is an envelope of $\mathscr{B}$ and $\mathscr{C}$ and is itself a part of every envelope of $\mathscr{B}$ and $\mathscr{C}$, then $\mathscr{A}$ is called the *join* of $\mathscr{B}$ and $\mathscr{C}$. This relation among bodies is denoted as follows:

$$\mathscr{A} = \mathscr{B} \vee \mathscr{C}. \tag{I.2-1}$$

Formally, this equation means that if ($\mathscr{B} \prec \mathscr{A}$ & $\mathscr{C} \prec \mathscr{A}$), then

$$(\mathscr{B} \prec \mathscr{D} \ \& \ \mathscr{C} \prec \mathscr{D}) \quad \Rightarrow \quad \mathscr{A} \prec \mathscr{D}. \tag{I.2-2}$$

Thus, the join of $\mathscr{B}$ and $\mathscr{C}$, if it exists, may be regarded as the *least* envelope of $\mathscr{B}$ and $\mathscr{C}$, since it is a part of every envelope of $\mathscr{B}$ and $\mathscr{C}$. Likewise, if $\mathscr{A} \prec \mathscr{B}$ & $\mathscr{A} \prec \mathscr{C}$ and

$$(\mathscr{D} \prec \mathscr{B},\ \mathscr{D} \prec \mathscr{C}) \quad \Rightarrow \quad \mathscr{D} \prec \mathscr{A}, \tag{I.2-3}$$

we write

$$\mathscr{A} = \mathscr{B} \wedge \mathscr{C} \tag{I.2-4}$$

and call $\mathscr{A}$ the *meet* of $\mathscr{B}$ and $\mathscr{C}$. If it exists, it is the *greatest* common part of $\mathscr{B}$ and $\mathscr{C}$, since every other common part of $\mathscr{B}$ and $\mathscr{C}$ is a part of it. Two bodies $\mathscr{B}$ and $\mathscr{C}$ may fail to have a meet or a join, or both, but, if they do have them, then plainly

$$\mathscr{B} \vee \mathscr{C} = \mathscr{C} \vee \mathscr{B}, \qquad \mathscr{B} \wedge \mathscr{C} = \mathscr{C} \wedge \mathscr{B}, \qquad \mathscr{B} \wedge \mathscr{C} \prec \mathscr{B} \prec \mathscr{B} \vee \mathscr{C}. \tag{I.2-5}$$

Also

$$\mathscr{B} \prec \mathscr{C} \quad \Leftrightarrow \quad \mathscr{B} \wedge \mathscr{C} = \mathscr{B} \quad \Leftrightarrow \quad \mathscr{B} \vee \mathscr{C} = \mathscr{C}, \tag{I.2-6}$$

whence

$$\mathscr{B} \wedge \mathscr{B} = \mathscr{B} \vee \mathscr{B} = \mathscr{B}, \tag{I.2-7}$$

and here, trivially, both the meet and the joint exist.

Exercise I.2.1.

$$\mathscr{B} \prec \mathscr{C} \qquad \Rightarrow \qquad (\mathscr{B} \wedge \mathscr{A} \prec \mathscr{C} \wedge \mathscr{A}) \;\&\; (\mathscr{B} \vee \mathscr{A} \prec \mathscr{C} \vee \mathscr{A}), \tag{I.2-8}$$

each indicated meet and join being assumed to exist.

If $\mathscr{B}_q$ are bodies from any collection, membership in which is indexed by subscripts q taken from some given set, meets and joins of the collection are defined in the same way and denoted by $\bigwedge\limits_q \mathscr{B}_q$ and $\bigvee\limits_q \mathscr{B}_q$. The former, for example, if it exists, is a body which is a part of each $\mathscr{B}_q$ and contains every other such body. For three bodies it is sometimes clearer to use the longer special notation $\mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D}$, but then we must recall that the order of considering $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{D}$ makes no difference, as is clear from the general notation and is illustrated in $(5)_{1,2}$.

To see that partial ordering does not ensure the existence of meets and joins, it suffices to consider the example of a universe $\mathbf{\Omega}$ consisting in all non-empty, half-open intervals $]a, b]$ and $[c, d[$ of real numbers, with $\prec$ defined as being inclusion in the sense of set theory. If $\mathscr{B} = [0, 2[$, $\mathscr{C} = [1, 4]$, and $\mathscr{D} = [3, 5[$, then $\mathscr{B}$ and $\mathscr{C}$ are common parts of infinitely many half-open intervals, yet they have no join, since if a certain half-open interval contains all the points of $\mathscr{B}$ and $\mathscr{C}$, we can find a shorter one that does so. The same may be said of the pair $\mathscr{C}$, $\mathscr{D}$. Nevertheless, $\mathscr{B} \vee \mathscr{C} \vee \mathscr{D} = [0, 5[$, which is a member of $\mathbf{\Omega}$. Note that $\mathscr{B} \vee \mathscr{D} = \mathscr{B} \vee \mathscr{C} \vee \mathscr{D} \neq \mathscr{B} \cup \mathscr{D}$.

Exercise I.2.2. If $\mathscr{B} \wedge \mathscr{C}$ and $\mathscr{C} \wedge \mathscr{D}$ exist, and if either $(\mathscr{B} \wedge \mathscr{C}) \wedge \mathscr{D}$ or $\mathscr{B} \wedge (\mathscr{C} \wedge \mathscr{D})$ exists, then both do, and so does $\mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D}$; also

$$(\mathscr{B} \wedge \mathscr{C}) \wedge \mathscr{D} = \mathscr{B} \wedge (\mathscr{C} \wedge \mathscr{D}) = \mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D}. \tag{I.2-9}$$

A body $\mathscr{O} \in \mathbf{\Omega}$ is called the ***null body*** if and only if it is a part of every body in $\mathbf{\Omega}$:

$$\mathscr{O} \prec \mathscr{B} \qquad \forall \mathscr{B} \in \mathbf{\Omega}. \tag{I.2-10}$$

$\mathbf{\Omega}$ need not contain such an element, but if it does, Axiom B2 makes that element unique. A body is called the ***universal body*** and denoted by ∞ if and only if every body in $\mathbf{\Omega}$ is a part of it:

$$\mathscr{B} \prec \infty \qquad \forall \mathscr{B} \in \mathbf{\Omega}. \tag{I.2-11}$$

Such a body, if it exists, is obviously unique.

If the set $\mathbf{\Omega}$ does not contain the null body or the universal body, we can formally adjoin either or both of these bodies so as to form the corresponding *closed universe* $\bar{\mathbf{\Omega}}$, defined as follows: $\bar{\mathbf{\Omega}} := \mathbf{\Omega} \cup \{\mathscr{O}, \infty\}$. By the following definitions we extend the partial order $\prec$ in $\mathbf{\Omega}$ so as to form a partial order in $\bar{\mathbf{\Omega}}$:

$$\begin{aligned} &\mathscr{O} \prec \mathscr{B} \quad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}, \qquad \mathscr{B} \prec \mathscr{O} \quad \Rightarrow \quad \mathscr{B} = \mathscr{O}, \\ &\mathscr{B} \prec \infty \quad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}, \qquad \infty \prec \mathscr{B} \quad \Rightarrow \quad \mathscr{B} = \infty. \end{aligned} \tag{I.2-12}$$

It is easy to verify that with the definitions (I.2-12) $\bar{\mathbf{\Omega}}$ becomes a partially ordered set with partial order $\prec$, and that $\mathscr{O}$ and ∞ are the null body and the universal body of $\bar{\mathbf{\Omega}}$.

In $\bar{\mathbf{\Omega}}$ clearly

$$\mathscr{B} \wedge \mathscr{O} = \mathscr{O}, \qquad \mathscr{B} \vee \mathscr{O} = \mathscr{B}, \qquad \mathscr{B} \wedge \infty = \mathscr{B}, \qquad \mathscr{B} \vee \infty = \infty. \tag{I.2-13}$$

Any two bodies $\mathscr{B}$ and $\mathscr{C}$ in $\bar{\mathbf{\Omega}}$ have at least one common part, namely $\mathscr{O}$. If they have no other common part, they are called *separate*. Thus $\mathscr{B}$ and $\mathscr{C}$ are separate if and only if

$$\mathscr{B} \wedge \mathscr{C} = \mathscr{O}. \tag{I.2-14}$$

Exercise I.2.3.

$$(\mathscr{B} \wedge \mathscr{C} = \mathscr{O})\ \&\ (\mathscr{D} \prec \mathscr{C}) \qquad \Rightarrow \qquad \mathscr{B} \wedge \mathscr{D} = \mathscr{O}. \tag{I.2-15}$$

Setting $\mathscr{D} = \mathscr{B}$ and using (7) shows that the only part of a body separate from that body is $\mathscr{O}$.

Next we need a concept of environment of a given body $\mathscr{B}$, so as to provide which we lay down a further axiom:

Axiom B4. *With each body $\mathscr{B}$ in $\bar{\mathbf{\Omega}}$ is associated a unique body $\mathscr{B}^e$, which is called the* exterior *of $\mathscr{B}$, such that*

$$\mathscr{B} \wedge \mathscr{B}^e = \mathscr{O}, \qquad \mathscr{B} \vee \mathscr{B}^e = \infty. \tag{I.2-16}$$

Thus $\mathscr{B}^e$ is separate from $\mathscr{B}$, and the only body that contains both $\mathscr{B}$ and $\mathscr{B}^e$ is ∞.

The example given just before Exercise I.2.2 shows that Axiom B4 cannot follow from Axioms B1, B2, and B3, since the points exterior to [0, 1[do not constitute a half-open interval.

Exercise I.2.4.

$$\mathscr{B} \prec \mathscr{B}^e \qquad \Rightarrow \qquad \mathscr{B} = \mathscr{O}. \tag{I.2-17}$$

By putting $\mathscr{B} = \mathscr{O}$ in $(13)_{3,4}$ and comparing the outcome with (16), we see that

$$\mathscr{O}^e = \infty, \qquad \infty^e = \mathscr{O}. \tag{I.2-18}$$

Likewise

$$(\mathscr{B}^e)^e = \mathscr{B}. \tag{I.2-19}$$

Also, putting $\mathscr{B} = \mathscr{C}^e$ in (15) shows that

$$\mathscr{D} \prec \mathscr{C} \qquad \Rightarrow \qquad \mathscr{D} \wedge \mathscr{C}^e = \mathscr{O}. \tag{I.2-20}$$

We now postulate that the converse of this proposition holds:

Axiom B5. *The only bodies separate from $\mathscr{C}^e$ are the parts of $\mathscr{C}$.*

While it has been proved[1] that Axiom B5 does not follow from Axioms B1–B4, I could find no simple example to illustrate this fact.

Formally, we may combine (20) with Axiom B5 as follows:

$$\mathscr{B} \prec \mathscr{C} \qquad \Leftrightarrow \qquad \mathscr{B} \wedge \mathscr{C}^e = \mathscr{O}. \tag{I.2-21}$$

By (19), then,

$$\mathscr{B} \prec \mathscr{C} \qquad \Leftrightarrow \qquad \mathscr{C}^e \wedge (\mathscr{B}^e)^e = \mathscr{O}. \tag{I.2-22}$$

[1] R. P. Dilworth, "Lattices with unique complements," *Transactions of the American Mathematical Society* **57** (1945): 123–154.

If we now replace $\mathscr{B}$ by $\mathscr{C}^{\mathrm{e}}$ and $\mathscr{C}$ by $\mathscr{B}^{\mathrm{e}}$ in (21) and compare the result with (22), we see that

$$\mathscr{B} \prec \mathscr{C} \qquad \Leftrightarrow \qquad \mathscr{C}^{\mathrm{e}} \prec \mathscr{B}^{\mathrm{e}}. \tag{I.2-23}$$

Now let q run over the elements of some specified collection. Then if $\bigvee_q \mathscr{B}_q$ exists, so does $\bigwedge_q (\mathscr{B}_q)^{\mathrm{e}}$, and

$$\bigwedge_q (\mathscr{B}_q)^{\mathrm{e}} = \left(\bigvee_q \mathscr{B}_q\right)^{\mathrm{e}}, \tag{I.2-24}$$

while if $\bigwedge_q \mathscr{B}_q$ exists, so does $\bigvee_q (\mathscr{B}_q)^{\mathrm{e}}$, and

$$\bigvee_q (\mathscr{B}_q)^{\mathrm{e}} = \left(\bigwedge_q \mathscr{B}_q\right)^{\mathrm{e}}. \tag{I.2-25}$$

Drawing a diagram will make evident the statement and proof of the following theorem, where all meets and joins indicated are assumed to exist:

Theorem. *If*

$$\mathscr{A}_1 \wedge \mathscr{B} \prec \mathscr{C}, \qquad \mathscr{A}_2 \wedge \mathscr{B} \prec \mathscr{C}, \qquad \mathscr{D} \prec \mathscr{B}, \qquad \mathscr{D} \prec \mathscr{A}_1 \vee \mathscr{A}_2, \tag{I.2-26}$$

then

$$\mathscr{D} \prec \mathscr{C}. \tag{I.2-27}$$

Proof. By $(26)_{1,2}$ and (21),

$$(\mathscr{A}_q \wedge \mathscr{B}) \wedge \mathscr{C}^{\mathrm{e}} = \mathscr{O}, \qquad q = 1, 2. \tag{I.2-28}$$

Let $\mathscr{E}$ be a common part of $\mathscr{D}$ and $\mathscr{C}^{\mathrm{e}}$:

$$\mathscr{E} \prec \mathscr{D}, \qquad \mathscr{E} \prec \mathscr{C}^{\mathrm{e}}. \tag{I.2-29}$$

Then $\mathscr{E} \prec \mathscr{B}$ by $(26)_3$. Let $\mathscr{F}_q$ be a common part of $\mathscr{E}$ and $\mathscr{A}_q$:

$$\mathscr{F}_q \prec \mathscr{E}, \qquad \mathscr{F}_q \prec \mathscr{A}_q, \qquad q = 1, 2. \tag{I.2-30}$$

Then $\mathscr{F}_q \prec \mathscr{B}$, and hence

$$\mathscr{F}_q \prec \mathscr{A}_q \wedge \mathscr{B}, \qquad q = 1, 2. \tag{I.2-31}$$

Using this conclusion and (28) in (15), we find that

$$\mathscr{F}_q \wedge \mathscr{C}^{\mathrm{e}} = \mathscr{O}. \tag{I.2-32}$$

But $\mathscr{F}_q \prec \mathscr{E} \prec \mathscr{C}^{\mathrm{e}}$, and so by the conclusion in Exercise I.2.3 we see that $\mathscr{F}_q = \mathscr{O}$. In view of the hypothesis (30), then,

$$\mathscr{E} \wedge \mathscr{A}_q = \mathscr{O}, \qquad q = 1, 2. \tag{I.2-33}$$

By (21), then, $\mathscr{A}_q \prec \mathscr{E}^{\mathrm{e}}$, and hence

$$\mathscr{A}_1 \vee \mathscr{A}_2 \prec \mathscr{E}^{\mathrm{e}}. \tag{I.2-34}$$

But by $(26)_4$ and $(29)_1$

$$\mathscr{E} \prec \mathscr{A}_1 \vee \mathscr{A}_2 . \tag{I.2-35}$$

Hence $\mathscr{E} \prec \mathscr{E}^{\mathrm{e}}$, so by (17) $\mathscr{E} = \mathscr{O}$. The hypothesis (29) has thus led to the conclusion that $\mathscr{D} \wedge \mathscr{C}^{\mathrm{e}} = \mathscr{O}$, and by (21) we obtain (27). △

The theorem enables us to prove the ***distributive laws*** of Boolean algebra. First, if $\mathscr{A}_1 \wedge \mathscr{B}$, $\mathscr{A}_2 \wedge \mathscr{B}$, and $\mathscr{A}_1 \vee \mathscr{A}_2$ exist, then

$$(\mathscr{A}_1 \vee \mathscr{A}_2) \wedge \mathscr{B} = (\mathscr{A}_1 \wedge \mathscr{B}) \vee (\mathscr{A}_2 \wedge \mathscr{B}), \tag{I.2-36}$$

provided either side exists. Indeed, the theorem tells us that any body which is a part of both $\mathscr{B}$ and $\mathscr{A}_1 \vee \mathscr{A}_2$ is also a part of any body of which $\mathscr{A}_1 \wedge \mathscr{B}$ and $\mathscr{A}_2 \wedge \mathscr{B}$ are parts. Thus the body on the right-hand side of (36), if it exists, contains every common part of $\mathscr{A}_1 \vee \mathscr{A}_2$ and $\mathscr{B}$. Since it is trivially a common part of $\mathscr{A}_1 \vee \mathscr{A}_2$ and $\mathscr{B}$, by the definition of "meet" it is $(\mathscr{A}_1 \vee \mathscr{A}_2) \wedge \mathscr{B}$. Similar reasoning applies if the body on the left-hand side is assumed to exist.

If we replace the bodies occurring in (36) by their exteriors and use (24) and (25), we obtain the second distributive law: If $\mathscr{A}_1 \vee \mathscr{B}$, $\mathscr{A}_2 \vee \mathscr{B}$, and $\mathscr{A}_1 \wedge \mathscr{A}_2$ exist, then

$$(\mathscr{A}_1 \wedge \mathscr{A}_2) \vee \mathscr{B} = (\mathscr{A}_1 \vee \mathscr{B}) \wedge (\mathscr{A}_2 \vee \mathscr{B}), \tag{I.2-37}$$

provided either side exists.

If in (36) we take $\mathscr{A}_1$ and $\mathscr{A}_2$ as being $\mathscr{A}$ and $\mathscr{A}^e$, we see that if $\mathscr{A} \wedge \mathscr{B}$ and $\mathscr{A}^e \wedge \mathscr{B}$ exist, then

$$\mathscr{B} = (\mathscr{A} \wedge \mathscr{B}) \vee (\mathscr{A}^e \wedge \mathscr{B}). \tag{I.2-38}$$

Finally, we have the ***basic decomposition theorem,*** which enables us to express any body $\mathscr{B}$ as the join of any one of its parts $\mathscr{A}$ with a certain, uniquely determined, separate body $\mathscr{C}$:

$$(\mathscr{B} = \mathscr{A} \vee \mathscr{C}) \;\&\; (\mathscr{A} \wedge \mathscr{C} = \mathscr{O}) \quad \Leftrightarrow \quad (\mathscr{A} \prec \mathscr{B}) \;\&\; (\mathscr{C} = \mathscr{B} \wedge \mathscr{A}^e). \tag{I.2-39}$$

To prove this implication, we assume first that the decomposition on the left-hand side exists. Then $\mathscr{A} \prec \mathscr{B}$, and by (19) and (21) $\mathscr{C} \prec \mathscr{A}^e$; equivalently, by $(6)_1$, $\mathscr{C} \wedge \mathscr{A}^e = \mathscr{C}$. By $(39)_1$ and (36), then,

$$\begin{aligned} \mathscr{B} \wedge \mathscr{A}^e &= (\mathscr{A} \vee \mathscr{C}) \wedge \mathscr{A}^e, \\ &= (\mathscr{A} \wedge \mathscr{A}^e) \vee (\mathscr{C} \wedge \mathscr{A}^e), \\ &= \mathscr{O} \vee \mathscr{C} = \mathscr{C}, \end{aligned} \tag{I.2-40}$$

so that the implication forward in (39) is proved. Now suppose, conversely, that $\mathscr{A} \prec \mathscr{B}$ and $\mathscr{C} = \mathscr{B} \wedge \mathscr{A}^e$. Then $\mathscr{A} \wedge \mathscr{B} = \mathscr{A}$, so that (38) yields $\mathscr{B} = \mathscr{A} \vee \mathscr{C}$. Since $\mathscr{C} \prec \mathscr{A}^e$, it follows that $\mathscr{A} \wedge \mathscr{C} = \mathscr{O}$. $\triangle$

The final axiom for bodies asserts the existence of the meet:

Axiom B6. *For any two bodies $\mathscr{B}$ and $\mathscr{C}$, the meet $\mathscr{B} \wedge \mathscr{C}$ exists.*

In the next section we shall see by example that Axiom B6 is not a consequence of Axioms B1–B5. By adopting it, we may omit the qualifications hitherto expressed regarding the existence of meets and joins, for (24) shows that $(\mathscr{A}^e \wedge \mathscr{B}^e)^e = \mathscr{A} \vee \mathscr{B}$.

There is a notation for the part of $\mathscr{A}$ that is not a part of $\mathscr{B}$:

$$\mathscr{A} \backslash \mathscr{B} := \mathscr{A} \wedge \mathscr{B}^e. \tag{I.2-41}$$

A basic representation theorem due to STONE asserts that every collection that is a Boolean algebra with respect to finite joins and meets is isomorphic to a *field of sets*[1]

[1] A field of sets is a non-empty collection of subsets of a given space that is closed with respect to the operations of finite union, intersection, and complementing.

in an appropriate topological space (see *e.g.* Theorem 8.2 of SIKORSKY's book, cited above in Footnote 2 on p. 7.) Thus, as far as finite Boolean operations are concerned, a universe and a collection of sets endowed with the usual operations of set theory are in essence equivalent. Nonetheless, such equivalence fails in general when we consider, as we do in continuum mechanics, infinite Boolean operations (see Chapter II of SIKORSKY's book).

3. Examples of Universes

We now mention two systems satisfying Axioms B1–B6 and one satisfying only Axioms B1–B5. In each, $\mathbf{\Omega}$ is a class of sets, and the symbol $\prec$ is taken as being $\subset$, the sign of inclusion, but only in the first one are $\wedge$ and $\vee$ the same as $\cap$ and $\cup$, the symbols of intersection and union. To avoid any possible confusion with the term "particle" as used in physics, in this book we shall call elements of the sets in $\mathbf{\Omega}$ *substantial points* henceforth until Section IV.2; there we shall adjoin further properties to those points, justifying our calling them thereafter *material points*.

Example 1. Let $\mathbf{\Omega}$ consist in all subsets of some set $\mathscr{Y}$. Then $\mathscr{A} \wedge \mathscr{B} = \mathscr{A} \cap \mathscr{B}$, $\mathscr{A} \vee \mathscr{B} = \mathscr{A} \cup \mathscr{B}$, and $\mathscr{A}^e$ is the complement of $\mathscr{A}$ in $\mathscr{Y}$. Also $\infty = \mathscr{Y}$ and $\varnothing = \mathscr{O}$, $\varnothing$ being the null set.

For example, $\mathscr{Y}$ may be a finite set, say $X_1, X_2, \ldots, X_n$. Universes of this kind are used in the classical dynamics of discrete systems. We shall develop the basic principles of that traditional mechanics in some of the succeeding sections of this chapter.

Example 2. Let $\mathbf{\Omega}_O$ consist in all regularly open sets[1] in a topological space $\mathscr{T}$. The exterior $\mathscr{B}^e$ of $\mathscr{B}$ is the interior of the complement of $\mathscr{B}$. For any collection of bodies $\mathscr{B}_k$ the meet, defined as follows:

$$\bigwedge_k \mathscr{B}_k := \operatorname{int} \operatorname{clo} \bigcap_k \mathscr{B}_k\,, \tag{I.3-1}$$

is a body of $\mathbf{\Omega}_O$. Thus $\mathbf{\Omega}_O$ is a universe.

The meet of a finite collection of bodies is simply the intersection, but for infinite collections such is not always the case.

[1] A set $\mathscr{A}$ is *regularly open* if $\operatorname{int}\operatorname{clo}\mathscr{A} = \mathscr{A}$. An example of an open set that is not regularly open is an open disc in the plane with one interior point removed.

Indeed, let $\mathscr{T}$ be the real line, and consider in $\mathbf{\Omega}_{\mathrm{O}}$ the particular bodies $\mathscr{B}_k :=$ $]-1/k, 1[$, $k = 1, 2, 3, \ldots$. Then $\bigcap_{k=1}^{\infty} \mathscr{B}_k = [0, 1[$, which does not belong to $\mathbf{\Omega}_{\mathrm{O}}$, but (1) shows that $\bigwedge_{k=1}^{\infty} \mathscr{B}_k =]0, 1[$, which does belong to $\mathbf{\Omega}_{\mathrm{O}}$.

The join of any collection of bodies $\mathscr{B}_k$ is given by

$$\bigvee_k \mathscr{B}_k = \operatorname{int}\operatorname{clo}\left(\bigcup_k \mathscr{B}_k\right). \tag{I.3-2}$$

Exercise I.3.1. The statement (2) is a consequence of the general definition of join, given in Section I.2.

Example 3. Let $\mathscr{T}$ be a Euclidean space $\mathscr{E}$. Consider the collection $\mathbf{\Omega}_{\mathrm{r}}$ of all regularly open sets in $\mathscr{E}$ that have piecewise smooth boundaries. $\mathbf{\Omega}_{\mathrm{r}}$ is a subcollection of $\mathbf{\Omega}_{\mathrm{O}}$. Thus, the meet in $\mathbf{\Omega}_{\mathrm{r}}$ is defined by (1). Nonetheless, $\mathbf{\Omega}_{\mathrm{r}}$ does not satisfy Axiom B6 because the meet of two sets of $\mathbf{\Omega}_{\mathrm{r}}$ need not belong to $\mathbf{\Omega}_{\mathrm{r}}$.

To see this last, we remark that the intersection of two sets with piecewise smooth boundaries need not itself have a piecewise smooth boundary. Suppose, for example, the elements of $\mathbf{\Omega}_{\mathrm{r}}$ be sets in the plane; let $\mathscr{B}_1$ be the open square $-1 < y < 0$, $0 < x < 1$, while $\mathscr{B}_2$ is the set of points such that $0 < x < 1$, $-1 < y < x^2 \sin x^{-1}$. Then $\mathscr{B}_1 \wedge \mathscr{B}_2 \notin \mathbf{\Omega}_{\mathrm{r}}$.

The student should recall the example just presented when he comes to Section II.1, in which we shall present a universe suitable for continuum mechanics.

4. Mass

Using Example 2 in Section I.3, we employ as $\bar{\mathbf{\Omega}}$ the closure of $\mathbf{\Omega}_{\mathrm{O}}$ obtained by adjunction of $\emptyset$ and $\mathscr{T}$ as the null body and the universal body. The bodies of interest in mechanics have mass; as we may say, they are *massy*. The massy bodies form a non-empty subclass $\mathbf{\Omega}_M$ of $\bar{\mathbf{\Omega}}$. The mass of $\mathscr{B}$ is the value $M(\mathscr{B})$ of a non-negative *mass function* M defined over $\mathbf{\Omega}_M$:

Axiom M1. $0 \leqq M(\mathscr{B}) \leqq \infty \quad \forall \mathscr{B} \in \mathbf{\Omega}_M$.

Further, we lay down

Axiom M2.

$$\mathscr{B} \in \mathbf{\Omega}_M \quad \Rightarrow \quad \mathscr{B}^e \in \mathbf{\Omega}_M,$$

$$\mathscr{B}_1 \ \& \ \mathscr{B}_2 \in \mathbf{\Omega}_M \quad \Rightarrow \quad \mathscr{B}_1 \vee \mathscr{B}_2 \in \mathbf{\Omega}_M.$$

That is, the exteriors[1] and joins of massy bodies also are massy bodies. In particular, $\mathscr{O}$ and ∞ are massy. Because $(\mathscr{B}_1^e \vee \mathscr{B}_2^e)^e = \mathscr{B}_1 \wedge \mathscr{B}_2$, it follows from Axiom M2 that

$$\mathscr{B}_1 \ \& \ \mathscr{B}_2 \in \mathbf{\Omega}_M \quad \Rightarrow \quad \mathscr{B}_1 \wedge \mathscr{B}_2 \in \mathbf{\Omega}_M. \tag{I.4-1}$$

Thus, the meet of two massy bodies is massy. Moreover, we assume that *mass is additive*:

Axiom M3. *If $\mathscr{B}_1$ and $\mathscr{B}_2$ are separate massy bodies, then*

$$M(\mathscr{B}_1 \vee \mathscr{B}_2) = M(\mathscr{B}_1) + M(\mathscr{B}_2).$$

Hence

$$M(\varnothing) = 0, \qquad M(\infty) = M(\mathscr{B}) + M(\mathscr{B}^e) \quad \forall \mathscr{B} \in \mathbf{\Omega}_M. \tag{I.4-2}$$

The mass assigned to the infinite body ∞ need not be ∞. If $M(\infty) = \infty$, then $M(\mathscr{B}) < \infty \Rightarrow M(\mathscr{B}^e) = \infty$. A body of mass 0 is called *massless*. Thus $\varnothing$ is massless, but of course there may be other massless bodies. That is, $M(\mathscr{B}) = 0 \nLeftarrow \mathscr{B} = \varnothing$. Also, by (I.2-39),

$$\mathscr{A} \prec \mathscr{B} \quad \Rightarrow \quad M(\mathscr{B}) = M(\mathscr{A}) + M(\mathscr{B} \wedge \mathscr{A}^e) \geqq M(\mathscr{A}). \tag{I.4-3}$$

Exercise I.4.1.

$$M(\mathscr{B} \vee \mathscr{C}) \leqq M(\mathscr{B}) + M(\mathscr{C}). \tag{I.4-4}$$

Though these properties reflect the obvious requirements of the idea of mass, they do not suffice to define it effectively. As is well known, if we are to obtain the convenient mathematical structure called *measure theory*, further

[1] If the requirement that $\mathscr{B}^e$ should be massy seems artificial, the student should recall that the possibility that $M(\mathscr{B}^e) = 0$ is not excluded.

assumptions must be laid down. While we assume that $\mathbf{\Omega}$ contains a massy part $\mathbf{\Omega}_M$, we shall not attempt to construct a measure based on $\mathbf{\Omega}_M$, for at present there seems to be no entirely satisfactory way of doing so in general.

While the notions of mass and electric charge, along with volume and area, were distilled to provide the basis of measure theory, that theory in its present state accounts satisfactorily only for the latter two, not for the former. Indeed, the mass function is a measure, but measure theory does not suffice for constructing a mass function. That is so because measure theory refers to *sets*, while, as we have seen in Section I.3, the notions of meet $\wedge$ and join $\vee$ of bodies generally are not the same as intersection $\cap$ and union $\cup$ in the algebra of sets, even in the case when bodies are indeed sets. A good mathematical theory of mass would be purely algebraic, assuming of bodies no more than the axioms B1–B6 (and preferably not the last).[1] The defect here is more one of clarity and elegance than application, since, as we shall see more clearly in Chapter II, the concepts of shape and motion enable us to use in continuum mechanics the theory of Lebesgue measure.

From now on we assume that the mass M defined over $\mathbf{\Omega}_M$ can be extended so as to be a Borel measure defined over all the Borel sets[2] of $\mathscr{T}$. There will be no confusion if we denote also this extended measure by M, even though a Borel set need not be a body. The assumption is more confining than it may appear at first glance, for if M is a measure on the Borel sets, it is additive on disjoint unions of them. Our basic Axiom M3 requires only that it be additive on the joins of separate bodies.

Once a non-negative mass function M be given, clearly KM is also a non-negative mass function if K is any positive constant. To any one particular body $\mathscr{B}$ that is not massless we may assign any positive mass we please, and the ratios

[1] A purely algebraic theory was developed by C. CARATHÉODORY in his last book, *Mass und Integral und ihre Algebraisierung*, Basel, Birkhäuser, 1956, translated as *Algebraic Theory of Measure and Integration*, Bronx, New York, Chelsea, 1963. While the σώματα over which CARATHÉODORY defines a measure formalize a concept of "body", he uses again and again the axiom that an enumerable collection of bodies have a join, which for applications in continuum mechanics is not always true.

[2] The collection of *Borel sets* in a topological space is the smallest σ-algebra that contains all of the open sets. A Boolean algebra whose elements are sets (with $\vee$ and $\wedge$ taken as $\cup$ and $\cap$) is a *σ-algebra* (or *Borel field*) if it includes every join of an enumerable collection of its elements. Thus all open sets, all closed sets, and all intersections of enumerable collections of open sets or closed sets are Borel sets.

The Borel sets suffice to define a measure on the topological space to which they belong, and every continuous map of that space is measurable. *Borel measure* serves to define the *Borel integral* of a real function whose values are not negative.

A brief and clear treatment of Borel sets and Borel measure is given by WALTER RUDIN, Chapters 1 and 2 of *Real and Complex Analysis*, 2nd edition, New York *etc.*, McGraw-Hill, 1974.

Because M is a Borel measure, the measure of every compact set is finite, a fact which is important for some of the arguments in Chapter III.

of the masses of bodies are unaffected by this choice. In physics the assignment of a particular mass to some one body in the universe is called "fixing the unit of mass."

Henceforth, apart from a few specified exceptions, we shall consider only $\mathbf{\Omega}_M$, not any greater universe $\bar{\mathbf{\Omega}}$, and we shall use the symbol $\bar{\mathbf{\Omega}}$ to denote $\mathbf{\Omega}_M$, thus excluding tacitly from our discourse any bodies that are not massy. Our assumptions enable us to write

$$M(\mathscr{A}) = \int_{\mathscr{A}} dM \qquad \text{if } \mathscr{A} \in \bar{\mathbf{\Omega}} \tag{I.4-5}$$

and the integral

$$\int_{\mathscr{A}} f\, dM \tag{I.4-6}$$

of any continuous function f can be defined in the way shown in books on the theory of measure and integration.

If $\mathbf{\Omega}_M$ consists in the subsets of a finite set whose elements are, say, $X_1, \ldots, X_n$, then a positive mass M_k is assigned to $\{X_k\}$:

$$M_k := M(\{X_k\}), \qquad k = 1, 2, \ldots, n, \tag{I.4-7}$$

and the masses of the other bodies in $\mathbf{\Omega}$ are obtained by addition of the masses of the separate elements composing them; for example, $M(\{X_1, X_2\}) := M_1 + M_2$. The substantial points $X_1, X_2, \ldots, X_n$ are called *mass-points*.

By assuming masses directly to the bodies of the universe we express a physical idea: *mass is conserved*.

This principle is nowadays considered appropriate to mathematical models for phenomena in which chemical or nuclear reactions may be neglected and the speeds associated to bodies are small in comparison with the speed of light. In theories of chemical reactions the principle still holds, but only for sufficiently large bodies, among the parts of which mass generally is exchanged.

5. Force

The general theory of systems of forces that we now present is NOLL's. So far, the whole refers to a fixed instant.

A *system of forces* on a universe $\mathbf{\Omega}$ is an assignment of vectors in some inner-product space $\mathscr{F}$ to all pairs of separate bodies of $\mathbf{\Omega}$. Vectors are denoted by bold-faced letters.

Let $(\mathbf{\Omega} \times \mathbf{\Omega})_0$ be the collection of such pairs. The first axiom of forces is

Axiom F1. $\mathbf{f}: (\mathbf{\Omega} \times \mathbf{\Omega})_0 \to \mathscr{F}$.

The vector $\mathbf{f}(\mathscr{B}, \mathscr{C})$ is called the *force exerted on* $\mathscr{B}$ *by* $\mathscr{C}$. Since we are here considering $\mathbf{\Omega}$ rather than $\bar{\mathbf{\Omega}}$, no force need be assigned to pairs one member of which is ∞ or $\mathscr{O}$. Moreover, the force exerted by two separate bodies on a third body separate from both is the sum of the forces exerted by each, and the force exerted by a body on the join of two separate parts of a separate body is the sum of the forces exerted on each. That is, the function $\mathbf{f}$ is additive in each of its variables:

Axiom F2. $\mathbf{f}(\mathscr{C}_1 \vee \mathscr{C}_2, \mathscr{B}) = \mathbf{f}(\mathscr{C}_1, \mathscr{B}) + \mathbf{f}(\mathscr{C}_2, \mathscr{B})$.

Axiom F3. $\mathbf{f}(\mathscr{B}, \mathscr{C}_1 \vee \mathscr{C}_2) = \mathbf{f}(\mathscr{B}, \mathscr{C}_1) + \mathbf{f}(\mathscr{B}, \mathscr{C}_2)$.

Both of these axioms refer to *pairwise separate* bodies $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{B}$.

If $\mathbf{f} \cdot \mathbf{g}$ denotes an inner product in $\mathscr{F}$, then $K\mathbf{f} \cdot \mathbf{g}$ is also an inner product if $K > 0$. Choice of a particular K is called "fixing the unit of force".

It is easy to extend $\mathbf{f}$ from $(\mathbf{\Omega} \times \mathbf{\Omega})_0$ to $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$, since Axioms F2 and F3 allow no other value but $\mathbf{0}$ for the force exerted by or on the null body. Thus we must set

$$\mathbf{f}(\mathscr{B}, \mathscr{O}) := \mathbf{f}(\mathscr{O}, \mathscr{B}) := \mathbf{0} \qquad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}. \tag{I.5-1}$$

The choices $\mathscr{B} = \infty$ or $\mathscr{B} = \mathscr{O}$ are not excluded here.

Since Axioms F2 and F3 are statements of additivity, we see that *if* $\mathbf{f}_1$ *and* $\mathbf{f}_2$ *are systems of forces on* $(\mathbf{\Omega} \times \mathbf{\Omega})_0$, *then for any numbers A and B the sum* $A\mathbf{f}_1 + B\mathbf{f}_2$ *is a system of forces.*

Since every body in $\bar{\mathbf{\Omega}}$ is separate from its exterior, Axiom F1 enables us to form $\mathbf{f}(\mathscr{B}, \mathscr{B}^e)$, the force exerted on $\mathscr{B}$ by its exterior. We call this particular force the *resultant force* on $\mathscr{B}$. Resultant forces are subject to a fundamental identity:

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{C}, \mathscr{B}) = \mathbf{f}(\mathscr{B}, \mathscr{B}^e) + \mathbf{f}(\mathscr{C}, \mathscr{C}^e) - \mathbf{f}(\mathscr{B} \vee \mathscr{C}, (\mathscr{B} \vee \mathscr{C})^e), \tag{I.5-2}$$

for all pairs of separate bodies $\mathscr{B}$ and $\mathscr{C}$ in $\bar{\mathbf{\Omega}}$. To prove the identity, suppose first that $\mathscr{C} = \mathscr{B}^e$. Then mere statement of (2) requires extension of $\mathbf{f}$ to $\bar{\mathbf{\Omega}}$ and hence leads to (1), whence (2) follows trivially. If $\mathscr{C} \neq \mathscr{B}^e$, extension of

$\mathbf{f}$ to $\bar{\mathbf{\Omega}}$ is not needed, and the following argument holds in $\mathbf{\Omega}$ as well as in $\bar{\mathbf{\Omega}}$, provided only $\mathscr{B}$ and $\mathscr{B} \vee \mathscr{C}$ have exteriors. Since $\mathscr{B} \wedge \mathscr{C} = \mathscr{O}$ by hypothesis, from (I.2-21), (I.2-38), and (I.2-24) we see that $\mathscr{B}^{\mathrm{e}}$ may be decomposed into separate parts as follows:

$$\mathscr{B}^{\mathrm{e}} = \mathscr{C} \vee (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}} \qquad \forall \mathscr{C} \prec \mathscr{B}^{\mathrm{e}}. \tag{I.5-3}$$

By Axiom F3

$$\begin{aligned} \mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) &= \mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{B}, (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}}), \\ \mathbf{f}(\mathscr{C}, \mathscr{C}^{\mathrm{e}}) &= \mathbf{f}(\mathscr{C}, \mathscr{B}) + \mathbf{f}(\mathscr{C}, (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}}), \end{aligned} \tag{I.5-4}$$

while by Axiom F2

$$\mathbf{f}(\mathscr{B} \vee \mathscr{C}, (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}}) = \mathbf{f}(\mathscr{B}, (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}}) + \mathbf{f}(\mathscr{C}, (\mathscr{B} \vee \mathscr{C})^{\mathrm{e}}). \tag{I.5-5}$$

Adding $(4)_1$ to $(4)_2$ and subtracting (5) from the sum yields (2). $\triangle$

If the force exerted by $\mathscr{C}$ on $\mathscr{B}$ is of magnitude equal and of sign opposite to that exerted by $\mathscr{B}$ on $\mathscr{C}$, that is,

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) = -\mathbf{f}(\mathscr{C}, \mathscr{B}) \qquad \forall (\mathscr{B}, \mathscr{C}) \in (\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0, \tag{I.5-6}$$

the system of forces $\mathbf{f}$ is said to be *pairwise equilibrated*. This term describes the meaning of the idea; *skew* is shorter but less suggestive. From (2) we may read off the following

Theorem (NOLL, GURTIN & WILLIAMS). *A system of forces is pairwise equilibrated if and only if the resultant force* $\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}})$, *regarded as a function of* $\mathscr{B}$, *is additive on the separate bodies of* $\bar{\mathbf{\Omega}}$.

A system of forces such that the resultant force on every body vanishes:

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = \mathbf{0} \qquad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}, \tag{I.5-7}$$

is *balanced*. Since the function whose value is $\mathbf{0}$ is additive, the above theorem has the following

Corollary (NOLL). *Every balanced system of forces is pairwise equilibrated.*

As is clear from the foregoing theorem, the converse of NOLL's corollary does not hold. Indeed, there are many systems of forces that are pairwise equilibrated but not balanced. One important example is presented later on in this section; another is furnished by the contact forces in continuum mechanics, as will be explained in Section III.1.

In the past, instances of (6) were often inferred from a vague "axiom" called the law of "action and reaction", which was regarded as expressing the content of NEWTON's Third Law of Motion: "To an action there is always a contrary and equal reaction; or, the actions of two bodies mutually upon one another are always equal and directed toward contrary parts." If, indeed, what NEWTON meant by "action" is what we here call "force", which is by no means clear from his own words or the contexts in which he applied them, then the above argument shows that axiom to be equivalent, as far as pairs of separate bodies are concerned, to additivity of resultant forces on separate bodies. This fact is independent of whatever relations there may be among forces and motions.

Axiom F2 states, among other things, that the forces exerted by the exterior $\mathscr{B}^e$ of a body $\mathscr{B}$ on the separate parts of that body are additive:

$$(\mathscr{P}_1 \prec \mathscr{B}) \;\&\; (\mathscr{P}_2 \prec \mathscr{B}) \;\&\; (\mathscr{P}_1 \wedge \mathscr{P}_2 = \mathscr{O})$$

$$\Rightarrow \quad \mathbf{f}(\mathscr{P}_1 \vee \mathscr{P}_2, \mathscr{B}^e) = \mathbf{f}(\mathscr{P}_1, \mathscr{B}^e) + \mathbf{f}(\mathscr{P}_2, \mathscr{B}^e). \quad \text{(I.5-8)}$$

This fact suggests that for every particular body $\mathscr{B}$ the forces exerted by $\mathscr{B}^e$ on a certain set of parts of $\mathscr{B}$ might define a vector-valued measure over $\mathscr{B}$, a measure which we could denote formally thus:

$$\mathbf{f}(\mathscr{A}, \mathscr{B}^e) = \int_{\mathscr{A}} d\mathbf{f}_{\mathscr{B}^e} \qquad \text{if } \mathscr{A} \prec \mathscr{B}. \quad \text{(I.5-9)}$$

It would be desirable to construct an abstract theory of integration with respect to systems of forces, as defined only by the above axioms and some further ones of a technical nature, but since no such general theory is presently available, we here simply assume that our systems of forces are of this kind:

Axiom F4. *For each $\mathscr{B}$ in $\bar{\boldsymbol{\Omega}}$, the function $\mathbf{f}(\cdot, \mathscr{B}^e)$ is a vector-valued measure over $\mathscr{B}$.*

Theorem. *If $\mathscr{A}$ and $\mathscr{B}$ are separate, then $\mathbf{f}(\cdot, \mathscr{B})$ is a measure over $\mathscr{A}$.*

Proof. By (I.2-19), every body $\mathscr{B}$ is the exterior of another one, namely, $\mathscr{B}^e$. By Axiom F4, $\mathbf{f}(\cdot, \mathscr{B})$ is a measure over $\mathscr{B}^e$. If $\mathscr{A}$ and $\mathscr{B}$ are separate, $\mathscr{A} \prec \mathscr{B}^e$ by Axiom B5, and so Axiom F4 yields the theorem at once. $\triangle$

In fact the theorem merely rephrases Axiom F4.

Axioms are used in two ways. First, they may serve as a mine, whence theorems are drawn by mathematical deduction. Secondly, they may express a criterion: A mathematical system, whether already constructed or in course of construction, may be proved conform with them. We shall use Axioms F1, F2, and F3 in the first way, but we will not call upon Axiom F4 as an assumption. Rather, *we shall demonstrate mathematically that the systems of forces in the two branches of mechanics developed in this book obey Axiom F4.* In the former branch, which is the analytical dynamics of mass-points, only finite sums occur, and the conclusion is obvious, as the student will see below through the steps (15)–(28). In the latter branch, which is the mechanics of three-dimensional continua, upon the forces one body exerts on another which is separate from it but in general contiguous we will impose as an axiom the physically immediate bound (III.1-10). Then we shall sketch a long and difficult analysis from which, among other important conclusions, such a system of forces will be proved conform with Axiom F4.

While the development thus far in this section applies generally to the class of universes discussed in Section I.2, now and henceforth we return to use of $\mathbf{\Omega}_M$ as specified in the preceding section, and again we use $\bar{\mathbf{\Omega}}$ to denote it. The mass M is Borel measure or an extension of it such as Lebesgue measure.

We may introduce the Stieltjes integral of a continuous real function h over $\mathscr{B}$ with respect to the measure $\mathbf{f}(\cdot, \mathscr{B}^e)$; we denote this integral by

$$\int h \, d\mathbf{f}_{\mathscr{B}^e} \tag{I.5-10}$$

and call it "the integral of h with respect to $\mathbf{f}_{\mathscr{B}^e}$".

For example, if every $\mathscr{B}$ is a subset of the set of mass-points $X_1, X_2, \ldots, X_n$,

$$\int_{\mathscr{B}} h \, d\mathbf{f}_{\mathscr{B}^e} = \sum_{k=1}^{n} h(X_k)\mathbf{f}(\{X_k\}, \mathscr{B}^e). \tag{I.5-11}$$

If

$$\mathbf{w}\colon \mathscr{B} \to \mathscr{F}, \tag{I.5-12}$$

so that $\mathbf{w}(X) \in \mathscr{F}$, and if $\mathbf{a} \in \mathscr{F}$, then $\mathbf{a}\cdot\mathbf{w}$ is a scalar field over $\mathscr{B}$, and $\int_{\mathscr{B}}(\mathbf{a}\cdot\mathbf{w})\,d\mathbf{f}_{\mathscr{B}^e}$, if it exists, is a linear function of $\mathbf{a}$. Consequently, there is a linear transformation on $\mathscr{F}$ whose value is $\int_{\mathscr{B}}(\mathbf{a}\cdot\mathbf{w})\,d\mathbf{f}_{\mathscr{B}^e}$. Denoting the transpose of this transformation by $\int_{\mathscr{B}} \mathbf{w} \otimes d\mathbf{f}_{\mathscr{B}^e}$, we have

$$\left[\int_{\mathscr{B}} \mathbf{w} \otimes d\mathbf{f}_{\mathscr{B}^e}\right]^{\mathsf{T}} \mathbf{a} = \int_{\mathscr{B}} (\mathbf{a}\cdot\mathbf{w})\,d\mathbf{f}_{\mathscr{B}^e}. \tag{I.5-13}$$

The trace of this linear transformation will be written as follows:

$$\int_{\mathscr{B}} \mathbf{w}\cdot d\mathbf{f}_{\mathscr{B}^e} := \operatorname{tr}\left[\int_{\mathscr{B}} \mathbf{w} \otimes d\mathbf{f}_{\mathscr{B}^e}\right]. \tag{I.5-14}$$

For example, if $\mathscr{B}$ is a finite set of substantial points X_k, then

$$\int_{\mathscr{B}} \mathbf{w}\cdot d\mathbf{f}_{\mathscr{B}^e} = \sum_{k=1}^{n} \mathbf{w}(X_k)\cdot\mathbf{f}(\{X_k\}, \mathscr{B}^e). \tag{I.5-15}$$

While the formulae (I.4-7), (11), and (15) are appropriate to bodies which are subsets of a finite set, they are merely instances of general conclusions. When it comes to systems of forces, the classical dynamics of mass-points offers a peculiar variant, to which we now turn for the nonce. In describing that dynamics we shall use X_k instead of $\{X_k\}$ to denote the set consisting in the one mass-point X_k, $k = 1, 2, \ldots, n$, and we shall adjoin one further body X_0, not necessarily massy, called "the environment" of the "system" $X_1, X_2, \ldots, X_n$. Thus

$$\infty = \bigvee_{k=0}^{n} X_k. \tag{I.5-16}$$

It is the usage of analytical dynamics to apply the word "body" only to subcollections of $\{X_1, X_2, \ldots, X_n\}$, excluding $\mathscr{O}$, X_0, and ∞. That is, the bodies treated are those defined as follows by a subset $\mathrm{s}_{\mathscr{B}}$ of $\{1, 2, \ldots, n\}$:

$$\mathscr{B} = \bigvee_{k \in \mathrm{s}_{\mathscr{B}}} X_k. \tag{I.5-17}$$

The traditional notations, more or less, are as follows:

$$\begin{aligned} \mathbf{f}_{kq} &:= \mathbf{f}(X_k, X_q), \\ \mathbf{f}_k^{\mathrm{e}} &:= \mathbf{f}(X_k, X_0); \end{aligned} \tag{I.5-18}$$

here k and q run from 1 to n. The forces $\mathbf{f}(X_k, X_q)$ are called *mutual*; the forces $\mathbf{f}(\mathscr{B}, X_0)$ are called *extrinsic*. Supposing assigned the quantities $\mathbf{f}_{kq}$ and $\mathbf{f}_k^{\mathrm{e}}$, we define the entire system of forces by the requirement that Axioms F1, F2, and F3 be satisfied. The resultant force $\mathbf{f}_k$ acting on X_k is given thus:

$$\mathbf{f}_k := \mathbf{f}(X_k, X_k^{\mathrm{e}}) = \mathbf{f}_k^{\mathrm{e}} + {\sum_{q=1}^{n}}' \mathbf{f}_{kq}; \tag{I.5-19}$$

the symbol $\sum'$ indicates a sum omitting the term for which $q = k$. If by $\mathrm{s}_{\mathscr{B}}^{-1}$ we denote the complement of $\mathrm{s}_{\mathscr{B}}$ in $\{1, 2, \ldots, n\}$, the resultant force on $\mathscr{B}$ is obtained as follows:

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = \sum_{k \in \mathrm{s}_{\mathscr{B}}} \left(\mathbf{f}_k^{\mathrm{e}} + \sum_{q \in \mathrm{s}_{\mathscr{B}}^{-1}} \mathbf{f}_{kq} \right). \tag{I.5-20}$$

The double sum is the *resultant mutual force* on $\mathscr{B}$; the single sum is the *resultant extrinsic force* on $\mathscr{B}$. In particular, if $\mathscr{B} = \{X_1, X_2, \ldots, X_n\}$, then $\mathrm{s}_{\mathscr{B}}^{-1}$ is empty, and so

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = \sum_{k=1}^{n} \mathbf{f}_k^{\mathrm{e}}. \tag{I.5-21}$$

If the system of forces is balanced, then (6) holds, so that in particular

$$\mathbf{f}_{kq} = -\mathbf{f}_{qk}, \qquad q \neq k; \tag{I.5-22}$$

then (20) may be written as

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = \sum_{k = \mathrm{s}_{\mathscr{B}}} \left(\mathbf{f}_k^{\mathrm{e}} + {\sum_{q=1}^{n}}' \mathbf{f}_{kq} \right) \tag{I.5-23}$$

because the terms by which the right-hand side differs from that of (20) cancel each other in pairs. By choosing $\mathscr{B}$ as X_k we conclude that

$$\mathbf{f}_k^{\mathrm{e}} + {\sum_{q=1}^{n}}' \mathbf{f}_{kq} = \mathbf{0}. \tag{I.5-24}$$

If, conversely, (24) and (22) hold when $k = 1, 2, \ldots, n$, then (20) shows that $\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = \mathbf{0}$. Thus *the conditions* (24) *and* (22) *are necessary and sufficient that the system of forces* $\mathbf{f}$ *on the universe of analytical dynamics be balanced*, provided we agree that also $\mathbf{f}(X_0, X_k) = -\mathbf{f}_k^{\mathrm{e}}$, $k = 1, 2, \ldots, n$. Either by summing (24) on k or by inspection of (21) we conclude that

$$\sum_{k=1}^{n} \mathbf{f}_k^{\mathrm{e}} = 0: \tag{I.5-25}$$

the total extrinsic force acting on the system is null. These simple theorems provide the standard structure of analytical dynamics.

It is easy to extend the foregoing to arbitrary pairs of bodies, which need not be distinct. If we introduce the *self-force* $\mathbf{f}_{kk}$ of X_k, the force exerted by X_k on itself, then we can define as follows the force exerted by the arbitrary body $\mathscr{C}$ on the arbitrary body $\mathscr{B}$:

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) := \sum_{\substack{h \in \mathrm{s}_{\mathscr{B}} \\ q \in \mathrm{s}_{\mathscr{C}}}} \mathbf{f}_{hq}, \tag{I.5-26}$$

$\mathrm{s}_{\mathscr{B}}$ and $\mathrm{s}_{\mathscr{C}}$ being the sets of integers that define $\mathscr{B}$ and $\mathscr{C}$ according to (17). When $\mathscr{B}$ and $\mathscr{C}$ are separate, this function $\mathbf{f}$ reduces to the $\mathbf{f}$ defined by the requirements F2 and F3 on the basis of (18). We may call $\mathbf{f}(\mathscr{B}, \mathscr{B})$ the *self-force* of $\mathscr{B}$. From (22) we see at once that *in a balanced system of forces,*

$$\mathbf{f}(\mathscr{B}, \mathscr{B}) = 0 \quad \forall \mathscr{B} \qquad \Leftrightarrow \qquad \mathbf{f}_{kk} = 0, \qquad k = 0, 1, \ldots, n. \tag{I.5-27}$$

Exercise I.5.1.

$$\mathbf{f}(\mathscr{B}, \infty) = \mathbf{f}(\mathscr{B}, \mathscr{B}) + \mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}). \tag{I.5-28}$$

In analytical dynamics it is customary to assume both that $\mathbf{f}_{kk} = \mathbf{0}$ and that the system of forces is balanced. From (27) and (28) it then follows that $\mathbf{f}(\mathscr{B}, \infty) = \mathbf{f}(\infty, \mathscr{B}) = \mathbf{0}\ \forall \mathscr{B}$. We may express this fact as a statement that the *universal body of analytical dynamics is* passive: The body ∞ exerts null force upon its parts.

As their statements suggest, (28) and the theorem stated just after it are not limited to discrete systems. RIZZO has proposed, in effect, the following axioms as a natural extension of NOLL's:

Axiom FE1. $\mathbf{f}\colon \bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}} \to \mathscr{F}$.

Axiom FE2.

$$\mathbf{f}(\mathscr{C}_1 \vee \mathscr{C}_2, \mathscr{B}) = \mathbf{f}(\mathscr{C}_1, \mathscr{B}) + \mathbf{f}(\mathscr{C}_2, \mathscr{B}) - \mathbf{f}(\mathscr{C}_1 \wedge \mathscr{C}_2, \mathscr{B}),$$
$$\mathbf{f}(\mathscr{O}, \mathscr{D}) = \mathbf{0}.$$

Axiom FE3.

$$\mathbf{f}(\mathscr{B}, \mathscr{C}_1 \vee \mathscr{C}_2) = \mathbf{f}(\mathscr{B}, \mathscr{C}_1) + \mathbf{f}(\mathscr{B}, \mathscr{C}_2) - \mathbf{f}(\mathscr{B}, \mathscr{C}_1 \wedge \mathscr{C}_2),$$

$$\mathbf{f}(\mathscr{D}, \mathscr{O}) = \mathbf{0}.$$

If $\mathscr{B}$, $\mathscr{C}_1$, and $\mathscr{C}_2$ are separate, these axioms reduce to NOLL's, and so any $\mathbf{f}$ that satisfies them is an extension from $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$ to $\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}}$ of an $\mathbf{f}$ that satisfies NOLL's axioms and (1). The formula (26) effects such an extension explicitly for a discrete universe. Axioms FE2 and FE3 are easy to motivate intuitively.

Exercise I.5.2.

$$\begin{aligned} \mathbf{f}(\mathscr{B}, \infty) &= \mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{B}, \mathscr{C}^{\mathrm{e}}), \\ \mathbf{f}(\infty, \mathscr{B}) &= \mathbf{f}(\mathscr{C}, \mathscr{B}) + \mathbf{f}(\mathscr{C}^{\mathrm{e}}, \mathscr{B}). \end{aligned} \tag{I.5-29}$$

Hence (28) holds,

$$\mathbf{f}(\infty, \mathscr{B}) = \mathbf{f}(\mathscr{B}, \mathscr{B}) + \mathbf{f}(\mathscr{B}^{\mathrm{e}}, \mathscr{B}), \tag{I.5-30}$$

and

$$\begin{aligned} \mathbf{f}(\mathscr{B}, \mathscr{B}) &= \mathbf{f}(\infty, \mathscr{B}) - \mathbf{f}(\mathscr{B}^{\mathrm{e}}, \mathscr{B}), \\ \mathbf{f}(\mathscr{B}, \mathscr{B}) &= \mathbf{f}(\mathscr{B}, \infty) - \mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}). \end{aligned} \tag{I.5-31}$$

From (28), now proved in generality, we see that if $\mathbf{f}(\mathscr{B}, \infty) = \mathbf{0}$, then

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}) = -\mathbf{f}(\mathscr{B}, \mathscr{B}): \tag{I.5-32}$$

If the universal body is passive, the resultant force on each body is the negative of its self-force. Thus, in such a universe, *the system of forces is balanced if and only if the self-force of every body is* $\mathbf{0}$. More generally, if $\mathbf{f}(\mathscr{B}, \infty) \neq \mathbf{0}$ for some $\mathscr{B}$, (32) does not hold, and we cannot easily infer anything about $\mathbf{f}$ on $\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}}$ from the statements obtained above about its restriction to $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$. In particular, it is not obvious how to infer (6) for pairs of bodies that are not separate.

As a first step in this direction, we find on the basis of the extended axioms FE1–FE3 a counterpart for the theorem of NOLL and GURTIN & WILLIAMS concerning pairs of separate bodies.

Lemma. *A system of forces on* $\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}}$ *is pairwise equilibrated for separate bodies if and only if the self-force* $\mathbf{f}(\mathscr{B}, \mathscr{B})$, *regarded as a function of* $\mathscr{B}$, *is additive on all bodies of* $\bar{\mathbf{\Omega}}$.

Proof. We apply Axiom FE2 when $\mathscr{B} = \mathscr{C}_1 \vee \mathscr{C}_2$ and $\mathscr{C}_1 \wedge \mathscr{C}_2 = \mathscr{O}$, then expand the conclusion by use of Axiom FE3. Then,

$$\begin{aligned}\mathbf{f}(\mathscr{B}, \mathscr{B}) &= \mathbf{f}(\mathscr{C}_1, \mathscr{C}_1 \vee \mathscr{C}_2) + \mathbf{f}(\mathscr{C}_2, \mathscr{C}_1 \vee \mathscr{C}_2),\\ &= \mathbf{f}(\mathscr{C}_1, \mathscr{C}_1) + \mathbf{f}(\mathscr{C}_2, \mathscr{C}_2) + \mathbf{f}(\mathscr{C}_1, \mathscr{C}_2) + \mathbf{f}(\mathscr{C}_2, \mathscr{C}_1). \qquad \text{(I.5-33)}\end{aligned}$$

The basic decomposition theorem following (I.2-38) assures us that for any $\mathscr{B}$ we may choose $\mathscr{C}_1$ as any of its parts. $\triangle$

Since $\mathbf{0}$ is an additive function, the lemma has the following

Corollary (RIZZO). *If the self-force of every body is* $\mathbf{0}$, *the system of forces is pairwise equilibrated for separate bodies.*

Exercise I.5.3.

$$\begin{aligned}\mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{C}, \mathscr{B}) = {}& \mathbf{f}(\mathscr{B} \vee \mathscr{C}, \mathscr{B} \wedge \mathscr{C}) + \mathbf{f}(\mathscr{B} \wedge \mathscr{C}, \mathscr{B} \vee \mathscr{C})\\ &+ \mathbf{f}(\mathscr{B} \vee \mathscr{C}, \mathscr{B} \vee \mathscr{C}) + \mathbf{f}(\mathscr{B} \wedge \mathscr{C}, \mathscr{B} \wedge \mathscr{C})\\ &- \mathbf{f}(\mathscr{B}, \mathscr{B}) - \mathbf{f}(\mathscr{C}, \mathscr{C}). \qquad \text{(I.5-34)}\end{aligned}$$

Theorem (RIZZO). *In order that*

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{C}, \mathscr{B}) = \mathbf{0} \qquad \forall(\mathscr{B}, \mathscr{C}) \in \bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}}, \qquad \text{(I.5-35)}$$

it is necessary and sufficient that

$$\mathbf{f}(\mathscr{B}, \mathscr{B}) = \mathbf{0} \qquad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}. \qquad \text{(I.5-36)}$$

Proof. Necessity is obvious. To prove sufficiency, we note that the bodies $\mathscr{B} \wedge \mathscr{C}$ and $(\mathscr{B} \vee \mathscr{C}) \wedge (\mathscr{B} \wedge \mathscr{C})^e$ are separate, and that

$$\mathscr{B} \vee \mathscr{C} = (\mathscr{B} \wedge \mathscr{C}) \vee [(\mathscr{B} \vee \mathscr{C}) \wedge (\mathscr{B} \wedge \mathscr{C})^e]. \qquad \text{(I.5-37)}$$

By use of Axioms FE2 and FE3, with the aid of (36) we show that

$$\begin{aligned}\mathbf{f}(\mathscr{B} \vee \mathscr{C}, \mathscr{B} \wedge \mathscr{C}) &= \mathbf{f}((\mathscr{B} \vee \mathscr{C}) \wedge (\mathscr{B} \wedge \mathscr{C})^{e}, \mathscr{B} \wedge \mathscr{C}),\\ \mathbf{f}(\mathscr{B} \wedge \mathscr{C}, \mathscr{B} \vee \mathscr{C}) &= \mathbf{f}(\mathscr{B} \wedge \mathscr{C}, (\mathscr{B} \vee \mathscr{C}) \wedge (\mathscr{B} \wedge \mathscr{C})^{e}).\end{aligned} \tag{I.5-38}$$

Substituting (38) into (34), shortened by use of (36), we conclude that

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) + \mathbf{f}(\mathscr{C}, \mathscr{B}) = \mathbf{f}(\mathscr{D}, \mathscr{E}) + \mathbf{f}(\mathscr{E}, \mathscr{D}), \qquad \mathscr{D} \wedge \mathscr{E} = \mathscr{O}. \tag{I.5-39}$$

The preceding corollary assures us that the right-hand side of the former of these equations equals naught. △

Systems of forces defined on pairs of bodies that are not separate will not be considered further in this book.

NOLL's corollary, derived above, asserts that a balanced system of forces is pairwise equilibrated on separate bodies. We may ask if the same holds for all pairs of bodies. The answer is no. From (28) we see that in a balanced system of forces $\mathbf{f}(\mathscr{B}, \infty) = \mathbf{f}(\mathscr{B}, \mathscr{B})$. Only if the universal body is passive does it follow that $\mathbf{f}(\mathscr{B}, \mathscr{B}) = \mathbf{0}\ \forall \mathscr{B}$. Since this last condition is necessary for the forces of $\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}}$ to be pairwise equilibrated, we conclude that *in order for a balanced system of forces to be pairwise equilibrated for all bodies, it is necessary and sufficient that the universal body be passive.*

Nothing said about forces in this section restricts the dimension of $\mathscr{F}$.

6. The Event World. Rigid Frames

In common life we regard ourselves and other objects as occupying places, which are sets of points in a three-dimensional space, the properties of which are given once and for all and are not altered by our presence or absence. Moreover, the changes we perceive in ourselves and in our environment we regard as occurring at specific instants, which are points in a one-dimensional space altogether independent of the space of places.

Places and instances are associated to *events*. We take an event as being a primitive entity like a point in geometry, not defined, but in some measure made clear by the mathematical properties we attribute to it.

We endeavor now to make this rough idea somewhat precise. There are several ways to do so.

We call the totality of events the *event-world* $\mathscr{W}$.

The event world is the blank canvas on which pictures of nature may be painted, the quarry for blocks from which statues of nature may be carved. This canvas, this quarry, must be chosen by the artist before he sets to work. It lays limitations upon his art, but it by no means determines the images he will fashion. Various kinds of mechanics rest upon use of different event-worlds. For example, the event-worlds for relativity and for the mechanics of oriented materials differ from the event-world we use in this book.

C.-S. MAN has kindly provided most of the text following in this section.

We presume that every event occurs at some definite instant. Instants have their being in and by themselves; they are elements of a given one-dimensional Euclidean space $\mathscr{T}$. Events that take place at the same instant are said to be *simultaneous*. We denote by $\mathscr{W}_t$ the totality of simultaneous events at the instant t. We assume that each $\mathscr{W}_t$ is a three-dimensional Euclidean space; moreover, we presume that physical means are available to compare distances in different $\mathscr{W}_t$ so that it is meaningful to say whether a bijection between $\mathscr{W}_t$ and $\mathscr{W}_{t'}$ $(t \neq t')$ is an isometry.[1]

Simultaneous events that occur some distance apart in $\mathscr{W}_t$ can be regarded as occurring at different places. For two events e and e' that are not simultaneous, to say whether or not they occur at the same place has no absolute meaning. A place by itself has no identity except at a specific instant. The identity of a place through different instants is assigned externally by an observer. One way to effect such assignments is by use of rigid frames, which we now introduce.

Let $\mathscr{E}$ be a three-dimensional Euclidean point space. Let

$$\oint: \mathscr{W} \to \mathscr{E} \times \mathscr{T}, \qquad \oint(e) = (\mathbf{x}, t), \tag{I.6-1}$$

be a bijection such that $\oint_t$, the restriction of $\oint$ to $\mathscr{W}_t$, is a bijection of $\mathscr{W}_t$ onto $\mathscr{E}$ for each instant t. Formally,

$$\oint_t: \mathscr{W}_t \to \mathscr{E}, \qquad \oint_t(e) = \mathbf{x} \qquad \text{if } \oint(e) = (\mathbf{x}, t). \tag{I.6-2}$$

Suppose we can compare distances in $\mathscr{W}_t$ and $\mathscr{E}$. If $\oint_t$ is an isometry for each instant t, we call the bijection $\oint$ a *rigid frame*; $\mathscr{E}$ is the *background* of $\oint$, and the elements $\mathbf{x}$ of $\mathscr{E}$ are the *places* in $\mathscr{E}$. Two events $e \in \mathscr{W}_t$ and $e' \in \mathscr{W}_{t'}$ that are not simultaneous occur at the same place in $\mathscr{E}$ if $\oint_t(e) = \oint_{t'}(e')$.

In informal speech the symbol $\oint$ may be pronounced "the reference" or "the observer".

Let us give an example to illustrate how a rigid frame may be realized (*cf.* Section I.3). Consider a universe $\mathbf{\Omega}$ with a supply of substantial points X such that the physical existence of each X at the instant t is marked by (or *traverses*) an event in $\mathscr{W}_t$. Henceforth, for simplicity, we shall refer to the substantial point X in $\mathscr{W}_t$ when we really mean the event traversed by X in $\mathscr{W}_t$. Suppose we can select four such substantial points X_i

[1] An isometry of two Euclidean spaces is a bijection that preserves distances.

($i = 1, 2, 3, 4$) that satisfy the following conditions:

(i) X_i traverses an event in $\mathscr{W}_t$ at each instant t;
(ii) the four substantial points X_i are not coplanar in any $\mathscr{W}_t$;
(iii) the distance between any two substantial points X_i and X_j, $i \neq j$, remains fixed for all instants t.

By the laws of Euclidean geometry, for each instant t the four substantial points X_i determine a unique three-dimensional Euclidean space $\mathscr{E}$ that includes the X_i among its elements. Like the substantial points X_i, each element of $\mathscr{E}$ exists independently of t and coincides with an event in $\mathscr{W}_t$ for each $t \in \mathscr{T}$. This relation of coincidence defines for each t a bijection $\oint_t : \mathscr{W}_t \to \mathscr{E}$, which is clearly an isometry. The bijections $\oint_t$ together provide a rigid frame $\oint: \mathscr{W} \to \mathscr{E}$. In the older literature[1] such a rigid frame is usually called a "frame of reference". There the non-coplanar substantial points that define a "frame of reference" are often so chosen as to define a cartesian co-ordinate system. Each element of $\mathscr{E}$ is identified by its cartesian co-ordinates, and the space $\mathscr{E}$ itself is identified with $\mathscr{R}^3$.

A rigid frame represents an observer. Just as there are many putative observers, there are many rigid frames. On the other hand, given two rigid frames $\oint: \mathscr{W} \to \mathscr{E} \times \mathscr{T}$ and $\oint^*: \mathscr{W} \to \mathscr{E}^* \times \mathscr{T}$, we can always pick an instant t and identify $\mathscr{E}^*$ and $\mathscr{E}$ through the isometry $\oint_t^* \circ \oint_t^{-1}: \mathscr{E} \to \mathscr{E}^*$. Since there is no real loss of generality in doing so, for convenience we shall choose one Euclidean space $\mathscr{E}$ once and for all and consider only rigid frames that have $\mathscr{E}$ as background.

As no confusion should arise, hereafter we shall often refer to rigid frames simply as "frames".

Some authors use the word "frame" to denote any bijection $\oint: \mathscr{W} \to \mathscr{E} \times \mathscr{T}$ such that $\oint_t : \mathscr{W}_t \to \mathscr{E}$ is a bijection for each $t \in \mathscr{T}$; here $\mathscr{E}$ is a three-dimensional Euclidean space. In that usage, $\oint_t$ need not be an isometry. For clarity let us call such bijections *generalized frames*. Two generalized frames $\oint$ and $\oint^*$ are *rigidly related* if the bijections $\oint_t^* \circ \oint_t^{-1}$ and $\oint_t \circ \oint_t^{*-1}$ of $\mathscr{E}$ onto $\mathscr{E}$ are isometries for each $t \in \mathscr{T}$. Every generalized frame $\oint$ gives rise to an equivalence class of generalized frames rigidly related to $\oint$. We call that class the *rigid class* of $\oint$. The family of all generalized frames is the disjoint union of such rigid classes. It is clear that all rigid frames belong to the same rigid class of generalized frames.

Much of what we shall discuss below about changes of rigid frames (Section I.9) and material frame-indifference (Section IV.2) remains valid if we replace rigid frames by

[1] See, for example, A. E. H. LOVE, *Theoretical Mechanics*, Cambridge, at the University Press, 1897. In the edition of 1921, reprinted 1987, the mathematical discussion of frames of reference, which is excellent, is on pp. 299–303. An excellent physical description of frames of reference may be found in the opening of Chapter X, "Relativistic Mechanics," of G. Joos, *Theoretical Physics*, 1932, corrected text translated by I. M. FREEMAN, New York, Stechert, 1934. A quotation from it is given below in Section I.13.

generalized frames that belong to the same rigid class. (*Cf.* the remark after (IV.2-6).) Nonetheless, for simplicity we shall use only rigid frames in the main text of this book.

Henceforth we regard the "unit of length" as a distance fixed once for all. The "physical distance" between two places in $\mathscr{E}$ is then represented by a real number, namely the ratio of that distance to the unit of length. Following customary usage in mathematics, we refer to such ratios also as "distances". A "distance" of five in the mathematical representation means a physical distance of five units.

We denote the translation space of $\mathscr{E}$ by $\mathscr{V}$. Henceforth the term *vector* will denote always an element $\mathbf{v}$ in $\mathscr{V}$, and $|\mathbf{v}|$ will denote the magnitude of $\mathbf{v}$. The inner product of vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathscr{V}$ will be denoted by $\mathbf{v}\cdot\mathbf{w}$, and linear transformations of $\mathscr{V}$ into itself, which we shall call *tensors* over $\mathscr{V}$, will be denoted by bold-faced letters $\mathbf{T}, \mathbf{S}, \ldots$. The notation $\mathbf{u} = \mathbf{T}\mathbf{v}$ is read, "The tensor $\mathbf{T}$ transforms $\mathbf{v}$ into $\mathbf{u}$." The Euclidean distance between the places $\mathbf{x}$ and $\mathbf{y}$ in $\mathscr{E}$ will be written as $|\mathbf{x}-\mathbf{y}|$, since $\mathbf{x}-\mathbf{y}$ is the vector in $\mathscr{V}$ that translates $\mathbf{y}$ into $\mathbf{x}$.

A transformation of $\mathscr{V}$ cannot preserve the inner products of all pairs of vectors in $\mathscr{V}$ unless it is a tensor. A necessary and sufficient condition for a tensor $\mathbf{Q}$ to preserve all inner products is

$$\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}. \tag{I.6-3}$$

Such a tensor is called *orthogonal.*[1]

We may assign a co-ordinate system to $\mathscr{T}$. The co-ordinate of an instant is called the *time* of that instant. Only those assignments of times to instants that preserve the orientation of $\mathscr{T}$ are allowed. The oriented distance between instants whose times are t_1 and t_2 is $t_2 - t_1$. It is called the *time interval* between those instants, and if that interval is positive, the time t_2 is said to be *later* than t_1, whereas t_1 is *earlier* than t_2. Choice of a particular co-ordinate system on $\mathscr{T}$ is called in physics "fixing the unit and origin of time".

Commonly the unit and origin of time in a frame ϕ are regarded as set in advance by the observer, and $\mathscr{T}$ is identified with the real line $\mathscr{R}$ according to this choice of co-ordinates and metric. That is, instants are confounded with

[1] From (3) we see that $\det\mathbf{Q} = \pm 1$. If $\det\mathbf{Q} = +1$, $\mathbf{Q}$ is a *rotation*. Every orthogonal tensor on a space of odd dimension is either a rotation or the product of a rotation by the central inversion $-\mathbf{1}$; that is, there is one and only one rotation $\mathbf{R}$ such that either $\mathbf{Q} = \mathbf{R}$ or $\mathbf{Q} = -\mathbf{R}$, and the only possible proper numbers of $\mathbf{Q}$ are $+1$ and -1. If, as we always suppose, $\dim\mathscr{V} = 3$, then 1 is a proper number of every $\mathbf{R}$, and the corresponding proper space is one-dimensional unless $\mathbf{R} = \mathbf{1}$. This last statement is the content of a famous theorem of EULER: Every non-identical rotation about a point is in fact a rotation about a single line. The *axis* of $\mathbf{Q}$ is the proper line of the one and only $\mathbf{R}$ to which $\mathbf{Q}$ is proportional.

times, which are the co-ordinates of instants. We shall follow that custom in this book. Moreover, for simplicity we shall assume that all observers adopt the same unit of time.

With these conventions we replace $(1)_1$ by

$$\oint: \mathscr{W} \to \mathscr{E} \times \mathscr{R}. \tag{I.6-4}$$

In practice we usually consider only an interval $\mathscr{I}$ of times in $\mathscr{R}$. The present time t is always an interior point of $\mathscr{I}$.

We have agreed that a rigid frame $\oint$: $\mathscr{W} \to \mathscr{E} \times \mathscr{R}$ represents an observer. Since forces are experienced by observers, we assume that the vector-space $\mathscr{F}$ to which forces belong is isomorphic to $\mathscr{V}$, the translation space of $\mathscr{E}$. Were we to consider a mathematical model in mechanics resting upon a different event-world, we might need to make a different choice of the vector-space $\mathscr{F}$.

I.6A. Newton's View of Time and Space

The scholion NEWTON put after the definitions and before the laws of motion in his *Principia* reads, in part, as follows.

> . . . Time, space, place, and motion are very well known to all. It must be noted nonetheless that the people may not conceive those quantities except through their relation to sensible objects. And thence arise certain prejudices for lifting which it is fitting to distinguish them into absolute and relative, true and apparent, mathematical and common.
>
> I. Absolute, true, and mathematical time, of itself and from its nature, flows equably without relation to anything external, and another name for it is duration: Relative, apparent, and common time is some sensible and external measure of duration through a motion . . . that the people use instead of true time, such as hour, day, month, year.
>
> II. Absolute space, by its nature, without relation to anything external, remains always like and immobile: Relative space is a measure of this absolute space or some sort of movable dimension which is determined by our senses through its place in respect of bodies and by the people is taken as the immovable space. Such is the dimension of the subterranean space, the aerial, the celestial, defined through its place in respect of the earth. The absolute and relative spaces are the same in kind and magnitude, but they do not remain the same in number. . . .
>
> III. Place is the part of space that a body occupies, and according to the space [used], it is either absolute or relative. I say, a part of space, not the location of a body . . . . Positions, properly speaking, have no quantity, nor are they so much places as the properties of places. . . .
>
> IV. Absolute motion is the translation of a body from one absolute place into another; relative motion, from one relative place into another. . . .

In this book we adopt NEWTON's "absolute time". His absolute space cannot now be accepted, though his view of "place" still serves essentially.

Our times as co-ordinates of instants are NEWTON's "relative times".

Our "unit of length" is NEWTON's "some sort of movable dimension" in "relative space". NEWTON's "part of space that a body occupies" is our "shape", to be defined mathematically and developed in the following section.

Much of the role of NEWTON's "absolute space" is taken over by the identification of "inertial frames" through NEWTON's First Law of Motion, which is presented below at the beginning of Section I.13. There the student will find a conceptual determination also of NEWTON's absolute time.

7. Motions

From now on until the end of Section I.8, we assume that a rigid frame ϕ: $\mathscr{W} \to \mathscr{E} \times \mathscr{R}$ is given, and we do not investigate the topological and differentiable structure of $\mathscr{W}$ that the bijections ϕ induce.

A *world-line* is a curve[1] in $\mathscr{W}$ whose image in $\mathscr{E} \times \mathscr{R}$ associates one place to each time, so that we may represent a world-line as follows:

$$\boldsymbol{\lambda}: \mathscr{I} \to \mathscr{E}, \tag{I.7-1}$$

$\mathscr{I}$ being an interval of $\mathscr{R}$. A collection of world-lines defined over $\mathscr{I}$ is a *world-tube*. The places on a world-tube at a fixed time t form a set $\mathscr{S}_t$, and for any two times t' and t'' in $\mathscr{I}$, every place in $\mathscr{S}_{t'}$ is connected with one or more places in $\mathscr{S}_{t''}$ by world-lines of the world-tube. Thus we may regard a world-tube $\boldsymbol{\tau}$ as a mapping of an interval of times into the set of all subsets of $\mathscr{E}$, which is commonly denoted by $P(\mathscr{E})$:

$$\begin{aligned} \boldsymbol{\tau}: \mathscr{I} &\to P(\mathscr{E}), \\ t &\mapsto \mathscr{S}_t . \end{aligned} \tag{I.7-2}$$

Intersections of world-lines represent collisions or the creation or destruction of bodies or elements of bodies. In specific mechanical theories such intersections are usually excluded altogether or allowed as exceptional cases subject to specified conditions.

Experiences are to be correlated with world-lines and world-tubes. We think of these as progressing "through" the event world $\mathscr{W}$ as time goes on.

[1] A curve is a piecewise differentiable, one-parameter family of events: $e = f(s)$, and s varies over some real interval.

A mapping $\boldsymbol{\mu}$ of the universe $\boldsymbol{\Omega}$ into the set $P(\mathscr{W})$ of all subsets of the event world $\mathscr{W}$,

$$\boldsymbol{\mu}:\ \boldsymbol{\Omega} \to P(\mathscr{W}), \tag{I.7-3}$$

is called a *motion* if for each body $\mathscr{B}$ in $\boldsymbol{\Omega}$, $\boldsymbol{\mu}(\mathscr{B})$ is a world-tube. Thus a motion may be represented alternatively as a mapping $\boldsymbol{\chi}_{\boldsymbol{\Omega}}$ of $\boldsymbol{\Omega} \times \mathscr{I}$ into $P(\mathscr{E})$:

$$\boldsymbol{\chi}_{\boldsymbol{\Omega}}:\ \boldsymbol{\Omega} \times \mathscr{I} \to P(\mathscr{E}). \tag{I.7-4}$$

$\mathscr{I}$ is again some interval in $\mathscr{R}$, such as for example $]-\infty, t_0[$ for some t_0. The value $\boldsymbol{\chi}_{\boldsymbol{\Omega}}(\mathscr{B}, t)$ of $\boldsymbol{\chi}_{\boldsymbol{\Omega}}$, which is a set in $\mathscr{E}$, is called the *shape*[1] of $\mathscr{B}$ at the time t. When thinking of t as being the present time we shall call $\boldsymbol{\chi}_{\boldsymbol{\Omega}}(\mathscr{B}, t)$ the *present shape* of $\mathscr{B}$.

As we have stated in Sections I.3 and I.4, we consider only massy bodies that are sets of points, which we call substantial points, in some topological space $\mathscr{T}$:

$$\mathscr{B} = \{X, Y, \ldots\}. \tag{I.7-5}$$

The motion of a body composed of substantial points is engendered by the motions of those points. Using the symbol $\boldsymbol{\chi}$ for this more detailed motion, we write

$$\boldsymbol{\chi}:\ \mathscr{B} \times \mathscr{I} \to \mathscr{E}, \tag{I.7-6}$$

and, explicitly,

$$\mathbf{x} = \boldsymbol{\chi}(X, t) \qquad \forall X \in \mathscr{B}, \quad \forall t \in \mathscr{I}. \tag{I.7-7}$$

In words, $\mathbf{x}$ is the place in $\mathscr{E}$ that the substantial point X *occupies* at the time t in the motion $\boldsymbol{\chi}$. Moreover, the shape of $\mathscr{B}$ at the time t is the set of places its substantial points occupy then:

$$\boldsymbol{\chi}_{\boldsymbol{\Omega}}(\mathscr{B}, t) = \{\boldsymbol{\chi}(X, t):\ X \in \mathscr{B}\}. \tag{I.7-8}$$

Each substantial point X is thus associated with a world-line, and the world-lines of all the points of $\mathscr{B}$ constitute the world-tube of $\mathscr{B}$.

[1] In the literature usually both a map from the set of bodies into the set of all subsets of $\mathscr{E}$ and the value of such a map for a given body are called "configurations".

The concept of "motion" embodied in (3) does not require assignment of a frame, but the "more detailed motion" (6) does. Henceforth the term "motion" is to be understood in the latter, special sense.

It is customary in mechanics, with specifically stated exceptions, to consider only such motions χ as are differentiable with respect to t at least twice and often as many times as desired for each substantial point X. Denoting the derivatives of χ with respect to t when X is held fixed by $\dot{\chi}, \ddot{\chi}, \ldots, \overset{(n)}{\chi}$, so that in particular $\dot{\chi} = \overset{(1)}{\chi}$ and $\ddot{\chi} = \overset{(2)}{\chi}$, we call the values of these derivatives the *velocity* $\mathbf{v}$, the *acceleration* $\mathbf{a}, \ldots$, the n^{th} *velocity* ${}_n\mathbf{v}$ of the substantial point at the time t:

$$\begin{aligned} \mathbf{v} &:= \dot{\chi}(X, t), \\ \mathbf{a} &:= \ddot{\chi}(X, t), \ldots, \\ {}_n\mathbf{v} &:= \overset{(n)}{\chi}(X, t). \end{aligned} \tag{I.7-9}$$

Thus ${}_1\mathbf{v} = \mathbf{v}$, and ${}_2\mathbf{v} = \mathbf{a}$. It is easy to show that, for any given χ, the velocities of a given substantial point are vectors:

$${}_n\mathbf{v} \in \mathscr{V}, \qquad n = 1, 2, 3, \ldots, \tag{I.7-10}$$

and therefore at each time t the function $\overset{(n)}{\chi}(\cdot, t)$ is a vector field defined over $\mathscr{B}$.

As was stated in the preceding section, the metric in the Euclidean point space $\mathscr{E}$ is determined by the inner product in the translation space $\mathscr{V}$; the metric in the space of instants is determined by the assignment of times to instants. We describe these facts by saying that "the units of ${}_n\mathbf{v}$ are those of (length)$\div$(time)^n."

While the restriction $\chi(X, \cdot)$ of the mapping χ to a particular substantial point X has been assumed smooth, nothing in the way of smoothness has been imputed to the restriction $\chi(\cdot, t)$ to a fixed time. For mechanics in its most general form, $\chi(\cdot, t)$ need not even be a one-to-one mapping of substantial points onto places in $\mathscr{E}$. Indeed, in the example furnished by analytical dynamics, the motion χ carries the several mass-points into a discrete set of places $\mathbf{x}_i$ at each time t, but the restricted mapping $\chi(\cdot, t)$ is not always one-to-one, for at a collision the world-lines of two or more mass-points intersect, and it is possible even that two world-lines coalesce for an interval of time and then split asunder again. In continuum mechanics, contrarily, the mapping $\chi(\cdot, t)$: $\mathscr{B} \to \chi_{\mathbf{Q}}(\mathscr{B}, t)$ is assumed bijective. This statement, which asserts that two distinct substantial points never come to occupy the same place at the same time, is sometimes called the *Axiom of Impenetrability.*

Of course it is possible to relax the Axiom of Impenetrability at singular points, curves, or surfaces so as to represent shock waves, slip sheets, tears, welds, and fractures, but in this book we do not consider those.

In a particular branch of mechanics a particular universe $\mathbf{\Omega}$ is laid down once and for all. Two examples have been provided above in Section I.3. When a choice of $\mathbf{\Omega}$ has been made, there is no danger of confusion if we write χ for $\chi_{\mathbf{\Omega}}$ in (4) while retaining also the sense (6).

In Section I.5 we have developed a mathematical theory of forces acting upon pairs of bodies. We have now introduced motions undergone by bodies. Putting these two theories together, we remark that the force $\mathbf{f}(\mathscr{A}, \mathscr{B})$ exerted by $\mathscr{B}$ upon $\mathscr{A}$ will generally be a function of t.

8. Linear Momentum. Rotational Momentum. Kinetic Energy. Working. Torque

We continue to suppose given a particular frame ϕ, in terms of which a motion χ of a body $\mathscr{B}$ is defined and is described by (I.7-6). The vector fields defined over $\mathscr{B}$ at the time t by means of the motion χ of $\mathscr{B}$ give rise to certain additive set functions, the values of integrals with respect to mass over $\mathscr{B}$. The most important of these are, first, the *linear momentum* of $\mathscr{B}$:

$$\mathbf{m}(\mathscr{B}; \chi(\cdot, t)) := \int_{\mathscr{B}} \dot{\chi}(\cdot, t)\, dM; \tag{I.8-1}$$

second, the *rotational momentum* of $\mathscr{B}$ with respect to the place $\mathbf{x}_0$:

$$\mathbf{M}(\mathscr{B}; \chi(\cdot, t))_{\mathbf{x}_0} := \int_{\mathscr{B}} (\chi(\cdot, t) - \mathbf{x}_0) \wedge \dot{\chi}(\cdot, t)\, dM = -\mathbf{M}(\mathscr{B}; \chi(\cdot, t))^{\mathsf{T}}_{\mathbf{x}_0}; \tag{I.8-2}$$

and, third, the *kinetic energy* of $\mathscr{B}$:

$$K(\mathscr{B}; \chi(\cdot, t)) := \frac{1}{2}\int_{\mathscr{B}} |\dot{\chi}(\cdot, t)|^2\, dM. \tag{I.8-3}$$

From the definitions of $\mathbf{m}$, $\mathbf{M}_{\mathbf{x}_0}$, and K we see that for a given motion χ of a given body $\mathscr{B}$ the values of these functions at a given time t are vectors, skew tensors, and scalars, respectively.

To lighten the notation we shall henceforth usually leave $\mathscr{B}$, χ, and t unwritten in formulae involving $\mathbf{m}$, $\mathbf{M}_{\mathbf{x}_0}$, and K. We shall always remember that these important functions of t are associated to $\mathscr{B}$ by a motion χ.

It is obvious from (2) that

$$\mathbf{M}_{\mathbf{x}_0} = \mathbf{M}_{\mathbf{x}_1} + (\mathbf{x}_1 - \mathbf{x}_0) \wedge \mathbf{m}. \qquad \text{(I.8-4)}$$

For a given motion χ of a given body $\mathscr{B}$, the quantities $\mathbf{m}$, $\mathbf{M}_{\mathbf{x}_0}$, and K are functions of time alone. Denoting the derivative with respect to time by a superimposed dot, we see that

$$\begin{aligned} \dot{\mathbf{m}} &= \int_{\mathscr{B}} \ddot{\chi}\, dM, \\ \dot{\mathbf{M}}_{\mathbf{x}_0} &= \int_{\mathscr{B}} (\chi - \mathbf{x}_0) \wedge \ddot{\chi}\, dM, \\ \dot{K} &= \int_{\mathscr{B}} \dot{\chi} \cdot \ddot{\chi}\, dM, \end{aligned} \qquad \text{(I.8-5)}$$

on the assumption that the indicated differentiations be permissible. Furthermore, in $(5)_2$ the place $\mathbf{x}_0$ is taken as a stationary one in the frame $\oint$.

Exercise I.8.1. Let the place $\mathbf{x}_0$ be stationary, and let $\mathbf{x}_1(\cdot)$ be any place-valued, differentiable function of time. Then

$$\begin{aligned} \dot{\mathbf{M}}_{\mathbf{x}_0} &= \dot{\mathbf{M}}_{\mathbf{x}_1} + (\mathbf{x}_1 - \mathbf{x}_0) \wedge \dot{\mathbf{m}} + \dot{\mathbf{x}}_1 \wedge \mathbf{m}, \\ &= \int_{\mathscr{B}} (\chi - \mathbf{x}_1) \wedge \ddot{\chi}\, dM + (\mathbf{x}_1 - \mathbf{x}_0) \wedge \dot{\mathbf{m}}. \end{aligned} \qquad \text{(I.8-6)}$$

These definitions and relations are introduced here for later convenience. The basic principles of mechanics relate the rates of change $\dot{\mathbf{m}}$, $\dot{\mathbf{M}}_{\mathbf{x}_0}$, and $\dot{K}$ to the forces acting on $\mathscr{B}$, as we shall explain in Sections I.12 and I.14.

In Section I.5 we have defined a system of forces and an integration over $\mathscr{B}$ with respect to the forces exerted on the parts of $\mathscr{B}$ by its exterior, $\mathscr{B}^e$.

At the end of Section I.6 we have agreed that forces belong to a vector space isomorphic to the inner-product space $\mathscr{V}$. By using a particular isomorphism we may form inner products of forces and other vectors such as velocities or accelerations. That there are infinitely many different isomorphisms of this kind, reflects the fact that units of force are not yet related to units of length and time. We shall consider any one isomorphism and by using the definition

(I.5-14) introduce as follows the *working* W of the system of forces $\mathbf{f}_{\mathscr{B}^e}$ in the motion χ of $\mathscr{B}$ at the time t:

$$W(\mathscr{B}; \chi(\cdot, t); \mathbf{f}_{\mathscr{B}^e}) := \int_{\mathscr{B}} \dot{\chi}(\cdot, t) \cdot d\mathbf{f}_{\mathscr{B}^e} . \tag{I.8-7}$$

The units of working are those of (force)(length)÷(time). When there is no fear of confusion, we shall drop from the notation the arguments of W.

Forces are conceived as acting upon bodies, and when those bodies undergo motions and hence take shapes in $\mathscr{E}$, the forces are carried over to those shapes in some specified way. Since the shapes themselves depend upon the choice of frame, so also must any transference to those shapes of the forces acting on bodies. Consequently the definition (7) of the working W rests also upon a particular choice of frame. In Section I.12 we shall impose as the basic axiom of mechanics the requirement that such dependence of W be only apparent: that is, that the working, although it is defined by (7) in terms of a frame ϕ, shall have the same value for all frames.

Since by means of the isomorphism selected we may in effect say that $\mathbf{f} \in \mathscr{V}$, we may define also the tensor product $\mathbf{v} \otimes \mathbf{f}$ and the exterior product $\mathbf{v} \wedge \mathbf{f}$, provided $\mathbf{v} \in \mathscr{V}$. In particular, the skew tensor $(\mathbf{x} - \mathbf{x}_0) \wedge \mathbf{f}$ is called the "moment at $\mathbf{x}$ of $\mathbf{f}$ with respect to $\mathbf{x}_0$." More generally, the moment $\mathbf{F}_{\mathbf{x}_0}$ of a system of forces $\mathbf{f}_{\mathscr{B}^e}$ on a part $\mathscr{A}$ of $\mathscr{B}$ in the motion of $\mathscr{B}$, with respect to $\mathbf{x}_0$, is defined thus:

$$\mathbf{F}(\mathscr{A}, \mathscr{B}^e; \chi(\cdot, t))_{\mathbf{x}_0} := \int_{\mathscr{A}} [\chi(\cdot, t) - \mathbf{x}_0] \wedge d\mathbf{f}_{\mathscr{B}^e} . \tag{I.8-8}$$

Although the moment is a special case of what is called a *torque*, in this book we shall regard the two terms as interchangeable and prefer to use the monosyllable. The particular torque $\mathbf{F}(\mathscr{B}, \mathscr{B}^e; \chi(\cdot, t))_{\mathbf{x}_0}$ is called the *resultant torque* of the system of forces on $\mathscr{B}$ with respect to $\mathbf{x}_0$ in the motion χ at the time t.

The moment of a system of forces acting on a body $\mathscr{B}$ is defined in terms of the shape of $\mathscr{B}$ in the motion χ, a particular frame ϕ being presupposed. The moment is a skew tensor having the dimensions of (force) × (length). More generally, any skew tensor having these dimensions is called a *torque*, and a torque-valued function $\mathbf{F}(\mathscr{B}, \mathscr{C})$ of pairs of bodies is called a *system of torques* if it satisfies axioms obtained from Axioms F1–F4 when $\mathbf{f}$ is replaced by $\mathbf{F}$ throughout. Torques that are not moments of forces are sometimes called *couples*. When all torques are moments of forces, as we shall assume in this book, the system of torques is called *simple*.

The importance of the resultant torque will appear in Section I.12. For the time being, we remark only that

$$\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0} = \mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_1} + (\mathbf{x}_1 - \mathbf{x}_0) \wedge \mathbf{f}(\mathscr{B}, \mathscr{B}^e). \tag{I.8-9}$$

That is, at the time t the resultant torque with respect to $\mathbf{x}_0$ differs from that with respect to $\mathbf{x}_1$ by the moment at $\mathbf{x}_1$, with respect to $\mathbf{x}_0$, of the resultant force on $\mathscr{B}$. Here we have dropped $\chi(\cdot, t)$ from the notation.

In Section I.5 we have defined a balanced system of forces as one in which the resultant force on each body is $\mathbf{0}$. By (9) we see that *if the system of forces is balanced, the resultant torque it exerts on any body is the same with respect to all places.*

In view of what has just been shown, the following definition makes sense: The torques arising from a balanced system of forces are said themselves to be *balanced* if the resultant torque on every body vanishes.

In a system of torques more generally, NOLL's corollary in Section I.5 applies with merely verbal changes, enabling us to conclude that *in a balanced system of torques*, $\mathbf{F}(\mathscr{B}, \mathscr{C}) = -\mathbf{F}(\mathscr{C}, \mathscr{B})$.

The general axioms of mechanics that we shall lay down in Section I.12 will imply that a certain, basic system of forces and torques be balanced.

In the universe of analytical dynamics (Section I.5, above), where a body $\mathscr{B}$ is defined by (I.5-17),

$$\begin{aligned}
\mathbf{m}(\mathscr{B}; \chi(\cdot, t)) &= \sum_{k \in s_{\mathscr{B}}} M_k \dot{\mathbf{x}}_k , \\
\mathbf{M}(\mathscr{B}; \chi(\cdot, t))_{\mathbf{x}_0} &= \sum_{k \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_0) \wedge M_k \dot{\mathbf{x}}_k , \\
K(\mathscr{B}; \chi(\cdot, t)) &= \tfrac{1}{2} \sum_{k \in s_{\mathscr{B}}} M_k |\dot{\mathbf{x}}_k|^2, \\
W(\mathscr{B}; \chi(\cdot, t); \mathbf{f}_{\mathscr{B}^e}) &= \sum_{k \in s_{\mathscr{B}}} \dot{\mathbf{x}}_k \cdot \left[\mathbf{f}_k^e + \sum_{q \in s_{\mathscr{B}}^{-1}} \mathbf{f}_{kq} \right].
\end{aligned} \tag{I.8-10}$$

Here we use as abbreviations the places and velocities given to the mass-points by their motions:

$$\mathbf{x}_k := \chi(X_k, t), \qquad \dot{\mathbf{x}}_k := \dot{\chi}(X_k, t), \tag{I.8-11}$$

and the other notations are those introduced in connection with analytical dynamics in Section I.5. To obtain $(10)_4$, we have used (I.5-15), (I.5-20), and the fact that $\mathscr{F}$ and $\mathscr{V}$ are isomorphic. Two instances of $(10)_4$ are of major interest. First, suppose $\mathscr{B}$ consists in X_k alone. Then

$$W(X_k;\, \chi(\cdot, t);\, \mathbf{f}_{X_k^e}) = \dot{\mathbf{x}}_k \cdot \mathbf{f}_k \tag{I.8-12}$$

By (I.5-24) we see that *if the system of forces is balanced,*

$$W(X_k;\, \chi(\cdot, t);\, \mathbf{f}_{X_k^e}) = 0. \tag{I.8-13}$$

Second, if $\mathscr{B} = \{X_1, X_2, \dots, X_n\}$, then $(10)_4$ reduces to

$$W(\mathscr{B};\, \chi(\cdot, t);\, \mathbf{f}_{\mathscr{B}^e}) = \sum_{k=1}^{n} \dot{\mathbf{x}}_k \cdot \mathbf{f}_k^e = -{\sum_{k,q=1}^{n}}' \dot{\mathbf{x}}_k \cdot \mathbf{f}_{kq}\,, \tag{I.8-14}$$

where the last expression holds if the system of forces is balanced. Thus, in general, the working of a system of forces on a dynamical system does not vanish.

The torque $\mathbf{F}(\mathscr{B}, \mathscr{C})_{\mathbf{x}_0}$ exerted by $\mathscr{C}$ on $\mathscr{B}$ with respect to $\mathbf{x}_0$ is defined by

$$\mathbf{F}(\mathscr{B}, \mathscr{C})_{\mathbf{x}_0} := \sum_{k \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_0) \wedge {\sum_{q \in s_{\mathscr{C}}}}' \mathbf{f}_{kq}\,, \tag{I.8-15}$$

in which $\mathscr{B}$ and $\mathscr{C}$ need not be separate. Likewise the torque exerted by the environment X_0 on the mass-point X_k is defined by

$$\mathbf{F}(X_k, X_0)_{\mathbf{x}_0} = (\mathbf{x}_k - \mathbf{x}_0) \wedge \mathbf{f}_k^e\,. \tag{I.8-16}$$

These definitions square with (8), and the resultant torque on $\mathscr{B}$ is given by

$$\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0} = \sum_{k \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_0) \wedge \left(\mathbf{f}_k^e + \sum_{q \in s_{\mathscr{B}}^{-1}} \mathbf{f}_{kq} \right). \tag{I.8-17}$$

The *self-torque* of $\mathscr{B}$ is the torque it exerts on itself. By (15),

$$\mathbf{F}(\mathscr{B}, \mathscr{B})_{\mathbf{x}_0} = \sum_{k \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_0) \wedge {\sum_{q \in s_{\mathscr{B}}}}' \mathbf{f}_{kq}\,. \tag{I.8-18}$$

In a balanced system of forces, (I.5-22) holds, and hence

$$\mathbf{F}(\mathscr{B}, \mathscr{B})_{\mathbf{x}_0} = \tfrac{1}{2} \sum_{k, q \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_q) \wedge \mathbf{f}_{kq}\,. \tag{I.8-19}$$

If the force $\mathbf{f}_{kq}$ exerted by X_q on X_k is parallel to the vector $\mathbf{x}_k - \mathbf{x}_q$ that translates the place $\mathbf{x}_q$ occupied by X_q into the place $\mathbf{x}_k$ occupied by X_k, the mutual forces are called *central*. For central mutual forces each summand in (19) vanishes, and we have the

Theorem (POISSON). *For a balanced system of forces on the universe of analytical dynamics, the self-torque of every body vanishes if the mutual forces are central.*

When the self-torque of $\mathscr{B}$ vanishes, the resultant torque (17) may be written in the form

$$\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0} = \sum_{k \in s_{\mathscr{B}}} (\mathbf{x}_k - \mathbf{x}_0) \wedge \left(\mathbf{f}_k^e + \sum_{q=1}^{n}{}' \mathbf{f}_{kq} \right). \tag{I.8-20}$$

That is, the resultant torque on $\mathscr{B}$ is the sum of the moments of the resultant forces acting on the mass-points that make up $\mathscr{B}$. In a balanced system of forces, each of those resultant forces vanishes, and so $\mathbf{F}(\mathscr{B}, \mathscr{B}^e) = \mathbf{0}$. That is, the system of torques is balanced.

Suppose, conversely, that the system of torques be balanced. Then by the analogue of (I.5-6),

$$\mathbf{F}(\mathscr{B}, \mathscr{C})_{\mathbf{x}_0} = -\mathbf{F}(\mathscr{C}, \mathscr{B})_{\mathbf{x}_0} \tag{I.8-21}$$

for all separate bodies $\mathscr{B}$ and $\mathscr{C}$. In particular, then,

$$\mathbf{F}(X_k, X_q)_{\mathbf{x}_0} = -\mathbf{F}(X_q, X_k)_{\mathbf{x}_0}. \tag{I.8-22}$$

That is,

$$(\mathbf{x}_k - \mathbf{x}_0) \wedge \mathbf{f}_{kq} = -(\mathbf{x}_q - \mathbf{x}_0) \wedge \mathbf{f}_{qk}. \tag{I.8-23}$$

If the forces are balanced, by (I.5-22) we obtain

$$(\mathbf{x}_k - \mathbf{x}_q) \wedge \mathbf{f}_{kq} = \mathbf{0}, \tag{I.8-24}$$

so that the mutual forces are central. In summary of the argument in this paragraph and the preceding one, we have the following

Theorem (NOLL). *If a system of forces on the universe of analytical dynamics is balanced, the corresponding system of torques is balanced if and only if the mutual forces are central.*

Exercise I.8.2. From (I.5-26) we see that in a balanced system of forces

$$\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0} + \mathbf{F}(\mathscr{B}, \mathscr{B})_{\mathbf{x}_0} = 0. \tag{I.8-25}$$

Hence the system of torques is balanced if and only if the self-torque of every body vanishes, and again NOLL's theorem follows.

Thus in analytical dynamics the balance of torques is *equivalent* to the hypothesis that the mutual forces are central, on the assumption that the system of forces is balanced. As should be plain from the arguments leading to NOLL's theorem, no such reduction of the balance of torques to the balance of forces can be expected in the more general and typical universes of mechanics. In continuum mechanics central forces, and indeed mutual forces, rarely appear.

The approach of analytical dynamics is untypical and next to useless in the general science of mechanics.

A position vector of a place $\mathbf{x}$ in $\mathscr{E}$ is a vector that translates some given origin $\mathbf{x}_0$ into $\mathbf{x}$. Thus, a *position vector field* $\mathbf{p}$ corresponding with the motion χ is given by

$$\mathbf{p}(X, t) := \chi(X, t) - \mathbf{x}_0\,. \tag{I.8-26}$$

Often the origin $\mathbf{x}_0$ is a fixed place. Then the time derivatives of $\mathbf{p}$ when X is held fixed equal the corresponding time derivatives of the motion itself:

$$\dot{\mathbf{p}} = \dot{\chi}, \qquad \ddot{\mathbf{p}} = \ddot{\chi}, \qquad etc. \tag{I.8-27}$$

The *center of mass* $\mathbf{x}_c$ of a body $\mathscr{B}$ of positive mass $M(\mathscr{B})$ in a shape $\chi_\Omega(\mathscr{B}, t)$ is that place whose position vector $\bar{\mathbf{p}}$ is the mean, in the sense of the mass, of the position vectors of all the substantial points of $\mathscr{B}$:

$$\bar{\mathbf{p}}(\mathscr{B}) := \frac{1}{M(\mathscr{B})}\int_{\mathscr{B}} \mathbf{p}\, dM. \tag{I.8-28}$$

Of course $\bar{\mathbf{p}}$ generally varies in time for a given body $\mathscr{B}$, but we do not indicate this fact in the notation. While $\bar{\mathbf{p}}$ depends upon the choice of the fixed place $\mathbf{x}_0$, its time derivative $\dot{\bar{\mathbf{p}}}$ does not, and by (1) we see that

$$\mathbf{m}(\mathscr{B}; \chi) = M(\mathscr{B})\dot{\bar{\mathbf{p}}}(\mathscr{B}). \tag{I.8-29}$$

Comparison with $(10)_1$ yields the following

Theorem (KELVIN & TAIT). *The linear momentum of a body $\mathscr{B}$ is the same as that of a mass-point having the same mass as $\mathscr{B}$ and moving so as always to occupy the center of mass of $\mathscr{B}$.*

Exercise I.8.3. Let the place $\mathbf{x}_0$ with respect to which the position vector is calculated be fixed. Then

$$\dot{\mathbf{M}}_{\mathbf{x}_0} = \dot{\mathbf{M}}_{\mathbf{x}_c} + \bar{\mathbf{p}} \wedge \dot{\mathbf{m}}. \tag{I.8-30}$$

9. Changes of Frame[1]

Once a rigid frame $\oint$ has been laid down as in Section I.6, we may wish to consider another one, $\oint^*$:

$$\oint^*: \mathscr{W} \to \mathscr{E} \times \mathscr{R}. \tag{I.9-1}$$

Since both $\oint$ and $\oint^*$ are bijections, the composition $\oint^* \circ \oint^{-1}$ is a bijection of $\mathscr{E} \times \mathscr{R}$ onto itself:

$$\oint^* \circ \oint^{-1}: \mathscr{E} \times \mathscr{R} \to \mathscr{E} \times \mathscr{R}, \qquad (\mathbf{x}, t) \mapsto (\mathbf{x}^*, t^*). \tag{I.9-2}$$

Because both t and t^* correspond to the same instant and because we have agreed that all observers adopt the same unit of time (*cf.* Section I.6), we can express (2) as

$$\mathbf{x}^* = \boldsymbol{\theta}(\mathbf{x}, t), \qquad t^* = t + a. \tag{I.9-3}$$

For a fixed t the mapping $\boldsymbol{\theta}(\cdot, t): \mathscr{E} \to \mathscr{E}$ is an isometry, because $\boldsymbol{\theta}(\cdot, t) = \oint^*_{t^*} \circ \oint_t^{-1}$ (*cf.* Section I.6). In $(3)_2$ a is a constant that depicts the possible difference in the origins of time for the frames $\oint$ and $\oint^*$.

Mappings $\oint^* \circ \oint^{-1}$ are called *changes of frame.*

We shall sometimes say "$\mathbf{x}^*$ and $\mathbf{x}$ are the places at which the same event occurs in $\oint^*$ and $\oint$, respectively." The student is expected to recognize this and like statements as pointing toward the interpretation of the mathematical structure in terms of experience. Similarly, the velocity of a substantial point defined with respect to the frame $\oint$ will be called its velocity *in* $\oint$, and a similar usage will be followed for all other quantities defined in terms of frames: acceleration, momentum, *etc.*

Let $\mathscr{V}$ be the *translation space* associated to $\mathscr{E}$. A mapping $\mathbf{h}: \mathscr{V} \to \mathscr{V}$ is an *isometry* if $|\mathbf{h}(\mathbf{u}) - \mathbf{h}(\mathbf{v})| = |\mathbf{u} - \mathbf{v}|$ for all $\mathbf{u}$ and $\mathbf{v}$ in $\mathscr{V}$. The following

[1] For the opening of this section in its present form I am indebted to C.-S. MAN.

representation theorem for an isometry is well known: *There is a uniquely determined, orthogonal, linear mapping* $\mathbf{Q}$: $\mathscr{V} \to \mathscr{V}$ *such that*

$$\mathbf{h}(\mathbf{v}) = \mathbf{Q}\mathbf{v} + \mathbf{h}(\mathbf{0}) \tag{I.9-4}$$

for all $\mathbf{v}$ in $\mathscr{V}$.

Exercise I.9.1. The analogue of (4) holds for two n-dimensional inner-product spaces.

Because $\mathbf{f}(\cdot, t)$: $\mathscr{E} \to \mathscr{E}$ is an isometry, use of (4) leads to the following

Theorem. *A change of frame* (3) *has the representation*

$$\left.\begin{aligned} \mathbf{x}^* &= \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\mathbf{x} - \mathbf{x}_0), \\ t^* &= t + a, \end{aligned}\right\} \qquad (\text{I.9-5})_{1,2}$$

in which $\mathbf{x}_0$ *is a fixed place in* $\oint$, $\mathbf{x}_0^*$ *maps times onto places in* $\oint^*$, $\mathbf{Q}$ *maps times onto orthogonal, linear mappings of* $\mathscr{V}$, *generally unique, and* a *is a constant.*

Proof. Choose and fix some $\mathbf{x}_0$ in $\mathscr{E}$. Then each vector $\mathbf{v}$ in $\mathscr{V}$ has the unique expression

$$\mathbf{v} = \mathbf{x} - \mathbf{x}_0 \qquad (\text{I.9-5})_3$$

for some $\mathbf{x}$ in $\mathscr{E}$. Let $\mathbf{h}_t$: $\mathscr{V} \to \mathscr{V}$ be defined by $\mathbf{h}_t(\mathbf{v}) = \mathbf{f}(\mathbf{x}, t) - \mathbf{f}(\mathbf{x}_0, t)$. By (4), there is a uniquely determined, orthogonal, linear mapping $\mathbf{Q}(t)$: $\mathscr{V} \to \mathscr{V}$ such that $\mathbf{h}_t(\mathbf{v}) = \mathbf{Q}(t)\mathbf{v} + \mathbf{h}_t(\mathbf{0})$. Since $\mathbf{h}_t(\mathbf{0}) = \mathbf{0}$, $\mathbf{x}^* = \mathbf{f}(\mathbf{x}, t)$ and $\mathbf{x}_0^*(t) = \mathbf{f}(\mathbf{x}_0, t)$, we conclude that $\mathbf{x}^* - \mathbf{x}_0^*(t) = \mathbf{Q}(t)(\mathbf{x} - \mathbf{x}_0)$. $\triangle$

Regarding $\mathbf{x}_0$ as the place of some one event as observed in $\oint$, we interpret $\mathbf{x}_0^*(t)$ as the place of the same event as observed in $\oint^*$, while $\mathbf{Q}$ maps all the lines through $\mathbf{x}_0$ as observed in $\oint$ isometrically and conformally onto lines through $\mathbf{x}_0^*$ as observed in $\oint^*$.

Note that the functions $\mathbf{x}_0^*(\cdot)$ and $\mathbf{Q}(\cdot)$ in $(5)_1$ can be arbitrary functions of the time t; in particular, they need not be continuous. If we use the frame $\oint$ to define a differentiable structure on $\mathscr{W}$, the frame $\oint^*$ need not be compatible with that structure. For example, a substantial point moving smoothly in $\oint$ may be hopping about discontinuously in $\oint^*$. In the rest of this section we shall restrict our discussion to changes of frame from $\mathscr{f}$ to $\mathscr{f}^*$ that are compatible with $\mathscr{f}$, thus accepting a single differentiable structure. For instance, when we talk about acceleration of a substantial point in $\mathscr{f}$, we have implicitly used $\mathscr{f}$ to define a C^2 structure on $\mathscr{W}$. Then, for changes of frame, we shall allow only those $\mathscr{f}^*$ for which the functions $\mathbf{x}_0^*(\cdot)$ and $\mathbf{Q}(\cdot)$ in $(5)_2$ belong to C^2, and so we can talk about the acceleration of the same substantial point as observed in $\mathscr{f}^*$. Since in mechanics we always deal with velocity, acceleration, *etc.*, we shall consider only those changes of frame for which the functions $\mathbf{x}_0^*(\cdot)$ and $\mathbf{Q}(\cdot)$ in $(5)_1$ are at least twice continuously differentiable.

Under the present interpretation the value $\mathbf{Q}(t)$ of $\mathbf{Q}$ is sometimes called the *relative orientation* of $\mathscr{f}^*$ with respect to $\mathscr{f}$ at time t. At a time t_0 when $\mathbf{Q}(t_0) = \mathbf{1}$ and $\mathbf{x}_0^*(t_0) = \mathbf{x}_0$, the two frames are said to *coincide*.

Since we have assumed that the orthogonal transformation $t \mapsto \mathbf{Q}(t)$ is continuously differentiable, $\det \mathbf{Q} = 1$ always or $\det \mathbf{Q} = -1$ always. Many authors prefer to keep the restriction $\det \mathbf{Q} = 1$ for admissible changes of frame. Such a restriction is obviously not required by the concepts of kinematics. In this book we study purely mechanical theories in which the current and past distances among the substantial points determine a body's current mechanical response (Section IV.2, below). We cannot use the mechanical responses of such a body as observed in two frames to distinguish them as long as both are equally fit as backgrounds to describe the class of kinematical processes that the body may undergo. In this context there is no reason to discard those changes of frame for which $\det \mathbf{Q} = -1$.

From (5) we see that one particular event, to which a place and time $(\mathbf{x}_0, t)$ are assigned by the frame $\mathscr{f}$, may be assigned an arbitrary place $\mathbf{x}_0^*(t)$ by some other frame $\mathscr{f}^*$. The vector that translates the fixed place $\mathbf{x}_0$ into a general place $\mathbf{x}$ in $\mathscr{f}$ is then rotated in $\mathscr{V}$, perhaps also reflected, into the vector that translates $\mathbf{x}_0^*(t)$ into the corresponding place $\mathbf{x}^*$ in $\mathscr{f}^*$, the rotation being the same for all places $\mathbf{x}$ at any one time.

If we like, we may picture a change of frame in terms of a motion (Section I.7). If we suppose a body to be given such a shape that one of its substantial points remains at the place $\mathbf{x}$ in $\mathscr{f}$, then (5) is the motion of that point in $\mathscr{f}^*$.

Exercise I.9.2. In (5) we may take as the constant place $\mathbf{x}_0$ in $\mathscr{f}$ any one we please, or, if we prefer, we may substitute for it any place-valued function of time: $\mathbf{x}_0(\cdot)$. Hence the class of all changes of frame forms a group.

From $(5)_1$ we see that the definition of "world-line" in Section I.7 is independent of the choice of frame.

We interpret the possibility of a change of frame as meaning that two observers who have chosen the same units of length and time may set their clocks differently and may be in arbitrary rigid motion with respect to one another, yet both are equally qualified to describe the phenomena represented by classical mechanics, whether they be right-handed or left-handed, and so any statement made by one is equivalent to some statement made by the other. For example, if a function $f(\mathbf{x}^*, t^*)$ is given, substitution of (5) into it yields a function $g(\mathbf{x}, t)$ with the same value $g(\mathbf{x}, t) = f(\mathbf{x}^*, t^*)$, and any two functions so related are regarded as equivalent under the change of frame. In fact both g and f represent, by use of the frames ϕ and ϕ^*, respectively, the same function defined on the event world $\mathscr{W}$.

A change of frame induces a transformation of the translation space $\mathscr{V}$. Indeed, suppose that

$$\mathbf{v} := \mathbf{x}_1 - \mathbf{x}_2 . \tag{I.9-6}$$

Then by (5)

$$\begin{aligned} \mathbf{v}^* := \mathbf{x}_1^* - \mathbf{x}_2^* &= \mathbf{Q}(t)(\mathbf{x}_1 - \mathbf{x}_2), \\ &= \mathbf{Q}(t)\mathbf{v}. \end{aligned} \tag{I.9-7}$$

Likewise, a change of frame induces a transformation of the tensor space over $\mathscr{V}$. If $\mathbf{w} \in \mathscr{V}$ and $\mathbf{v} \in \mathscr{V}$, and if there is a tensor $\mathbf{T}$ such that

$$\mathbf{w} = \mathbf{T}\mathbf{v}, \tag{I.9-8}$$

then by (7)

$$\mathbf{w}^* = \mathbf{Q}(t)\mathbf{w} = \mathbf{Q}(t)\mathbf{T}\mathbf{Q}(t)^{\mathsf{T}}\mathbf{v}^*; \tag{I.9-9}$$

that is, $\mathbf{w}^* = \mathbf{T}^*\mathbf{v}^*$, and

$$\mathbf{T}^* = \mathbf{Q}(t)\mathbf{T}\mathbf{Q}(t)^{\mathsf{T}}. \tag{I.9-10}$$

Rules of just the same form are induced for vector-valued and tensor-valued functions of time $\mathbf{v}(t)$ and $\mathbf{T}(t)$, respectively.

A change of frame (5) induces also a change in the motion (I.7-7) of a substantial point X of a body $\mathscr{B}$. Namely, in ϕ^* the place $\mathbf{x}^*$ occupied by X at the time t^* is given by the relations

$$\begin{aligned} \mathbf{x}^* &= \chi^*(X, t^*) = \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\chi(X, t) - \mathbf{x}_0), \\ t^* &= t + a. \end{aligned} \tag{I.9-11}$$

We shall regard χ^* and χ as being the *same* motions as observed in $\mathcal{f}^*$ and $\mathcal{f}$, respectively. We regard $\mathbf{x}_0$ and t as the place and time assigned by $\mathcal{f}$ to some particular event and $\mathbf{x}_0^*(t)$ and t^* as the place and time assigned by $\mathcal{f}^*$ to that same event. When we need to emphasize the role of a frame, we shall call χ *the motion of $\mathscr{B}$ in $\mathcal{f}$*, and χ^* *the same motion of $\mathscr{B}$ but in $\mathcal{f}^*$*, as explained above. We shall refer to the transformation (11), which relates the motion in $\mathcal{f}$ with that in $\mathcal{f}^*$, by the same name as the transformation (5) of the frame $\mathscr{E} \times \mathscr{R}$, namely, a *change of frame*.

If a certain prescription defines vectors in terms of a frame, and if the prescription itself is independent of the choice of frame, it will deliver vectors $\mathbf{v}^*$ and $\mathbf{v}$, respectively, according as $\mathcal{f}^*$ or $\mathcal{f}$ is used, and generally these two vectors will not be the same. Consider, for example, the prescriptions $(\text{I}.7\text{-}9)_{1,2}$, which define the velocity and the acceleration in any frame:

$$\begin{aligned} \mathbf{v} &:= \dot{\chi}(X, t), & \mathbf{v}^* &:= \dot{\chi}^*(X, t^*), \\ \mathbf{a} &:= \ddot{\chi}(X, t), & \mathbf{a}^* &:= \ddot{\chi}^*(X, t^*). \end{aligned} \tag{I.9-12}$$

The dots in the second column indicate derivatives with respect to t^*, and the motion χ^* in $\mathcal{f}^*$ is related to the motion χ in $\mathcal{f}$ by (11). Thus

$$\dot{\chi}^*(X, t^*) = \dot{\mathbf{x}}_0^*(t) + \mathbf{Q}(t)\dot{\chi}(X, t) + \dot{\mathbf{Q}}(t)(\chi(X, t) - \mathbf{x}_0). \tag{I.9-13}$$

Therefore the velocity $\dot{\chi}^*$ in $\mathcal{f}^*$ is related to the velocity $\dot{\chi}$ in $\mathcal{f}$ by the formula

$$\dot{\chi}^* - \mathbf{Q}\dot{\chi} = \dot{\mathbf{x}}_0^* + \mathbf{A}(\chi^* - \mathbf{x}_0^*), \tag{I.9-14}$$

in which

$$\mathbf{A} := \dot{\mathbf{Q}}\mathbf{Q}^{\mathsf{T}} = -\mathbf{A}^{\mathsf{T}}. \tag{I.9-15}$$

The skew tensor $\mathbf{A}$ is the *spin*[1] of $\mathcal{f}$ with respect to $\mathcal{f}^*$. In (11) we regarded $\mathbf{x}_0^*(t)$ and t^* as the place and time assigned by $\mathcal{f}^*$ to a certain reference event to which $\mathcal{f}$ assigns the place $\mathbf{x}_0$ and the time t. Thus the value of the function $\dot{\mathbf{x}}_0^*$ is the rate of change of the place $\mathbf{x}_0^*(t)$ assigned by $\mathcal{f}^*$ to that reference event. On the other hand by (13), the velocity in $\mathcal{f}^*$ of the substantial point that occupies the place $\mathbf{x}_0$ at the time t is $\dot{\mathbf{x}}_0^*(t) + \mathbf{Q}(t)\dot{\chi}(\chi^{-1}(\mathbf{x}_0, t), t)$, which reduces to $\dot{\mathbf{x}}_0^*(t)$ if and only if the substantial point is at rest in $\mathcal{f}$.

[1] The old term "angular velocity" is gradually falling out of use, since not only is it an awkward polysyllable but also it suggests we should look for angles, which in general considerations we are better advised not to do.

Heretofore we have regarded the relative orientation $\mathbf{Q}$ as a known, differentiable function of t. Suppose instead we know the spin $\mathbf{A}$, which we assume to be a continuous function whose values are skew tensors. Considering the first-order linear differential equation

$$\dot{\mathbf{Y}} - \mathbf{A}\mathbf{Y} = \mathbf{0}, \qquad \text{(I.9-16)}$$

we observe first that it has a unique solution $\mathbf{Y}$ such that $\mathbf{Y}(t_0)$ assumes an assigned value.

Exercise I.9.3. If $\mathbf{Z} := \mathbf{Y}\mathbf{Y}^{\mathsf{T}}$ and $\mathbf{Y}$ satisfies (16), then

$$\dot{\mathbf{Z}} = \mathbf{A}\mathbf{Z} - \mathbf{Z}\mathbf{A}. \qquad \text{(I.9-17)}$$

Appeal to the uniqueness theorem for ordinary differential equations shows that a solution $\mathbf{Y}$ of (16) which is orthogonal when $t = t_0$ is orthogonal for all t. Likewise, if $\mathbf{Y}$ is a rotation when $t = t_0$, it is a rotation for all t.

The argument completed in the foregoing exercise is summarized in the following

Theorem. *Let the spin* $\mathbf{A}$ *of* $\oint$ *with respect to* $\oint^*$ *be a continuous function of time, and let the relative orientation* $\mathbf{Q}(t)$ *be prescribed at some one time* t_0. *Then a unique change of frame is determined by assignment of the place* $\mathbf{x}_0^*(t)$ *occupied in* $\oint^*$ *at the time* t *by some one place* $\mathbf{x}_0$ *in* $\oint$.

Exercise I.9.4. If $\mathbf{A}^*$ denotes the spin of $\oint^*$ with respect to $\oint$, then

$$\mathbf{A}^* = -\mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q}. \qquad \text{(I.9-18)}$$

More generally, if $\mathbf{Q}_1$ and $\mathbf{Q}_2$ correspond with changes of frame from $\oint$ to $\oint_1$ and from $\oint_1$ to $\oint_2$, respectively, and if $\mathbf{Q}_3$ corresponds with the change from $\oint$ to $\oint_2$, then

$$\begin{aligned} \mathbf{Q}_3 &= \mathbf{Q}_2\mathbf{Q}_1, \\ \mathbf{A}_3 &= \mathbf{A}_2 + \mathbf{Q}_2\mathbf{A}_1\mathbf{Q}_2^{\mathsf{T}}. \end{aligned} \qquad \text{(I.9-19)}$$

Hence if the framings $\oint_2$ and $\oint_1$ coincide at some instant, the spin of $\oint_2$ with respect to $\oint$ is the sum of its spin with respect to $\oint_1$ and the spin of $\oint_1$ with respect to $\oint$ at that instant.

The conclusions (19) are commonly described as asserting that while rotations are multiplicative, spins are additive.

The axis of the $\mathbf{Q}(t)$ in the change of frame (5) is the *axis of rotation* of that change of frame at the time t. The *angle of rotation* of the change of frame at time t is the angle of rotation $\theta(t)$ of the rotation $\mathbf{R}(t)$ such that $\mathbf{Q} = \mathbf{R}$ or $\mathbf{Q} = -\mathbf{R}$ (*cf.* Section App. IIA.14). Since $\mathbf{A}(t)$ is skew, its null space, likewise, is a single line except in the trivial instance $\mathbf{A}(t) = \mathbf{0}$. This line is called the *axis of spin*. The corresponding proper number of $\mathbf{A}(t)$ is 0. Since $\mathbf{A}\cdot\mathbf{A} = \dot{\mathbf{Q}}\cdot\dot{\mathbf{Q}}$, the magnitude of $\mathbf{A}$ is the same as the magnitude of $\dot{\mathbf{Q}}$. The value of $|\mathbf{A}|/\sqrt{2}$ is called the *angular speed* ω at which $\oint$ is rotating with respect to $\oint^*$ at the time t.

Exercise I.9.5. If the axis of rotation is independent of t, then it is also the axis of spin. Hence if the angle of rotation is $\theta(t)$, then $\omega = |\dot{\theta}(t)|$.

With the aid of a convention of sign, we can define a vector $\boldsymbol{\omega}$ such that $\boldsymbol{\omega} \times \mathbf{b} = \mathbf{A}\mathbf{b}$ for all vectors $\mathbf{b}$. Here the sign $\times$ denotes the cross-product of 3-dimensional vector analysis. The vector $\boldsymbol{\omega}$ is called the *angular velocity* of the rotation $\mathbf{R}$. Of course $|\boldsymbol{\omega}| = \omega$.

Exercise I.9.6. (Galletto). If $\theta \neq 0$ and if $\mathbf{e}$ is a suitably selected unit vector in the axis of rotation,

$$\boldsymbol{\omega}\cdot\mathbf{e} = \dot{\theta}. \tag{I.9-20}$$

We turn now to the acceleration. If we differentiate (14) with respect to t, we obtain by $(12)_2$ and (15) the following relation between the acceleration $\ddot{\chi}^*$ in $\oint^*$ and the acceleration $\ddot{\chi}$ in $\oint$:

$$\begin{aligned}\ddot{\chi}^* - \mathbf{Q}\ddot{\chi} &= \dot{\mathbf{Q}}\dot{\chi} + \ddot{\mathbf{x}}_0^* + \dot{\mathbf{A}}(\chi^* - \mathbf{x}_0^*) + \mathbf{A}(\dot{\chi}^* - \dot{\mathbf{x}}_0^*),\\ &= \ddot{\mathbf{x}}_0^* + 2\mathbf{A}(\mathbf{Q}\dot{\chi}) + (\dot{\mathbf{A}} + \mathbf{A}^2)(\chi^* - \mathbf{x}_0^*);\end{aligned} \tag{I.9-21}$$

the second right-hand side follows by substituting (14) into $(21)_1$. Here, and sometimes later, we use abbreviated notation as in (14). The first term on the right-hand side is the acceleration of the place in $\oint^*$ assigned at the time t to the place $\mathbf{x}_0$ in $\oint$. The second term, named after Coriolis, is the acceleration in $\oint^*$ that corresponds with the velocity $\dot{\chi}$ in $\oint$ (as this vector is seen by $\oint^*$, that is, rotated by $\mathbf{Q}$) and to the spin of $\oint$ with respect to $\oint^*$. The third term has two parts, the first of which, named after Euler, corresponds with the rate of change of the angular velocity, while the second, called the *centripetal acceleration*, expresses the acceleration caused by the pure transport of substantial points with respect to $\oint^*$.

Exercise I.9.7. For any field $\mathbf{A}$ of skew tensors depending on t only,

$$-\mathbf{A}^2\mathbf{p} = \nabla(-\tfrac{1}{2}\mathbf{p}\cdot\mathbf{A}^2\mathbf{p}), \tag{I.9-22}$$

∇ being the gradient operator and $\mathbf{p}$ a position vector. This statement can be interpreted in terms of the centripetal acceleration. ($\mathbf{A}^2$ is a symmetric tensor; for the particular $\mathbf{A}$ defined by (15) $\mathbf{A}^2$ has the axis of spin as its nullspace and has $-\omega^2$ and 0 as its proper numbers.)

The linear momentum, rotational momentum, and kinetic energy of a body depend likewise upon the frame. The transformations of these quantities and their rates of change induced by a change of frame are easy to calculate by substituting (14) and (21) into appropriate formulae of Section I.8.

10. Rigid Motion

A motion of a body is called *rigid* if there is a frame $\oint^*$ such as to make its velocity field vanish. The frame $\oint^*$ is called a *rest frame* for that motion. To calculate the velocity field of a rigid motion in a general $\oint$, we need only set $\dot{\chi}^* = \mathbf{0}$ in (I.9-14), generalized to allow $\mathbf{x}_0$ to depend on t as in Exercise I.9.2, and then by use of (I.9-15) obtain the following

Theorem (EULER). *A motion is rigid if and only if its velocity field in any, and hence every, frame $\oint$ is of the form*

$$\begin{aligned} \dot{\chi} &= \dot{\mathbf{x}}_0 - \mathbf{Q}^{\mathsf{T}}\dot{\mathbf{x}}_0^* - \mathbf{Q}^{\mathsf{T}}\mathbf{A}(\chi^* - \mathbf{x}_0^*), \\ &= \mathbf{c} + \mathbf{W}(\chi - \mathbf{x}_0); \end{aligned} \qquad \text{(I.10-1)}$$

here $\mathbf{x}_0(t)$ is a place in $\mathscr{E}$, $\mathbf{c}(t)$ is a vector, and $\mathbf{W}(t)$ is a skew tensor.

Of course $\mathbf{W} = \mathbf{A}^*$, the spin of a rest frame $\oint^*$ with respect to $\oint$, related to $\mathbf{A}$ through (I.9-18). We use the special symbol $\mathbf{W}$ to remind the reader that we refer to a particular kind of motion of a body, or, if we like, a particular frame, while $\mathbf{A}$ is defined for any pair of frames, irrespective of whatever motion of a body may be taking place with respect to them. We call $\mathbf{W}$ the *spin* of the rigid motion.

The nullspace of $\mathbf{W}(t)$ is called the *axis* of the rigid motion in $\oint$ at the time t. Substantial points lying upon a line through $\mathbf{x}_0$ and parallel to the axis of the motion are moving with the common velocity $\mathbf{c}(t)$.

Exercise I.10.1. For a given rigid motion at a given time, $\mathbf{W}$ is unique if and only if the shape of $\mathscr{B}$ is not part of a straight line.

In Section I.9 we showed that the function $\mathbf{A}$ determines the function $\mathbf{Q}$ uniquely if $\mathbf{Q}(t_0)$ is prescribed. By (I.9-18) we may use the function $\mathbf{W}$ to determine $\mathbf{Q}$ in the same way, as we may see equally well by writing the differential equation (I.9-16) in the form

$$\dot{\mathbf{Y}} + \mathbf{Y}\mathbf{W} = \mathbf{0} \tag{I.10-2}$$

and seeking the solution $\mathbf{Y}$ that assumes the value $\mathbf{Q}(t_0)$ when $t = t_0$.

From $(1)_2$ we see that the vector $\mathbf{c}(t)$ is the velocity of the substantial point currently occupying the place $\mathbf{x}_0$ in $\oint$; from $(1)_1$, that $\mathbf{c}$ is expressed as follows in terms of the functions $\mathbf{x}_0$ and $\mathbf{x}_0^*$ in (I.9-11):

$$\mathbf{c} = \dot{\mathbf{x}}_0 - \mathbf{Q}^{\mathsf{T}}\dot{\mathbf{x}}_0^* . \tag{I.10-3}$$

Thus, once $\mathbf{Q}$ has been determined and $\mathbf{x}_0$ assigned, $\mathbf{c}$ determines the function $\mathbf{x}_0^*$ to within an arbitrary constant place, on the presumption that the function $\mathbf{c}$ is continuous. If we choose as $\mathbf{x}_0(t)$ the place occupied in $\oint$ by a certain substantial point X_0, then $\dot{\mathbf{x}}_0(t)$ is its velocity in $\oint$, and, since $\oint^*$ is a rest frame, $\dot{\chi}^*(X_0, t) = \mathbf{0}$, and so in this way we recover the conclusion with which this paragraph began.

Exercise I.10.2. Directly from (1), without use of the general concepts and framework of the earlier sections but with $\mathbf{x}_0(t)$ chosen as $\chi(X_0, t)$ for some substantial point X_0, it follows that if $\mathbf{p}_2$ and $\mathbf{p}_1$ are the position vectors with respect to $\mathbf{x}_0$ of the substantial points X_2 and X_1 in a rigid motion at time t, then in fact $\mathbf{p}_2 \cdot \mathbf{p}_1$ is constant in time.

In summary of the foregoing argument we have the following

Theorem (EULER). *Let the motion of single point X_0 of $\mathscr{B}$ be given as a differentiable function of time in $\oint$, and let $\mathbf{W}$ be a continuous function of time whose values are skew tensors. Then choice of the relative orientation $\mathbf{Q}(t_0)$ at some one time t_0 determines a unique rest frame and hence a unique rigid motion of $\mathscr{B}$ corresponding to the spin $\mathbf{W}$. If $\mathbf{W} = \mathbf{0}$, all points of $\mathscr{B}$ move with the same velocity as does X_0. Otherwise, the only points to share the velocity of X_0 at the time t are those lying on the single line through the place $\mathbf{x}_0(t)$ occupied by X_0 and parallel to the axis of $\mathbf{W}(t)$.*

In rough terms, EULER's theorem states that a rigid motion of $\mathscr{B}$ is composed instantaneously of a translation of $\mathscr{B}$ with the velocity of any one of its points and a rotation of $\mathscr{B}$ about a certain, generally time-dependent, axis through that point.

As we shall see below, the rotational momentum of a body undergoing rigid motion has an especially simple form in terms of the *Euler tensor* $\mathbf{E}_{\mathbf{x}_0}$ with respect to $\mathbf{x}_0$, defined as follows:

$$\mathbf{E}_{\mathbf{x}_0} := \int_{\mathscr{B}} \mathbf{p} \otimes \mathbf{p} \, dM, \tag{I.10-4}$$

$\mathbf{p}$ being the position vector field (I.8-26). $\mathbf{E}_{\mathbf{x}_0}$ is symmetric. If, as we shall assume now, the subbodies of $\mathscr{B}$ that have positive mass are not presently confined to a single plane, $\mathbf{E}_{\mathbf{x}_0}$ is positive. Then $\mathbf{E}_{\mathbf{x}_0}$ has at least one orthonormal triad of proper vectors, the directions of which are called the *principal axes of inertia* of $\mathscr{B}$ with respect to $\mathbf{x}_0$ in χ, and the proper numbers E_k corresponding with them are positive. This statement is ***Segner's Theorem***. The sum of the three latent roots E_k, namely $\operatorname{tr} \mathbf{E}_{\mathbf{x}_0}$, is the *polar moment of inertia of $\mathscr{B}$ about* $\mathbf{x}_0$, and $\operatorname{tr} \mathbf{E}_{\mathbf{x}_0} - E_k$, which is positive, is the *moment of inertia*[1] about the k^{th} principal axis through $\mathbf{x}_0$.

If the motion is rigid, the position vector $\mathbf{p}^*(X, t)$ of a substantial point X does not change in a rest frame $\oint^*$. The corresponding tensor $\mathbf{E}^*_{\mathbf{x}_0^*}$ is then constant in time. It is determined once and for all by the mass function and by the shape of $\mathscr{B}$.

Exercise I.10.3. If $\mathbf{S}$ is a tensor function of t only,

$$\int_{\mathscr{B}} \mathbf{p} \wedge \mathbf{S}\mathbf{p} \, dM = \mathbf{E}_{\mathbf{x}_0}\mathbf{S}^{\mathsf{T}} - \mathbf{S}\mathbf{E}_{\mathbf{x}_0}. \tag{I.10-5}$$

We consider first the case in which $\dot{\mathbf{x}}_0 = \dot{\mathbf{x}}_0^* = \mathbf{0}$. Then $\mathbf{c} = \mathbf{0}$ by (3), and so substitution of $(1)_2$ into (I.8-2), followed by use of (5), yields the following

Theorem (EULER). *Let a body undergo a rigid motion such that in $\oint$ one of its substantial points remains at rest at the place* $\mathbf{x}_0$. *Then in $\oint$*

$$\mathbf{M}_{\mathbf{x}_0} = -\mathbf{E}_{\mathbf{x}_0}\mathbf{W} - \mathbf{W}\mathbf{E}_{\mathbf{x}_0}. \tag{I.10-6}$$

Equivalently,

$$\mathbf{Q}\mathbf{M}_{\mathbf{x}_0}\mathbf{Q}^{\mathsf{T}} = \mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A} + \mathbf{A}\mathbf{E}^*_{\mathbf{x}_0^*}, \tag{I.10-7}$$

$\mathbf{E}^*_{\mathbf{x}_0^*}$ *being calculated in a rest frame.*

[1] Traditionally the tensor $(\operatorname{tr} \mathbf{E}_{\mathbf{x}_0})\mathbf{1} - \mathbf{E}_{\mathbf{x}_0}$ is called the *tensor of inertia*, and Segner's Theorem is expressed in terms of it. In the notation used in (9), its determinant is $\mathrm{I} \cdot \mathrm{II} - \mathrm{III}$.

The theorem at the end of Section App. IIA.12 shows that if $\mathbf{M}_{\mathbf{x}_0}$ is given, a unique $\mathbf{W}$ is determined by (6). Thus for all rigid motions of a given body *the spin and the rotational momentum determine each other uniquely.*

Exercise I.10.4. If both $\mathbf{x}_0$ and $\mathbf{x}_0^*$ are functions of t, in a rigid motion of a body whose mass is M

$$\begin{aligned} \mathbf{M}_{\mathbf{x}_0} &= M\bar{\mathbf{p}} \wedge \mathbf{c} - \mathbf{E}_{\mathbf{x}_0}\mathbf{W} - \mathbf{W}\mathbf{E}_{\mathbf{x}_0}\,, \\ \mathbf{Q}\mathbf{M}_{\mathbf{x}_0}\mathbf{Q}^{\mathsf{T}} &= M\bar{\mathbf{p}}^* \wedge (\mathbf{Q}\dot{\mathbf{x}}_0 - \dot{\mathbf{x}}_0^*) + \mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A} + \mathbf{A}\mathbf{E}^*_{\mathbf{x}_0^*}\,, \end{aligned} \tag{I.10-8}$$

$\bar{\mathbf{p}}$ and $\bar{\mathbf{p}}^*$ being the position vectors of the center of mass with respect to $\mathbf{x}_0$ in $\oint$ and with respect to $\mathbf{x}_0^*$ in $\oint^*$, respectively. The solution of $(8)_1$ for $\mathbf{W}$, with the subscript $\mathbf{x}_0^*$ understood, is

$$\begin{aligned} \mathbf{W} = \frac{1}{I \cdot II - III}\{&[I^2 - II][M\bar{\mathbf{p}} \wedge \mathbf{c} - \mathbf{M}] \\ &- [\mathbf{E}^2_{\mathbf{x}_0}(M\bar{\mathbf{p}} \wedge \mathbf{c} - \mathbf{M}) + (M\bar{\mathbf{p}} \wedge \mathbf{c} - \mathbf{M})\mathbf{E}^2_{\mathbf{x}_0}]\}, \end{aligned} \tag{I.10-9}$$

in which I, II, and III are the principal invariants of $\mathbf{E}_{\mathbf{x}_0}$ (Section App. IIA.10).

EULER's theorem presupposes that no substantial-point remains at rest in $\oint$. Whether or not such a restriction is imposed, we can always choose $\mathbf{x}_0^*$ as the center of mass of $\mathscr{B}$ in a rest frame. Then, in general, $\dot{\mathbf{x}}_0 \neq \mathbf{0}$, but of course $\dot{\mathbf{x}}_0^* = \mathbf{0}$, $\bar{\mathbf{p}}^* = \mathbf{0}$, and (I.9-7) shows that $\bar{\mathbf{p}} = \mathbf{0}$, and so again (6) and (7) follow. The second of these conclusions is important because $\mathbf{E}^*_{\mathbf{x}_0^*}$, being calculated in the rest frame, does not change in time.

Now regarding (6), we consider a vector $\mathbf{e}$ that lies upon a principal axis of inertia in $\oint$. Then $\mathbf{E}_{\mathbf{x}_0}\mathbf{e} = E_k\mathbf{e}$, E_k being the proper number corresponding with the principal axis upon which $\mathbf{e}$ lies, and hence $(\mathbf{E}_{\mathbf{x}_0}\mathbf{W} + \mathbf{W}\mathbf{E}_{\mathbf{x}_0})\mathbf{e} = (\mathbf{E}_{\mathbf{x}_0} + E_k\mathbf{1})\mathbf{W}\mathbf{e}$. Since $\mathbf{E}_{\mathbf{x}_0} + E_k\mathbf{1}$ is positive, in order that $(\mathbf{E}_{\mathbf{x}_0} + E_k\mathbf{1})\mathbf{W}\mathbf{e} = \mathbf{0}$ it is necessary and sufficient that $\mathbf{W}\mathbf{e} = \mathbf{0}$. Similar reasoning may be applied to $\mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A} + \mathbf{A}\mathbf{E}^*_{\mathbf{x}_0^*}$.

Thus we have the following generalization of EULER's theorem.

Theorem. *Let a body $\mathscr{B}$ undergo a rigid motion, and let $\mathbf{x}_0(t)$ be either the place occupied in $\oint$ at the time t by the center of mass of $\mathscr{B}$, or the place occupied in $\oint$ by some substantial point of $\mathscr{B}$ that remains at rest in $\oint$. Then any two of the following three properties of a line imply the third*:

1. *It is a principal axis of inertia of $\mathscr{B}$ at $\mathbf{x}_0(t)$.*
2. *It is the axis of spin.*
3. *It is the axis of rotational momentum with respect to $\mathbf{x}_0(t)$.*

Exercise I.10.5.

$$\dot{\mathbf{W}} = -\mathbf{Q}^{\mathsf{T}}\dot{\mathbf{A}}\mathbf{Q}, \qquad \mathbf{W}^2 = \mathbf{Q}^{\mathsf{T}}\mathbf{A}^2\mathbf{Q}. \tag{I.10-10}$$

We shall now calculate the acceleration field of a rigid motion. Supposing that $\mathcal{f}^*$ be a rest frame for that motion, we could set $\ddot{\chi}^* = \mathbf{0}$ in (I.9-21) after generalizing it so as to allow $\mathbf{x}_0$ to depend on t, but it is easier to differentiate $(1)_2$ instead. Doing so, we obtain

$$\begin{aligned}\ddot{\chi} &= \dot{\mathbf{c}} + \mathbf{W}(\mathbf{c} - \dot{\mathbf{x}}_0) + (\dot{\mathbf{W}} + \mathbf{W}^2)(\chi - \mathbf{x}_0),\\ &= \ddot{\mathbf{x}}_0 + \mathbf{Q}^{\mathsf{T}}[-\ddot{\mathbf{x}}_0^* + 2\mathbf{A}\dot{\mathbf{x}}_0^* - (\dot{\mathbf{A}} - \mathbf{A}^2)(\chi^* - \mathbf{x}_0^*)],\end{aligned} \tag{I.10-11}$$

the second step being a consequence of (10) and (3).

Again supposing first that one substantial point of $\mathscr{B}$ remain fixed at the place $\mathbf{x}_0$ in $\mathcal{f}$, we calculate the rate of change of rotational momentum with respect to that place. To this end we need only set $\dot{\mathbf{x}}_0 = \dot{\mathbf{x}}_0^* = \ddot{\mathbf{x}}_0 = \ddot{\mathbf{x}}_0^* = \mathbf{0}$ in (11), substitute the result into $(\text{I.8-5})_2$, and use (10). Thus we obtain the following

Theorem (EULER). *Let a body undergo a rigid motion such that in $\mathcal{f}$ one of its substantial points remains at rest at the place $\mathbf{x}_0$. Then in $\mathcal{f}$*

$$\begin{aligned}\dot{\mathbf{M}}_{\mathbf{x}_0} &= -\mathbf{E}_{\mathbf{x}_0}\dot{\mathbf{W}} - \dot{\mathbf{W}}\mathbf{E}_{\mathbf{x}_0} + \mathbf{E}_{\mathbf{x}_0}\mathbf{W}^2 - \mathbf{W}^2\mathbf{E}_{\mathbf{x}_0},\\ \mathbf{Q}\dot{\mathbf{M}}_{\mathbf{x}_0}\mathbf{Q}^{\mathsf{T}} &= \mathbf{E}^*_{\mathbf{x}_0^*}\dot{\mathbf{A}} + \dot{\mathbf{A}}\mathbf{E}^*_{\mathbf{x}_0^*} + \mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A}^2 - \mathbf{A}^2\mathbf{E}^*_{\mathbf{x}_0^*}.\end{aligned} \tag{I.10-12}$$

Exercise I.10.6. Differentiating (7), noting that $\mathbf{E}^*_{\mathbf{x}_0^*}$ is constant, and then using $(8)_2$ delivers (12).

An axis of rotation that is constant in time is called a *steady axis of rotation*. If $\mathbf{We} = \mathbf{0}$ and $\dot{\mathbf{e}} = \mathbf{0}$, then $\dot{\mathbf{W}}\mathbf{e} = \mathbf{0}$, and so we may apply essentially the same reasoning to $\dot{\mathbf{M}}_{\mathbf{x}_0}$ as we did to $\mathbf{M}_{\mathbf{x}_0}$ and conclude the following

Corollary (EULER). *Let a body undergo a rigid motion such that in $\mathcal{f}$ one of its substantial points remains at rest at the place $\mathbf{x}_0$. Then in $\mathcal{f}$ a steady axis of rotation is an axis of the rate of change of rotational momentum with respect to $\mathbf{x}_0$ if and only if it is a principal axis of inertia at $\mathbf{x}_0$.*

Exercise I.10.7. In rotation about a steady axis $\dot{\mathbf{W}} + \mathbf{W}^2$ has the same nullspace as does $\mathbf{W}$, namely, the axis of spin. Hence the proof of the foregoing theorem is completed. $\triangle$

In general, no substantial point will remain at rest in $\mathcal{f}$. A statement of simple form may be found even so by taking moments and position vectors with respect to the center of mass $\mathbf{x}_c$ of $\chi(\mathscr{B}, t)$, so that $\bar{\mathbf{p}} = \mathbf{0}$. If we choose $\mathbf{x}_0$ in $(8)_1$ as $\mathbf{x}_c$ and then differentiate the result with respect to t, we obtain

$$\dot{\mathbf{M}}_{\mathbf{x}_c} = -\mathbf{E}_{\mathbf{x}_c}\dot{\mathbf{W}} - \dot{\mathbf{W}}\mathbf{E}_{\mathbf{x}_c} + \mathbf{E}_{\mathbf{x}_c}\mathbf{W}^2 - \mathbf{W}^2\mathbf{E}_{\mathbf{x}_c}. \tag{I.10-13}$$

Returning to use of a fixed place $\mathbf{x}_0$, by substituting (13) into (I.8-30) we prove that

$$\dot{\mathbf{M}}_{\mathbf{x}_0} = \bar{\mathbf{p}} \wedge \dot{\mathbf{m}} - \mathbf{E}_{\mathbf{x}_c}\dot{\mathbf{W}} - \dot{\mathbf{W}}\mathbf{E}_{\mathbf{x}_c} + \mathbf{E}_{\mathbf{x}_c}\mathbf{W}^2 - \mathbf{W}^2\mathbf{E}_{\mathbf{x}_c}. \tag{I.10-14}$$

Exercise I.10.8 (König, Euler). The kinetic energy of a body $\mathscr{B}$ in rigid motion is given by

$$K = \tfrac{1}{2}M|\mathbf{c}|^2 + M\mathbf{c}\cdot\mathbf{W}\bar{\mathbf{p}} - \tfrac{1}{2}\mathbf{W}^2\cdot\mathbf{E}_{\mathbf{x}_0}. \tag{I.10-15}$$

If $\mathbf{x}_0(t)$ is taken as the place $\mathbf{x}_c(t)$ occupied by the center of mass of $\mathscr{B}$ at the time t, then the kinetic energy of $\mathscr{B}$ may be decomposed into translational and rotational parts as follows:

$$K = \tfrac{1}{2}M|\dot{\mathbf{x}}_c|^2 - \tfrac{1}{2}\mathbf{A}^2\cdot\mathbf{E}^*_{\mathbf{x}_0^*}. \tag{I.10-16}$$

The first summand is the kinetic energy of a mass-point whose mass M is that of $\mathscr{B}$ and which moves with the speed of the center of mass of $\mathscr{B}$, while the second term is the kinetic energy that would correspond to the spin and shape of $\mathscr{B}$ if the center of mass of $\mathscr{B}$ were at rest in $\mathcal{f}$.

A body insusceptible of any motions other than rigid ones is a *rigid body*. Except for a remark below in Section IV.7, this book will not treat further of rigid bodies.

11. Frame-Indifference

While the event world $\mathscr{W}$ is the seat of phenomena, we may apprehend these only through the intermediary of a frame, since we always report observations in terms of places and times. A phenomenon, of course, is independent of frame, though a description of it in one frame generally differs from a description of it in another. The same phenomenon is reported differently by different observers. Thus arises the question how to relate statements about one and the same phenomenon made in terms of different frames.

First, we may always make a statement with respect to one frame $\mathcal{f}$ and then simply translate it into a statement with respect to any other frame $\mathcal{f}^*$. We have seen an example in the case of a motion χ of a body. If χ is given with respect to $\mathscr{f}$, we *define* χ^* in $\mathscr{f}^*$ by (I.9-11), which simply reflects our interpretation of the concepts of motion and change of frame. We may do the same thing with other quantities we regard as intrinsic to the event world. The other principal example in mechanics is mass (Section I.4). We say that such quantities are *frame-indifferent*. We shall discuss the frame-indifference of mass, force, and torque in the next section.

Second, and more commonly, we shall encounter a prescription or definition that delivers a particular quantity in each frame. The prescription or definition itself is *frame-indifferent* in the sense that it is equally effective in all frames. We have seen examples already, namely, the velocity and the acceleration, which are calculated from the motion by rules that make no mention of frames and hence apply for any choice of $\mathscr{f}$. These particular rules have been stated as (I.9-12). We have then been able to express the velocity and acceleration in $\mathscr{f}^*$ in terms of their counterparts in $\mathscr{f}$, with the aid, of course, of the functions $\mathbf{Q}$, $\mathbf{x}_0$, and $\mathbf{x}_0^*$ that specify the change (I.9-5) from $\mathscr{f}$ to $\mathscr{f}^*$. The conclusions so obtained have been stated as (I.9-14) and (I.9-21). It is clear from them that the velocity and acceleration as observed by $\mathscr{f}$ and $\mathscr{f}^*$ are not simply functions defined on the event world $\mathscr{W}$ and then referred to frames, for if they were, under change of frame their values, which are vectors, would have to follow the transformation such a change induces on the translation space $\mathscr{V}$ of $\mathscr{E}$, and this transformation, as we have seen, is (I.9-7). Because in general $\dot{\chi}^* \neq \mathbf{Q}\dot{\chi}$ and $\ddot{\chi}^* \neq \mathbf{Q}\ddot{\chi}$, we say that velocity and acceleration are *not frame-indifferent*. This example makes it clear that *a frame-indifferent prescription or definition leads in general to a quantity that is not frame-indifferent.*

Of course some prescriptions, although stated in terms of a frame, do lead to quantities intrinsic to $\mathscr{W}$. Such quantities we shall call *frame-indifferent*, since in principle they could have been introduced abstractly without use of any frame. Suppose certain prescriptions deliver in $\mathscr{f}$ and $\mathscr{f}^*$ the scalar fields A and A^*, respectively. If

$$A^*(\mathbf{x}^*, t^*) = A(\mathbf{x}, t) \qquad \text{(I.11-1)}$$

when $(\mathbf{x}^*, t^*)$ is related to $(\mathbf{x}, t)$ through (I.9-5), we shall say that A and A^* represent a *frame-indifferent scalar*. Loosely, we shall refer to the value of A, which of course is a number assigned to a place and time in $\mathscr{f}$, as being itself a frame-indifferent scalar. Likewise, the vector field $\mathbf{v}$ and the tensor field $\mathbf{T}$ will be called *frame-indifferent* if for all t

$$\begin{aligned} \mathbf{v}^*(\mathbf{x}^*, t^*) &= \mathbf{Q}(t)\mathbf{v}(\mathbf{x}, t), \\ \mathbf{T}^*(\mathbf{x}^*, t^*) &= \mathbf{Q}(t)\mathbf{T}(\mathbf{x}, t)\mathbf{Q}(t)^{\mathsf{T}}, \end{aligned} \qquad \text{(I.11-2)}$$

respectively, where again $(\mathbf{x}^*, t^*)$ is related to $(\mathbf{x}, t)$ through (I.9-5), and $\mathbf{Q}(t)$ is the relative orientation of $\oint^*$ and $\oint$ at the time t. The first of these requirements asserts that $\mathbf{v}^*$ and $\mathbf{v}$ are the same "arrow" at the same event as observed in different frames. The second asserts, as we have seen in Section I.9, that $\mathbf{T}^*$ and $\mathbf{T}$ are at the same event the same linear transformations of such arrows. For details the reader may refer back to the discussion between (I.9-6) and (I.9-10), but it is even clearer and not more difficult to demonstrate (2) directly by referring to the abstract quantities that $\mathbf{v}$, $\mathbf{v}^*$, $\mathbf{T}$, $\mathbf{T}^*$, *etc.*, represent, as we shall do now.

Let T be the instant that corresponds to the times t and t^* in $\oint$ and $\oint^*$, respectively. Let $\mathscr{V}_T$ be the translation space of $\mathscr{W}_T$ (*cf.* Section I.6). The mapping $D\oint_T : \mathscr{V}_T \to \mathscr{V}$, which is the derivative of the mapping $\oint_T : \mathscr{E}_T \to \mathscr{E}$, is linear. Let $\mathbf{v}_T$ be the vector in $\mathscr{V}_T$ that becomes $\mathbf{v}$ when observed in the frame $\oint$. Then $\mathbf{v} = (D\oint_T)\mathbf{v}_T$. Similarly, $\mathbf{v}^* = (D\oint_T^*)\mathbf{v}_T$. Thence

$$\mathbf{v}^* = (D\oint_T^*)(D\oint_T)^{-1}\mathbf{v} = D(\oint_T^* \circ \oint_T^{-1})\mathbf{v} = \mathbf{Q}\mathbf{v}. \qquad \text{(I.11-2A)}$$

Exercise I.11.1. Reference to the linear transformation on $\mathscr{V}_T$ that has $\mathbf{T}^*$ and $\mathbf{T}$ as representatives in the frames in question leads to proof of $(2)_2$.

The position vector $\mathbf{p}$ of $\mathbf{x}$ with respect to $\mathbf{x}_0$, defined by (I.8-26), is obviously a frame-indifferent vector. Hence the Euler tensor $\mathbf{E}_{\mathbf{x}_0}$, defined by (I.10-4), is a frame-indifferent tensor, and the principal moments of inertia are frame-indifferent scalars.

Most of the fields we encounter in mechanics are not frame-indifferent. The examples of velocity and acceleration suggest, nevertheless, that if we restrict attention to a subgroup $\mathscr{g}$ of changes of frame from a particular $\oint$, we may obtain conclusions of the forms (1) or (2). In such a case we may say that a particular scalar, vector, or tensor is *frame-indifferent in* $\mathscr{g}$ from the particular $\oint$. For example, from $(\text{I.9-21})_2$ we see that $\ddot{\mathbf{x}}^* = \mathbf{Q}\ddot{\mathbf{x}}$ for all motions if and only if $\ddot{\mathbf{x}}_0^* = \mathbf{0}$ and $\mathbf{A} = \mathbf{0}$, so that $\dot{\mathbf{x}}_0^* =$ const. and $\mathbf{Q} =$ const. This subgroup of changes of frame from $\oint$, consisting in those under which the acceleration is frame-indifferent, is called the group of *galilean transformations* of $\oint$. These transformations interconvert the frames of observers moving at uniform velocities with respect to one another and with no change of relative orientation in time.

Any quantity that is frame-indifferent under galilean transformations of $\oint$ is called a *galilean invariant* of $\oint$.

The term "galilean" is merely traditional and should not be regarded as an attribution to GALILEO.

Exercise I.11.2. The rate of change $\dot{\mathbf{m}}$ of the linear momentum of a body $\mathscr{B}$ in a motion χ in $\oint$ is frame-indifferent in the subgroup of galilean transformations of $\oint$.

The set of all frames obtainable from $\oint$ by galilean transformations is the *galilean class* of $\oint$.

Specifically, from (I.9-14) we see that $\dot{\chi}^* = \mathbf{Q}\dot{\chi}$ for all motions if and only if $\dot{\mathbf{x}}_0^* = \mathbf{0}$ and $\mathbf{A} = \mathbf{0}$. Thus $\mathbf{x}_0$ and $\mathbf{Q}$ are constants. This subgroup of the galilean group of changes of frame from $\oint$, consisting in those under which the velocity is frame-indifferent, is called the group of *constant rigid transformations* of $\oint$. These transformations interconvert the frames of observers at rest with respect to one another. The class of all frames obtainable from $\oint$ by constant rigid transformations is the *constant rigid class* of $\oint$ (*cf.* Section I.6). In Section I.10 a rigid motion was defined as one whose velocity field vanishes in some $\oint^*$. We now see that the velocity field of a rigid motion is independent of place in all frames belonging to the constant rigid class of $\oint^*$, and only in such frames. In particular, all rest frames for a rigid motion are obtained from any given one by changes of frame in which $\mathbf{x}_0^* = \text{const.}$, $\mathbf{Q} = \text{const.}$ As is plain from the concept of rigid motions, these frames may be obtained from one another by time-independent translations and rotations. These also constitute a subgroup, the *rest class* of the given rigid motion.

Exercise I.11.3. The gradient of a frame-indifferent scalar field is a frame-indifferent vector field; the proper numbers, trace, and determinant of a frame-indifferent tensor field are frame-indifferent scalar fields; the proper vectors of such a tensor are frame-indifferent vector fields; the scalar product of two frame-indifferent vector fields is a frame-indifferent scalar field; and the tensor product and exterior product of frame-indifferent vector fields are frame-indifferent tensor fields.

Exercise I.11.4. An oriented unit normal field to a surface is a frame-indifferent vector field.

At the beginning of this section we remarked that prescription of a quantity in one particular $\oint$ can always be extended trivially to form the definition of a corresponding frame-indifferent quantity. So as to illustrate this fact, we now consider the acceleration $\ddot{\chi}$ of some substantial point of $\mathscr{B}$ in $\oint$. If

$$\boldsymbol{\alpha}_{\oint} := \ddot{\chi}^* - \ddot{\mathbf{x}}_0^* - 2\mathbf{A}(\dot{\chi}^* - \dot{\mathbf{x}}_0^*) - (\dot{\mathbf{A}} - \mathbf{A}^2)(\chi^* - \mathbf{x}_0^*), \qquad \text{(I.11-3)}$$

$\ddot{\chi}^*$ and $\dot{\chi}^*$ being the acceleration and the velocity in $\oint^*$, and $\mathbf{A}$ being the spin of $\oint$ with respect to $\oint^*$, then by (I.9-21) we recognize $\boldsymbol{\alpha}_{\oint}$ as being that frame-indifferent vector field over $\mathscr{B}$ which in $\oint$ is the acceleration field of $\mathscr{B}$. Of course, to within multiplication by a constant, orthogonal tensor, it is the acceleration field in all frames in the galilean class of $\mathfrak{f}$. The frame-indifferent vector field $\boldsymbol{\alpha}_{\mathfrak{f}}$ is of central importance in dynamics.

12. Axioms of Mechanics

Mechanics relates the motions of bodies to the masses assigned to them and the forces that act on them. Bodies are encountered only in their shapes. Masses and forces, therefore, can be correlated with experience in nature only when they are assigned to the shapes of bodies. Indeed, the value of the mass of a body is a real number, and we may simply associate that number to all shapes of that body: *Mass is frame-indifferent.* We may state this fact formally as

Axiom A1.

$$M^* = M, \tag{I.12-1}$$

the notation being that used in Section I.11.

We dignify A1 by the title "axiom" since such it would have to be, had we chosen to describe everything in terms of frames from the start.

Since a force is a vector in $\mathscr{V}$, the translation space of $\mathscr{E}$, assignment of forces presumes that a frame has already been assigned. If forces are to have primary meaning, the transport of them to the shapes of bodies must be independent of the observer. The forces acting upon the shapes of $\mathscr{B}$ in $\mathfrak{f}$ and $\mathfrak{f}^*$ at the times t and t^* should be related by the transformation the change of frame from $\mathfrak{f}$ to $\mathfrak{f}^*$ induces in $\mathscr{V}$. In other words, *we require that all forces be frame-indifferent.* Formally, we lay down

Axiom A2.

$$\mathbf{f}^* = \mathbf{Q}\mathbf{f}, \tag{I.12-2}$$

the notation being again that of Section I.11.

Here and for the rest of this section the time is not indicated in the notation.

Axiom A1 is part of the assumption commonly called "the principle of conservation of mass"; the other part, which asserts that the mass of a body is the same in all shapes of the body, is implied by our Axiom M1 in Section I.4, according to which mass is assigned to bodies with no mention of any shapes they may assume. Axiom A2, until recently, was left to be inferred from the context and hence was not given a name.

Without exception, the various traditional ways of presenting the foundations of mechanics leave the concept of force in the shadows of intuition. Some even foster the illusion that force is a derived concept, the existence of which follows from some mysterious legerdemain with potential functions and variational principles and magical δs. Assumptions must be made about forces in these treatments, since nothing comes from nothing, but the assumptions are tacit, even struthious. Modern fundamental thought in mechanics has reverted to the viewpoint of NEWTON and EULER: Forces are basic, *a priori* concepts in mechanics. While NEWTON and EULER left forces, as they did many other things, largely unformalized, today we apply to mechanics the requirement of HILBERT, now universally accepted in the rest of mathematics: An object which enters a mathematical structure must be described by explicit, formal axioms specifying mathematical properties which make it possible to set and solve mathematical problems. If one such axiomatic basis suffices, so do infinitely many others. The one we adopt in this book is close to the ideas used informally and successfully by engineers for over a century.

Axioms A1 and A2 require that mass and force as observed in ϕ and in ϕ^* be assigned the same units, just as the change of frame (I.9-5) leaves the units of length and time unchanged. Of course, a fully general formulation, while allowing it to be *possible* that different observers use the same units, *i.e.*, to choose the same metrics in $\mathscr{E}$, $\mathscr{R}$, and $\mathscr{V}$, would not require them to do so. The generality so obtained is merely apparent and is not worth the complication it introduces into the mathematics at this level. It may be achieved, if desired, by simply allowing free change of units afterward in all frames, once the requirements of frame-indifference shall have been satisfied, if they can be, by one choice of units.

Exercise I.12.1. Axiom A2 implies that $\int_{\mathscr{A}}(\chi - \mathbf{x}_0) \otimes d\mathbf{f}_{\mathscr{B}^e}$ is frame-indifferent. Hence, in particular, the resultant torque is frame-indifferent.

In a more general system of mechanics allowing for couples as well as moments of forces, an additional axiom is needed: The torques are frame-indifferent.

In Section I.5 we have remarked that a linear combination $A\mathbf{f}_1 + B\mathbf{f}_2$ of two systems of forces $\mathbf{f}_1$ and $\mathbf{f}_2$ is a system of forces. If, as is natural, we require the scalar coefficients A and B to be frame-indifferent, then Axiom A2 is satisfied also by $A\mathbf{f}_1 + B\mathbf{f}_2$. Thus, even after the imposition of Axiom A2, a linear combination of two systems of forces is a system of forces. Conversely, if $\mathbf{f}$ and $\mathbf{g}$ are systems of forces, the trivial decomposition $\mathbf{f} = \mathbf{g} + (\mathbf{f} - \mathbf{g})$ allows us to regard $\mathbf{f}$ as the sum of $\mathbf{g}$ and another system of forces. To justify this decomposition, we cannot take for $\mathbf{g}$ simply any function that satisfies the axioms of forces listed in Section I.5. Rather, we must be sure that $\mathbf{g}$ is frame-indifferent, since Axiom A2 requires that *all forces be frame-indifferent*.

We are now in a position to impose requirements relating forces to motions, or, in looser terms, to state the effects of forces in producing motions. Specifically, we lay down

NOLL'S **Axiom.** *For every assignment of forces to bodies, the working of a system of forces acting on each body is frame-indifferent, no matter what be the motion.*

Formally, in the notations (I.8-7) and (I.11-1),

Axiom A3.

$$W^* = W \qquad \forall \mathscr{B} \in \bar{\mathbf{\Omega}}, \quad \forall \chi. \tag{I.12-3}$$

On the assumption that A2 is satisfied, we can demonstrate that (3) expresses a necessary and sufficient condition for the resultant force and torque on each body $\mathscr{B}$ to vanish. Indeed, by applying (I.9-13) to the definition (I.8-7) we see that, for given $\mathscr{B}$ and χ,

$$\begin{aligned} W^* - W &= \int_{\mathscr{B}} (\dot{\chi}^* \cdot d\mathbf{f}^*_{\mathscr{B}^e} - \dot{\chi} \cdot d\mathbf{f}_{\mathscr{B}^e}) \\ &= \int_{\mathscr{B}} [\dot{\mathbf{x}}_0^* + \dot{\mathbf{Q}}(\chi - \mathbf{x}_0) + \mathbf{Q}\dot{\chi}] \cdot \mathbf{Q}\, d\mathbf{f}_{\mathscr{B}^e} - \int_{\mathscr{B}} \dot{\chi} \cdot d\mathbf{f}_{\mathscr{B}^e} \\ &= \mathbf{Q}^{\mathsf{T}} \dot{\mathbf{x}}_0^* \cdot \int_{\mathscr{B}} d\mathbf{f}_{\mathscr{B}^e} - \mathbf{Q}^{\mathsf{T}} \dot{\mathbf{Q}} \cdot \int_{\mathscr{B}} (\chi - \mathbf{x}_0) \otimes d\mathbf{f}_{\mathscr{B}^e} \\ &= \mathbf{Q}^{\mathsf{T}} \dot{\mathbf{x}}_0^* \cdot \mathbf{f}(\mathscr{B}, \mathscr{B}^e) - \tfrac{1}{2} \mathbf{Q}^{\mathsf{T}} \dot{\mathbf{Q}} \cdot \mathbf{F}(\mathscr{B}; \mathscr{B}^e)_{\mathbf{x}_0}. \end{aligned} \tag{I.12-4}$$

By Axiom A3 the right-hand side of this equation must vanish for all choices of the functions $\mathbf{Q}$ and $\dot{\mathbf{x}}_0^*$. We consider a particular time t and choose $\mathbf{Q}$ such that $\dot{\mathbf{Q}}(t) = \mathbf{0}$. Since $\mathbf{Q}(t)^{\mathsf{T}} \dot{\mathbf{x}}_0^*(t)$ may be any vector whatever, Axiom A3 requires that

$$\mathbf{f}(\mathscr{B}, \mathscr{B}^e) = \mathbf{0}. \tag{I.12-5}$$

This being so, Axiom A3 again applied to (4) shows that in the space of skew tensors $\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0}$ must be perpendicular to every tensor of the form $\mathbf{Q}(t)^{\mathsf{T}} \dot{\mathbf{Q}}(t)$, the values of $\mathbf{Q}(t)^{\mathsf{T}}$ being orthogonal tensors. If $\mathbf{W}$ is a constant skew tensor, and if $\mathbf{Q}(t) := \mathbf{e}^{(t-t_0)\mathbf{W}}$, then $\mathbf{Q}(t_0) = \mathbf{1}$ and $\dot{\mathbf{Q}}(t_0) = \mathbf{W}$, and so $\mathbf{Q}(t_0)^{\mathsf{T}} \dot{\mathbf{Q}}(t_0) = \mathbf{W}$. Thus the skew tensor $\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0}$ must be perpendicular to every skew tensor. Therefore

$$\mathbf{F}(\mathscr{B}, \mathscr{B}^e)_{\mathbf{x}_0} = \mathbf{0}. \tag{I.12-6}$$

Conversely, (5) and (6) suffice for the truth of Axiom A3, it being presumed always that Axiom A2 holds. Thus we have established the following

Theorem (NOLL). *The working of a system of forces is frame-indifferent if and only if that system and its associated system of torques are both balanced.*

Exercise I.12.2 Axiom A3 $\Rightarrow$ Axiom A2

Examination of (4) shows that to prove the necessity of (5) and (6) we need not assume that $W^* = W$ for all orthogonal $\mathbf{Q}$ but only for proper rotations that are affine functions of t.

The reader accustomed to the usual treatments of mechanics needs to be reminded that here forces of all kinds are included. The common and useful separation of forces into "applied" forces and "inertial" forces will be made in the succeeding section.

As a consequence of NOLL's theorem here, NOLL's corollary in Section I.5, and the counterpart for torques mentioned in Section I.8, Axioms A2 and A3 imply the

Corollary (Principle of Action and Reaction). *For each pair of separate bodies $\mathscr{B}$ and $\mathscr{C}$*

$$\begin{aligned} \mathbf{f}(\mathscr{B}, \mathscr{C}) &= -\mathbf{f}(\mathscr{C}, \mathscr{B}), \\ \mathbf{F}(\mathscr{B}, \mathscr{C})_{\mathbf{x}_0} &= -\mathbf{F}(\mathscr{C}, \mathscr{B})_{\mathbf{x}_0}. \end{aligned} \tag{I.12-7}$$

While, as we have seen in Section I.8, the special assumptions of analytical dynamics, once the system of forces is assumed balanced, reduce the balance of torques to the hypothesis of central forces, in more general and typical universes of mechanics the balance of torques is independent of the balance of forces. The proof of NOLL's theorem makes it clear that the balance of forces expresses the invariance of the working under translations, while the balance of torques expresses the invariance of the working under rotations. Since rotations and translations may be chosen independently in a change of frame, no relation between the two principles can be expected except in degenerate cases.

We have made the existence of a rest frame $\oint^*$ the definition of a rigid motion (Section I.10). In a rest frame, directly from the definition (I.8-7) we see that $W^* = 0$. By Axiom A3, therefore, $W = 0$ in any frame $\oint$: *The working of any system of forces vanishes in a rigid motion.* This is the

work theorem of the dynamics of rigid motions. Thus for a rigid motion the value of the quantity whose frame-indifference NOLL's Axiom asserts is in fact 0.

Any motion of a single mass-point is rigid. Thus we may obtain again the trivial conclusion (I.8-13). Our earlier proof assumed the system of forces to be balanced, which NOLL's theorem ensures.

Work theorems similar to that just stated hold in some other special branches of mechanics also, but by no means in all of them. For example, if in the analytical dynamics of a system of three or more bodies, we consider the body consisting in X_1 and X_2, for it the working does not generally vanish. Likewise, if $\mathscr{B}$ is the join of two parts, each of which is in rigid motion, W does not generally vanish unless both parts have the same spin.

13. The Axioms of Inertia. Euler's Laws of Motion

Thus far we have considered an armature on which models of mechanical occurrences may be constructed: all the massy bodies in the universe, set in motion through the entire event world. By its nature, human experience can never use with profit, let alone test the worth of so embracing a picture, for human experience is limited to a portion of the event world and to those bodies which have occupied that portion within a limited period of time. This subset of the universe may be a small one, this interval of time a short one; at most, the former represents all bodies whose existence has so far been seen or inferred by man, and the latter, the total length of time through which human experience is known to have existed or can be shrewdly extrapolated. Whatever be the limitation chosen, some limitation there must be, for otherwise we could not isolate a class of putative phenomena from all the rest so as to form models for experiments or for the future course of nature.

On the other hand, we cannot simply disregard the existence of all bodies but those in the subcollection or *great system* Σ in Ω that we choose to isolate for attention, since such further bodies as may exist will generally exert forces upon those we do consider. The idea of "isolation" requires merely that the forces among members of the excluded set of bodies, and the consequent motions of those bodies, need not be known. If $\mathscr{B} \in \Sigma$, and if we denote by Σ^e the join of all the bodies exterior to Σ, then we consider $\mathbf{f}(\mathscr{B}, \Sigma^e)$ and disregard whatever forces the parts of Σ^e may exert upon each other and whatever motions those parts may undergo.

We might also limit the event world and the space of instants, but in classical mechanics it is not usual to do so explicitly.

In classical mechanics in its most general form, the great system Σ is characterized by two ***axioms of inertia***.

Axiom I1. *There is a frame such that if* $\mathbf{m}(\mathscr{B}, \chi)$ *is constant over an open interval of time, then in that interval* $\mathbf{f}(\mathscr{B}, \Sigma^{e}) = \mathbf{0}$, *and conversely.* Equivalently, by (I.8-29), *there is a frame such that the center of mass* $\bar{\mathbf{p}}$ *of* $\mathscr{B}$ *moves along a straight line at uniform speed in that frame if and only if* Σ^{e} *exerts no force on* $\mathscr{B}$.

The frame whose existence Axiom I1 posits is called an *inertial frame.*

The First Axiom of Inertia, while it asserts the existence of a particular frame, is itself a frame-indifferent statement in that the condition it lays down restricts but does not depend upon the assignment of a frame to the event world. Moreover, it does not depend upon what system of forces is being considered. Axiom A2 asserts that all forces are frame-indifferent. Therefore, no matter what be the function $\mathbf{f}$, so long as it satisfies the axioms imposed on systems of forces, the force exerted by Σ^{e} on $\mathscr{B}$ vanishes in one frame if and only if it vanishes in all frames.

We can express the First Axiom of Inertia in another way. The exterior $\mathscr{B}^{e}$ of $\mathscr{B}$ may be decomposed into two separate parts: Σ^{e} and the join $\mathscr{B}^{e}_{\Sigma}$ of all bodies of Σ separate from $\mathscr{B}$:

$$\mathscr{B}^{e} := \mathscr{B}^{e}_{\Sigma} \vee \Sigma^{e}. \tag{I.13-1}$$

If $\mathscr{A} \prec \mathscr{B}$, by Axiom F3 in Section I.5

$$\begin{aligned} \mathbf{f}(\mathscr{A}, \mathscr{B}^{e}) &= \mathbf{f}(\mathscr{A}, \mathscr{B}^{e}_{\Sigma}) + \mathbf{f}(\mathscr{A}, \Sigma^{e}), \\ d\mathbf{f}_{\mathscr{B}^{e}} &= d\mathbf{f}_{\mathscr{B}^{e}_{\Sigma}} + d\mathbf{f}_{\Sigma^{e}}\,; \end{aligned} \tag{I.13-2}$$

the second equation refers to the vector-valued measures provided by Axiom F4 in Section I.5. By (I.8-8) we have a similar decomposition of the torque with respect to $\mathbf{x}_0$:

$$\begin{aligned} \mathbf{F}(\mathscr{A}, \mathscr{B}^{e})_{\mathbf{x}_0} &= \mathbf{F}(\mathscr{A}, \mathscr{B}^{e}_{\Sigma})_{\mathbf{x}_0} + \mathbf{F}(\mathscr{A}, \Sigma^{e})_{\mathbf{x}_0}\,, \\ (d\mathbf{F}_{\mathscr{B}^{e}})_{\mathbf{x}_0} &= (d\mathbf{F}_{\mathscr{B}^{e}_{\Sigma}})_{\mathbf{x}_0} + (d\mathbf{F}_{\Sigma^{e}})_{\mathbf{x}_0}\,. \end{aligned} \tag{I.13-3}$$

In particular, the resultant force and resultant torque have such decompositions:

$$\begin{aligned} \mathbf{f}(\mathscr{B}, \mathscr{B}^{e}) &= \mathbf{f}(\mathscr{B}, \mathscr{B}^{e}_{\Sigma}) + \mathbf{f}(\mathscr{B}, \Sigma^{e}), \\ \mathbf{F}(\mathscr{B}, \mathscr{B}^{e})_{\mathbf{x}_0} &= \mathbf{F}(\mathscr{B}, \mathscr{B}^{e}_{\Sigma})_{\mathbf{x}_0} + \mathbf{F}(\mathscr{B}, \Sigma^{e})_{\mathbf{x}_0}\,. \end{aligned} \tag{I.13-4}$$

In each statement the first term on the right-hand side, since it depends only on the bodies within the great system Σ, is accessible in principle to observation and measurement. We call these terms the *applied force* on $\mathscr{B}$ and the *applied torque* on $\mathscr{B}$, respectively, and we denote the corresponding functions of $\mathscr{B}$ by $\mathbf{f}^a$ and $\mathbf{F}^a_{\mathbf{x}_0}$:

$$\mathbf{f}^a(\mathscr{B}) := \mathbf{f}(\mathscr{B}, \mathscr{B}^e_\Sigma), \qquad \mathbf{F}^a(\mathscr{B})_{\mathbf{x}_0} := \mathbf{F}(\mathscr{B}, \mathscr{B}^e_\Sigma)_{\mathbf{x}_0} . \tag{I.13-5}$$

While Axiom I1 may be imposed independently of the general axioms of mechanics laid down in Section I.12, we shall of course wish to adopt those axioms also. Then in virtue of NOLL's theorem in Section I.12 the left-hand sides of $(4)_1$ and $(4)_2$ vanish, and so (4) becomes

$$\begin{aligned} \mathbf{f}^a(\mathscr{B}) &= -\mathbf{f}(\mathscr{B}, \Sigma^e), \\ \mathbf{F}^a(\mathscr{B})_{\mathbf{x}_0} &= -\mathbf{F}(\mathscr{B}, \Sigma^e)_{\mathbf{x}_0} . \end{aligned} \tag{I.13-6}$$

Accordingly, we may express Axiom I1 in the following, equivalent forms, provided we grant Axioms A1–A3 in Section I.12:

1. *There is a frame in which the linear momentum of $\mathscr{B}$ is constant if and only if no applied force acts on $\mathscr{B}$.*
2. *There is a frame in which the center of mass of $\mathscr{B}$ moves along a straight line at uniform speed if and only if no applied force acts on $\mathscr{B}$.*

NEWTON set forth in 1687 three Laws of Motion. The first of these was, "Every body perseveres in its state of rest or of uniform motion straight ahead, unless it be compelled to change that state by forces impressed upon it." In the generality maintained in modern mechanics, this axiom is not always valid, for a body may be subject to internal or external constraints not expressed in terms of a system of forces. For example, a rigid body subject to no applied force spins about some axis through its center of mass; its parts, which also are bodies, move in such a way that their centers of mass describe circles about that axis. NEWTON himself did not specify any mathematical properties of bodies or forces, and so his intentions must be inferred by the reader, and in the course of time different readers have read different meanings into his words. Our Axiom I1 may be regarded as including one interpretation of NEWTON's First Law.

JOOS[1] wrote the following physical justification of inertial frames in classical physics and their employment to construct NEWTON's absolute time.

[1]See Section X.2 of JOOS's book, cited above in the footnote on p. 31.

We start with the empirical fact that there exist reference frames for which computations based on NEWTON's Second Law are in complete accord with experiment. Consider, for example, celestial mechanics, which employs a stationary reference frame at the centre of gravity of the solar system. We do not now say with Newton that this frame is at rest in absolute space, or that it is in uniform rectilinear motion with respect to it . . .; we content ourselves merely with giving such a frame in which Newton's laws are valid a name. Because of the validity of the Law of Inertia [NEWTON's First Law], we call such a system an *inertial frame*. Evidently the confirmation of our calculations depends also upon a reasonable measurement of time. How may we obtain a criterion as to whether our frame of reference, together with our clock, represents an inertial system of space and time? For this purpose we perform, at least in thought, the simplest of mechanical experiments—rectilinear motion of a particle subject to no forces. If we divide the line of motion into equal segments, we can take the time between the passing of two successive marks as the unit of time. However, *one* direction is not sufficient. If our measured path were accelerated with respect to a true inertial frame, we would obtain a non-uniform clock which would give impossible results for other experiments. It is readily seen that the necessary and sufficient condition for an inertial frame is that *three* particles projected in non-coplanar directions describe straight paths. Then, by dividing the path of any one of the particles into equal intervals, we can obtain an inertial measure of the time.

In Section I.11 a galilean class was defined and seen to be the set of all frames obtainable from a given one by galilean transformations. If the acceleration of a certain substantial point vanishes in one frame, it vanishes in all frames belonging to the same galilean class. In view of (I.8-5), then, the linear momentum of a body is constant in one frame if and only if it is constant in all frames of the same galilean class. Accordingly, Axiom I1 requires that if for all $\mathscr{B}$ the system of forces $\mathbf{f}$ be such that $\mathbf{f}(\mathscr{B}, \Sigma^e) = \mathbf{0}$ in $\oint$, then $\mathbf{f}^*(\mathscr{B}, \Sigma^e) = \mathbf{0}$ in every $\oint^*$ belonging to the galilean class containing $\oint$. Thus, finally, *the galilean class of an inertial frame is the set of all inertial frames* for a given Σ.

By Axiom I1 alone, the inertial frames that pertain to two different great systems Σ need not belong to the same galilean class. It is customary, nonetheless, to assume that there is but one single galilean class of inertial frames for all great systems. This galilean class of inertial frames defines an affine structure on the event world $\mathscr{W}$. The world-lines of body-points of a body at rest in an inertial frame are parallel straight lines in the affine space-time $\mathscr{W}$.

Instead of introducing the affine structure on $\mathscr{W}$ here, we might postulate it even before talking about frames.[1] Indeed we should have done so, had logical efficiency been our only guide. The somewhat long path that we have taken has its merits nevertheless. First, it shows clearly that the mathematical structure on $\mathscr{W}$ is nothing sacred or divine:

[1] An example of such an approach is furnished by T. MATOLCSI, "On material frame-indifference," *Archive for Rational Mechanics and Analysis* **91** (1985/86): 99–118.

that structure describes only the physics that we mortals comprehend. The mathematical structures that reflect the requirements of kinematics, dynamics, and frame-indifference of material response are different. Secondly, our approach above is in fact more flexible. For instance, we may elect, following CARTAN, to incorporate Newtonian gravitation into our theory by allowing $\mathscr{W}$ to have curvature.[1] To this end we should have to modify the Axioms of Inertia I1 and I2, but we might still leave the discussions in Sections I.6, I.9 and I.11 intact. Had we postulated $\mathscr{W}$ to be affine at the outset, CARTAN's approach to Newtonian gravitation would seem inconceivable.

According to astronomers, certain of the most distant stars seem to be nearly at rest with respect to one another. It is customary to interpret the class of inertial frames in the theory as being those that are obtained by uniform translation of one in which those "fixed stars" are stationary. The theory itself, however, merely assumes that there are inertial frames and does not enter into the question of how they should be interpreted in nature.

In Section I.9 we saw that two general rigid frames need not be compatible in terms of the differentiable structures they define on $\mathscr{W}$. In view of the affine structure on $\mathscr{W}$, we shall henceforth consider only those rigid frames that are compatible with the differentiable structure defined by the galilean class of inertial frames.

Once a frame satisfying Axiom I1 is given, we may ask what forces are exerted upon a body $\mathscr{B}$ experiencing general motion with respect to it. These forces are restricted by the following conditions:

1. Since $\mathbf{f}$ is a function of pairs of separate bodies, $\mathbf{f}(\mathscr{B}, \Sigma^e)$ should depend upon the motions of bodies at most through the motion of $\mathscr{B}$ and the motion of Σ^e.
2. Since we know nothing about the nature of Σ^e or its motion, $\mathbf{f}(\mathscr{B}, \Sigma^e)$ should depend upon $\mathscr{B}$ and its motion alone.
3. For consistency with Axiom I1, $\mathbf{f}(\mathscr{B}, \Sigma^e)$ should vanish if $\mathbf{m}(\mathscr{B}, \chi) =$ const.

Classical mechanics rests upon what seems to be the simplest assumption consistent with these three requirements, namely,

Axiom I2 (NEWTON, EULER, and others). *In an inertial frame*

$$\mathbf{f}(\mathscr{B}, \Sigma^e) = -\dot{\mathbf{m}}(\mathscr{B}; \chi). \tag{I.13-7}$$

[1] For an exposition of CARTAN's theory of Newtonian gravitation, see Chapter 12 of *Gravitation* by C. W. MISNER, K. S. THORNE, & J. A. WHEELER, San Francisco, W. H. Freeman and Co., 1973. For an attempt to incorporate CARTAN's theory of gravitation into continuum mechanics, see the paper by P. G. APPLEBY & N. KADIANAKIS, "A frame-independent description of the principles of classical mechanics," *Archive for Rational Mechanics and Analysis* **95** (1986): 1–22.

Up to now, the units of length, time, mass, and force have been independent, and Axiom I1 does not require there to be any relation among them, since it merely asserts that a certain force vanishes when a certain acceleration vanishes. Before it becomes legitimate even to state Axiom I2, we must assume that *forces can be specified in mechanical units*—in particular, that the dimensions of force are the dimensions of (mass) $\times$ (acceleration), which are (mass)(length)(time)$^{-2}$.

The origin of this assumption seems not to have been any particular experiment or observation but rather the fact that at first only special forces, namely, *weights*, were recognized. Weight was seen in time to be proportional to mass, and indeed in early studies of mechanics force, weight, and mass seem to have been confused often. That the units of force are of the special kind required in order that we be allowed even to consider Axiom I2 as a possible assumption in a theory of natural phenomena, should be accessible to test by experiment.[1] While no specific experiment seems ever to have been proposed, let alone effected, so as to test this assumption, it seems to be universally accepted.

Axiom I2 is consistent with Axiom F4 in Section I.5, since, as shown by $(\text{I.8-5})_1$, the rate of change of linear momentum of a part of $\mathscr{B}$ is the value of a measure over $\mathscr{B}$. Specifically, for smooth motions $(\text{I.8-5})_1$ enables us to express (7) in the form

$$\begin{aligned} d\mathbf{f}_{\Sigma^e} &= -\ddot{\chi}\, dM, \\ (d\mathbf{F}_{\Sigma^e})_{\mathbf{x}_0} &= -(\chi - \mathbf{x}_0) \wedge \ddot{\chi}\, dM. \end{aligned} \tag{I.13-8}$$

The second assertion follows from the first because we have assumed (I.8-8). In a more general system of mechanics, we should have to lay down $(8)_2$, or some other axiom, independently of $(8)_1$.

The forces and torques given by (8) are called *inertial*. Provided the frame $\oint$ be an inertial one, these forces and torques are those exerted upon the bodies of the great system Σ by the bodies, whatever they may be, that are outside Σ. When we choose instead to use a general frame $\oint^*$, we think of the unknown motions of the exterior Σ^e as being subjected to the same change from the frame $\oint$ to the frame $\oint^*$ as are the motions of Σ. Therefore, the second axiom of inertia, while it refers to a particular class of frames, expresses a frame-indifferent principle. While we follow tradition in stating it as we have, in terms of an inertial frame, we need not do so. Axiom A2 of Section I.12 asserts that

[1] The question is somewhat similar to that underlying the "first law of thermodynamics", which allows flow of heat to be measured in mechanical units.

all forces are frame-indifferent. Thus the quantity on the left-hand side of $(8)_1$ is frame-indifferent. Accordingly, a frame-indifferent statement that reduces to $(8)_1$ in an inertial frame is

$$d\mathfrak{f}(\mathscr{B}, \Sigma^e) = -\boldsymbol{\alpha}_{\mathscr{I}}\, dM, \tag{I.13-9}$$

$\boldsymbol{\alpha}_{\mathscr{I}}$ being that frame-indifferent vector field over $\mathscr{B}$ which in the inertial frame $\mathscr{I}$ reduces to $\ddot{\boldsymbol{\chi}}$. We have already calculated $\boldsymbol{\alpha}_{\mathscr{I}}$ and recorded it in (I.11-3). The student should recall the role in $\boldsymbol{\alpha}_{\mathscr{I}}$ played by $\mathbf{A}$, which is the spin of $\mathscr{I}$ with respect to the general rigid frame $\mathscr{I}^*$.

In the remainder of this book we shall follow the tradition of mechanics in assuming tacitly that *the frame used is an inertial one*, and so (8) holds.

Our use of an inertial frame rests on more than respect for tradition. An essential feature of classical mechanics is *the existence of special frames in which the relation between forces and the motions they produce is especially simple.* Since we have these felicitous frames, it would be simply foolish not to use them. When for purposes of interpretation in a particular application we need to employ some frame that is not inertial, as for example in problems referred to a rotating earth, we formulate the laws of mechanics first in an inertial frame and then transform them to the other frame of interest. Such is the traditional approach, which derives from CLAIRAUT and EULER. In replacing (8) by (9) we formulate that approach in general terms.

Recalling that the general axioms of mechanics imply (6), from Axiom I2 we see that

$$\mathbf{f}^a(\mathscr{B}) = \dot{\mathbf{m}}(\mathscr{B}; \boldsymbol{\chi}), \qquad \mathbf{F}^a(\mathscr{B})_{\mathbf{x}_0} = \dot{\mathbf{M}}(\mathscr{B}, \boldsymbol{\chi})_{\mathbf{x}_0}, \tag{I.13-10}$$

where to obtain the second statement we have used $(8)_2$. *That is, the applied force on $\mathscr{B}$ equals the rate of change of the linear momentum of $\mathscr{B}$ in an inertial frame, and the applied torque equals the rate of change of rotational momentum of $\mathscr{B}$ in the same frame*, both torque and rotational momentum being taken with respect to a place $\mathbf{x}_0$ that is stationary in the inertial frame. These two statements are ***Euler's Laws of Motion***. The formal treatment in the rest of this book is based upon them rather than upon the more general ideas from which we have developed them. Usually we shall write them in the shorter notation

$$\mathbf{f}^a = \dot{\mathbf{m}}, \qquad \mathbf{F}^a = \dot{\mathbf{M}}. \tag{I.13-11}$$

If for a given body in an inertial frame

$$\mathbf{f}^a = \mathbf{0}, \qquad \mathbf{F}^a = \mathbf{0}, \tag{I.13-12}$$

that body is *isolated*. From Euler's Laws (11) we see that *the linear and rotational momenta of a body remain constant if and only if that body is isolated*. The theorem given at the end of Section I.8 shows that *in an inertial frame the center of mass of an isolated body moves along a straight line at constant speed*.

We have seen that in a general rigid frame, $(8)_1$ must be replaced by (9). The corresponding replacement in $(8)_2$ has to be treated with care, since in it $\mathbf{x}_0$ is a fixed place in an inertial frame. The resulting general forms of EULER's Laws (10) are

$$\begin{aligned} \mathbf{f}^a(\mathscr{B}) &= \int_{\mathscr{B}} \boldsymbol{\alpha}_{\oint} dM, \\ \mathbf{F}^a(\mathscr{B})_{\mathbf{y}_0} &= \int_{\mathscr{B}} (\boldsymbol{\chi} - \mathbf{x}_0) \wedge \boldsymbol{\alpha}_{\oint} dM. \end{aligned} \tag{I.13-13}$$

Here $\mathbf{y}_0$ is a fixed place in an inertial frame $\oint$, and $\mathbf{x}_0$ is the corresponding place in the general frame $\oint^*$. Only if also $\oint^*$ is inertial are the right-hand sides of (13) equal to the rates of change of linear momentum and rotational momentum, respectively.

Returning to use of an inertial frame, as we shall do henceforth in this book, we note from (I.8-29) that EULER's First Law (11) can be written in terms of the motion of the center of mass $\bar{\mathbf{p}}$ of $\mathscr{B}$:

$$\mathbf{f}^a = M\ddot{\bar{\mathbf{p}}}. \tag{I.13-14}$$

Thus, in an inertial frame, the applied force on a body equals the mass of that body times the acceleration of its center of mass.

This last is one of the oldest of the commonly accepted principles of mechanics, used again and again, with or without explicit statement, in the eighteenth century. It is sometimes regarded as expressing the Second Law of NEWTON: "The change of motion is proportional to the impressed motive force, and it is made in the direction of the right line along which that force is impressed."

The point $\mathbf{x}_0$ with respect to which torques and rotational momenta entering EULER's Second Law are calculated is a fixed point in an inertial frame. We may use EULER's two laws together so as to calculate the effect of the applied loads upon the rotational momentum with respect to the center of mass $\mathbf{x}_c$. In (I.8-9) we replace $\mathscr{B}^e$ by $\mathscr{B}^e_\Sigma$. Then EULER's Second Law $(10)_2$ makes the left-hand side of the outcome equal the left-hand side of (I.8-30). If in the former we take $\mathbf{x}_c$ for $\mathbf{x}_1$, by use of EULER's First Law $(10)_1$ we conclude at once that $(10)_2$ holds with $\mathbf{x}_c$ replacing $\mathbf{x}_0$. *That is, the applied torque on $\mathscr{B}$ equals the rate* *of change of rotational momentum of $\mathscr{B}$ in an inertial frame when both torque and rotational momentum are taken with respect to the center of mass.*

The student interested only in continuum mechanics may pass now to the next section.

In analytical dynamics (above, Sections I.3, I.5, I.8) the "environment" X_0 is considered to have two separate parts, one inside the system Σ and the other being Σ^e:

$$X_0 = X_e \vee \Sigma^e, \tag{I.13-15}$$

say, so that

$$\begin{aligned} \mathbf{f}_k^e &= \mathbf{f}(X_k, X_0) = \mathbf{f}(X_k, X_e) + \mathbf{f}(X_k, \Sigma^e), \\ &= \mathbf{f}_k^0 - M_k \ddot{\mathbf{x}}_k, \end{aligned} \tag{I.13-16}$$

and

$$\mathbf{f}_k^0 := \mathbf{f}(X_k, X_e), \qquad \ddot{\mathbf{x}}_k := \ddot{\chi}(X_k, \cdot). \tag{I.13-17}$$

The force $\mathbf{f}_k^0$ is called the *external* or *extrinsic* applied force acting upon X_k. In terms of it and the mutual forces $\mathbf{f}_{kq}$, EULER's First Law $(10)_1$ assumes the form

$$\mathbf{f}_k^a = M_k \ddot{\mathbf{x}}_k, \qquad \mathbf{f}_k^a := \mathbf{f}_k^0 + {\sum_{q=1}^{n}}' \mathbf{f}_{kq}, \qquad k = 1, 2, \ldots, n, \tag{I.13-18}$$

as may be seen also from putting (16) into (I.5-24). Equations of this form are often called "Newtonian", though they occur nowhere in the writings of NEWTON.

Exercise I.13.1 (NOLL). The axioms of inertia when applied to analytical dynamics do not alter the requirement (I.5-22) and NOLL's theorem at the end of Section I.8. Thus in analytical dynamics EULER's Second Law is *equivalent*, the first being presumed imposed already, to the statement that the mutual forces are central.

Moreover, for the entire system of n mass-points

$$\begin{aligned} \sum_{k=1}^{n} \mathbf{f}_k^0 &= \frac{d}{dt} \sum_{k=1}^{n} M_k \dot{\mathbf{x}}_k = M\ddot{\mathbf{p}}, \\ \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{x}_0) \wedge \mathbf{f}_k^0 &= \frac{d}{dt} \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{x}_0) \wedge M_k \dot{\mathbf{x}}_k ; \end{aligned} \tag{I.13-19}$$

the second relation, while its form suggests the principle of rotational momentum, is a simple consequence of (18) and (I.8-24). These are the *theorems of linear and rotational momentum* of analytical dynamics.

Comparison of $(19)_1$ with (14) shows that *in an inertial frame, the motion of the center of mass of a body $\mathscr{B}$ is the same as that of a mass-point having the same mass as $\mathscr{B}$, located at the center of mass of $\mathscr{B}$, and subject to the resultant applied force on $\mathscr{B}$*. Thus if we are satisfied with knowing no more about the motion of a body than the motion of its center of mass, and if we can determine the applied force on that body, we need enter no more deeply into mechanics than the level of analytical dynamics. As HAMEL wrote in 1909, "what is understood in practice as the mechanics of points is neither more nor less than the theorem on the center of gravity." This fact goes far to explain the pragmatic success of analytical dynamics. In particular, use of it does not require that the body $\mathscr{B}$ really occupy no more than a discrete set of points in space, but only that our curiosity be slaked by determining the motions of such a set of points. The standard example here is furnished by the sun and its planets and comets. It is a typical example in that whether or not analytical dynamics be sufficient to describe its motion depends on how far we choose to inquire into it. For certain problems or in certain refined cases we need to take account of the spins and even the shapes of the bodies, and then analytical dynamics, as embodied in (18), (I.5-22), and (I.8-24), no longer suffices.

From (14) we see that the motion of the center of mass of any body is determined, to within arbitrary assigned position and velocity at some one time, if the resultant force on that body is a known function of time. For a rigid motion still more can be said. Consider first a rigid motion of a body one of whose substantial points remains constantly at a fixed place, say $\mathbf{x}_0$, in an inertial frame. Substitution of $(\mathrm{I}.10\text{-}12)_2$ into $(11)_2$ then yields ***Euler's Differential Equation*** for such a motion:

$$\mathbf{Q}\mathbf{F}^{\mathrm{a}}_{\mathbf{x}_0}\mathbf{Q}^{\mathsf{T}} = \mathbf{F}^{\mathrm{a}^*}_{\mathbf{x}_0^*} = \mathbf{E}^*_{\mathbf{x}_0^*}\dot{\mathbf{A}} + \dot{\mathbf{A}}\mathbf{E}^*_{\mathbf{x}_0^*} + \mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A}^2 - \mathbf{A}^2\mathbf{E}^*_{\mathbf{x}_0^*}\,. \qquad (\mathrm{I}.13\text{-}20)$$

Here $\mathbf{F}^{\mathrm{a}^*}_{\mathbf{x}_0^*}$ is the applied torque in the rest frame with respect to the stationary place $\mathbf{x}_0^*$ occupied in that frame by the substantial point that remains at rest at the one place $\mathbf{x}_0$ in an inertial frame. Since $\mathbf{E}^*_{\mathbf{x}_0^*}$ is a known, constant tensor, (20) is a differential equation of first order for the spin $\mathbf{A}$ of the rest frame ϕ^* with respect to the inertial frame ϕ, on the presumption that the resultant torque $\mathbf{F}^{\mathrm{a}^*}_{\mathbf{x}_0^*}$ be known.

Even if no substantial point remains at rest in an inertial frame, we may appeal to the italicized theorem above, just before the remarks on analytical dynamics, and so by use of (I.10-13) conclude that

$$\mathbf{F}^{\mathrm{a}}_{\mathbf{x}_c} = -\mathbf{E}_{\mathbf{x}_c}\dot{\mathbf{W}} - \dot{\mathbf{W}}\mathbf{E}_{\mathbf{x}_c} + \mathbf{E}_{\mathbf{x}_c}\mathbf{W}^2 - \mathbf{W}^2\mathbf{E}_{\mathbf{x}_c}\,, \qquad (\mathrm{I}.13\text{-}21)$$

or, equivalently,

$$\mathbf{F}^{a^*}_{\mathbf{x}^*_c} = \mathbf{E}^*_{\mathbf{x}^*_c}\dot{\mathbf{A}} + \dot{\mathbf{A}}\mathbf{E}^*_{\mathbf{x}^*_c} + \mathbf{E}^*_{\mathbf{x}^*_c}\mathbf{A}^2 - \mathbf{A}^2\mathbf{E}^*_{\mathbf{x}^*_c}, \tag{I.13-22}$$

$\mathbf{x}^*_c$ being the place occupied by the center of mass in the rest frame $\oint^*$. This statement is of the same form as (20) and can be interpreted similarly.

If $\mathbf{F}^{a^*}_{\mathbf{x}^*_0}$ is a continuous function of time, there is a unique solution $\mathbf{A}$ of (20) corresponding to any given initial value $\mathbf{A}(t_0)$, and if that initial value is skew, so is $\mathbf{A}(t)$ for all t, as the student will easily verify. A theorem given in Section I.9 states conditions under which the spin $\mathbf{A}$ determines the relative orientation $\mathbf{Q}$. Similar reasoning may be applied to (22).

Summarizing all these conclusions, we have the following

Theorem. *Let a body $\mathscr{B}$ be in rigid motion, and in a rest frame let $\mathbf{E}^*_{\mathbf{x}^*_0}$ be its Euler tensor with respect to the place $\mathbf{x}^*_0$. Suppose that either:*

A. *the place $\mathbf{x}^*_0$ in $\oint^*$ is occupied by a substantial point which remains at rest at the place $\mathbf{x}_0$ in the inertial frame $\oint$, or*
B. *the place $\mathbf{x}^*_0$ in $\oint^*$ is occupied by the center of mass of $\mathscr{B}$.*

In Case B, suppose that the place $\mathbf{x}_0(t)$ occupied in the inertial frame $\oint$ by the center of mass of $\mathscr{B}$ be known, e.g. *by integration of* (14).

Then the assignment of the initial orientation $\mathbf{Q}(t_0)$ of $\oint^*$ with respect to $\oint$ determines a unique rest frame $\oint^*$ and hence a unique rigid motion of $\mathscr{B}$.

Roughly, if one substantial point of $\mathscr{B}$ remains at rest in an inertial frame, or if the motion of the center of mass of $\mathscr{B}$ with respect to an inertial frame is known, a rigid motion of $\mathscr{B}$ is determined by an assigned resultant torque, to within inessential constants.

This theorem enables us to refine, if we so desire, the bare skeleton of mechanics furnished by analytical dynamics. If we are content to regard the motion of a body as rigid, we may calculate that motion from the resultant torque, once the existence of a fixed point or the motion of the center of mass has been determined. For this purpose we need to know about the body itself only its Euler tensor $\mathbf{E}^*_{\mathbf{x}^*_0}$ with respect to an appropriate place $\mathbf{x}^*_0$ in a rest frame.

As we noticed above, to apply the mechanics of mass-points we need not assume that the shape of $\mathscr{B}$ be a single place; rather, we must simply be content with determining the motion of the center of mass of $\mathscr{B}$, leaving unknown such motion relative to that center as the remaining points of $\mathscr{B}$ may have. Likewise, to apply the theory of rigid motions, we need not assume that the body $\mathscr{B}$ be susceptible *only* of such motions; rather, we must simply remain content with specifying some one shape of $\mathscr{B}$ and supposing that in *some* frame, that shape shall remain unchanged to within a rigid motion. In rough terms, analytical dynamics and the theory of rigid motions determine certain *aspects* of the motions of all bodies, whether or not they be mass-points or rigid bodies.

If we cast back one look at the purely kinematical corollary on steady rotations at the end of Section I.10, by use of the axioms of inertia we may now obtain from it a major proposition of dynamics. An *axis of free rotation* is a line whose direction is steady in a rest frame and about which a body subject to no resultant torque with respect to some point on that axis may spin. Such an axis is necessarily a steady axis of rotation and, of course, an axis of rotational momentum. Thus we have the following

Theorem (EULER). *The axes of free rotation through the center of mass of a body, or through the place of a body-point which is at rest in an inertial frame, are the principal axes of inertia.*

In particular, *a body of a given shape cannot spin freely about any line that is not one of its axes of inertia.* Since $\mathbf{E}^*_{\mathbf{x}_0^*}$ is positive and symmetric, there are either exactly three such axes, which are orthogonal to one another, or infinitely many. In the latter case, either every line is a principal axis of inertia, or the principal axes of inertia are one certain line and all lines in a plane perpendicular to it.

Exercise I.13.2 (EULER). For a particular $\mathscr{B}$, let $\mathbf{e}$, $\mathbf{f}$, $\mathbf{g}$ be an orthonormal triad of proper vectors of $\mathbf{E}^*_{\mathbf{x}_0^*}$ in a rest frame, and suppose $\mathscr{B}$ to rotate about the principal axis of inertia defined by $\mathbf{e}$, so that

$$\mathbf{A} = \omega \mathbf{f} \wedge \mathbf{g}. \qquad \text{(I.13-23)}$$

Then $\omega = |\mathbf{A}|/\sqrt{2}$. If $\mathbf{x}_0^* = \mathbf{x}_0$, a place on the axis, then EULER's equation (22) reduces to

$$\mathbf{F}^{a*}_{\mathbf{x}_0^*} = F \mathbf{f} \wedge \mathbf{g}, \qquad \text{(I.13-24)}$$

and

$$F = I\dot{\omega}, \qquad I = E_2 + E_3, \qquad \text{(I.13-25)}$$

E_2 and E_3 being the proper numbers of the Euler tensor $\mathbf{E}^*_{\mathbf{x}_0^*}$ corresponding with the proper vectors $\mathbf{f}$ and $\mathbf{g}$.

14. Power. Kinetic Energy. Potential Energy

In Section I.12 we have imposed the requirement that the working W be frame-indifferent. In an inertial frame the working has an especially useful interpretation. Namely, if we substitute $(\text{I}.13\text{-}8)_1$ into (I.13-2) and then substitute the outcome into the definition (I.8-7) of W, by comparison with $(\text{I}.8\text{-}5)_3$ we see that

$$W = P - \dot{K}, \tag{I.14-1}$$

P being the *power*, namely, the working of the forces exerted on $\mathscr{B}$ by the exterior bodies in the great system Σ alone:

$$P = \int_{\mathscr{B}} \dot{\boldsymbol{\chi}} \cdot d\mathbf{f}(\mathscr{B}, \mathscr{B}^{\mathrm{e}}_{\Sigma}), \tag{I.14-2}$$

and K being the kinetic energy of $\mathscr{B}$. We have eased the writing by leaving arguments such as $\mathscr{B}$, $\boldsymbol{\chi}$, and t unwritten. The statement (1) asserts that the working W is the power of the forces exerted upon $\mathscr{B}$ by the exterior of $\mathscr{B}$ in the great system Σ, less the rate of increase of the kinetic energy of $\mathscr{B}$, in an inertial frame. We may say equally that the working of the inertial forces is $-\dot{K}$.

If in an inertial frame all work done is converted into kinetic energy, $P = \dot{K}$, so that

$$W = 0, \tag{I.14-3}$$

and the term *mechanically perfect* is applied. That term may refer to the body, to the system of forces, or to the motion, whichever of these we choose to regard as being restricted by the statement. The condition (3) is frame-indifferent, and so it may be imposed on all bodies, all motions, or all systems of forces, in any combination we please. In Section I.12 we have proved that a rigid motion of any body and all motions of a single mass-point are mechanically perfect. This statement is a consequence of NOLL's Axiom in Section I.12 and does not require the Axioms of Inertia. These latter, however, enable us to interpret the statement as follows: In an inertial frame the working of the forces on $\mathscr{B}$ is balanced by increase of the kinetic energy of $\mathscr{B}$.

Theories of mechanically perfect motions or bodies are untypical of general mechanics because they permit us to study and determine the effects of external forces without having to take up effects of deformation and dissipation. Examples are furnished by any motion of a mass-point and by the rigid motion of any body, since, as we have seen in Section I.12, for both of these $W = 0$ always.

With some special kinds of bodies and special systems of forces we may associate a *potential energy*. This term has somewhat different meanings in different special theories. For illustration we shall select analytical dynamics. In Sections III.6 and IV.7 and in Volume 3 we shall obtain conclusions of the same kind for fluids and elastic solids.

First, from $(\text{I.8-10})_3$ and $(\text{I.8-11})_2$ we see that for a motion of a system of n mass-points

$$\dot{K} = \sum_{k=1}^{n} M_k \dot{\mathbf{x}}_k \cdot \ddot{\mathbf{x}}_k, \qquad \ddot{\mathbf{x}}_k := \ddot{\chi}(X_k, t). \tag{I.14-4}$$

By use of $(\text{I.13-18})_1$ it follows that

$$\dot{K} = \sum_{k=1}^{n} \dot{\mathbf{x}}_k \cdot \mathbf{f}_k^{\mathrm{a}}. \tag{I.14-5}$$

In the simplest examples of the analytical dynamics of systems of mass-points the force $\mathbf{f}_k^{\mathrm{a}}$ is the value at $\mathbf{x}_k$ of a field that when $\chi(X_k, t) = \mathbf{x}_k$ acts upon the mass-point X_k. The simplest such fields are those that derive from a *potential function* $U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, defined and continuously differentiable on an open set of $\mathscr{E} \times \mathscr{E} \times \cdots \times \mathscr{E}$ that contains all the places occupied by the mass-points $X_1, X_2, \ldots, X_n$ in the course of the motion. Then the following definition makes sense for each motion χ:

$$U(t) := U(\chi(X_1, t), \chi(X_2, t), \ldots, \chi(X_n, t)), \tag{I.14-6}$$

and so

$$\dot{U}(t) = \sum_{k=1}^{n} \dot{\mathbf{x}}_k \cdot \partial_{\mathbf{x}_k} U|_{\mathbf{x}_r = \chi(X_r, t)}, \qquad r = 1, 2, \ldots, n, \tag{I.14-7}$$

in which $\partial_{\mathbf{x}_k} U$ is the partial derivative of U with respect to $\mathbf{x}_k$. Recalling $(\text{I.13-18})_1$, we assume that

$$\mathbf{f}_k^{\mathrm{a}} = -\partial_{\mathbf{x}_k} U|_{\mathbf{x}_r = \chi(X_r, t)}, \qquad r = 1, 2, \ldots, n, \tag{I.14-8}$$

and so conclude that

$$\dot{U} = -\sum_{k=1}^{n} \dot{\mathbf{x}}_k \cdot \mathbf{f}_k^{\mathrm{a}}. \tag{I.14-9}$$

Such systems of forces are called *conservative* because putting (5) and (9) together yields the ***Energy Theorem***:

$$\dot{K} + \dot{U} = 0; \tag{I.14-10}$$

that is, *for each motion χ of a dynamical system subject to the forces* $\mathbf{f}_1^{\mathrm{a}}, \mathbf{f}_2^{\mathrm{a}}, \ldots, \mathbf{f}_n^{\mathrm{a}}$ *derived from U through* (8) *the quantity $K + U$ remains constant.* The value $U(t)$ given by (6) for a given χ is called the *potential energy* of the system at the time t in the motion χ.

A particular kind of potential function U that delivers forces as the sum of extrinsic applied forces $\mathbf{f}_k^0$ and of mutual forces $\mathbf{f}_{qk}$ as in (I.13-18) is given as follows in terms of differentiable real functions U_k^0 and U_{qk}, $q = 1, 2, \ldots, n$:

$$U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \sum_{k=1}^{n} U_k^0(\mathbf{x}_k) + \sum_{\substack{q,k=1\\ k>q}}^{n} U_{qk}(\mathbf{x}_q, \mathbf{x}_k). \tag{I.14-11}$$

Without loss of generality $U_{kq}(\mathbf{x}_k, \mathbf{x}_q) := U_{qk}(\mathbf{x}_q, \mathbf{x}_k)$ when $k < q$, $U_{qq} := 0$, $q = 1, 2, \ldots, n$; then (11) can be rewritten as

$$U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \sum_{k=1}^{n} U_k^0(\mathbf{x}_k) + \frac{1}{2} \sum_{r,q=1}^{n} U_{rq}(\mathbf{x}_r, \mathbf{x}_q). \tag{I.14-12}$$

From (12) we find that

$$-\partial_{\mathbf{x}_k} U = -\partial_{\mathbf{x}_k} U_k^0 - \frac{1}{2} \sum_{q=1}^{n} (\partial_{\mathbf{x}_k} U_{kq} + \partial_{\mathbf{x}_k} U_{qk}) = -\partial_{\mathbf{x}_k} U_k^0 - \sum_{q=1}^{n} \partial_{\mathbf{x}_k} U_{kq}. \tag{I.14-13}$$

Applying (8) yields

$$\mathbf{f}_k^{\mathrm{a}} = -\partial_{\mathbf{x}_k} U_k^0 \big|_{\mathbf{x}_k = \chi(X_k, t)} - \sum_{q=1}^{n} \partial_{\mathbf{x}_k} U_{kq} \big|_{\mathbf{x}_k = \chi(X_k, t),\, \mathbf{x}_q = \chi(X_q, t)}. \tag{I.14-14}$$

The first term on the right-hand side refers to the place $\mathbf{x}_k$ only; the summand in the second term, to the places $\mathbf{x}_k$ and $\mathbf{x}_q$ only, k, $q = 1, 2, \ldots, n$. Thus we may, if we like, regard the functions U_k^0 as the potentials of the external forces $\mathbf{f}_k^0$ in (I.13-18), the functions U_{kq} as the potentials of the mutual forces $\mathbf{f}_{kq}$. Alternatively, we might start with those potentials and argue backward to (13) and (12).

Thus far we have abused terms in calling the various quantities $\mathbf{f}$ with subscripts "forces". All forces are frame-indifferent. When we apply that requirement, we substantially restrict the various functions U that we have called potentials. The following exercise indicates the restrictions resulting for the functions U_k^0 and U_{kq} when they are assumed to be frame-indifferent.

Exercise I.14.1. Under a change of frame (I.9-5),

$$U_{qk}(\mathbf{x}_q^*, \mathbf{x}_k^*) = U_{qk}(\mathbf{x}_q, \mathbf{x}_k) \tag{I.14-15}$$

if and only if $U_{qk}(\mathbf{x}_q, \mathbf{x}_k)$ depends on $\mathbf{x}_q$ and $\mathbf{x}_k$ through $|\mathbf{x}_q - \mathbf{x}_k|$ alone. That is, (15) holds if and only if there is a real function V_{qk} such that

$$U_{qk}(\mathbf{x}_q, \mathbf{x}_k) = V_{qk}(|\mathbf{x}_q - \mathbf{x}_k|). \tag{I.14-16}$$

If U_{qk} is the potential of the force exerted by X_k upon X_q in the motion $\mathbf{x}_r = \chi(X_r, t)$, $r = 1, 2, \ldots, n$, the mutual forces are central and pairwise equilibrated. Hence *balance of rotational momentum* (Exercise I.13.1) and *frame-indifference of the potentials* U_{kq} in (14) are equivalent for the special systems of mass-points considered here. If $U_k^0(\mathbf{x}_k) = V_k^0(\mathbf{x}_k, \mathbf{x}_0)$, $\mathbf{x}_0$ fixed (*cf.* (I.13-16)), then V_k^0 is frame-indifferent if and only if $\partial_{\mathbf{x}_k} V_k^0(\mathbf{x}_k, \mathbf{x}_0)$ is parallel to $\mathbf{x}_k - \mathbf{x}_0$. If, more generally, $U_k^0(\mathbf{x}_k) = V_k^0(\mathbf{x}_k - \mathbf{x}_0, \mathbf{n})$ and $\mathbf{n}$ is a frame-indifferent unit vector (*e.g.* the unit normal to a surface through $\mathbf{x}_k$), then V_k^0 is frame-indifferent if and only if $\partial_{\mathbf{x}_k} V_k^0(\mathbf{x}_k - \mathbf{x}_0, \mathbf{n})$ equals the sum of a vector parallel to $\mathbf{x}_k - \mathbf{x}_0$ and a vector parallel to $\mathbf{n}$.

The passages on analytical dynamics included in this chapter are designed only to help the student grasp the general principles of mechanics through their reduction to very familiar concepts and theorems.

The quantity $K + U$ is sometimes called the "total energy" of the system, but this usage misleads. Potential energy is not a fundamental concept of mechanics. When a potential energy exists, it is useful in solving special problems, but that it should exist, is only a fortunate accident to the body, for not all forces occurring in nature are represented well by mathematical forces deriving from a potential energy. The succeeding section explains and defines total energy.

15. Internal Energy

The contents of this section will not be used in the following text except for an exercise at the end of Section III.6. I include these statements as the beginning of an answer to the student's natural question, "How does mechanics fit into the scheme of natural phenomena?" In what follows now are some remarks about heat. The next step would bring in electromagnetism, but even the elements of that subject defy clear summary in one short section.

In nature the exercise of forces need not give rise to motion but may in whole or part be consumed in production of heat, and, *vice versa*, heating a body may set it in motion, as we may see from the example of compressing a gas by a piston, or allowing the gas by expanding to move such a piston and so cool itself by working. Also there are circumstances in which a body may be heated or cooled with no consequent effects recognizable as being motions. Common experience requires that the mathematical theory of continua be broadened to include the effects of heating. While we shall not treat those in this book, here we sketch the basis of the general theory that extends mechanics so as to subsume the simplest thermodynamics.

The structure of a *system of heatings* is much like that of a system of forces (above, Section I.5). The elements of the system are real functions Q defined on pairs of separate bodies. They are subject to axioms just like F1–F4 in Section I.5. $Q(\mathscr{B}, \mathscr{C})$ is the ***heating*** effected upon $\mathscr{B}$ by $\mathscr{C}$. The ***resultant heating*** of $\mathscr{B}$ is $Q(\mathscr{B}, \mathscr{B}^e)$, namely, the heating of $\mathscr{B}$ effected by its surroundings.

The classical thermodynamics of homogeneous processes, begun as a somewhat mathematical science, relates heat given to a body and the work done by that body in undergoing a process. Only through its changes of volume is the deformation of such a body taken into account in that theory. CLAUSIUS in developing it proved the existence of the *internal energy* E, which compensates any excess or defect of heat and work. While we shall not in this book present a theory of thermomechanics, at this point the student may find it helpful to recall some of the main assumptions and conclusions of the classical thermodynamics of homogeneous processes.[1]

[1] Textbooks on thermodynamics by and for physicists and engineers are more likely to obscure the basic assumptions and the logical structure of classical thermodynamics than to enlighten a critical student familiar with rigorous calculus and with the orderly, deductive arguments used in mechanics. Otherwise respectable books by mathematicians include among the applications adduced to illustrate the pure mathematics developed (*e.g.*, CARATHÉODORY's theorem on Pfaffian forms) passages of silly babble on thermodynamics. The only trustworthy elementary book I have seen is D. R. OWEN's *A First Course in the Mathematical Foundations of Thermodynamics*, New York, *etc.* Springer-Verlag, 1984. Unfortunately, as OWEN writes on p. 3, the theory he presents does not cover the anomalous behavior of water.

For many years I strove to correct, recast, and integrate the works of the pioneers—CARNOT, KELVIN, RANKINE, CLAUSIUS, REECH—and so construct classical thermodynamics as a mathematical science, clearly and rigorously developed. My final presentation is Appendix 1A, "Thermodynamics for beginners," *Rational Thermodynamics*, Second Edition, corrected and enlarged, New York, *etc.*, Springer-Verlag, 1984. I have circulated a corrected, fuller text (1988), revising that published by the Accademia dei Lincei in 1986: "Classical thermodynamics cleansed and cured," pp. 265–291 of *Meeting on Finite Thermoelasticity* (1985), Contributi del centro Linceo interdisciplinare di Scienze Matematiche e loro applicazioni, No. 76. Further information may be found in the book I wrote in collaboration with BHARATHA, *The Concepts and Logic of Classical Thermodynamics* ..., second, corrected printing, New York, *etc.*, Springer-Verlag, 1988, and my *The Tragicomical History of Thermodynamics, 1822–1854*, New York, *etc.*, Springer-Verlag, 1980. While, like the works of discovery, those just cited treat only systems with two independent variables, M. PITTERI has extended the analysis to systems with three or more variables: "Classical thermodynamics of homogeneous systems based upon Carnot's general axiom," *Archive for Rational Mechanics and Analysis* **80** (1982):333–385. In all these works the anomalous behavior of water is explained naturally and easily as a special instance of thermodynamic behavior.

First, in conditions such that no work is done on or by $\mathscr{B}$, the resultant heating of $\mathscr{B}$ is stored within it as internal energy:

$$\dot{\mathrm{E}}(\mathscr{B}) = Q(\mathscr{B}, \mathscr{B}^{\mathrm{e}}). \tag{I.15-1}$$

Circumstances in which this equation holds are called *energetically perfect.*

Exercise I.15.1. On the assumption that (1) holds, $\dot{\mathrm{E}}$ is an additive set function if and only if $Q(\mathscr{B}, \mathscr{C}) = -Q(\mathscr{C}, \mathscr{B})$ for all pairs of separate bodies $\mathscr{B}$ and $\mathscr{C}$.

Since the functions Q are defined over pairs of bodies, they may be transferred to the shapes of bodies. Since there is no basis for assigning preference to one frame rather than another in considerations of heat, we assume that the Q are frame-indifferent:

Axiom E1.

$$Q^* = Q. \tag{I.15-2}$$

The occasional connection between heating and the action of forces, mentioned above, shows that the conditions (I.14-3) and (1) cannot be general. It suggests that W, $\dot{\mathrm{E}}$, and Q may be related, but it does not dictate any particular relation. The units assigned to Q and hence determined for E by (1) were specified originally in terms of conditions in which forces and motions were absent. These units are called "thermal", and they are still in wide use today.

The pioneers of thermodynamics made scant use of the principles and theorems of mechanics. They regarded the condition of a body capable of absorbing and emitting heat as being specified sufficiently by its volume V and its temperature θ on some accepted scale, and they assumed that Q was a linear function of $\dot{V}$ and $\dot{\theta}$ with coefficients depending upon both V and θ. Beginning with Carnot, they considered mainly changes such that V and θ arrived finally at the values initially given them: "cyclic processes". Clausius assumed (in effect) that if a body so specified undergoes a cyclic process in the interval of time $[0, T]$, then

$$\int_0^T W\,dt = \mathrm{J}\int_0^T Q\,dt, \tag{I.15-3}$$

and the constant J is the same for all bodies and all cyclic processes; of course it depends upon the systems of mechanical and thermal units employed. Nowadays this statement can be proved mathematically as a consequence of physically natural assumptions.[1] It is called the *Principle of Equivalence of Heat and*

[1] Here the student may consult any of the works cited in the footnote on the preceding page.

Work in Cyclic Processes. From it we see that despite the dissimilarity of the original concepts of heating from those of mechanics, *heating may be measured in units of working*. It suggests that we may consider heating Q and working W as co-operating to produce energy. CLAUSIUS and KELVIN, each in his own way, arrived at the following *Balance of Energy*,[1] usually called "The First Law of Thermodynamics":

Axiom E2.

$$\dot{\mathrm{E}} = W + Q. \tag{I.15-4}$$

The arguments $\mathscr{B}$, $\mathscr{B}^{e}$, χ, *etc*., are omitted for ease of writing, and the units are so chosen that $\mathrm{J} = 1$.

Since Q and W are frame-indifferent, so is E.

By comparing (3) with (I.14-3) and (1) we see that the term *mechanically perfect* is applicable now only to *energetically perfect* circumstances, and conversely.

Theories of mechanically and energetically perfect motions or bodies are untypical of thermomechanics since they permit us to study and determine the effects of forces without specifying or even mentioning heating, or the effects of heating without specifying or even mentioning forces. Examples of the former kind are furnished by any motion of a mass-point and by the rigid motion of any body, since as we have seen in Section I.12, in both these cases $W = 0$ always. An example of the latter is furnished by the classical theory of the conduction of heat, which assumes at its very start that $\dot{\mathrm{E}} = Q$. If in that theory it is assumed that all bodies are at rest, obviously $W = 0$, but in fact the theory in its classical form is consistent with Axiom E2 *only* when $W = 0$.

Axiom E2 makes no use of the Axioms of Inertia, since W is the net working, which may include the working of inertia. In an inertial frame we have the relation (I.14-1) between the power P of the forces within the great system Σ and the kinetic energy K, and so Axiom E2 yields

$$\dot{K} + \dot{\mathrm{E}} = P + Q. \tag{I.15-5}$$

The sum $K + \mathrm{E}$ is called the *total energy* of $\mathscr{B}$ in its actual shape $\chi_{\Omega}(\mathscr{B}, t)$. Thus (5) states that *in an inertial frame, the sum of the heating of $\mathscr{B}$ and the power of the forces within the great system acting on $\mathscr{B}$ equals the rate of increase of the total energy of $\mathscr{B}$.*

[1]Here the student may consult any of the works cited above on p. 80.

Potential energy, which has been considered above in Section I.14, is a quality of certain special systems, which, while often useful, are not at all typical of the conditions mechanics and thermomechanics design to model. Potential energy must never be confused with internal energy.

Only in Exercise III.6.6 shall we again in this book refer to internal energy.

General References

W. Noll, "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* (Colloquium at Stanford, 1957), Amsterdam, North-Holland Publ., 1959. Reprinted in W. Noll, *The Foundations of Mechanics and Thermodynamics*, New York, Heidelberg, and Berlin, Springer-Verlag, 1974.

W. Noll, "La mécanique classique, basée sur un axiome d'objectivité," pp. 47–56 of *La Méthode Axiomatique dans les Mécaniques Classiques et Modernes* (Colloque International à Paris, 1959), Paris, Gauthier-Villars, 1963. Reprinted along with the preceding.

W. Noll, "Euclidean geometry and Minkowskian chronometry," *American Mathematical Monthly* **71**, 129–144 (1964). Reprinted along with the preceding.

W. Noll, "Lectures on the foundations of continuum mechanics and thermodynamics," *Archive for Rational Mechanics and Analysis* **52**, 62–92 (1973). Reprinted along with the preceding.

This Page Intentionally Left Blank

Chapter II

Kinematics

These theorems render the forms of motion . . .at least approachable in concept.

HELMHOLTZ
On the integrals of the
hydrodynamical equations that
correspond with vortex motion
Journal für die Reine und
Angewandte Mathematik **55**
(1858): 25–55.

The great clarity which geometrical investigation lends to the study of the dynamics of solids leads us to expect significant success in hydrodynamics through a study of the kinematics of deformable systems.

ZHUKOVSKI
Кинематика Жидказо Тюла (1876)

The theory of these general phenomena of motion in continuous media has a yet unbounded scope of development. Nevertheless, it is necessary to approach them entirely without prejudice . . . .

JAUMANN
Introduction to *Die Grundlagen*
der Bewegungslehre von einem
modernen Standpunkte aus
Leipzig (1905)

1. Placements. Universes of Shapes

In Section I.4 we have agreed that by the term "body" $\mathscr{B}$ we mean a regularly open set in some topological space over which a non-negative Borel measure M, called mass, is defined.[1] In Section I.3 the elements X of $\mathscr{B}$ are called substantial points. Here and henceforth, unless the contrary is stated, we presume given a rigid frame, and by $\mathbf{x}$ we denote a place in it (*cf.* Section I.6). In Section I.7 we defined a motion χ of $\mathscr{B}$, namely, a mapping of the substantial points comprised by $\mathscr{B}$ onto points of a three-dimensional Euclidean space $\mathscr{E}$ at the time t:

$$\mathbf{x} = \chi(X, t) \qquad \forall X \in \mathscr{B}, \quad \forall t \in \mathscr{I}. \tag{I.7-7)$_{\mathrm{r}}$ (}$$

For each t the mapping $\chi(\cdot, t)$ is a *placement* of the substantial points of $\mathscr{B}$; the place $\mathbf{x}$ is occupied by the substantial point X at the time t in the motion χ. The range of the placement is the shape assumed by $\mathscr{B}$ at the time t (Section I.7). When regarding t as the present time we call the shape of $\mathscr{B}$ its present shape.

Without fear of confusion we may write $\chi(\mathscr{B}, t)$ for the shape of $\mathscr{B}$ at the time t, thus using the symbol χ in two different though related senses: as a mapping of substantial points onto places and as a mapping of the bodies they constitute onto regions of space.

All properties we shall posit for $\chi(\cdot, t)$ will allow the possibility that χ be a function of X alone, the same function for all t. Examples of constant placements will be encountered below in Sections II.2 and II.3 and in Volume 3.

While in physical experience bodies are available to us only in some shape or other, the shapes are not to be confused with the bodies themselves. In analytical dynamics (Section I.3, Example 1) the substantial points stand in one-to-one correspondence with the numbers $1, 2, \ldots, n$, and the placements of bodies are discrete sets of points in $\mathscr{E}$. Nobody ever confuses the sixth substantial point with the number 6, or with the place the sixth substantial point happens to occupy at some time. The number 6 is merely a label attached to the substantial-point, and other labels would do just as well. Similarly, in continuum mechanics a body may assume infinitely many different shapes.

We shall refer to the subbodies $\mathscr{P}$ of a given body $\mathscr{B}$ as the parts of $\mathscr{B}$ (Section 1.2). The student should here reread Example 2 in Section I.3.

Henceforth, *we consider only continua*. We assume that *for each t the mapping $\chi(\cdot, t)$ is a homeomorphism of $\mathscr{B}$ onto its shape $\chi(\mathscr{B}, t)$*; we as-

[1] "Borel set", "Borel measure", and "σ-algebra" are defined above in Footnote 2 on p. 18.

sume further that *that homeomorphism carries $\mathscr{B}$ and its parts, which are regularly open sets in a topological space $\mathscr{T}$, into regularly open sets in $\mathscr{E}$.* To that end it is sufficient to assume that *the homeomorphism $\chi(\cdot, t)$ derives from a homeomorphism of $\mathscr{T}$ onto $\mathscr{E}$ by restriction of both its domain and its co-domain.* The shapes of all bodies constitute a collection $\mathbf{\Omega}_S$ with the structure of a universe of bodies if $\prec$ is taken to be inclusion $\subset$ and the meet is defined by (I.3-1). We call $\mathbf{\Omega}_S$ a ***universe of shapes***; it is an image of the universe of bodies $\mathbf{\Omega}_O$ defined in Section I.3 with $\mathscr{T}$ taken as $\mathscr{E}$.

The student must recall always that *a body in assuming various shapes never loses its identity and the properties assigned to it.* Its main properties are its mass distribution (Section I.4) and the material or materials assigned to its substantial points. The theory of materials is presented below in Chapter IV.

For volume in $\mathscr{E}$ we use Lebesgue measure[1] and denote it by V. Every element of $\mathbf{\Omega}_S$ is a Lebesgue-measurable set, and if it is not empty, its measure is positive.

The homeomorphism $\chi(\cdot, t)$ has an inverse $\chi^{-1}(\cdot, t)$, defined over the shape $\chi(\mathscr{B}, t)$:

$$X = \chi^{-1}(\mathbf{x}, t) \qquad \forall \mathbf{x} \in \chi(\mathscr{B}, t), \quad \forall t \in \mathscr{I}. \tag{II.1-1}$$

In three-dimensional continuum mechanics, as the student will see abundantly in the succeeding chapters, the *integral-gradient theorem*, which is the basis of "Green's transformation", often called "the divergence theorem", is a tool of central importance. All the shapes of bodies should be such as to make the integral-gradient theorem apply whenever the fields integrated are smooth to the degrees ordinarily assumed. That is not so for all the elements in $\mathbf{\Omega}_S$. For our purposes $\mathbf{\Omega}_S$ includes too many sets.

We wish to narrow the class of placements of bodies so as to make the collection of their shapes a universe fit for continuum mechanics. We might think that the subcollection $\mathbf{\Omega}_r$ would be suitable, and in fact the integral-gradient theorem does hold for the sets in it, but, as we have stated in the warning at the end of Section I.3, pairs of elements of $\mathbf{\Omega}_r$ do not always have a meet that lies in it. If we adjoin sets such as to render satisfied Axiom B6 of Section I.2, we cannot be sure that adjoined sets will make the integral-gradient theorem hold. To overcome these two difficulties, we need some interesting, rather advanced mathematics, the main course of which we outline, without proofs. A development of traditional kinematics begins in Section II.2.

To help seek proper subcollections of $\mathbf{\Omega}_S$ that might serve as universes of shapes for continuum mechanics, we begin with a definition.

[1] See, for example, Theorem 2.20 of the book by RUDIN cited above in Footnote 2 on p. 18.

Sets of finite perimeter. If $\mathscr{C}$ is a Borel set in a Euclidean space $\mathscr{E}$, its *perimeter*

$$\operatorname{per}\mathscr{C} := \sup\left\{\int_{\mathscr{C}} \operatorname{div}\mathbf{g}\,dV\colon\ |\mathbf{g}(\mathbf{x})| \leqq 1,\ \mathbf{g}\in C_0^1(\mathscr{E},\mathscr{V})\right\}. \qquad \text{(II.1-2)}$$

As the notation indicates, the supremum is taken over all continuously differentiable functions $\mathbf{g}$ that map $\mathscr{E}$ into its translation space $\mathscr{V}$, are of compact support, and have values nowhere longer than 1. If per $\mathscr{C}$ is finite, $\mathscr{C}$ is a *set of finite perimeter*.[1] Such sets are often called "Caccioppoli sets".

Exercise II.1.1. Let $\partial\mathscr{C}$ have a tangent plane at each point, and suppose that the normal to that plane be a continuous function of position on $\partial\mathscr{C}$. Then per $\mathscr{C} = A(\partial\mathscr{C})$, the symbol A denoting "area of", which is taken to be two-dimensional Hausdorff measure.[2]

Theorem. *If the subsets $\mathscr{C}$ and $\mathscr{D}$ of $\mathscr{E}$ have finite perimeter, then so do $\mathscr{C}\cap\mathscr{D}$ and $\mathscr{C}\cup\mathscr{D}$. Moreover,*

$$\operatorname{per}(\mathscr{C}\cap\mathscr{D}) + \operatorname{per}(\mathscr{C}\cup\mathscr{D}) \leqq \operatorname{per}\mathscr{C} + \operatorname{per}\mathscr{D}; \qquad \text{(II.1-3)}$$

this inequality is sharp.

The first statement in the theorem makes the sets of finite perimeter a Boolean algebra with respect to intersection and union.

There is a theorem that relates sets of finite perimeter directly to the integral-gradient theorem. The concept on which the connection rests is the *outer normal* of a set of finite perimeter. A point $\mathbf{x}$ of $\partial\mathscr{C}$ has an outer normal if there is a plane through $\mathbf{x}$ that lies essentially to one side of $\partial\mathscr{C}$ near $\mathbf{x}$. Formally, let

[1] An essentially elementary discussion of sets of finite perimeter and the associated concept of functions of bounded variation may be found in Chapters 4 and 5 of the book by A. I. Vol'pert and S. I. Hudjaev, *Analysis in Classes of Discontinuous Functions and Equations of Mathematical Physics*, Dordrecht *etc.*, Martinus Nijhoff, 1985. Be it noted that this clear, excellent, and compact book is written by and for engineers.

For applications to continuum mechanics see the paper by M. E. Gurtin, W. O. Williams, & W. P. Ziemer, "Geometric measure theory and the axioms of continuum thermodynamics," *Archive for Rational Mechanics and Analysis* **92** (1986), 1–22, and also the papers by Ziemer, "Cauchy flux and sets of finite perimeter," *ibid.* **84** (1983): 189–201 and M. Šilhavý, "General Cauchy fluxes," *ibid.* **90** (1985): 195–212.

[2] For "Hausdorff measure" see Section 22 of F. J. Almgren's *Plateau's Problem*, N.Y. and Amsterdam, Benjamin, 1966 and pages 169–171 of H. Federer's *Geometric Measure Theory*, Berlin *etc.*, Springer-Verlag, 1969.

$\mathscr{S}_r(\mathbf{x})$ denote a ball of radius r centered at $\mathbf{x}$; given a unit vector $\mathbf{n}$, let

$$\begin{aligned}\mathscr{S}_r^+(\mathbf{x};\,\mathbf{n}) &:= \mathscr{S}_r(\mathbf{x}) \cap \{\mathbf{y}\colon\ (\mathbf{y}-\mathbf{x})\cdot\mathbf{n} \geqq 0\},\\ \mathscr{S}_r^-(\mathbf{x};\,\mathbf{n}) &:= \mathscr{S}_r(\mathbf{x}) \cap \{\mathbf{y}\colon\ (\mathbf{y}-\mathbf{x})\cdot\mathbf{n} \leqq 0\}.\end{aligned} \tag{II.1-4}$$

Definition. Let $\mathscr{C}$ be a set of finite perimeter. Then $\mathbf{n}$ is an *outer normal* to $\mathscr{C}$ at $\mathbf{x}$ if

$$\begin{aligned}\lim_{r\to 0}\frac{V(\mathscr{S}_r^+(\mathbf{x};\,\mathbf{n})\cap\mathscr{C})}{V(\mathscr{S}_r^+(\mathbf{x};\,\mathbf{n}))} &= 0,\\ \lim_{r\to 0}\frac{V(\mathscr{S}_r^-(\mathbf{x};\,\mathbf{n})\cap(\mathscr{E}\backslash\mathscr{C}))}{V(\mathscr{S}_r^-(\mathbf{x};\,\mathbf{n}))} &= 0.\end{aligned} \tag{II.1-5}$$

The points at which an outer normal exists constitute the *reduced boundary* $\partial^*\mathscr{C}$.

It is easily seen that $\partial^*\mathscr{C}$ is a subset of the topological boundary $\partial\mathscr{C}$, and that if $\partial\mathscr{C}$ has a tangent plane everywhere, then $\partial\mathscr{C} = \partial^*\mathscr{C}$.

The concept of reduced boundary, introduced by DE GIORGI, is central to the following development.

Theorem. For each $\mathbf{x} \in \partial^*\mathscr{C}$ there is *only one* outer normal to $\mathscr{C}$ at $\mathbf{x}$.

Hence we may define the mapping $\mathbf{n}_{\mathscr{C}}$ which assigns to each $\mathbf{x} \in \partial^*\mathscr{C}$ the unique outer normal $\mathbf{n}_{\mathscr{C}}(\mathbf{x})$ to $\mathscr{C}$ at $\mathbf{x}$. We call $\mathbf{n}_{\mathscr{C}}$ the *outer normal field* of $\mathscr{C}$.

Theorem. The reduced boundary $\partial^*\mathscr{C}$ has finite area, and

$$A(\partial^*\mathscr{C}) = \operatorname{per}\mathscr{C}. \tag{II.1-6}$$

The following theorem puts in better perspective the statement in Exercise II.1.1.

Theorem (DE GIORGI). *The reduced boundary $\partial^*\mathscr{C}$ differs from the union of a countable collection of compact subsets of C^1-surfaces only by a set of null area.*

Integral-Gradient Theorem. *For every continuous function f: $\operatorname{clo}\mathscr{C} \to \mathscr{R}$ for which $f|_{\mathscr{C}}$, the restriction of f to $\mathscr{C}$, is differentiable and*

$\nabla f|_{\mathscr{C}}$ *is integrable on* $\mathscr{C}$,

$$\int_{\mathscr{C}} \nabla f|_{\mathscr{C}}\, dV = \int_{\partial^* \mathscr{C}} f \mathbf{n}_{\mathscr{C}}\, dA. \tag{II.1-7}$$

For a set $\mathscr{C}$ of finite perimeter $A(\partial\mathscr{C}\backslash\partial^*\mathscr{C}) \neq 0$ in general, and so in (7) the reduced boundary $\partial^*\mathscr{C}$ cannot be replaced by the topological boundary $\partial\mathscr{C}$.

Proofs of these theorems lie beyond the mathematical resources expected of students for whom this book is designed; comprehension of their meaning and importance for continuum mechanics does not.

In a major memoir NOLL & VIRGA[1] propose for a universe of shapes a subcollection of sets of finite perimeter. Their *fit regions* are subsets of $\mathscr{E}$ that are *regularly open*, *bounded*, *of finite perimeter*, and with *negligible boundary*.[2] The reduced boundary of a fit region $\mathscr{C}$ has the remarkable property $\partial\mathscr{C} = \mathrm{clo}(\partial^*\mathscr{C})$.

Fit regions provide examples of universes of shapes. For instance, NOLL & VIRGA proved that $\mathbf{\Omega}(\mathscr{C})$, the collection of all fit regions that are subsets of a given fit region $\mathscr{C}$, satisfies Axioms B1–B6 of Section I.2. Also $\bar{\mathbf{\Omega}}(\mathscr{C}) := \mathbf{\Omega}(\mathscr{C}) \cup \{\mathscr{C}, \varnothing\}$.

This book will concern mainly local analysis of the equilibrium and motion of continuous media. The student will rarely need to refer to the matters discussed just above. Nevertheless it would be dishonest as well as misleading to omit them, for otherwise he might gain the false notion that modern continuum mechanics lacks a precise mathematical formulation.

The conditions we have laid down for the shapes bodies may take on imply restrictions upon the structure of the bodies themselves. Because in experience we encounter bodies only in their shapes, specification of those suffices for efficient practice of the mathematical theory.

In this book we assume that the successive shapes of a body are bounded. With some technical detail it is possible to include also bodies whose shapes fill infinite regions. We shall sometimes describe motions of such bodies, as for example in the case of flow of a fluid body filling all of space or the region between parallel planes, but in the mathematical treatment we shall confine attention to some part whose shapes in some finite interval of time remain bounded, or we shall carry out a limit process with such parts. Unless the

[1] W. NOLL & E. G. VIRGA, "Fit regions and functions of bounded variation," *Archive for Rational Mechanics and Analysis* **102** (1988): 1–21.

[2] A subset $\mathscr{D}$ of $\mathscr{E}$ is *negligible* if for every positive real number ϵ it can be covered by a finite collection of balls, the sum of whose volumes does not exceed ϵ.

contrary is stated explicitly, from now on the term "body" will be taken to refer only to such parts.

Henceforth we shall restrict attention to shapes that are *fit regions*. Precisely, *we shall assume that each placement of a body* $\mathscr{B}$ *maps* $\mathscr{B}$ *onto a fit region of* $\mathscr{E}$, *and that, if* $\chi_1(\cdot, t)$ *and* $\chi_2(\cdot, t)$ *are placements of* $\mathscr{B}$, *then* $\chi_1 \circ \chi_2^{-1}$ *is a* C^1*-diffeomorphism*[1] *of* $\chi_2(\mathscr{B}, t)$ *onto* $\chi_1(\mathscr{B}, t)$. Since the class of all fit regions in $\mathscr{E}$ is invariant under C^1-diffeomorphism (see the paper cited above in Footnote 1 on p. 90), the above requirements on placements are consistent.

We have assumed also in Section I.7 that χ is differentiable as often as need be with respect to t, and we have defined the velocity $\mathbf{v}$, the acceleration $\mathbf{a}$, and the n^{th} velocity ${}_n\mathbf{v}$, $n \geqq 1$, as the values of the successive time derivatives for a given substantial point X at a given time t:

$$\begin{aligned} \mathbf{v} &= \dot{\chi}(X, t), \\ \mathbf{a} &= \ddot{\chi}(X, t), \ldots, \\ {}_n\mathbf{v} &= \overset{(n)}{\chi}(X, t); \end{aligned} \tag{I.7-9$_\mathrm{r}$}$$

the vector fields $\dot{\chi}(\cdot, t)$, $\ddot{\chi}(\cdot, t), \ldots$ are defined over $\mathscr{B}$ and have values in $\mathscr{V}$, the translation space of $\mathscr{E}$.

For most of our analysis it will suffice to assume that functions occurring in kinematical statements are twice continuously differentiable; sometimes once is enough.

As explained in Example 2 of Section I.3, the operations $\vee$ and $\wedge$ in the Boolean algebra of bodies are defined in continuum mechanics as follows:

$$\bigwedge_k \mathscr{B}_k := \operatorname{int}\operatorname{clo}\left(\bigcap_k \mathscr{B}_k\right), \tag{I.3-1$_\mathrm{r}$}$$

$$\bigvee_k \mathscr{B}_k := \operatorname{int}\operatorname{clo}\left(\bigcup_k \mathscr{B}_k\right), \tag{I.3-2$_\mathrm{r}$}$$

for any collection of bodies $\mathscr{B}_k$ in $\mathbf{\Omega}_{\mathrm{O}}$. Equivalently,

$$\chi\left(\bigwedge_k \mathscr{B}_k, t\right) = \operatorname{int}\operatorname{clo}\left(\bigcap_k \chi(\mathscr{B}_k, t)\right), \tag{II.1-8}$$

$$\chi\left(\bigvee_k \mathscr{B}_k, t\right) = \operatorname{int}\operatorname{clo}\left(\bigcup_k \chi(\mathscr{B}_k, t)\right), \tag{II.1-9}$$

for all $t \in \mathscr{I}$ provided $\chi(\cdot, t)$ be defined on $\bigvee_k \mathscr{B}_k$.

[1] A C^1-diffeomorphism is a continuously differentiable homeomorphism whose inverse is also continuously differentiable.

From (8) we see that the shapes of separate bodies are disjoint fit regions.

For the portions of the foregoing section that differ from the text of the first edition I am deeply indebted to E. VIRGA and W. O. WILLIAMS.

2. Mass-Density

Since in continuum mechanics $\mathscr{B}$ is a regularly open set, unless it is empty it contains infinitely many distinct substantial points. However, the assignment of mass M is left arbitrary so far and might be discrete, or partially so. Of primary interest in continuum mechanics are masses which are absolutely continuous functions of volume. To assume M absolutely continuous is to assume that if a part takes a shape having sufficiently small volume, then that part has arbitrarily small mass. Thus, formally, concentrated masses are excluded, and analytical dynamics will not emerge directly as a special case of continuum mechanics (though the two are always related through the theorem of KELVIN & TAIT in Section I.8 and through (I.13-14)).

Let $\boldsymbol{\sigma}$ be a placement of $\mathscr{B}$. By the Radon–Nikodym Theorem,[1] the mass of any massy part $\mathscr{P}$ of $\mathscr{B}$ may be expressed as the Lebesgue integral of a non-negative *mass-density* ρ_{σ} over the shape $\boldsymbol{\sigma}(\mathscr{P})$:

$$M(\mathscr{P}) = \int_{\boldsymbol{\sigma}(\mathscr{P})} \rho_{\sigma}\, dV. \tag{II.2-1}$$

The density ρ_{σ} exists and is unique almost everywhere in $\boldsymbol{\sigma}(\mathscr{P})$.

The existence of a mass-density expresses a relation between the body $\mathscr{B}$ and such shapes as it may assume. At almost every place $\mathbf{x}$ in $\boldsymbol{\sigma}(\mathscr{P})$ the density is the ultimate ratio of mass to volume in the following sense: If $\mathscr{P}_k$ is a suitably chosen sequence of nested parts, $\mathscr{P}_{k+1} \subset \mathscr{P}_k$, such that all the $\mathscr{P}_k$ have but the single substantial point $\boldsymbol{\sigma}^{-1}(\mathbf{x})$ in common, and that $V(\boldsymbol{\sigma}(\mathscr{P}_k)) \to 0$ as $k \to \infty$, then

$$\rho_{\sigma}(\mathbf{x}) = \lim_{k\to\infty} \frac{M(\mathscr{P}_k)}{V(\boldsymbol{\sigma}(\mathscr{P}_k))}. \tag{II.2-2}$$

In all its shapes a part $\mathscr{P}$ has the same mass $M(\mathscr{P})$. We have made this assumption plain by assigning masses directly to the massy parts of $\mathscr{B}$. To each shape of $\mathscr{B}$ we may apply (1). Thus,

$$M(\mathscr{P}) = \int_{\boldsymbol{\sigma}_1(\mathscr{P})} \rho_{\sigma_1}\, dV = \int_{\boldsymbol{\sigma}_2(\mathscr{P})} \rho_{\sigma_2}\, dV. \tag{II.2-3}$$

[1] For example, Theorem 6.9 of W. RUDIN's book cited above in Footnote 2 on p. 18.

Since both $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are placements of $\mathscr{B}$, $\boldsymbol{\sigma}_2 \circ \boldsymbol{\sigma}_1^{-1}$ is a C^1-diffeomorphism of $\boldsymbol{\sigma}_1(\mathscr{B})$ onto $\boldsymbol{\sigma}_2(\mathscr{B})$. If $\boldsymbol{\gamma} := \boldsymbol{\sigma}_2 \circ \boldsymbol{\sigma}_1^{-1}$, the chain rule of differential calculus shows that $(\nabla\boldsymbol{\gamma})(\nabla\boldsymbol{\gamma}^{-1}) = \mathbf{1}$, and so $\det\nabla\boldsymbol{\gamma} \times \det\nabla\boldsymbol{\gamma}^{-1} = 1$. Therefore, neither $\det\nabla\boldsymbol{\gamma}$ nor $\det\nabla\boldsymbol{\gamma}^{-1}$ can vanish. Thus if

$$J := |\det\nabla\boldsymbol{\gamma}|, \tag{II.2-4}$$

then

$$J > 0. \tag{II.2-5}$$

A theorem of integral calculus[1] tells us that for a measurable function f

$$\int_{\boldsymbol{\sigma}_2(\mathscr{P})} f(\mathbf{x})\,dV(\mathbf{x}) = \int_{\boldsymbol{\sigma}_1(\mathscr{P})} f(\boldsymbol{\gamma}(\mathbf{X}))J(\mathbf{X})\,dV(\mathbf{X}) \tag{II.2-6}$$

for each massy part $\mathscr{P}$ of $\mathscr{B}$. By applying this statement to (3) we obtain an equation relating the two densities almost everywhere:

$$\rho_{\boldsymbol{\sigma}_2}(\boldsymbol{\gamma}(\mathbf{X}))J(\mathbf{X}) = \rho_{\boldsymbol{\sigma}_1}(\mathbf{X}), \qquad \mathbf{X} \in \boldsymbol{\sigma}_1(\mathscr{P}). \tag{II.2-7}$$

Thus the mass-density field over one shape of $\mathscr{B}$ determines the mass-density fields over all others to within a set of null volume. These qualities apply to the shapes $\mathscr{B}$ assumes when it undergoes a motion.

As in Section I.4, integration with respect to mass is defined on the massy parts $\mathscr{P}$ of a body $\mathscr{B}$. In continuum mechanics the assumption that M is an absolutely continuous function of V enables us to replace all integrals so defined by counterparts taken over regions of $\mathscr{E}$. Thus,

$$\int_{\mathscr{P}} f\,dM = \int_{\chi(\mathscr{P},\,t)} \rho f\,dV. \tag{II.2-8}$$

On the left-hand side, f stands for $f(X, t)$, while on the right-hand side, f stands for $f(\chi^{-1}(\mathbf{x}, t), t)$ and ρ is written for ρ_χ. Alternatively, we may start

[1] The formula (6), inferred by a formal or pictorial argument, derives from EULER's researches on hydrodynamics in the middle of the eighteenth century. A clear and simple statement and proof within the theory of Riemann integration is given in Theorems 3.13 and 3.14 of M. SPIVAK, *Calculus on Manifolds*, New York, Benjamin, 1965; for Lebesgue integrals, in Theorem 8.26 in the book of W. RUDIN, cited in Footnote 2 on p. 18. In both cases the integrand is merely assumed integrable, and $\boldsymbol{\gamma}$ is assumed to be a bijective, differentiable mapping of an open set of $\mathscr{E}_n$ into $\mathscr{E}_n$. For the former theorem, $\boldsymbol{\gamma}$ is assumed continuously differentiable; for the latter, $\boldsymbol{\gamma}^{-1}$ is assumed continuous, and Range $\boldsymbol{\gamma}$ is assumed open and bounded. While both of these assumptions allow J to vanish on a set of measure 0, the conditions we assume deliver (II.2-5).

with the right-hand side and regard f as standing for $f(\mathbf{x}, t)$, so that on the left-hand side f stands for $f(\chi(X, t), t)$. This abbreviated notation, which is common in continuum mechanics, will be developed further in Sections II.4 and II.6.

In this book we shall always assume not only that mass is an absolutely continuous function of volume but even that mass is ultimately bounded by volume: For any placement $\boldsymbol{\sigma}$ of $\mathscr{B}$ there is a constant K, which depends upon $\boldsymbol{\sigma}$, such that *if* $V(\boldsymbol{\sigma}(\mathscr{P}))$ *is sufficiently small, then*

$$M(\mathscr{P}) \leqq KV(\boldsymbol{\sigma}(\mathscr{P})). \tag{II.2-9}$$

Equivalently, $\rho_{\boldsymbol{\sigma}}$ is essentially bounded. Any additive set function that is bounded with respect to volume is bounded also with respect to mass, and conversely. In passages where the manipulations of differential calculus are brought to bear, we shall presume the still stronger assumption that $\rho_{\boldsymbol{\sigma}}$ is a continuously differentiable function of its arguments at all places and times we may choose to consider.

3. Reference Placement. Transplacement

Often it is convenient to select the placement of $\mathscr{B}$ at some one time t in some putative motion, not necessarily the motion χ being studied, and to refer everything concerning $\mathscr{B}$ and its motion to that placement, which we shall call the *reference placement*. We denote by $\mathbf{X}$ the place given to the substantial point X by the reference placement $\boldsymbol{\kappa}$:

$$\mathbf{X} = \boldsymbol{\kappa}(X). \tag{II.3-1}$$

Since $\boldsymbol{\kappa}$, by assumption, is invertible,

$$X = \boldsymbol{\kappa}^{-1}(\mathbf{X}), \tag{II.3-2}$$

and both $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}^{-1}$ are continuous. Hence the motion (I.7-7) may be written in the form

$$\mathbf{x} = \chi(\boldsymbol{\kappa}^{-1}(\mathbf{X}), t) =: \chi_{\boldsymbol{\kappa}}(\mathbf{X}, t). \tag{II.3-3}$$

In the description furnished by this equation, the motion is expressed as a mapping $\chi_{\boldsymbol{\kappa}}$ of the *reference shape* $\boldsymbol{\kappa}_{\Omega}(\mathscr{B})$ onto the actual shapes $\chi(\mathscr{B}, t)$ as t

progresses or regresses. Thus the motion, which is defined by (I.7-7) as mapping substantial points onto places in space, is now represented as mapping spatial regions onto spatial regions. A reference shape is introduced so as to allow us to employ immediately the apparatus of Euclidean geometry. The mapping $\boldsymbol{\chi}_{\boldsymbol{\kappa}}$ is the *transplacement* of the substantial points of $\mathscr{B}$ from their reference places $\mathbf{X}$ into their actual places $\mathbf{x}$ at the time t. By the assumptions on placements made in Section II.1, the transplacements are homeomorphisms of the reference shape onto present shapes. Now we assume more: each $\boldsymbol{\chi}_{\boldsymbol{\kappa}}$ is a C^1-diffeomorphism.

With no fear of confusion we may drop the subscript $\boldsymbol{\Omega}$ and write $\boldsymbol{\kappa}(\mathscr{B})$ for the reference shape of $\mathscr{B}$, just as we have already written $\boldsymbol{\chi}(\mathscr{B}, t)$ for $\boldsymbol{\chi}_{\boldsymbol{\Omega}}(\mathscr{B}, t)$. As in ordinary language a body is "deformed" when its shape changes, we shall say that the transplacement $\boldsymbol{\chi}_{\boldsymbol{\kappa}}$ *deforms* the reference shape $\boldsymbol{\kappa}(\mathscr{B})$ into the actual shape $\boldsymbol{\chi}(\mathscr{B}, t)$. While "strain" is commonly used to denote deformation or some aspect of it, in this book we do not give any precise meaning to the word, but we shall use it descriptively from time to time.

The choice of reference placement, like the choice of a co-ordinate system, is arbitrary. The reference placement, which may be any smooth mapping of $\mathscr{B}$ into $\mathscr{E}$, need not be the value of the motion $\boldsymbol{\chi}(\mathscr{B}, \cdot)$ at any time t_0. If it is, then $\boldsymbol{\chi}(\cdot, t_0) = \boldsymbol{\kappa}$. In the treatment of surface waves in Volume 2 we shall encounter a classic example in hydrodynamics for which it is preferable to use a reference placement that is never occupied by the body considered.

For each different $\boldsymbol{\kappa}$, a different transplacement $\boldsymbol{\chi}_{\boldsymbol{\kappa}}$ for the same motion $\boldsymbol{\chi}$ is defined by (3). Thus one motion of the body is represented by infinitely many different mappings of parts of space in the course of time, one for each choice of $\boldsymbol{\kappa}$. For some choice of $\boldsymbol{\kappa}$ we may get a particularly simple description, just as in geometry one choice of co-ordinates may lead to a simple equation for a particular figure while another may not, but the reference placement itself has nothing to do with such motions as it may be used to describe, just as the co-ordinate system has nothing to do with geometrical features themselves. A reference placement is introduced so as to allow the use of mathematical apparatus familiar in other contexts. Again there is an analogy to co-ordinate geometry, where co-ordinates are introduced, not because they are natural or germane to geometry, but because they allow the familiar apparatus of algebra to be applied at once.

4. Descriptions of Motion: Substantial, Referential, and Spatial

There are four methods of describing the motion of a body: the substantial, the referential, the spatial, and the relative. Because of our hypotheses of smoothness, all are equivalent.

In the *substantial description* we deal directly with the substantial points X. This description extends the only one used in analytical dynamics, where we always speak of the first, second, ..., n^{th} substantial points, which are usually called masses. To be precise, there we should say, "the mass-point X_q whose mass is M_q," but commonly this expression is abbreviated to "the mass q" or "the body M_q," *etc.* In continuum mechanics every body $\mathscr{B}$ comprises infinitely many substantial points X. The substantial description employs as independent variables X and t, the substantial point and the time. While the substantial description is the most natural in concept, it was not mentioned in continuum mechanics until a few decades ago, was then called "material", and is still used little. With the substantial description for continua, strictly interpreted, few analytical tools are at hand. For some time the term "material description" was used to denote another and older description often confused with it, the description to which we turn next.

The *referential description* employs some assigned reference placement $\boldsymbol{\kappa}$. Thus it describes the motion χ by means of the transplacement $\chi_{\boldsymbol{\kappa}}$. We must always bear in mind that the choice of $\boldsymbol{\kappa}$ is ours, that $\boldsymbol{\kappa}(\mathscr{B})$ is merely some shape that $\mathscr{B}$ has occupied or might occupy, and that it must be possible always to state hypotheses and equations in forms valid for any choice of $\boldsymbol{\kappa}$, although for one choice of $\boldsymbol{\kappa}$ the corresponding transplacement $\chi_{\boldsymbol{\kappa}}$ may show the important properties of some particular motion more easily than do the transplacements corresponding with other choices of $\boldsymbol{\kappa}$. Any motion of a body has infinitely many different referential descriptions, equally valid.

For the purposes of this book, and for most purposes in mechanics, the substantial description and the referential description may be confused, at least locally, as they long have been. To see that they are in principle different, and that the referential description may not always suffice, we need only consider the two-body problem of analytical dynamics. No one would find it convenient to use as labels for the first and second mass-points the places they occupied at some particular time. If that time were one at which the two mass-points collided, such names would not distinguish those two bodies. Since analytical dynamics always envisions the chance that collisions may occur, the distinction between substantial and referential descriptions is not a matter of purism or mere abstraction, and in fact nobody has ever employed the referential description in analysis of the motions of discrete systems. The referential description is useful only for systems in which it is convenient to use a place as a name for an element of an abstract manifold. Such naming is indeed convenient in continuum mechanics.

In the mid-eighteenth century EULER introduced the description that hydrodynamicists still call "Lagrangean". This is a particular referential description, in which the cartesian co-ordinates of the position $\mathbf{X}$ of the body-point X at the time $t = 0$ are used as a label for that body-point. It was recognized that such labelling by initial co-ordinates was arbitrary; writers on the foundations of hydrodynamics have often mentioned that the essential conclusions must be and are independent of the choice of the initial time, and some have remarked that the parameters of any triple system of surfaces moving in such a way as to be at all times the sites of the substantial points on them at any one time would do just as well. The referential description, taking $\mathbf{X}$ and t as independent variables, includes all these possibilities. Some form of it is always used in classical elasticity theory, and the best studies of the foundations of classical hydrodynamics from EULER's day to the present have employed it almost without fail. It is the description commonly used in modern works on continuum mechanics, and we shall use it in this book.

In view of (II.3-2), any function $F(X, t)$ may be replaced by a function $F_{\boldsymbol{\kappa}}(\mathbf{X}, t)$ that has the same value at corresponding arguments X and $\mathbf{X}$, for given $\boldsymbol{\kappa}$:

$$F(X, t) = F(\boldsymbol{\kappa}^{-1}(\mathbf{X}), t) =: F_{\boldsymbol{\kappa}}(\mathbf{X}, t). \tag{II.4-1}$$

In (II.3-3) we have already encountered a special instance. Moreover,

$$\partial_t F = \partial_t F_{\boldsymbol{\kappa}} \tag{II.4-2}$$

at the respective arguments X and $\boldsymbol{\kappa}(X)$. We shall employ a superimposed dot to denote also time derivatives of functions of the *referential variables* $\mathbf{X}$ and t. Thus, by differentiating (II.3-3) and using the definitions (I.7-7) and (II.3-2), we see that for each choice of $\boldsymbol{\kappa}$ and at each time

$$\dot{\chi} = \dot{\chi}_{\boldsymbol{\kappa}}, \qquad \ddot{\chi} = \ddot{\chi}_{\boldsymbol{\kappa}}, \qquad \ldots, \qquad \overset{(n)}{\chi} = \overset{(n)}{\chi}_{\boldsymbol{\kappa}}, \tag{II.4-3}$$

the arguments of the functions on the left-hand sides being X and t, those of the functions on the right-hand sides being the corresponding $\mathbf{X}$ and t.

In the *spatial description*, attention is directed to the present shape of the body. This description, which was introduced by DANIEL BERNOULLI and D'ALEMBERT, is called "Eulerian" by the hydrodynamicists. The place $\mathbf{x}$ and the time t are taken as independent variables. In view of (I.7-7), any function $F(X, t)$ may be replaced by a function of the *spatial variables*, $\mathbf{x}$ and t, that has the same value at corresponding arguments X and $\mathbf{x}$:

$$F(X, t) = F[\chi^{-1}(\mathbf{x}, t), t] =: f(\mathbf{x}, t). \tag{II.4-4}$$

The function f, moreover, is unique. Thus, while there are infinitely many referential descriptions of a given motion, there is only one spatial description, just as there is only one substantial description. In the spatial description the velocity field is defined on $\chi(\mathscr{B}, t)$. As t changes, generally the shape of $\mathscr{B}$ changes. With the spatial description, we watch what occurs in a fixed region of space that remains within the successive shapes of $\mathscr{B}$. This description seems perfectly suited to studies of fluids, where often a rapidly deforming mass comes no-one knows whence and goes no-one knows whither, so that we may prefer to consider what happens here and now before our eyes. In many problems of hydrodynamics the boundary $\partial\chi(\mathscr{B}, t)$ remains fixed, making the spatial description especially suitable.

However convenient kinematically, the spatial description is awkward for questions of principle in mechanics, since in fact the laws of dynamics refer to what is suffered by the body, not by the region of space the body momentarily occupies. Some relations obvious and easy to derive in the substantial or referential descriptions seem to require contorted reasoning if approached by the strictly spatial standpoint sometimes adopted by specialists in applied hydrodynamics.

According to (4), the value of any function of the substantial points of $\mathscr{B}$ at the time t is given also by a field defined over the actual shape $\chi(\mathscr{B}, t)$. In this way, for example, we obtain from (I.7-9) the *velocity field* $\dot{\mathbf{x}}$, the *acceleration field* $\ddot{\mathbf{x}}$, and the n^{th} *velocity field* $\overset{(n)}{\mathbf{x}}$:

$$\mathbf{v} = \dot{\mathbf{x}}(\mathbf{x}, t), \qquad \mathbf{a} = \ddot{\mathbf{x}}(\mathbf{x}, t), \qquad \ldots, \qquad {}_n\mathbf{v} = \overset{(n)}{\mathbf{x}}(\mathbf{x}, t). \tag{II.4-5}$$

The fields $\overset{(n)}{\mathbf{x}}$ and $\overset{(n)}{\chi}$ have the common value ${}_n\mathbf{v}$ at arguments related through the motion:

$$\begin{aligned} {}_n\mathbf{v} &= \overset{(n)}{\mathbf{x}}(\chi(X, t), t) = \overset{(n)}{\chi}(X, t), \qquad X \in \mathscr{B} \\ &= \overset{(n)}{\mathbf{x}}(\mathbf{x}, t) = \overset{(n)}{\chi}(\chi^{-1}(\mathbf{x}, t), t), \qquad \mathbf{x} \in \chi_{\mathfrak{g}}(\mathscr{B}, t). \end{aligned} \tag{II.4-6}$$

In Section I.11 we have calculated the frame-indifferent field $\boldsymbol{\alpha}_{\oint}$ that reduces in the inertial frame $\oint$ to the acceleration. While as given by the right-hand side of (I.11-3) this field is defined over the body $\mathscr{B}$, of course we may convert it into a field over the actual shape $\chi^*(\mathscr{B}, t)$ in the general rigid frame $\oint^*$. Calling that field $\mathbf{a}_{\oint}$, we calculate it as follows from (I.11-3):

$$\mathbf{a}_{\oint} = \ddot{\mathbf{x}}^* - \ddot{\mathbf{x}}_0^* - 2\mathbf{A}(\dot{\mathbf{x}}^* - \dot{\mathbf{x}}_0^*) - (\dot{\mathbf{A}} - \mathbf{A}^2)(\mathbf{x}^* - \mathbf{x}_0^*); \tag{II.4-7}$$

here $\mathbf{x}^*$, $\dot{\mathbf{x}}^*$, and $\ddot{\mathbf{x}}^*$ are the spatial fields of place, velocity, and acceleration over the actual shape $\boldsymbol{\chi}^*(\mathscr{B}, t)$ in the frame $\oint^*$, and $\mathbf{A}$ is the spin of the inertial frame $\oint$ with respect to $\oint^*$.

The velocity field of a motion is often called a *flow*.

In the spatial description a function of place alone is called a *steady* function. For example, a velocity field $\dot{\mathbf{x}}$ that is independent of t is called a *steady flow*. Other flows are called *unsteady*. A steady flow may or may not have a steady density.

A flow that is steady in one frame generally fails to be steady in another. The property of steadiness is not even a galilean invariant. It is not a simple matter to determine whether a given flow that is not steady in the frame in which it is defined be steady in some other frame. *Cf.* CFT, Section 146.

A point at which $\dot{\mathbf{x}} = \mathbf{0}$ is called a *stagnation point*.

In the spatial description we may *superpose* the flows $\dot{\mathbf{x}}_1$ and $\dot{\mathbf{x}}_2$ in their common domain at each t, so obtaining a new flow:

$$\dot{\mathbf{x}}_{1+2} := \dot{\mathbf{x}}_1 + \dot{\mathbf{x}}_2 = \dot{\mathbf{x}}_2 + \dot{\mathbf{x}}_1 \, . \tag{II.4-8}$$

We may think of superposition as arising from adding the values of the vectors $\dot{\boldsymbol{\chi}}_1(X, t)$ and $\dot{\boldsymbol{\chi}}_2(X, t)$ for one and the same substantial point X at the time t, but it is not usual to do so.

The fourth common description of motion, called "relative", we shall develop in Section II.8.

5. Transplacement Gradient

The gradient of the transplacement $\boldsymbol{\chi}_{\boldsymbol{\kappa}}$ at a given t is called the *transplacement gradient*[1] $\mathbf{F}$:

$$\mathbf{F} := \mathbf{F}_{\boldsymbol{\kappa}}(\mathbf{X}, t) := \nabla\boldsymbol{\chi}_{\boldsymbol{\kappa}}(\mathbf{X}, t). \tag{II.5-1}$$

It is the linear approximation to the mapping (II.3-3) in a neighborhood of $\mathbf{X}$. More precisely, we should call it the gradient of the transplacement from $\boldsymbol{\kappa}$ to $\boldsymbol{\chi}$, but when, as is usual, a single reference placement $\boldsymbol{\kappa}$ is laid down once and for all, no confusion should result from failure to remind ourselves that the very concepts of transplacement and transplacement gradient presume use of a reference placement. If, as we may, we select independently co-ordinates X^α

[1]The reader ought not confuse $\mathbf{F}$ or $\mathbf{F}_{\boldsymbol{\kappa}}$ with the torque $\mathbf{F}_{\mathbf{x}_0}$ of a system of forces with respect to $\mathbf{x}_0$.

and x^m in the reference shape and the actual shape, respectively, the motion (II.3-3) is expressed as follows:

$$x^m = \chi_\kappa^m(X^1, X^2, X^3, t), \qquad m = 1, 2, 3, \tag{II.5-2}$$

and then the components of $\mathbf{F}$ are simply the nine partial derivatives of the functions χ_κ^m with respect to the X^α, *viz*

$$F^m_\alpha = x^m_{,\alpha} = \partial_{X^\alpha}\chi_\kappa^m(X^1, X^2, X^3, t), \qquad m = 1, 2, 3, \quad \alpha = 1, 2, 3. \tag{II.5-3}$$

There would be no loss in logical strictness were we to write out everything, as the older authors on continuum mechanics did, in cartesian co-ordinates. In practice, abstract notations are easier to understand and more efficient to manipulate, once they be grown familiar, and proofs using them are easier to follow. Particular applications often refer to certain particular directions and hence suggest use of a particular basis, which need not be the natural basis of any co-ordinate system (*cf.* the end of Section App. IIC.7). Thus it is to our advantage to express all the principles of our science directly in terms of the concepts of algebra and geometry, without the complicating intermediacy of co-ordinate systems.

In Section II.2 we have introduced the mass-density ρ_σ that corresponds to the placement $\sigma(\mathscr{B})$. Henceforth we shall write simply ρ for ρ_χ ; thus ρ is the mass-density field over the actual shape $\chi(\mathscr{B}, t)$. Choosing for σ_1 in (II.2-7) the reference placement κ, we obtain

$$\rho J = \rho_\kappa , \tag{II.5-4}$$

on the understanding that when the argument of ρ_κ is $\mathbf{X}$, the arguments of ρ and J are $\chi_\kappa(\mathbf{X}, t)$ and t, and that

$$J := |\det \mathbf{F}|. \tag{II.5-5}$$

We shall use J in the sense just defined rather than in the more general one expressed by (II.2-4). The relation (4) is EULER's *referential equation for the mass-density*. If $\mathscr{B}$ occupies at some time the reference placement selected for it, then the value of $\mathbf{F}$ at that time is $\mathbf{1}$, and so $\det \mathbf{F} = 1$ then. In that case $\det \mathbf{F} > 0$ always, and the bars may be dropped from (5).

The student will recall that

$$J > 0. \tag{II.2-5$_\mathrm{r}$}$$

While (4) is often called "the Lagrangean equation of continuity", that name is doubly misleading, since if the transplacement is smooth enough, (4) holds, but if the transplacement is not differentiable, let alone not continuous, J cannot be defined at all, and so (4) cannot even be stated, let alone used. Obviously (4) is neither more nor less than a formula that delivers the actual density ρ, once the transplacement gradient $\mathbf{F}$ and the reference density ρ_κ be known.

In the older literature (4) is sometimes related to an "axiom of impenetrability", according to which two distinct substantial points never come to occupy the same place, and thus no body enters into the shape of another body at the same time (*cf.* Section I.7). In truth, on the contrary, a formal condition such as (4) does not express that axiom but rather presumes that some such axiom has been laid down already.

Exercise II.5.1 (EULER, LIOUVILLE).

$$\operatorname{div}\dot{\mathbf{x}} = \dot{J}/J = \frac{d(\det \mathbf{F})}{dt} \Big/ \det \mathbf{F}, \tag{II.5-6}$$

in which the superimposed dot on the middle member denotes the time derivative, and $\operatorname{div}\dot{\mathbf{x}}$ is the divergence of the velocity field $(\text{II}.4\text{-}5)_1$.

In (6), as in (4), the field on the right-hand side is a referential one, while that on the left-hand side is a spatial one. Both conclusions assert that fields of these two kinds have at time t the same values at the places $\mathbf{X}$ and $\mathbf{x}$, respectively, selected so as to correspond with each other through the referential description (II.3-3) of the motion. It leads to less awkward statements if in such cases we simply presume that any referential field is replaced by the corresponding spatial one. For example, if we differentiate (4) with respect to time and then use (6), we obtain D'ALEMBERT and EULER's *spatial equation for the density*:

$$\dot{\rho} + \rho \operatorname{div}\dot{\mathbf{x}} = 0; \tag{II.5-7}$$

in which we follow the convention, as we did for (4), that allows us to interpret ρ as being the mass-density field over the present shape $\chi(\mathscr{B}, t)$. This equation has exactly the same meaning as (4), which, conversely, may be gotten from it by use of (6) followed by integration.

Exercise II.5.2 (LAGRANGE). If (7) is taken as a first-order differential equation for ρ in the spatial description, integration by the method of characteristics yields (4).

Exercise II.5.3 (D'ALEMBERT, EULER). A motion of $\mathscr{B}$ is called *isochoric* if the volume $V(\chi(\mathscr{P}, t))$ of the shape of each part $\mathscr{P}$ of $\mathscr{B}$ remains constant in time. Any one of the following three conditions is necessary and sufficient for isochoric motion:

1. $\operatorname{div}\dot{\mathbf{x}} = 0.$ (II.5-8)
2. There is a reference placement $\boldsymbol{\kappa}$ such that

$$\rho = \rho_{\boldsymbol{\kappa}} . \tag{II.5-9}$$

3. There is a reference placement $\boldsymbol{\kappa}$ such that

$$J = 1. \tag{II.5-10}$$

In *plane flow* the velocity is everywhere parallel to a given plane and is the same at all points on each line normal to that plane. To study plane flow, it suffices to confine attention to the fields of velocity and acceleration restricted to some one plane.

Exercise II.5.4 (D'ALEMBERT, NOLL). If the boundary of the region on which a plane flow $\dot{\mathbf{x}}$ is defined is the union of a finite number of curves in rigid motion, and if (for an infinite region) there is no flux into or out of the region of the plane beyond some sufficiently large circle, the general solution of (8) is given in terms of a *stream function* q by

$$\dot{\mathbf{x}} = (\nabla q)^{\perp} , \tag{II.5-11}$$

∇ denoting the gradient operator in the plane and $\perp$ denoting rotation counter-clockwise through a right angle about the normal to the plane. The velocity $\dot{\mathbf{x}}(\mathbf{x}, t)$ is tangent to the curve $q(\cdot, t) = \text{const.}$ through $\mathbf{x}$ at each t.

Exercise II.5.5. Given a vector field $\mathbf{v}$ defined on $\chi(\mathscr{B}, t)$, let it be desired to find a vector field $\mathbf{v}_{\boldsymbol{\kappa}}$ such that

$$\int_{\mathscr{S}} \mathbf{v}_{\boldsymbol{\kappa}} \cdot \mathbf{n}_{\boldsymbol{\kappa}} \, dA = \int_{\chi_{\boldsymbol{\kappa}}(\mathscr{S}, t)} \mathbf{v} \cdot \mathbf{n} \, dA \tag{II.5-12}$$

for any surface $\mathscr{S}$ in $\boldsymbol{\kappa}(\mathscr{B})$. Then

$$\mathbf{v}_{\boldsymbol{\kappa}} = J\mathbf{F}^{-1}\mathbf{v}. \tag{II.5-13}$$

Points at which $\rho = 0$ are generally unusual. From now on we shall often tacitly assume that

$$\rho > 0 \tag{II.5-14}$$

and hence be able to use freely the *specific volume* υ:

$$\upsilon := \frac{1}{\rho}, \tag{II.5-15}$$

which also is positive.

6. Substantial Time Rates and Gradients in the Spatial Description. Substantial Surfaces. Kinematic Boundaries

In continuum mechanics the need to distinguish a vast number of quantities often deprives us of the clarity gained by using for a function a symbol different from that for its value, as logically we ought to do. If two functions of different variables have the same value for properly corresponding arguments, and if *both* are denoted by that value, when we come to effect some functional operation it is not clear which function is intended. The distinction, which of course is essential, is traditionally made by introducing different symbols for the differential operators. Henceforth

$$\dot{f} \qquad \text{and} \qquad \operatorname{Grad} f$$

shall denote the partial time derivative and the gradient of the function $G(\mathbf{X}, t)$ at a given t such that

$$f = G(\mathbf{X}, t), \tag{II.6-1}$$

while

$$f' \qquad \text{and} \qquad \operatorname{grad} f$$

shall denote the partial time derivative and the gradient of the function $g(\mathbf{x}, t)$ at a given t that has the same value as G, namely,

$$f = g(\mathbf{x}, t) = G(\chi_\kappa^{-1}(\mathbf{x}, t), t), \tag{II.6-2}$$

by (II.3-3). If we apply the chain rule to the equation $G(\mathbf{X}, t) = g(\chi_\kappa(\mathbf{X}, t), t)$ and then denote by f *both* functions G and g, we obtain the classical formulae of Euler:

$$\begin{aligned} \dot{f} &= f' + (\operatorname{grad} f)\cdot\dot{\mathbf{x}}, \\ \dot{\mathbf{f}} &= \mathbf{f}' + (\operatorname{grad} \mathbf{f})\dot{\mathbf{x}}; \end{aligned} \tag{II.6-3}$$

the values of f are scalars, of $\mathbf{f}$, vectors, and an analogous rule holds for functions whose values are tensors. In particular, the acceleration field $\ddot{\mathbf{x}}$ is calculated from the velocity field $\dot{\mathbf{x}}$ by the D'ALEMBERT–EULER formula

$$\ddot{\mathbf{x}} = \dot{\mathbf{x}}' + (\operatorname{grad}\dot{\mathbf{x}})\dot{\mathbf{x}}. \tag{II.6-4}$$

The dot operator as defined by (3) is called the *substantial derivative.*[1] We have already agreed to use the dot to denote the time derivative in the substantial and referential descriptions, and the definition (3) has been framed so as to render the two usages consistent with each other.

Likewise,

$$\operatorname{Grad} f = \mathbf{F}^{\mathsf{T}} \operatorname{grad} f. \tag{II.6-5}$$

The notations div and Div shall stand for the traces of grad and Grad, respectively.

We have already introduced an instance of these conventions in (II.5-6) and (II.5-7). For example, by $(3)_1$ the latter equation may be written explicitly in the forms

$$\rho' + (\operatorname{grad}\rho)\cdot\dot{\mathbf{x}} + \rho\operatorname{div}\dot{\mathbf{x}} = 0, \qquad \rho' + \operatorname{div}(\rho\dot{\mathbf{x}}) = 0. \tag{II.6-6}$$

For a motion with steady density $(6)_2$ reduces to $\operatorname{div}(\rho\dot{\mathbf{x}}) = 0$.

In Section II.2 we have shown how to convert integration with respect to mass on $\mathscr{B}$ into integration with respect to volume on the shape of $\mathscr{B}$. If $f(X, t)$ is continuously differentiable with respect to t, the theorem on differentiation of an integral with respect to a parameter assures us that

$$\frac{d}{dt}\int_{\mathscr{P}} f\,dM = \int_{\mathscr{P}} \partial_t f\,dM. \tag{II.6-7}$$

We may now use (II.2-8) to convert the right-hand side into an integral over $\chi(\mathscr{P}, t)$. According to the convention of notation just established, the function of $\mathbf{x}$ and t whose value is $\partial_t f(X, t)$ at $\chi^{-1}(\mathbf{x}, t)$ is denoted by $\dot{f}$, and (3) provides us means to calculate $\dot{f}$ from data in the spatial description. Thus

$$\frac{d}{dt}\int_{\mathscr{P}} f\,dM = \int_{\chi(\mathscr{P}, t)} \rho\dot{f}\,dV. \tag{II.6-8}$$

[1] The substantial derivative is only one of many rates that may be calculated on the basis of a given time-dependent field such as a velocity field. Others are introduced below by (II.13-4) and (II.13-7). The problem is discussed from a general point of view by H. BOLDER, "Deformation of tensor fields described by time-dependent mappings," *Archive for Rational Mechanics and Analysis* **35** (1969): 321–341.

More generally, if Ψ denotes a tensor field of any order,

$$\frac{d}{dt}\int_{\mathscr{P}} \Psi \, dM = \frac{d}{dt}\int_{\chi(\mathscr{P},t)} \rho \Psi \, dV = \int_{\chi(\mathscr{P},t)} \rho \dot{\Psi} \, dV, \qquad \text{(II.6-9)}$$

and $\dot{\Psi}$ is to be calculated by an appropriate rule of the type (3). (The central expression, which involves an undefined operation d/dt, is to be regarded only as a suggestive way of writing the left-hand expression.) The commutation formula (9) is used so often in continuum mechanics that it is taken for granted without special reference. It expresses the time-rate of change of the integral of Ψ over a body $\mathscr{B}$ as that body moves through space, in terms of an integral over the present shape $\chi(\mathscr{B}, t)$ of $\mathscr{B}$.

Exercise II.6.1. Simple rearrangement of (9), supplemented by use of (II.5-7), delivers the *Reynolds Transport Theorem*: For a given part $\mathscr{P}$ of $\mathscr{B}$,

$$\left(\int_{\chi(\mathscr{P},t)} \mathfrak{U}\, dV\right)^{\cdot} = \left(\int_{\chi(\mathscr{P},t)} \mathfrak{U}\, dV\right)' + \int_{\partial\chi(\mathscr{P},t)} \mathfrak{U}\dot{\mathbf{x}}\cdot\mathbf{n}\, dA, \qquad \text{(II.6-10)}$$

the notations being defined as follows:

$$\begin{aligned} \left(\int_{\chi(\mathscr{P},t)} \mathfrak{U}\, dV\right)^{\cdot} &:= \frac{d}{dt}\int_{\mathscr{P}} \frac{\mathfrak{U}}{\rho(\chi(X,t),t)}\, dM, \\ \left(\int_{\chi(\mathscr{P},t)} \mathfrak{U}\, dV\right)' &:= \int_{\chi(\mathscr{P},t)} \mathfrak{U}'\, dV. \end{aligned} \qquad \text{(II.6-11)}$$

Thus a substantial derivative is expressed in terms of a local time-derivative and flux through a boundary. In particular,

$$\begin{aligned} \dot{\mathbf{m}} &= \mathbf{m}' + \int_{\partial\chi(\mathscr{P},t)} \rho\dot{\mathbf{x}}\,\dot{\mathbf{x}}\cdot\mathbf{n}\, dA, \\ \dot{\mathbf{M}}_{\mathbf{x}_0} &= \mathbf{M}'_{\mathbf{x}_0} + \int_{\partial\chi(\mathscr{P},t)} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\dot{\mathbf{x}}\,\dot{\mathbf{x}}\cdot\mathbf{n}\, dA. \end{aligned} \qquad \text{(II.6-12)}$$

A stationary surface $\mathscr{S}_\kappa$ in the reference shape $\kappa(\mathscr{B})$ is described by an equation of the form $f(\mathbf{X}) = 0$, and hence

$$\dot{f} = 0. \qquad \text{(II.6-13)}$$

Conversely, if (13) is satisfied by a function $f(\mathbf{X}, t)$, then in fact the surface $f = 0$ is a stationary surface in the reference shape, provided of course that

$\mathbf{X} \in \boldsymbol{\kappa}(\mathscr{B})$. At the time t the substantial points that make up $\mathscr{S}_{\kappa}$ constitute a certain surface $\mathscr{S}$ in the shape assumed by $\mathscr{B}$ in its motion at the time t. These surfaces are the successive forms of a single *substantial surface*. In accord with the convention we have established, we write f also for the function of $\mathbf{x}$ and t whose value at $\boldsymbol{\chi}_{\kappa}(\mathbf{X}, t)$ is $f(\mathbf{X})$, and so in order for the locus $f = 0$ to represent a substantial surface we have the necessary and sufficient condition (13), where now the operation signified by a dot is defined by $(3)_1$. Thus in the spatial description this requirement becomes EULER's condition:

$$f' + (\operatorname{grad} f)\cdot\dot{\mathbf{x}} = 0. \tag{II.6-14}$$

If $\mathbf{n}$ is the oriented unit normal to the surface $f = 0$, where of course f now stands for the function such that $f(\mathbf{x}, t) = 0$ is the locus of $\mathscr{S}$, then (14) may be written alternatively in the form

$$S_n = \mathbf{n}\cdot\dot{\mathbf{x}}, \tag{II.6-15}$$

provided S_n, which is called the *speed of displacement* of $\mathscr{S}$, be the speed at which that surface advances in the direction normal to itself in space:

$$S_n = \frac{-f'}{|\operatorname{grad} f|}. \tag{II.6-16}$$

EULER's condition (14) thus asserts that the speed of displacement of $\mathscr{S}$ at $(\mathbf{x}, t)$ is just the same as the speed at which the substantial point now occupying $(\mathbf{x}, t)$ is moving in the direction normal to $\mathscr{S}$.

Exercise II.6.2. Let a surface $\mathscr{S}$ have parametric representation $\mathbf{x} = \mathbf{g}(\mathbf{A}, t)$, the parameter $\mathbf{A}$ being an ordered pair of real parameters. If $\mathbf{A}$ is regarded as permanently denoting a particular point on $\mathscr{S}$ as $\mathscr{S}$ moves, calculation of its velocity $\mathbf{u}$ shows that $\mathbf{n}\cdot\mathbf{u} = S_n$. If $\mathscr{S}$ is represented by some spatial equation, say $h(\mathbf{x}, t) = 0$, the same field S_n is obtained in this way. This fact justifies the name "speed of displacement".

Exercise II.6.3 (LAGRANGE). If (14) is regarded as a partial-differential equation for f in the spatial description, integration by the method of characteristics yields $F(\boldsymbol{\chi}_{\kappa}^{-1}, f) = 0$. Thus, the substantial points that lie upon $f(\mathbf{x}, t) = \text{const.}$ at any one time lie always upon its image under the motion.

A *kinematic boundary* is a surface that separates permanently two parts of $\mathscr{B}$, one of them being possibly the null body. Thus a kinematic boundary is a substantial surface, and conversely. The special term "boundary" is introduced so as to distinguish particular substantial surfaces, usually assigned in advance like a wall or at least given some special role such as a surface separating two parts having different properties. The simplest example is a *stationary wall,* a surface $f(\mathbf{x}) = \text{const}$. In order for such a surface to be substantial and hence a possible kinematic boundary for a given motion of $\mathscr{B}$, by (15) we have the following necessary and sufficient condition relating the unit normal $\mathbf{n}$ to the velocity:

$$\mathbf{n}\cdot\dot{\mathbf{x}} = 0. \tag{II.6-17}$$

That is, *the velocity field on the wall is tangential,* as is obvious. More generally, if the places on a wall have assigned velocities $\mathbf{u}$, then at those places

$$\mathbf{n}\cdot\dot{\mathbf{x}} = \mathbf{n}\cdot\mathbf{u}. \tag{II.6-18}$$

Sometimes a stronger kinematic condition is imposed, that of *adherence.* The body is then constrained to move with the kinematic boundary. If the places on the wall have an assigned velocity $\mathbf{v}$, then on that wall

$$\dot{\mathbf{x}} = \mathbf{v}. \tag{II.6-19}$$

In the case of a stationary wall this condition becomes

$$\dot{\mathbf{x}} = \mathbf{0}. \tag{II.6-20}$$

Exercise II.6.4. Let the surface $\mathscr{S}$ whose equation is $g(\mathbf{x}, t) = 0$ in $\chi(\mathscr{B}, t)$ be the image of the surface $\mathscr{S}_\kappa$ whose equation is $G(\mathbf{X}, t) = 0$ in the reference shape $\kappa(\mathscr{B})$. (Note that $\mathscr{S}_\kappa$, in contradistinction with the substantial surfaces discussed above, generally moves with respect to $\kappa(\mathscr{B})$.) With the conventions of notation set at the beginning of this section, the oriented unit normals $\mathbf{n}_\kappa$ and $\mathbf{n}$ to these two surfaces are related by

$$\mathbf{n}_\kappa = \frac{|\operatorname{grad} g|}{|\operatorname{Grad} g|}\mathbf{F}^{\mathsf{T}}\mathbf{n}; \tag{II.6-21}$$

the speed of advance S_κ of the surface $\mathscr{S}_\kappa$ in the direction normal to itself in $\kappa(\mathscr{B})$ is given by

$$S_\kappa = -\frac{\dot{g}}{|\operatorname{Grad} g|}; \tag{II.6-22}$$

and

$$S_\kappa = \frac{|\operatorname{grad} g|}{|\operatorname{Grad} g|}(S_n - \mathbf{n}\cdot\dot{\mathbf{x}}). \tag{II.6-23}$$

The speed S_κ given by (22) is called the *speed of propagation* of the surface $\mathscr{S}_\kappa$ in $\kappa(\mathscr{B})$. It is the normal speed of advance of $\mathscr{S}_\kappa$ in $\kappa(\mathscr{B})$. Its reciprocal, S_κ^{-1}, is the *slowness* of $\mathscr{S}_\kappa$, and the vector $S_\kappa^{-1}\mathbf{n}_\kappa$ is the *slowness vector* of that surface.

When, as we may, we take the actual shape as being also $\kappa(\mathscr{B})$, the corresponding speed of propagation is denoted by S and called the *intrinsic speed of propagation* of $\mathscr{S}$. At $(\mathbf{x}, t)$ it is the speed at which the surface is advancing in the direction normal to itself and relative to the velocity of the substantial point instantaneously situate upon it. The *intrinsic slowness vector* of $\mathscr{S}$ is $S^{-1}\mathbf{n}$.

The intrinsic speed of propagation of $\mathscr{S}$ is related as follows to the speed of displacement of $\mathscr{S}$:

$$S = S_n - \mathbf{n}\cdot\dot{\mathbf{x}}; \tag{II.6-24}$$

this formula is an instance of (23). Finally, comparison with (23) shows that

$$S = \frac{|\mathrm{Grad}\, g|}{|\mathrm{grad}\, g|} S_\kappa\,. \tag{II.6-25}$$

Exercise II.6.5. If $\dot{\mathbf{x}}_{1+2}$ is defined by (II.4-8), then

$$\ddot{\mathbf{x}}_{1+2} = \ddot{\mathbf{x}}_1 + \ddot{\mathbf{x}}_2 + (\mathrm{grad}\,\dot{\mathbf{x}}_2)\dot{\mathbf{x}}_1 + (\mathrm{grad}\,\dot{\mathbf{x}}_1)\dot{\mathbf{x}}_2\,. \tag{II.6-26}$$

7. Change of Reference Placement

Let the same motion (II.1-1) be described alternatively by transplacements χ_{κ_1} and χ_{κ_2} with respect to two different reference placements, κ_1 and κ_2:

$$\begin{aligned} \chi_{\kappa_1}&: \kappa_1(\mathscr{B}) \to \chi(\mathscr{B}, t),\\ \chi_{\kappa_2}&: \kappa_2(\mathscr{B}) \to \chi(\mathscr{B}, t). \end{aligned} \tag{II.7-1}$$

The transplacements χ_{κ_1} and χ_{κ_2} have gradients $\mathbf{F}_1$ and $\mathbf{F}_2$ at (X, t). Let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the places occupied by the substantial point X in κ_1 and κ_2:

$$\mathbf{X}_1 = \kappa_1(X), \qquad \mathbf{X}_2 = \kappa_2(X). \tag{II.7-2}$$

Thus,

$$\mathbf{X}_2 = \kappa_2 \circ \kappa_1^{-1}(\mathbf{X}_1) =: \boldsymbol{\lambda}(\mathbf{X}_1), \tag{II.7-3}$$

say. The transplacement from κ_1 to χ can be effected in two ways: either straight off by use of χ_{κ_1} itself, or by using λ to get to κ_2 and then using χ_{κ_2} to get to χ. Thus

$$\chi_{\kappa_1} = \chi_{\kappa_2} \circ \lambda. \tag{II.7-4}$$

Because this relation holds among the three mappings, we see that their linear approximations, the gradients, compose in the corresponding order:

$$\mathbf{F}_{\kappa_1} = \mathbf{F}_{\kappa_2}\mathbf{P}, \qquad \mathbf{P} := \nabla\lambda. \tag{II.7-5}$$

This multiplication can be expressed also as a chain rule:

$$\partial_{X^\alpha}\chi^m_{\kappa_1} = (\partial_{X^A}\chi^m_{\kappa_2})\partial_{X^\alpha}\lambda^A; \tag{II.7-6}$$

in this notation X^α are the co-ordinates of the place occupied by X in κ_1, X^A are the co-ordinates of the place occupied by X in κ_2, the co-ordinate systems are arbitrary, and the summation convention is followed.

8. Present Placement as Reference

To serve as a reference, a placement need only be a homeomorph of $\mathscr{B}$. So far, we have employed a reference placement independent of time, but we could just as well use a varying one. Thus one motion may be described in terms of any other. The only varying placement often useful as a reference placement is the present one. If we take the present placement as reference, we describe the past and future as they seem to an observer fixed to the substantial point X that now occupies the place $\mathbf{x}$. The corresponding description is called *relative*.

To see how such a description is constructed, we consider places that are values of the motion of X at the two times t and τ:

$$\begin{aligned} \boldsymbol{\xi} &= \chi(X, \tau), \\ \mathbf{x} &= \chi(X, t). \end{aligned} \tag{II.8-1}$$

That is, $\boldsymbol{\xi}$ is the place occupied at the time τ by the substantial point that at the time t occupies $\mathbf{x}$:

$$\begin{aligned} \boldsymbol{\xi} &= \chi(\chi^{-1}(\mathbf{x}, t), \tau), \\ &=: \chi_t(\mathbf{x}, \tau), \end{aligned} \tag{II.8-2}$$

say. The function χ_t here defined is called the *relative transplacement*.

Sometimes we shall wish to calculate the relative transplacement when the motion is given to us only through the spatial description of the velocity field:

$$\mathbf{v} = \dot{\mathbf{x}}(\mathbf{x}, t). \tag{II.4-5$_r$}$$

By $(1)_1$

$$\partial_\tau \boldsymbol{\xi} = \dot{\mathbf{x}}(\boldsymbol{\xi}, \tau). \tag{II.8-3}$$

Since the right-hand side is a given function, we thus have a differential equation to integrate. The initial condition to be satisfied by the integral $\boldsymbol{\xi} = \boldsymbol{\chi}_t(\mathbf{x}, \tau)$ is

$$\boldsymbol{\xi}|_{\tau=t} = \boldsymbol{\chi}_t(\mathbf{x}, t) = \mathbf{x}. \tag{II.8-4}$$

When the motion is described by (2), we shall use a subscript t to denote quantities derived from the relative transplacement $\boldsymbol{\chi}_t$. Thus $\mathbf{F}_t$, the function of $\mathbf{x}$ and τ defined by

$$\mathbf{F}_t := \operatorname{grad} \boldsymbol{\chi}_t, \tag{II.8-5}$$

is the *relative transplacement gradient*. Of course

$$\mathbf{F}_t(t) = \mathbf{1}. \tag{II.8-6}$$

By (II.7-5), at $\mathbf{X}$

$$\mathbf{F}(\tau) = \mathbf{F}_t(\tau)\mathbf{F}(t). \tag{II.8-7}$$

As the fixed reference placement with respect to which $\mathbf{F}(\tau)$ and $\mathbf{F}(t)$ are taken we may select the placement of the body at the time t'. Then (7) yields

$$\mathbf{F}_{t'}(\tau) = \mathbf{F}_t(\tau)\mathbf{F}_{t'}(t), \tag{II.8-8}$$

a formula which, like (II.7-5), expresses a chain rule of differential calculus. In (6), (7), and (8) the argument $\mathbf{x}$ is understood and not written.

9. Stretch and Rotation

Since the transplacement χ_κ is invertible, so is its gradient $\mathbf{F}$, and the polar decomposition theorem of CAUCHY[1] yields two expressions for $\mathbf{F}$ in terms of an orthogonal tensor $\mathbf{R}$ and positive symmetric tensors $\mathbf{U}$ and $\mathbf{V}$, all three unique:

$$\mathbf{F} = \mathbf{RU} = \mathbf{VR}. \tag{II.9-1}$$

$\mathbf{R}$ is orthogonal but need not be proper-orthogonal: $\mathbf{RR}^{\mathsf{T}} = \mathbf{1}$, and so $\det \mathbf{R} = +1$ or -1, and $\det \mathbf{R}$ maintains either the one value or the other for all $\mathbf{X}$ and t, by continuity. Thus $\det \mathbf{U} = \det \mathbf{V} = |\det \mathbf{F}| = J$. $\mathbf{R}$ is called the *rotation tensor*[2]; $\mathbf{U}$ and $\mathbf{V}$, which satisfy the obvious relation

$$\mathbf{V} = \mathbf{RUR}^{\mathsf{T}}, \tag{II.9-2}$$

are called the *right* and *left stretch tensors*, respectively. These tensors, like $\mathbf{F}$ itself, are to be interpreted as comparing aspects of the present shape of $\mathscr{B}$ with their counterparts in the reference shape. Just how they do so, we shall proceed to show.

First, since $\mathbf{U}$ is symmetric, it has at least one orthogonal triad of principal axes; the members of any such triad are called *principal axes of strain at* $\mathbf{X}$ in the reference shape $\kappa(\mathscr{B})$. Likewise, $\mathbf{V}$ has an orthogonal triad of principal axes, which are called *principal axes of strain at* $\mathbf{x}$ in the present shape $\chi(\mathscr{B}, t)$. By

[1] This theorem is proved in any book on linear algebra, *e.g.*, in Section 83 of P. R. HALMOS, *Finite-Dimensional Vector Spaces*, 2nd ed., Princeton, Toronto, and London, Van Nostrand, 1958. It was discovered by CAUCHY in the present context; he proved it by geometrical arguments in $\mathscr{E}^3$.

[2] To reconcile the term with the definition, we could have imposed from the start the requirement that only reference placements such that $\det \mathbf{F} > 0$ be allowed, which would have implied that $\det \mathbf{R} = 1$ and have made $\mathbf{R}$ a rotation in the usual sense of that term. Since there is no reason to do so other than the convenience of language, we take advantage of that convenience without imposing the restriction. That is, in the text above we leave to the student such trivial changes of wording as may be needed when $-\mathbf{R}$ rather than $\mathbf{R}$ is proper. From the remarks made regarding (II.5-5) the student will recall that if the body in question ever occupies its reference placement, then $\mathbf{R} = \mathbf{1}$ at that time, and therefore $\mathbf{R}$ is always proper.

MARTINS & PODIO-GUIDUGLI, extending work of GRIOLI, have established the polar decomposition theorem through the following problem of minimization: for a given tensor $\mathbf{F}$, to find orthogonal tensors $\mathbf{R}$ such as to render $|\mathbf{F} - \mathbf{Q}|$ a minimum when $\mathbf{Q}$ varies over all orthogonal tensors. They prove that such $\mathbf{R}$ exist and that $\mathbf{FR}^{\mathsf{T}}$ is unique and not negative. If $\mathbf{V} := \mathbf{FR}^{\mathsf{T}}$, then $\mathbf{F} = \mathbf{VR}$, and if $\mathbf{U} := \mathbf{R}^{\mathsf{T}}\mathbf{VR}$, then $\mathbf{U}$ is unique and not negative. Hence $\mathbf{F} = \mathbf{RU} = \mathbf{VR}$. If $\mathbf{F}$ is invertible, $\mathbf{R}$ is unique, and $\mathbf{U}$ and $\mathbf{V}$ are positive. Thus the local rotation $\mathbf{R}$ in the polar decomposition of $\mathbf{F}$ is the unique orthogonal tensor closest to $\mathbf{F}$. *Cf.* L. C. MARTINS & P. PODIO-GUIDUGLI, "A variational approach to the polar decomposition theorem," *Accademia Nazionale dei Lincei, classe di Scienze Fisiche, Matematiche, e Naturali, Rendiconti* (6) **66** (1975): 487–493.

(2), $\mathbf{U}$ and $\mathbf{V}$ have their proper numbers in common. Indeed, if $\mathbf{e}_k$ is a proper vector of $\mathbf{U}$ corresponding to the proper number v_k, then

$$\mathbf{U}\mathbf{e}_k = v_k \mathbf{e}_k , \tag{II.9-3}$$

and so by (1) and (2)

$$\mathbf{V}(\mathbf{R}\mathbf{e}_k) = (\mathbf{R}\mathbf{U}\mathbf{R}^{\mathsf{T}})(\mathbf{R}\mathbf{e}_k) = v_k(\mathbf{R}\mathbf{e}_k). \tag{II.9-4}$$

Thus the rotation $\mathbf{R}$ carries principal axes of strain at $\mathbf{X}$ into principal axes of strain at $\mathbf{x}$. (Since $\mathbf{R}$ is unique but the principal axes of strain need not be, we cannot always use this property as a definition of $\mathbf{R}$.) If $\mathbf{e}_k$ points along the k^{th} principal axis of strain at $\mathbf{X}$ in $\boldsymbol{\kappa}(\mathscr{B})$, then v_k is the ratio of the length of the image $\mathbf{F}\mathbf{e}_k$ in $\chi(\mathscr{B}, t)$ to the length of the original $\mathbf{e}_k$. Thus, the v_k are called the *principal stretches*. Because $\mathbf{U}$ and $\mathbf{V}$ are positive, $v_k > 0$. When $\mathbf{R} = \mathbf{1}$, the transplacement is called a *pure stretch* at $\mathbf{X}$, t. In a pure stretch, $\mathbf{U} = \mathbf{V}$; the principal axes of strain at $\mathbf{X}$ and $\mathbf{x}$ coincide; and we may visualize the transplacement as being effected by stretching elements along those axes in the ratios v_1, v_2, v_3. If $\mathbf{U} = \mathbf{V} = \mathbf{1}$, the transplacement is called a *rotation* at $\mathbf{X}$, t. CAUCHY's decomposition tells us that the transplacement gradient may be obtained by effecting a pure stretch with principal stretches v_k along three suitable, mutually orthogonal directions $\mathbf{e}_k$, followed by a rotation of those directions, or by performing the same rotation first and then effecting the same stretches along the resulting directions.

The *right* and *left Cauchy–Green tensors*, $\mathbf{C}$ and $\mathbf{B}$, are defined as follows:

$$\begin{aligned} \mathbf{C} &:= \mathbf{U}^2 = \mathbf{F}^{\mathsf{T}}\mathbf{F}, \\ \mathbf{B} &:= \mathbf{V}^2 = \mathbf{F}\mathbf{F}^{\mathsf{T}} = \mathbf{R}\mathbf{C}\mathbf{R}^{\mathsf{T}}. \end{aligned} \tag{II.9-5}$$

While the fundamental decomposition (1) plays the major part in the proof of general theorems, calculation of $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{R}$ from $\mathbf{F}$ for particular transplacements may be awkward, since irrational operations are usually required. $\mathbf{C}$ and $\mathbf{B}$, nonetheless, are calculated by mere multiplication of $\mathbf{F}$ and $\mathbf{F}^{\mathsf{T}}$. *E.g.*, if g_{km} and $g^{\alpha\beta}$ are the covariant and contravariant metric components in arbitrarily selected co-ordinate systems in $\mathscr{E}$ and in $\boldsymbol{\kappa}(\mathscr{B})$, respectively, components of $\mathbf{C}$ and $\mathbf{B}$ are[1]

$$\begin{aligned} C_{\alpha\beta} &= F^k_{\alpha} F^m_{\beta} g_{km} , \\ B^{km} &= F^k_{\alpha} F^m_{\beta} g^{\alpha\beta} , \end{aligned} \tag{II.9-6}$$

[1] If both systems of co-ordinates (x^k) and (X^α) are cartesian, (6) follows at once from (5) and (II.5-3). To derive (6) in general coordinates it suffices to observe that $(6)_1$ and $(6)_2$ are tensorial equations which reduce in cartesian co-ordinates to the equations already demonstrated in the case when those co-ordinates are used.

in which $F^k_\alpha = x^k_{,\alpha} := \partial_{X^\alpha}\chi^k_\kappa(X^1, X^2, X^3, t)$. The proper numbers of $\mathbf{C}$ and $\mathbf{B}$ are the squares v_i^2 of the principal stretches. The *principal invariants* of $\mathbf{C}$ and $\mathbf{B}$ are given by

$$
\begin{aligned}
I &:= \operatorname{tr}\mathbf{B} = \operatorname{tr}\mathbf{C} = v_1^2 + v_2^2 + v_3^2\,,\\
II &:= \tfrac{1}{2}[(\operatorname{tr}\mathbf{B})^2 - \operatorname{tr}\mathbf{B}^2] = \tfrac{1}{2}[(\operatorname{tr}\mathbf{C})^2 - \operatorname{tr}\mathbf{C}^2] = v_1^2v_2^2 + v_2^2v_3^2 + v_3^2v_1^2\,, \qquad \text{(II.9-7)}\\
III &:= \det\mathbf{B} = \det\mathbf{C} = J^2 = v_1^2v_2^2v_3^2\,.
\end{aligned}
$$

Any symmetric function of v_1, v_2, and v_3 equals a function of I, II, and III.

The formulae obtained so far in this section apply to any invertible tensor, making no use of the fact that $\mathbf{F}$ is the gradient of χ_κ, which implies that it must satisfy the condition of compatibility $\operatorname{skw}\operatorname{grad}\mathbf{F} = \mathbf{0}$. Sometimes the relation of the values of $\mathbf{F}$ at different arguments $\mathbf{X}$ must be taken into account. One such example is furnished by the chain rule (II.8-7). For another, we note from $(7)_7$ and (II.5-4) that

$$
\rho_\kappa/\rho = \det\mathbf{U} = \det\mathbf{V} = \sqrt{III}. \qquad \text{(II.9-8)}
$$

Another example is furnished by the following exercise.

Exercise II.9.1 (MICHAL). If $\kappa(\mathscr{B})$ is connected, a transplacement whose gradient is orthogonal at each point is either a rigid rotation or the product of one by a central inversion. (If $\mathscr{B}$ ever occupies κ, a central inversion is excluded.) If χ_κ and $\bar{\chi}_\kappa$ are transplacements, for $\bar{\chi}_\kappa \circ \chi_\kappa^{-1}$ to preserve the distances between substantial points it is necessary and sufficient that $\bar{\mathbf{U}} = \mathbf{U}$.

If we begin with the gradient $\mathbf{F}_t$ of the relative transplacement, defined by (II.8-5), and apply to it the polar decomposition theorem, we obtain the *relative rotation tensor* $\mathbf{R}_t$, the *relative stretch tensors* $\mathbf{U}_t$ and $\mathbf{V}_t$ and the *relative Cauchy-Green tensors* $\mathbf{C}_t$ and $\mathbf{B}_t$:

$$
\mathbf{F}_t = \mathbf{R}_t\mathbf{U}_t = \mathbf{V}_t\mathbf{R}_t\,, \qquad \mathbf{C}_t = \mathbf{U}_t^2\,, \qquad \mathbf{B}_t = \mathbf{V}_t^2\,. \qquad \text{(II.9-9)}
$$

The uniqueness of a polar decomposition enables us to see from (II.8-6) that $\mathbf{U}_t(t) = \mathbf{V}_t(t) = \mathbf{R}_t(t) = \mathbf{1}$.

Exercise II.9.2. It follows from (II.8-7) that

$$
\mathbf{C}(\tau) = \mathbf{F}(t)^{\mathsf{T}}\mathbf{C}_t(\tau)\mathbf{F}(t). \qquad \text{(II.9-10)}
$$

When a transplacement is laid down for study, it is a trivial matter to calculate from it the tensors **B** and **C**. We consider here two examples, both of which will be useful later. In a *simple shear* each member of a family of parallel planes is transplaced tangentially a distance proportional to its distance from a particular plane in that family. If we let the particular plane be $X_1 = 0$, and if we let the direction of the shear be that of the co-ordinate X_2, then a simple shear is described in the co-ordinate system X_1, X_2, X_3 by the following components of transplacement:

$$\begin{aligned} x_1 &= X_1, \\ x_2 &= X_2 + KX_1, \\ x_3 &= X_3. \end{aligned} \tag{II.9-11}$$

The constant K is called the *amount* of shear.

Since

$$[\mathbf{F}] = \begin{Vmatrix} 1 & 0 & 0 \\ K & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \tag{II.9-12}$$

it follows that

$$\begin{aligned} [\mathbf{B}] = [\mathbf{F}\mathbf{F}^{\mathrm{T}}] &= \begin{Vmatrix} 1 & K & 0 \\ K & 1+K^2 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \\ [\mathbf{B}^{-1}] &= \begin{Vmatrix} 1+K^2 & -K & 0 \\ -K & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \end{aligned} \tag{II.9-13}$$

$$I = \operatorname{tr}\mathbf{B} = 3 + K^2 = II = \operatorname{tr}\mathbf{B}^{-1}, \qquad III = 1.$$

Exercise II.9.3 (Kelvin & Tait). In simple shear the principal stretches are expressed as follows in terms of the amount of shear:

$$\begin{aligned} v_1^2 &= 1 + \tfrac{1}{2}K^2 + K\sqrt{1 + \tfrac{1}{4}K^2}, \\ v_2^2 &= 1 + \tfrac{1}{2}K^2 - K\sqrt{1 + \tfrac{1}{4}K^2} = \frac{1}{v_1^2}, \\ v_3 &= 1. \end{aligned} \tag{II.9-14}$$

The angle θ through which the principal axes of strain in $\boldsymbol{\kappa}(\mathscr{B})$ are rotated so as to become the principal axes of strain in $\chi(\mathscr{B})$ is given by $\tan\theta = \frac{1}{2}K$.

An example illustrating the use of curvilinear co-ordinate systems is provided by the following components of transplacement in cylindrical polar co-ordinates:

$$r = \sqrt{AR^2 + B}, \qquad \theta = \Theta + DZ, \qquad z = FZ, \qquad AF = 1, \tag{II.9-15}$$

A, B, D, and F being constants. The cylinders R = const. are mapped into the cylinders r = const., and choice of the constants A and B allows an arbitrary expansion or contraction as well as an eversion of these cylinders. At the same time, there is a stretch F in the direction of the axis of the cylinders, so adjusted as to make the transplacement isochoric. Finally the planes Z = const. are rotated about the axis through angles proportional to their distance from the particular plane $Z = 0$. Thus a torsion of amount D/F is superimposed upon the isochoric expansion or contraction of the cylinders.

Exercise II.9.4. Use of $(6)_2$ shows that for the transplacement (15)

$$\|B^{km}\| = \begin{Vmatrix} A^2R^2/r^2 & 0 & 0 \\ 0 & R^{-2} + D^2 & DF \\ 0 & DF & F^2 \end{Vmatrix}. \tag{II.9-16}$$

To calculate $(B^{-1})_{km}$, the matrix (16) may be inverted; alternatively,

$$(B^{-1})_{km} = X^{\alpha}_{,k} X^{\beta}_{,m} g_{\alpha\beta}, \tag{II.9-17}$$

and hence

$$\|(B^{-1})_{km}\| = \begin{Vmatrix} r^2/(A^2R^2) & 0 & 0 \\ 0 & R^2 & -ADR^2 \\ 0 & -ADR^2 & A^2(1 + D^2R^2) \end{Vmatrix},$$

$$I = \operatorname{tr}\mathbf{B} = g_{km}B^{km} = \frac{A^2R^2}{r^2} + r^2\left(\frac{1}{R^2} + D^2\right) + F^2, \tag{II.9-18}$$

$$II = \operatorname{tr}\mathbf{B}^{-1} = g^{km}(B^{-1})_{km} = \frac{r^2}{A^2R^2} + \frac{R^2}{r^2} + A^2(1 + D^2R^2),$$

$$III = 1.$$

Exercise II.9.5. In *simple torsion* $A = F = 1, B = 0$. Comparing (16) with $(13)_1$ shows that simple torsion may be regarded as effecting on each cylinder R = const., when cut along a generator and developed onto a plane, a simple shear of amount of DR.

The transplacements (15) are members of a family that will be analysed in greater detail below in Section IV.15. There we shall encounter also transplacements conveniently described by cartesian co-ordinates in the reference placement, polar co-ordinates in the present placement.

Exercise II.9.6. In the notation used in Section II.6,

$$|\mathrm{Grad}\, f|^2 = \mathrm{grad}\, f \cdot \mathbf{B}\, \mathrm{grad}\, f, \tag{II.9-19}$$

and hence (II.6-21) can be written in the form

$$\mathbf{n}_\kappa = \frac{1}{\sqrt{\mathbf{n}\cdot\mathbf{B}\mathbf{n}}}\mathbf{F}^{\mathsf{T}}\mathbf{n}; \tag{II.9-20}$$

likewise, (II.6-23) becomes

$$S_\kappa = \frac{1}{\sqrt{\mathbf{n}\cdot\mathbf{B}\mathbf{n}}}(S_n - \mathbf{n}\cdot\dot{\mathbf{x}}). \tag{II.9-21}$$

10. Histories

Let Ψ denote a function of time whose value is a scalar, a vector, or a tensor. We shall often wish to consider the restriction of Ψ to present and past times only. For convenience, if t is the present time, we shall represent the past time τ by the positive quantity $s := t - \tau$. The *history of* Ψ *up to time* t is denoted by Ψ^t, the value of which is $\Psi^t(s)$:

$$\Psi^t(s) := \Psi(t - s), \qquad t \text{ fixed}, \quad s \geqq 0. \tag{II.10-1}$$

For each t the history Ψ^t is defined on $[0, \infty[$. The history Ψ^t, as its name suggests, is the portion of a function of time which corresponds to the present and past times only. Histories turn out to be of major importance in mechanics because it is circumstances at the present and past times that should determine future occurrences.

In this notation $\mathbf{C}_t^t$, for example, is the history of the relative right Cauchy–Green tensor $\mathbf{C}_t$ up to the time t. Of course

$$\psi^t(0) = \psi(t), \qquad \mathbf{F}_t^t(0) = \mathbf{1}, \quad etc. \tag{II.10-2}$$

11. Stretching and Spin

For the instantaneous time derivative of a tensor defined from the relative transplacement, for example $\mathbf{F}_t$, we introduce the notation[1]

$$\dot{\mathbf{F}}_t(t) := \partial_u \mathbf{F}_t(u)|_{u=t} = -\partial_s \mathbf{F}_t^t(s)|_{s=0}, \tag{II.11-1}$$

$\mathbf{X}$ being held constant, and we lay down the following definitions:

$$\begin{aligned} \mathbf{G} &:= \dot{\mathbf{F}}_t(t), \\ \mathbf{D} &:= \dot{\mathbf{U}}_t(t) = \dot{\mathbf{V}}_t(t), \\ \mathbf{W} &:= \dot{\mathbf{R}}_t(t). \end{aligned} \tag{II.11-2}$$

$\mathbf{D}$, which is called the *stretching*, is the rate of change of the stretch at the place of X in the shape at time $t+\epsilon$ with respect to that at time t, in the limit as $\epsilon \to 0$. Likewise, $\mathbf{W}$, which is called the *spin*, is the ultimate rate of change of the rotation at $\mathbf{x}$ from the present shape to one the body had just before or will have just afterward. Since $\mathbf{U}_t$ is symmetric, so is $\mathbf{D}$, being its derivative with respect to a parameter:

$$\mathbf{D}^{\mathsf{T}} = \mathbf{D}, \tag{II.11-3}$$

but $\mathbf{D}$, unlike $\mathbf{U}_t$, generally fails to be positive. Since $\mathbf{D}(\mathbf{x}, t)$ is symmetric, its proper numbers are real, and it has at least one orthogonal triad of proper vectors. The latent roots of $\mathbf{D}(\mathbf{x}, t)$ are called the *principal stretchings* $d_k(\mathbf{x}, t)$, $k = 1, 2, 3$; the directions of a corresponding orthogonal triad are called *principal axes of stretching*.

If we differentiate the relation $\mathbf{R}_t(u)\mathbf{R}_t(u)^{\mathsf{T}} = \mathbf{1}$ with respect to u, put $u = t$, and use $(2)_4$, we find that $\mathbf{W}$ is skew:

$$\mathbf{W}^{\mathsf{T}} + \mathbf{W} = \mathbf{0}. \tag{II.11-4}$$

[1]This notation could not be confused with the substantial derivative, introduced in Section II.6, since $\mathbf{F}_t(t) = \mathbf{1}$ and $\dot{\mathbf{1}} = \mathbf{0}$.

From its definition $(2)_1$, $\mathbf{G}$ is the ultimate rate of change of $\mathbf{F}_t$, but that is not all, for by differentiating (II.8-7) with respect to τ and then putting $\tau = t$ we obtain

$$\mathbf{G} = \dot{\mathbf{F}}\mathbf{F}^{-1}. \tag{II.11-5}$$

Differentiation of (II.5-1) with respect to t yields

$$\dot{\mathbf{F}} = \operatorname{Grad} \dot{\chi}_\kappa = (\operatorname{grad} \dot{\mathbf{x}})\mathbf{F}; \tag{II.11-6}$$

in view of (II.5-3), the last step follows by the chain rule of differential calculus. Substitution into (5) yields

$$\mathbf{G} = \operatorname{grad} \dot{\mathbf{x}}. \tag{II.11-7}$$

We have shown that the tensor $\mathbf{G}$, which we defined by $(2)_1$, is in fact the spatial *velocity gradient*.

If we differentiate the polar decomposition $(\text{II}.9\text{-}9)_1$ with respect to τ and then put $\tau = t$, we find that

$$\mathbf{G} = \mathbf{D} + \mathbf{W}. \tag{II.11-8}$$

This conclusion, showing that $\mathbf{D} = \operatorname{sym} \operatorname{grad} \dot{\mathbf{x}}$ and $\mathbf{W} = \operatorname{skw} \operatorname{grad} \dot{\mathbf{x}}$, expresses the fundamental ***Euler–Cauchy–Stokes Decomposition*** of the instantaneous motion at $\mathbf{x}$, t into the sum of a pure stretching along three mutually orthogonal axes, a spin, and a translation. The stretching $\mathbf{D}$ must not be confused with a rate of change of a stretch such as $\dot{\mathbf{U}}$ or $\dot{\mathbf{V}}$, and the spin is not generally the rate of change of a finite rotation.

The definitions $(2)_3$ and $(2)_4$ make the different kinematic meanings of $\mathbf{D}$ and $\mathbf{W}$ clear and suggest that both tensors will be useful in the description and classification of motions.

Of course, we could have defined $\mathbf{G}$ by (7) as the velocity gradient and $\mathbf{W}$ and $\mathbf{D}$ by (8) as the symmetric and skew parts of $\mathbf{G}$. We should then have had to prove $(2)_{2,4}$ as theorems so as to interpret $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{D}$ kinematically. Most writers on hydrodynamics prefer the argument in this order.

Exercise II.11.1 (LAGRANGE).

$$\ddot{\mathbf{x}} = \dot{\mathbf{x}}' + 2\mathbf{W}\dot{\mathbf{x}} + \operatorname{grad}(\tfrac{1}{2}\dot{x}^2). \tag{II.11-9}$$

Motions in which $\mathbf{W} = \mathbf{0}$ are called *irrotational*. They form the main subject of study in classical hydrodynamics. Motions in which $\mathbf{W} \neq \mathbf{0}$ are called *rotational*. Of course both conditions are local: a motion may be irrotational in one part of its domain and rotational in another. In Volume 2 we shall encounter the famous example called a *spherical vortex*.

Exercise II.11.2. Writing (II.6-26) as

$$\ddot{\mathbf{x}}_{1+2} = \ddot{\mathbf{x}}_1 + \ddot{\mathbf{x}}_2 + \mathbf{G}_1\dot{\mathbf{x}}_2 + \mathbf{G}_2\dot{\mathbf{x}}_1\,, \tag{II.11-10}$$

we see that if $\dot{\mathbf{x}}_1$ and $\dot{\mathbf{x}}_2$ are isochoric, then

$$\operatorname{div}\ddot{\mathbf{x}}_{1+2} = \operatorname{div}\ddot{\mathbf{x}}_1 + \operatorname{div}\ddot{\mathbf{x}}_2 + 2\mathbf{D}_1\cdot\mathbf{D}_2 + 2\mathbf{W}_1\cdot\mathbf{W}_2\,. \tag{II.11-11}$$

Because $\mathbf{W}$ is skew, it may be represented to within a convention of sign by the axial vector $\operatorname{curl}\dot{\mathbf{x}}$, which is called the "vorticity vector" in hydrodynamics: $\mathbf{w} := \operatorname{curl}\dot{\mathbf{x}}$.[1] Nowadays it seems more convenient not to introduce this vector but instead to use the tensor $\mathbf{W}$. We shall nevertheless use the letter w to denote the magnitude of $\operatorname{curl}\dot{\mathbf{x}}$:

$$w := \sqrt{2}|\mathbf{W}| = |\operatorname{curl}\dot{\mathbf{x}}|, \tag{II.11-12}$$

and it is the scalar field w that in this book we shall call the *vorticity*. In a rigid motion $\frac{1}{2}w = \omega$, the angular speed (*cf.* Section I.9). In the plane normal to the axis of spin at $\mathbf{x}$ we can choose orthogonal unit vectors $\mathbf{e}$ and $\mathbf{f}$ such that

$$\mathbf{W} = \tfrac{1}{2}w\mathbf{e}\wedge\mathbf{f}, \tag{II.11-13}$$

and so

$$\begin{gathered}\mathbf{We} = -\tfrac{1}{2}w\mathbf{f}, \qquad \mathbf{Wf} = \tfrac{1}{2}w\mathbf{e},\\ \mathbf{W} = -\mathbf{e}\otimes\mathbf{We} + \mathbf{We}\otimes\mathbf{e}.\end{gathered} \tag{II.11-14}$$

Another important scalar is the *expansion* E, defined as follows:

$$E := \dot{J}/J = \operatorname{div}\dot{\mathbf{x}} = \operatorname{tr}\mathbf{G} = \operatorname{tr}\mathbf{D} = \dot{v}/v = -\dot{\rho}/\rho; \tag{II.11-15}$$

[1] LAGRANGE's form (9) for the acceleration may be written as

$$\ddot{\mathbf{x}} = \dot{\mathbf{x}}' + \mathbf{w}\times\dot{\mathbf{x}} + \operatorname{grad}(\tfrac{1}{2}\dot{x}^2),$$

in which $\mathbf{w} := \operatorname{curl}\dot{\mathbf{x}}$.

the third and fourth expressions follow by use of (7) and (8), and the last by use of (II.5-7). E is the local rate of increase of volume of a substantial region, referred to unit volume. We remark again that necessary and sufficient conditions for isochoric motion are

$$0 = E = \dot{J} = \operatorname{div}\dot{\mathbf{x}} = \operatorname{tr}\mathbf{G} = \operatorname{tr}\mathbf{D} = \dot{v} = \dot{\rho}. \tag{II.11-16}$$

The velocity field of a rigid motion is given by (I.10-1). Taking the gradient of that equation yields $\mathbf{G} = \mathbf{W}$. Thus $\mathbf{D} = \mathbf{0}$ in a rigid motion, and the spin as defined by $(2)_4$ for a general motion reduces in a rigid motion to a field having as its value everywhere what we have called in Section I.10 the spin of that motion. Thus we may regard the spin field as a generalization of the spin of a rigid motion—in rough language, a local velocity of infinitesimal rotation.

Exercise II.11.3 (EULER). If $\mathbf{D} = \mathbf{0}$ is regarded as a differential equation for the spatial velocity field in a connected open set, integrating it yields (I.10-1).

These observations establish the following theorem: *The condition* $\mathbf{D} = \mathbf{0}$ *in a region at an instant is necessary and sufficient that the motion be rigid in that region at that instant.* In view of the interpretation of $\mathbf{D}$ given just after its definition, the theorem is obvious.

Clearly the spin $\mathbf{W}$ is generally something quite different from $\dot{\mathbf{R}}$, the time-rate of the rotation tensor, as the following two examples show.

In a *steady, simple shearing* cartesian velocity components are

$$\dot{x}_1 = 0, \qquad \dot{x}_2 = \kappa x_1, \qquad \dot{x}_3 = 0, \qquad \kappa = \text{const.}; \tag{II.11-17}$$

κ is the *shearing*; and each substantial point moves ahead at constant speed along a straight line parallel to the x_2-axis, yet unless $\kappa = 0$, the motion is rotational. In a *steady, simple vortex*, the cylindrical polar, contravariant velocity components are

$$\dot{r} = 0, \qquad \dot{\theta} = \omega(r), \qquad \dot{z} = 0, \tag{II.11-18}$$

and each substantial point rotates steadily about the polar axis on a circle r = const., z = const., at the angular speed $\omega(r)$, yet if $\omega(r) = Kr^{-2}$, the motion is irrotational. K, which is the magnitude of the rotational momentum per unit mass-density of the body undergoing the motion, is the *strength* of the irrotational vortex. More generally, the vorticity is given by $rw = (r^2\omega)'$, the prime denoting the derivative with respect to r.

The simple distinction of "rotational" and "irrotational" does not tell us how to decide whether a given rotational motion is strongly or weakly so.

To say that vorticity or stretching is small, has no meaning by itself, for the physical dimension of these quantities is the reciprocal of the dimension of time. Dimensionless measures of intervals of time, speeds, frequencies, *etc.*, are provided by ratios. For example, one intuitive concept of the rotationality of a motion may be made precise by defining a numerical degree called the *vorticity number*, namely the field $\mathfrak{W}$ defined as follows at all points where the motion is not rigid:

$$\mathfrak{W} := \frac{|\mathbf{W}|}{|\mathbf{D}|}. \tag{II.11-19}$$

If $\mathbf{D} = \mathbf{0}$ but $\mathbf{W} \neq \mathbf{0}$, we may choose to say that $\mathfrak{W} = \infty$. Then the value of $\mathfrak{W}$ is a degree of rotationality determined at each time at every interior point of the present shape of a body that does not presently lie in a neighborhood undergoing pure translation; the degree is 0 for an irrotational motion and ∞ for a rigid motion other than a state of rest. When $\mathfrak{W} = 1$, spin and stretching are precisely balanced. Such is the case in some flows commonly used to illustrate the effects of viscous friction, for example the steady, simple shearing (17).

Exercise II.11.4 (TRUESDELL). For the steady, simple vortex (18)

$$\mathfrak{W}^2 = 1 + \frac{4\omega(r\omega)'}{r^2\omega'^2}. \tag{II.11-20}$$

Thus $\mathfrak{W} > 1$ in regions where the linear speed $r\omega$ increases with r, while $\mathfrak{W} < 1$ in regions where $r\omega$ decreases with r. In regions where $\omega \propto r^{-2}$ the simple vortex is irrotational: $\mathfrak{W} = 1$. If $\omega \propto r^{-n}$, then

$$\mathfrak{W} = \left|1 - \frac{2}{n}\right| = \text{const.} \tag{II.11-21}$$

Thus $\mathfrak{W}$ in an isochoric motion may take on any value in $[0, \infty[$, and that value may be the same at all points even though neither $\mathbf{D}$ nor $\mathbf{W}$ need be so.

Exercise II.11.5 (TRUESDELL). If $\mathbf{W} \neq \mathbf{0}$, then

$$\mathfrak{W} = \left(1 + \frac{\operatorname{tr}\mathbf{G}^2}{|\mathbf{W}|^2}\right)^{-1/2} = \left(1 + \frac{\operatorname{div}\ddot{\mathbf{x}} - \dot{E}}{|\mathbf{W}|^2}\right)^{-1/2}. \tag{II.11-22}$$

$\mathfrak{W} = 1$ *in any accelerationless, isochoric, and rotational flow*. Indeed, for the last statement to hold it suffices (but is not necessary) that the expansion be substantially constant and that the acceleration field be solenoidal.

Exercise II.11.6 (TRUESDELL). Let there be given an isochoric motion of spin $\mathbf{W}_0$ and vorticity number $\mathfrak{W}_0$; let there be given also a rigid motion of spin $\mathbf{W}_r$; and let these two motions be superposed (Section II.4) to produce a motion whose spin $\mathbf{W} = \mathbf{W}_0 + \mathbf{W}_r$. Suppose that $\mathbf{W}_0 \neq \mathbf{0}$ and $\mathbf{W} \neq \mathbf{0}$. Then the vorticity number $\mathfrak{W}$ of the combined motion is determined as follows:

$$|\mathbf{W}|^2\left(1 - \frac{1}{\mathfrak{W}^2}\right) = |\mathbf{W}_0|^2\left(1 - \frac{1}{\mathfrak{W}_0^2}\right) + |\mathbf{W}_r|^2 + 2\mathbf{W}_r \cdot \mathbf{W}_0 . \tag{II.11-23}$$

If $\mathfrak{W}_0 = 1$, then

$$\mathfrak{W} \geqq 1 \quad \Leftrightarrow \quad |\mathbf{W}_r|^2 \geqq -2\mathbf{W}_r \cdot \mathbf{W}_0 . \tag{II.11-24}$$

In three dimensions the right-hand inequality becomes, if $\mathbf{W}_r \neq \mathbf{0}$,

$$|\mathbf{W}_r| \geqq -2|\mathbf{W}_0| \cos\theta , \tag{II.11-25}$$

θ being the least angle between the oriented axes of $\mathbf{W}_r$ and $\mathbf{W}_0$. In particular, the condition $\theta \leqq \frac{1}{2}\pi$ is sufficient to ensure that $\mathfrak{W} \geqq 1$ if $\mathfrak{W}_0 = 1$. If $\mathbf{W}_0 = -\mathbf{W}_r$, then $\theta = \pi$ and $\mathbf{W} = \mathbf{0}$.

Further enlightenment of the difference between stretch and stretching and between rotation and spin is furnished by the following exercise.[1]

Exercise II.11.7 (E. & F. COSSERAT, COLEMAN & TRUESDELL).

$$\begin{aligned} \dot{\mathbf{C}} &= 2\mathbf{F}^{\mathsf{T}}\mathbf{D}\mathbf{F}, \\ \mathbf{W} &= \dot{\mathbf{R}}\mathbf{R}^{\mathsf{T}} + \tfrac{1}{2}\mathbf{R}(\dot{\mathbf{U}}\mathbf{U}^{-1} - \mathbf{U}^{-1}\dot{\mathbf{U}})\mathbf{R}^{\mathsf{T}}, \\ \mathbf{D} &= \tfrac{1}{2}\mathbf{R}(\dot{\mathbf{U}}\mathbf{U}^{-1} + \mathbf{U}^{-1}\dot{\mathbf{U}})\mathbf{R}^{\mathsf{T}}, \end{aligned} \tag{II.11-26}$$

where $\mathbf{R}$ and $\mathbf{U}$ have their usual meanings as the rotation and right stretch tensors with respect to a fixed reference placement. Also $\dot{\mathbf{B}}|_{\mathbf{F}=\mathbf{1}} = 2\mathbf{D}$.

Various higher rates of change of stretch and rotation may be defined.

Exercise II.11.8. Including and generalizing (1) and $(2)_1$, set

$$\mathbf{G}_n := \overset{(n)}{\mathbf{F}_t}(t) := \partial_u^n \mathbf{F}_t(u)|_{u=t} , \qquad n = 1, 2, \ldots . \tag{II.11-27}$$

[1] $\dot{\mathbf{R}}$ as a function $\mathbf{W}$, $\mathbf{R}$, $\mathbf{U}$, and $\dot{\mathbf{U}}$ can be read off from $(26)_2$. GUO ZHONG-HENG, "Rates of stretch tensors," *Journal of Elasticity* **14** (1984): 263–267, determines $\dot{\mathbf{R}}$ and $\dot{\mathbf{V}}$ as functions of $\mathbf{V}$, $\mathbf{R}$, and $\mathbf{G}$; also he determines $\dot{\mathbf{U}}$ as a function of $\mathbf{U}$, $\mathbf{R}$, and $\mathbf{G}$. *Cf.* also A. HOGER & D. CARLSON, "On the derivative of the square root of a tensor and Guo's rate theorems," *ibid.* 329–336.

Differentiating (II.8-7) n times with respect to τ and setting $\tau = t$ shows that

$$\overset{(n)}{\mathbf{F}}\mathbf{F}^{-1} = \mathbf{G}_n , \tag{II.11-28}$$

and hence by use of the chain rule

$$\mathbf{G}_n = \operatorname{grad} \overset{(n)}{\mathbf{x}} . \tag{II.11-29}$$

In particular

$$\mathbf{G}_2 = \ddot{\mathbf{F}}\mathbf{F}^{-1} = \dot{\mathbf{G}} + \mathbf{G}^2 . \tag{II.11-30}$$

The most useful higher rates are the *Rivlin–Ericksen tensors* $\mathbf{A}_n$. They are defined as follows in terms of a notation like (27):

$$\mathbf{A}_n := \overset{(n)}{\mathbf{C}_t}(t) \tag{II.11-31}$$

and hence are symmetric. In particular, $\mathbf{A}_1 = 2\mathbf{D}$.

Exercise II.11.9 (DUPONT, RIVLIN & ERICKSEN).

$$\mathbf{A}_n = \mathbf{G}_n + \mathbf{G}_n^{\mathsf{T}} + \sum_{j=1}^{n-1} \binom{n}{j} \mathbf{G}_j^{\mathsf{T}} \mathbf{G}_{n-j} \tag{II.11-32}$$

and

$$\mathbf{A}_{n+1} = \dot{\mathbf{A}}_n + \mathbf{A}_n \mathbf{G} + (\mathbf{A}_n \mathbf{G})^{\mathsf{T}} . \tag{II.11-33}$$

Exercise II.11.10 (RIVLIN, TRUESDELL & NOLL). Differentiating the relation $\det \mathbf{C}_t(u) = 1$ repeatedly with respect to u and then putting $u = t$ shows that in an isochoric motion

$$\begin{aligned} \operatorname{tr} \mathbf{A}_1 &= 0, \\ \operatorname{tr} \mathbf{A}_2 &= \operatorname{tr} \mathbf{A}_1^2 , \\ \operatorname{tr} \mathbf{A}_3 &= -2 \operatorname{tr} \mathbf{A}_1^3 + 3 \operatorname{tr}(\mathbf{A}_2 \mathbf{A}_1), \end{aligned} \tag{II.11-34}$$

and in general $\operatorname{tr} \mathbf{A}_n$ is a linear combination of traces of products formed from $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{n-1}$.

We now consider restrictions imposed upon $\mathbf{W}$ by boundaries. To do so, we first appeal to KELVIN's transformation ("STOKES's theorem")[1] of an integral over a surface $\mathscr{S}$ into a line integral around the *border* of $\mathscr{S}$, which we shall denote by $\mathscr{C}$. ($\mathscr{S}$ is a compact set, $\mathscr{C} := \operatorname{clo}\mathscr{S}\backslash\mathscr{S}$, and $\mathscr{C}$ is a closed line, often called a *circuit*.) The *circulation* $\mathrm{C}(\mathscr{C})$ was introduced by KELVIN as a measure of the summed tangential speeds of the substantial points lying presently upon $\mathscr{C}$. Assuming that $\dim\mathscr{E} = 3$, we suppose the surface $\mathscr{S}$ to be given by a mapping $\mathbf{x} = \mathbf{f}(a, b)$ on a domain $\mathscr{D}$ of the parameters a and b. Then, with the usual convention of sign and on the assumption that the fields and the surface be sufficiently smooth,

$$\mathrm{C}(\mathscr{C}) := \int_{\mathscr{C}} \dot{\mathbf{x}}\cdot d\mathbf{x} = \int_{\mathscr{D}} \mathbf{W}\cdot(\partial_a\mathbf{x}\wedge\partial_b\mathbf{x})\, da\, db. \qquad \text{(II.11-35)}$$

For our first use of this statement, we apply it to a surface $\mathscr{S}$ that is normal to the velocity field $\dot{\mathbf{x}}$. Then $\mathrm{C}(\mathscr{C}) = 0$, and so the right-hand side of (35) vanishes. The same holds for every subsurface of $\mathscr{S}$. If $\mathbf{W}$ and $\partial_a\mathbf{x}\wedge\partial_b\mathbf{x}$ are continuous, then everywhere on $\mathscr{S}$

$$\mathbf{W}\cdot(\partial_a\mathbf{x}\wedge\partial_b\mathbf{x}) = 0. \qquad \text{(II.11-36)}$$

We have proved the following theorem: *At a point on a surface normal to the velocity field, either* $\mathbf{W} = \mathbf{0}$ *or the axis of* $\mathbf{W}$ *lies in the tangent plane.* Therefore, if $\mathbf{n}$ is a unit normal field to $\mathscr{S}$, we can take $\mathbf{n}$ for $\mathbf{e}$ in $(14)_3$ and conclude that at a point on $\mathscr{S}$

$$\mathbf{W} = -\mathbf{n}\otimes\mathbf{W}\mathbf{n} + \mathbf{W}\mathbf{n}\otimes\mathbf{n}. \qquad \text{(II.11-37)}$$

The foregoing statements hold *a fortiori* on a stationary boundary to which a body adheres.

Exercise II.11.11 (WEATHERBURN, BERKER, CASWELL, TRUESDELL). Interpretation of the gradient in terms of the directional derivative shows that if $\mathbf{k}$ is any vector in the tangent plane at the place $\mathbf{x}$ on a stationary wall to which a body adheres, then at $\mathbf{x}$

$$\mathbf{G}\mathbf{k} = \mathbf{0}. \qquad \text{(II.11-38)}$$

[1] A surface is a compact, oriented, two-dimensional manifold with boundary in a three-dimensional Euclidean space. The velocity field $\dot{\mathbf{x}}$ is assumed to be differentiable in an open set properly containing $\mathscr{S}$. A brief statement and a rigorous proof of KELVIN's transformation are given by M. SPIVAK at the end of his book cited above in the footnote on p. 93.

Hence at $\mathbf{x}$

$$\mathbf{D} = E\mathbf{n} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{W}\mathbf{n} + \mathbf{W}\mathbf{n} \otimes \mathbf{n}, \qquad \text{(II.11-39)}$$

and the principal stretchings are given by

$$\begin{aligned} 2D_1 &= E + \sqrt{E^2 + w^2} \geqq 0, \\ D_2 &= D_{(\mathbf{e})} = 0, \\ 2D_3 &= E - \sqrt{E^2 + w^2} \leqq 0. \end{aligned} \qquad \text{(II.11-40)}$$

Exercise II.11.12 (TRUESDELL). If a body undergoing an isochoric motion which is not rigid adheres to a stationary surface, then $\mathfrak{W} = 1$ on that surface.

Exercise II.11.13 (CAUCHY). If

$$\mathbf{W}_{\mathbf{a}} := \operatorname{skw} \mathbf{G}_2 = \operatorname{skw} \operatorname{grad} \ddot{\mathbf{x}}, \qquad \text{(II.11-41)}$$

then

$$(\mathbf{F}^{\mathsf{T}}\mathbf{W}\mathbf{F})^{\cdot} = \mathbf{F}^{\mathsf{T}}\mathbf{W}_{\mathbf{a}}\mathbf{F}. \qquad \text{(II.11-42)}$$

Hence a necessary and sufficient condition that $\mathbf{F}^{\mathsf{T}}\mathbf{W}\mathbf{F}$ remain constant for each substantial point X in the course of its motion is

$$\mathbf{W}_{\mathbf{a}} = \mathbf{0}. \qquad \text{(II.11-43)}$$

If (43) holds, then

$$\mathbf{F}^{\mathsf{T}}\mathbf{W}\mathbf{F} = \mathbf{f}, \qquad \text{(II.11-44)}$$

a function of place $\mathbf{X}$ in the reference shape. Because $\mathbf{F} = \mathbf{1}$ throughout that shape, from (44) we conclude that $\mathbf{f} = \mathbf{W}_{\kappa}$, the spin that X would have, were it to be at $\mathbf{X}$. In particular, (43) is satisfied by an irrotational flow.

The condition (43) is of central importance in classical fluid dynamics. There it is applied in a region, not merely to a single substantial point. It is called the ***D'Alembert–Euler condition***. A convenient way to express it is

$$\operatorname{skw} \operatorname{grad} \ddot{\mathbf{x}} = \mathbf{0}; \qquad \text{(II.11-45)}$$

because of (30), equivalently

$$\operatorname{skw}\ddot{\mathbf{F}}\mathbf{F}^{-1} = \operatorname{skw}(\dot{\mathbf{G}} + \mathbf{G}^2) = \mathbf{0}. \qquad \text{(II.11-46)}$$

We shall learn further consequences of this condition in Sections II.13, IV.14 and in Volume 2. For the time being we remark only that according to a familiar theorem on lamellar fields, in a simply connected region the field $\ddot{\mathbf{x}}$ satisfies (43) if and only if there is an *acceleration-potential* $P_{\mathbf{a}}$:

$$\ddot{\mathbf{x}} = -\operatorname{grad} P_{\mathbf{a}}. \qquad \text{(II.11-47)}$$

Exercise II.11.15 (D'Alembert, Euler, Beltrami).

$$\mathbf{W}_{\mathbf{a}} = \dot{\mathbf{W}} + \mathbf{D}\mathbf{W} + \mathbf{W}\mathbf{D}. \qquad \text{(II.11-48)}$$

Exercise II.11.16 (Appell). If $\dim\mathscr{E} = 3$,

$$(\tfrac{1}{2}|J\mathbf{W}|^2)^{\cdot} = J^2(\mathbf{W}\cdot\mathbf{W}_{\mathbf{a}} + |\mathbf{W}|^2\mathbf{n}\cdot\mathbf{D}\mathbf{n}), \qquad \text{(II.11-49)}$$

$\mathbf{n}$ being either unit vector in the nullspace of $\mathbf{W}$. Hence w satisfies the differential equation

$$(Jw)^{\cdot} = Jw\mathbf{n}\cdot\mathbf{D}\mathbf{n} \qquad \text{(II.11-50)}$$

if and only if

$$\mathbf{W}\cdot\mathbf{W}_{\mathbf{a}} = 0. \qquad \text{(II.11-51)}$$

Exercise II.11.17. A rigid motion has an acceleration-potential if and only if its spin is steady, and then

$$-P_{\mathbf{a}} = \tfrac{1}{2}\mathbf{p}\cdot\mathbf{W}^2\mathbf{p} + [\dot{\mathbf{c}} + \mathbf{W}(\mathbf{c} - \dot{\mathbf{x}}_0)]\cdot\mathbf{p}; \qquad \text{(II.11-52)}$$

here $\mathbf{p} := \mathbf{x} - \mathbf{x}_0$. If ω denotes the angular speed and r the distance from the axis of spin,

$$\mathbf{p}\cdot\mathbf{W}^2\mathbf{p} = -\tfrac{1}{4}w^2r^2 = -\omega^2r^2. \qquad \text{(II.11-53)}$$

As we have seen above in this section, the condition $\mathbf{W} = \mathbf{0}$ defines an irrotational motion. Consequently, a motion is irrotational in a simply connected region if and only if it has there a *velocity-potential* P_v:

$$\dot{\mathbf{x}} = -\operatorname{grad} P_v . \tag{II.11-54}$$

For that reason irrotational motions are often called *potential flows*. The potential P_v may depend upon t as well as $\mathbf{x}$. The surfaces $P_v(t, \mathbf{x}) = \text{const.}$, t fixed, are called *equipotentials*. The velocity is normal to the equipotential on which it lies. A system of equipotentials determined by giving to P_v successively equal, constant increments, say c, divides the region of flow into laminae, and hence an irrotational flow is sometimes called *lamellar*. If the constant c is very small, so also are the values of the function d which delivers the normal distances between the equipotentials, and $|\dot{\mathbf{x}}| \approx c/d$.

If an irrotational motion is also isochoric, then, as EULER remarked, (II.5-8) reduces to the linear partial-differential equation later to be called "Laplace's":

$$\Delta P_v = 0. \tag{II.11-55}$$

Solutions, which are called *harmonic functions*, are easy to obtain. The sum of two harmonic functions is a harmonic function, and so the outcome of superposing two isochoric, irrotational flows is likewise an isochoric, irrotational flow, and complicated flows may be built up from simple ones in this way. In the nineteenth century many general properties of them were discovered, and general methods for calculating solutions of (55) such as to satisfy (II.6-17) on given boundaries were constructed. The corpus of these properties is called "potential theory". The problem of determining an isochoric, irrotational flow within or about assigned boundaries is purely kinematical; it can be phrased with no reference to mechanics.

A disquieting property of isochoric, irrotational flows is revealed by a theorem in the theory of the "Laplacian" equation: The boundary condition (II.6-17), applied to the boundary of a closed, bounded, simply connected region, determines a unique velocity field in that region. Were the fluid to adhere to some bounding wall, there we should have to prescribe $\dot{\mathbf{x}}$, not merely $\mathbf{n}\cdot\dot{\mathbf{x}}$. A standard theorem of potential theory may be interpreted as follows: *If at a certain time a body undergoing isochoric, irrotational flow adheres to a not void, open set on a surface, that whole body must be at rest at that time.*

Neither isochoric motion nor the condition of adherence nor the restriction to a bounded domain is necessary to render impossible an irrotational motion other than a state of rest, as is shown by the following, purely kinematical

Theorem of Kelvin and Helmholtz. *Let an irrotational flow in a stationary, simply connected region be such that*

1. *It is isochoric, or its density is steady.*

2. *On all finite boundaries* $\dot{\mathbf{x}} \cdot \mathbf{n} = 0$.
3. *In any part of the region that lies outside of a sphere of arbitrarily large radius* r, *if the motion is isochoric then*

$$P_{\mathrm{v}} \partial_r P_{\mathrm{v}} = o\left(\frac{1}{r^2}\right) \qquad \text{as } r \to \infty, \tag{II.11-56}$$

while if the density is steady

$$\rho P_{\mathrm{v}} \partial_r P_{\mathrm{v}} = o\left(\frac{1}{r^2}\right) \qquad \text{as } r \to \infty. \tag{II.11-57}$$

Then $\dot{\mathbf{x}} = \mathbf{0}$ *everywhere.*

Of course the two main conditions, those of isochoric motion and of steady density, are not mutually exclusive, for it is easily possible that both (56) and (57) hold.

Proof. If $\operatorname{div}(A\dot{\mathbf{x}}) = 0$, then $\operatorname{div}(AP_{\mathrm{v}}\dot{\mathbf{x}}) = A\dot{\mathbf{x}} \cdot \operatorname{grad} P_{\mathrm{v}} = -A|\operatorname{grad} P_{\mathrm{v}}|^2$. If we integrate this equation over any finite region $\mathscr{R}$ in which P_{v} exists, we obtain

$$-\int_{\mathscr{R}} A|\operatorname{grad} P_{\mathrm{v}}|^2 \, dV = \int_{\partial\mathscr{R}} AP_{\mathrm{v}}\dot{\mathbf{x}} \cdot \mathbf{n} \, dA. \tag{II.11-58}$$

Condition 2 makes the integral over $\partial\mathscr{R}$ vanish. Condition 1 makes the conclusion apply if we put $A = 1$ for an isochoric flow, $A = \rho$ for a flow with steady density; we use (II.5-9) and (II.6-6)$_2$, respectively, so as to conclude that $A|\operatorname{grad} P_{\mathrm{v}}|^2 = 0$ throughout $\mathscr{R}$. If the flow is defined over a region that is not bounded, we choose $\mathscr{R}$ as the part of the region that lies within a sphere of large radius r. The boundary $\partial\mathscr{R}$ then consists partly of points where $\dot{\mathbf{x}} \cdot \mathbf{n} = 0$ and partly of points on the sphere. The former contribute nothing to the surface integral in (58). As $r \to \infty$, the integral on the portion of the large sphere tends to 0 because of (56) or (57). Thus the integral on the right-hand side of (58) converges to 0, and again it follows that $A|\operatorname{grad} P_{\mathrm{v}}|^2 = 0$ throughout the region where P_{v} exists. $\triangle$

In a multiply connected region a potential flow that is not a state of rest may exist. An example is provided by the irrotational, simple vortex included in (18). For it $P_{\mathrm{v}} = -A\theta$.

Exercise II.11.18 (Cisotti). The kinetic energy of a body undergoing irrotational flow in a stationary, bounded, simply connected region $\mathscr{R}$ on the boundary of which

$\dot{\mathbf{x}} \cdot \mathbf{n} = 0$ is given by

$$2K = -\int_{\mathscr{R}} P_{\mathrm{v}} \rho' \, dV. \qquad \text{(II.11-59)}$$

A flow is *complex-lamellar* if it is non-trivially proportional to a lamellar flow: There are scalar fields A and P, neither of them constant, such that $\dot{\mathbf{x}} = A\nabla P$. The surfaces $P(t, \mathbf{x}) = \text{const.}$, like the equipotentials of an irrotational flow, are normal to the field $\dot{\mathbf{x}}$, but a complex-lamellar flow is rotational. Thus the axis of spin at $\mathbf{x}_0$ lies in the tangent plane of the surface $P(t, \mathbf{x}) = \text{const.}$ containing $\mathbf{x}_0$, and $\mathbf{W}$ satisfies (37). *Cf.* Section App. IIC.5.

12. Homogeneous Transplacement

A transplacement χ_{κ} of the reference placement κ is said to be *homogeneous* if the substantial points occupying each straight line segment in $\kappa(\mathscr{B})$ are carried into some straight line segment in $\chi(\mathscr{B}, t)$. By a theorem of geometry, any such transplacement χ_{κ} must be affine at each time t. Thus a homogeneous transplacement of $\kappa(\mathscr{B})$ is of the form

$$\chi_{\kappa}(\mathbf{X}, t) = \mathbf{x}_0(t) + \mathbf{F}(t)(\mathbf{X} - \mathbf{X}_0), \qquad \det \mathbf{F}(t) \neq 0. \qquad \text{(II.12-1)}$$

In this formula $\mathbf{X}_0$ is a fixed place in $\kappa(\mathscr{B})$; $\mathbf{x}_0$ is a place-valued function of time; and $\mathbf{F}$ is a tensor-valued function of time. By (II.5-1) we see that $\mathbf{F}$ is the transplacement gradient, and that at any one time t it has the same value at all places in $\chi(\mathscr{B}, t)$. This property explains the name "homogeneous transplacement": A transplacement is homogeneous if and only if its gradient is uniform at each time.

For a given reference placement κ the composition of two homogeneous transplacements is a homogeneous transplacement. For each fixed t the transplacements homogeneous with respect to κ are restrictions of members of the affine group.

If κ_1 and κ_2 are two different reference placements, a motion that gives rise to a transplacement homogeneous with respect to κ_1 generally fails to do the same with respect to κ_2. The class of motions that give rise to transplacements homogeneous with respect to κ_1 coincides with the corresponding class of κ_2 if and only if the differentiable homeomorphism $\kappa_2 \circ \kappa_1^{-1}$ has a constant gradient.

Homogeneous transplacements are most easily visualized as mappings of one vector space into another. Let $\mathbf{p}_{\kappa}$ denote the field of position vectors in $\kappa(\mathscr{B})$

with respect to the origin $\mathbf{X}_0$, and let $\mathbf{p}$ denote the field of position vectors in $\chi(\mathscr{B}, t)$ with respect to $\mathbf{x}_0(t)$. That is, $\mathbf{p}_\kappa := \mathbf{X} - \mathbf{X}_0$, and $\mathbf{p} := \mathbf{x} - \mathbf{x}_0$. Then (1) may be written in the form

$$\mathbf{p} = \mathbf{F}\mathbf{p}_\kappa , \tag{II.12-2}$$

and $\mathbf{F}$ is a function of time only.

Let the two particular position vectors $\mathbf{m}_\kappa$ and $\mathbf{n}_\kappa$ in $\kappa(\mathscr{B})$ be mapped onto $\mathbf{m}$ and $\mathbf{n}$, respectively, by (2). Then

$$\mathbf{m}\cdot\mathbf{n} = \mathbf{m}_\kappa\cdot\mathbf{C}\mathbf{n}_\kappa , \tag{II.12-3}$$

$\mathbf{C}$ being the right Cauchy–Green tensor $(\text{II}.9\text{-}5)_1$. Likewise,

$$\mathbf{m}_\kappa\cdot\mathbf{n}_\kappa = \mathbf{m}\cdot\mathbf{B}^{-1}\mathbf{n}, \tag{II.12-4}$$

$\mathbf{B}$ being the left Cauchy–Green tensor $(\text{II}.9\text{-}5)_3$. The student will recall that $\mathbf{B}$ and $\mathbf{C}$ are symmetric and positive. All vectors parallel to $\mathbf{n}_\kappa$ are increased in length in the same ratio. In particular, if $\mathbf{n}_\kappa$ is a unit vector, generally the $\mathbf{n}$ corresponding with it through (2) has some length other than 1. This ratio of lengths is called the *stretch* $v_{(\mathbf{n}_\kappa)}$ in the direction of $\mathbf{n}_\kappa$. It may be calculated as follows:

$$v_{(\mathbf{n}_\kappa)} = \sqrt{\mathbf{n}_\kappa\cdot\mathbf{C}\mathbf{n}_\kappa}. \tag{II.12-5}$$

Two orthogonal vectors $\mathbf{m}_\kappa$ and $\mathbf{n}_\kappa$ in $\kappa(\mathscr{B})$ are mapped, generally, onto vectors $\mathbf{m}$ and $\mathbf{n}$ in $\chi(\mathscr{B}, t)$ that are not orthogonal. This phenomenon is called *shear*, and there are various ways to report it. The angle $\theta_{(\mathbf{n}_\kappa, \mathbf{m}_\kappa)}$ between the images in $\chi(\mathscr{B}, t)$ of two unit vectors $\mathbf{n}_\kappa$ and $\mathbf{m}_\kappa$ in $\kappa(\mathscr{B})$ is one measure of shear. It is determined by the relation

$$\cos\theta_{(\mathbf{n}_\kappa, \mathbf{m}_\kappa)} = \frac{1}{v_{(\mathbf{n}_\kappa)}v_{(\mathbf{m}_\kappa)}}\mathbf{n}_\kappa\cdot\mathbf{C}\mathbf{m}_\kappa . \tag{II.12-6}$$

The sphere $|\mathbf{m}_\kappa|^2 = \text{const.}$ in $\kappa(\mathscr{B})$ is mapped onto an ellipsoid in $\chi(\mathscr{B}, t)$, and the sphere $|\mathbf{m}|^2 = \text{const.}$ in $\chi(\mathscr{B}, t)$ is the image of an ellipsoid in $\kappa(\mathscr{B})$.

Exercise II.12.1 (CAUCHY). The principal axes of strain, as defined in Section II.9, are the principal axes of the ellipsoids just constructed; the principal stretches are the stretches in the directions of those axes; and these particular stretches are extremal. Thus a homogeneous transplacement is resolved into a translation and a rotation of one set of principal axes into the other, followed or preceded by pure stretches along those axes. The shear of each pair of principal axes is null.

Exercise II.12.2. Let the linearly independent vectors $\mathbf{p}_\kappa$, and $\mathbf{q}_\kappa$, and $\mathbf{r}_\kappa$ determine a parallelepiped of volume V_κ. Let the volume of the parallelepiped onto which it is mapped by (1) be V. Then $J = V/V_\kappa$. Let A_κ be the area of the parallelogram determined by $\mathbf{q}_\kappa$ and $\mathbf{r}_\kappa$; let A be the area of the parallelogram onto which it is mapped; let θ_κ and θ be the angles subtended upon those parallelograms by $\mathbf{p}_\kappa$ and $\mathbf{p}$, respectively. Then $V/V_\kappa = (A/A_\kappa)(|\mathbf{p}|/|\mathbf{p}_\kappa|)(\sin\theta/\sin\theta_\kappa)$.

The terms "stretching" and "shearing" in general refer to the rates of change of stretch and shear when these latter are defined with respect to the present shape as reference. We may discuss stretching and shearing just as we have discussed stretch and shear, starting from homogeneous transplacements. If we differentiate (2) with $\mathbf{X}$ held constant, then use (II.11-5), and then use (2) again, we obtain

$$\dot{\mathbf{p}} = \dot{\mathbf{F}}\mathbf{p}_\kappa = \mathbf{G}\mathbf{F}\mathbf{p}_\kappa = \mathbf{G}\mathbf{p}. \tag{II.12-7}$$

Hence by use of (II.11-8) we derive EULER's relation

$$|\mathbf{p}||\mathbf{p}|^{\cdot} = \mathbf{p}\cdot\dot{\mathbf{p}} = \mathbf{p}\cdot\mathbf{D}\mathbf{p}, \tag{II.12-8}$$

$\mathbf{D}$ being the stretching tensor; equivalently, if $\mathbf{p} \neq \mathbf{0}$ and if $\mathbf{n}$ is a unit vector in the direction of $\mathbf{p}$, then

$$(\log|\mathbf{p}|)^{\cdot} = \mathbf{n}\cdot\mathbf{D}\mathbf{n}. \tag{II.12-9}$$

Thus the component $\mathbf{n}\cdot\mathbf{D}\mathbf{n}$ of $\mathbf{D}$ is the rate of increase of length, per unit length, of a linear segment in $\kappa(\mathscr{B})$ presently parallel to $\mathbf{n}$ in $\chi(\mathscr{B}, t)$, and this rate is called the *stretching* in the direction of $\mathbf{n}$. The three principal stretchings, which were defined in Section II.11, are the extremal stretchings.

Exercise II.12.3 (EULER). For two orthogonal unit vectors $\mathbf{n}_\kappa$ and $\mathbf{m}_\kappa$, differentiating (6) shows that

$$-\dot{\theta}_{(\mathbf{n},\mathbf{m})}|_{\mathbf{F}=\mathbf{1}} = 2\mathbf{n}\cdot\mathbf{D}\mathbf{m}. \tag{II.12-10}$$

This statement has an interpretation in terms of shearing.

Let $\varphi_{(\mathbf{p},\mathbf{m}_\kappa)}$ denote the angle between the position vector $\mathbf{p}$ of $\mathbf{x}$ with respect to $\mathbf{x}_0$ in $\chi(\mathscr{B}, t)$ and the unit vector $\mathbf{m}_\kappa$ in $\kappa(\mathscr{B})$. Then by (2),

$$|\mathbf{p}| \cos \varphi = \mathbf{m}_\kappa \cdot \mathbf{F}\mathbf{p}_\kappa \,, \tag{II.12-11}$$

in which for simplicity we do not write the subscript $(\mathbf{p}, \mathbf{m}_\kappa)$. Differentiating (11) with respect to t yields

$$\begin{aligned} \cos \varphi \, |\mathbf{p}|^{\cdot} - |\mathbf{p}| \dot{\varphi} \sin \varphi &= \mathbf{m}_\kappa \cdot \dot{\mathbf{F}}\mathbf{p}_\kappa \,, \\ &= \mathbf{m}_\kappa \cdot \mathbf{G}\mathbf{F}\mathbf{p}_\kappa \,, \end{aligned} \tag{II.12-12}$$

by (II.11-5). If we now let the value of $\mathbf{p}_\kappa$ be a unit vector orthogonal to $\mathbf{m}_\kappa$, say $\mathbf{n}_\kappa$, and then take the present shape as the reference shape, so that the corresponding value of $\mathbf{p}$ also is $\mathbf{n}$, we find that

$$\dot{\varphi}_{(\mathbf{n}, \mathbf{m})}|_{\mathbf{F}=\mathbf{1}} = -\mathbf{m} \cdot \mathbf{G}\mathbf{n}. \tag{II.12-13}$$

This formula gives the angular rate at which a line segment in $\kappa(\mathscr{B})$ presently parallel to $\mathbf{n}$ turns *away* from the stationary unit vector $\mathbf{m}$ in $\chi(\mathscr{B}, t)$. Likewise, the rate at which a line segment in $\kappa(\mathscr{B})$ presently parallel to $\mathbf{m}$ is turning *toward* the stationary unit vector $\mathbf{n}$ is given by

$$\dot{\varphi}_{(\mathbf{m}, -\mathbf{n})}|_{\mathbf{F}=\mathbf{1}} = +\mathbf{n} \cdot \mathbf{G}\mathbf{m}. \tag{II.12-14}$$

By adding these formulae and using (II.11-8) we obtain

$$\tfrac{1}{2}(\dot{\varphi}_{(\mathbf{n}, \mathbf{m})} + \dot{\varphi}_{(\mathbf{m}, -\mathbf{n})})|_{\mathbf{F}=\mathbf{1}} = \mathbf{n} \cdot \mathbf{W}\mathbf{m}, \tag{II.12-15}$$

$\mathbf{W}$ being the spin. Thus we have proved a ***fundamental theorem*** of CAUCHY: *The component* $\mathbf{n} \cdot \mathbf{W}\mathbf{m}$ *of* $\mathbf{W}$ *corresponding to the orthogonal unit vectors* $\mathbf{n}$ *and* $\mathbf{m}$ *is the arithmetic mean of the rates of right-handed rotation of a line in* $\kappa(\mathscr{B})$ *presently parallel to* $\mathbf{n}$ *with respect to the direction of* $\mathbf{m}$ *in* $\chi(\mathscr{B}, t)$ *and of a line in* $\kappa(\mathscr{B})$ *presently parallel to* $\mathbf{m}$ *with respect to the direction of* $\mathbf{n}$ *in* $\chi(\mathscr{B}, t)$.

Exercise II.12.4. Because of (3)

$$(\mathbf{m} \cdot \mathbf{n})^{(k)} = \mathbf{m} \cdot \mathbf{A}_k \mathbf{n}, \qquad k = 1, 2, \ldots . \tag{II.12-16}$$

For a general motion the transplacement gradient $\mathbf{F}$ provides a local linear approximation to the transplacement χ_κ. We may say that to within an error that is $\mathbf{o}(\mathbf{X}-\mathbf{X}_0)$ as $\mathbf{X}-\mathbf{X}_0 \to \mathbf{0}$ the transplacement χ_κ is approximated at $\mathbf{X}_0$ in $\kappa(\mathscr{B})$ and hence at $\mathbf{x}_0$ in $\chi(\mathscr{B}, t)$, with an error $\mathbf{o}(\mathbf{x}-\mathbf{x}_0)$ as $\mathbf{x}-\mathbf{x}_0 \to \mathbf{0}$, by the homogeneous transplacement (1) that is defined by $\mathbf{F}(\mathbf{X}_0, t)$. Thus the conclusions reached in this section for homogeneous transplacements may be interpreted in general motions as first-order local approximations to counterparts for the present transplacement of $\kappa(\mathscr{B})$. In loose language, the conclusions valid for all lines in homogeneous transplacement are valid for infinitesimal line segments in any smooth transplacement.

For reference we record here also the velocity field and the acceleration field of the homogeneous transplacement (1):

$$\begin{aligned}\dot{\mathbf{x}} &= \dot{\mathbf{x}}_0 + \dot{\mathbf{F}}\mathbf{F}^{-1}(\mathbf{x}-\mathbf{x}_0) = \dot{\mathbf{x}}_0 + \mathbf{G}(\mathbf{x}-\mathbf{x}_0),\\ \ddot{\mathbf{x}} &= \ddot{\mathbf{x}}_0 + \ddot{\mathbf{F}}\mathbf{F}^{-1}(\mathbf{x}-\mathbf{x}_0) = \ddot{\mathbf{x}}_0 + (\dot{\mathbf{G}} + \mathbf{G}^2)(\mathbf{x}-\mathbf{x}_0).\end{aligned} \tag{II.12-17}$$

That these fields are given by affine functions of place, should be obvious without calculation and may be verified also by a glance at (II.11-5) and (II.11-30).

Suppose, conversely, that the velocity field of a transplacement of $\mathscr{B}$ be affine: For each place $\mathbf{x}$ in $\chi(\mathscr{B}, t)$ and each t in some interval,

$$\dot{\mathbf{x}} = \mathbf{c} + \mathbf{K}(\mathbf{x}-\bar{\mathbf{x}}), \tag{II.12-18}$$

in which $\mathbf{c}$ and $\mathbf{K}$ are functions of t alone and $\bar{\mathbf{x}}$ is a fixed place. Then $\mathbf{K} = \operatorname{grad}\dot{\mathbf{x}}$, and so from (II.11-7) and (II.11-5) we see that a transplacement gradient $\mathbf{F}$ from which $\mathbf{K}$ derives must satisfy the differential equation $\dot{\mathbf{F}}\mathbf{F}^{-1} = \mathbf{K}$. Solutions $\mathbf{F}$ may be functions of $\mathbf{x}$ as well as of t, but there is a solution that is a function of t alone, unique to within an initial value. *Cf.* (I.9-16).

13. Rates of Change of Integrals over Substantial Lines, Surfaces, and Regions. Substantial Vector Lines. The Vorticity Theorems of Helmholtz and Kelvin

In Section II.1 we introduced the notation $\chi(\mathscr{B}, t)$ for the shape of a body $\mathscr{B}$ at the time t. Likewise, any subset of the substantial points comprised by $\mathscr{B}$ will be given in time a sequence of shapes by χ. In Section II.3 we introduced a reference placement κ to assign a place $\mathbf{X}$ in $\mathscr{E}$ to each substantial point X, in terms of which we defined through (II.3-3) the transplacement χ_κ of the substantial points of $\mathscr{B}$ from their places $\mathbf{X}$ in the reference placement into the places $\mathbf{x}$ they occupy at the time t. Our assumptions of smoothness should suffice to ensure that χ_κ preserve the nature of the subsets of $\mathscr{B}$. For example, a surface $\mathscr{S}$ in $\kappa(\mathscr{B})$ should be mapped at the time t into a surface in $\chi_\kappa(\mathscr{B}, t)$. This sequence of surfaces $\chi_\kappa(\mathscr{S}, t)$ provides the successive loci of a *substantial surface* under χ, for at each t one and the same set of substantial points occupies $\chi_\kappa(\mathscr{S}, t)$. The same idea can be applied also to a line $\mathscr{L}$ in $\kappa(\mathscr{B})$; the sequence $\chi_\kappa(\mathscr{L}, t)$ provides the successive loci of a *substantial line*. If $\mathscr{A}$ is a subbody of $\mathscr{B}$, then $\chi_\kappa(\mathscr{A}, t)$ is the shape at t of a *substantial region*.

In Section II.6 we have shown how to calculate substantial time-rates for quantities given in the spatial description, and we have provided the criterion (II.6-13) to determine whether a set of points in $\mathscr{E}$ that satisfies $f(\mathbf{x}, t) = 0$ do or do not provide the successive shapes of a substantial surface under χ. Also in (II.6-8) and (II.6-10) we see the rule for calculating the substantial derivative of an integral over a substantial volume. As those examples show, the value of an integral of a spatial field over a substantial set will generally change in time for two reasons: first, because the field itself changes, and, second, because the domain of integration in $\mathscr{E}$ is changing in consequence of the motion.

We now enter into the details concerning integrals over substantial lines and substantial surfaces. Before going ahead, the student would do well to refresh his knowledge of the contents of Section II.3 and the first half of Section II.6.

We have mentioned that formulae valid strictly for homogeneous transplacements serve as first-order approximations in general. Since only the first-order terms affect the value of an integral, we may derive in this way exact formulae for the time-rate of change of integrals. For example, if $\mathscr{C}$ is a given curve in $\kappa(\mathscr{B})$, the time derivative of a line integral along its shape $\chi(\mathscr{C}, t)$ is obtained by supposing that the substantial rate of change $\dot{\overline{d\mathbf{x}}}$ of the element of arc $d\mathbf{x}$ is $\mathbf{G}\,d\mathbf{x}$, as (II.12-7) suggests. Thus we infer the following formula, due to KELVIN:

$$\begin{aligned}\frac{d}{dt}\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x} &= \int_{\mathscr{C}} [\dot{\mathbf{f}}\cdot d\mathbf{x} + \mathbf{f}\cdot(\mathbf{G}\,d\mathbf{x})],\\ &= \int_{\mathscr{C}} (\dot{\mathbf{f}} + \mathbf{G}^{\mathsf{T}}\mathbf{f})\cdot d\mathbf{x}.\end{aligned} \tag{II.13-1}$$

The abbreviated notation $\int_{\mathscr{C}}$ denotes integration over the parametric interval of the function $\mathbf{k}$ that defines $\mathscr{C}$ in the reference shape: say $\mathbf{X} = \mathbf{k}(l)$, $l \in [0, 1]$. The student should clear the details by solving the following exercise. They will be made obvious anyway by the treatment for the analogous but more complicated problem for surface integrals which we shall give a little further on.

Exercise II.13.1. Transforming line integrals along $\chi(\mathscr{C}, t)$ back into integrals along the stationary curve $\mathscr{C}$ in $\boldsymbol{\kappa}(\mathscr{B})$ delivers a formal proof of (1).

Exercise II.13.2. Use of (II.12-7) to calculate the rate of change of the volume of a substantial region in a homogeneous transplacement provides another proof of (II.6-8).

Exercise II.13.3. If $\int_{\mathscr{C}} \cdots ds$ denotes integration with respect to arc length along a substantial curve $\mathscr{C}$, and if $\mathbf{t}(s)$ is either of the two continuous fields of unit vectors tangent to the present shape of $\mathscr{C}$ at s, then

$$\frac{d}{dt}\int_{\mathscr{C}} f\,ds = \int_{\mathscr{C}} (\dot{f} + f\mathbf{t}\cdot\mathbf{D}\mathbf{t})\,ds. \tag{II.13-2}$$

Now suppose that a substantial surface $\mathscr{S}$ has the parametric representation $\mathbf{X} = \mathbf{H}(a, b)$ in $\boldsymbol{\kappa}(\mathscr{B})$, a and b being real parameters varying separately in some interval, say $[0, 1]$. The present shape of this substantial surface is $\mathbf{x} = \chi_{\kappa}(\mathbf{H}(a, b), t) =: \mathbf{h}(a, b, t)$. Let $\partial_a\mathbf{X}$ and $\partial_b\mathbf{X}$ denote the partial derivatives of $\mathbf{H}$; let $\partial_a\mathbf{x}$ and $\partial_b\mathbf{x}$ denote the partial derivatives of $\mathbf{h}$. Then by use of the rule for differentiating composite functions, followed by use of (II.11-5) and properties of the exterior product, we find that[1]

$$\begin{aligned} \partial_a\mathbf{x} &= \mathbf{F}\partial_a\mathbf{X}, \\ (\partial_a\mathbf{x})^{\cdot} &= \dot{\mathbf{F}}\partial_a\mathbf{X} = \dot{\mathbf{F}}\mathbf{F}^{-1}\partial_a\mathbf{x} = \mathbf{G}\partial_a\mathbf{x}, \\ (\partial_a\mathbf{x}\wedge\partial_b\mathbf{x})^{\cdot} &= \mathbf{G}(\partial_a\mathbf{x}\wedge\partial_b\mathbf{x}) + (\partial_a\mathbf{x}\wedge\partial_b\mathbf{x})\mathbf{G}^{\mathsf{T}}. \end{aligned} \tag{II.13-3}$$

Let $\mathbf{S}$ be a skew tensor field. By use of (3) we quickly obtain LAMB's formula

$$\begin{gathered} \frac{d}{dt}\int_{\mathscr{S}} \mathbf{S}\cdot(\partial_a\mathbf{x}\wedge\partial_b\mathbf{x})\,da\,db = \int_{\mathscr{S}} \mathbf{S}^{\mathrm{c}}\cdot(\partial_a\mathbf{x}\wedge\partial_b\mathbf{x})\,da\,db, \\ \mathbf{S}^{\mathrm{c}} := \dot{\mathbf{S}} + \mathbf{S}\mathbf{G} + \mathbf{G}^{\mathsf{T}}\mathbf{S}. \end{gathered} \tag{II.13-4}$$

The integral on the left-hand side is called the *flux* of $\mathbf{S}$ through the present shape $\chi(\mathscr{S}, t)$ of the substantial surface $\mathscr{S}$. From (4) we read off ***Zorawski's criterion:*** *In order for the flux of a skew tensor field* $\mathbf{S}$ *to remain constant in time for each substantial surface, it is necessary and sufficient that*

$$\mathbf{S}^{\mathrm{c}} = \mathbf{0}. \tag{II.13-5}$$

[1] Decompositions of these and other kinematical rates are given by J. CASEY, "Connections between kinematics of line, area, and volume elements," *Journal of Elasticity* **17**(1987): 71–74.

In this notation the theorem (II.11-48) of D'ALEMBERT, EULER, and BELTRAMI appears as

$$\mathbf{W}^{c} = \mathbf{W}_{a}. \tag{II.13-6}$$

If for $\mathbf{S}$ we take $\mathbf{W}$, the integral upon the left-hand side of (4) becomes the *flux of vorticity* through $\chi(\mathscr{S}, t)$. From ZORAWSKI's criterion (5) we then read off a classic vorticity theorem: *In order that the flux of vorticity through each substantial surface shall remain constant in time, it is necessary and sufficient that*

$$\mathbf{W}_{a} = \mathbf{0}. \qquad \text{(II.11-43)}_{r}$$

The statement that the D'Alembert–Euler condition (II.11-43) is sufficient for constant flux is ***Helmholtz's Third Vorticity Theorem***.

The substantial derivative $\dot{\mathbf{f}}$, defined by (II.6-3), reflects use of the Euclidean parallel transport. Following the path of a substantial point from the place it occupies at the time t to the place it occupies at the time $t + h$, we use the Euclidean parallel transport to translate the value of $\mathbf{f}$ at the latter point back to the former point, subtract from it the value of $\mathbf{f}$ there, divide by h, and pass to the limit to obtain $\dot{\mathbf{f}}$. If we do just the same thing but use the parallelism induced by the motion of the deforming body, we obtain[1]

$$\begin{aligned} \mathbf{f}^{c} &:= \mathbf{f}' + (\operatorname{grad}\mathbf{f})\dot{\mathbf{x}} - \mathbf{G}\mathbf{f}, \\ &= \dot{\mathbf{f}} - \mathbf{G}\mathbf{f}. \end{aligned} \tag{II.13-7}$$

Thus $\mathbf{f}^{c} = \mathbf{0}$ if and only if $\mathbf{f}$ obeys in all motions the same relation as does the position vector of a substantial point in a homogeneous transplacement, namely, (II.12-7)_3. The quantity $\mathbf{S}^{c}$ defined by $(4)_2$ has a similar interpretation. We may refer to $\mathbf{f}^{c}$ and $\mathbf{S}^{c}$ as the *convected time-fluxes* of $\mathbf{f}$ and $\mathbf{S}$, respectively.

The *vector lines* of a non-vanishing vector field $\mathbf{f}$ are the curves everywhere tangent to $\mathbf{f}$. Generally these curves move and deform in the course of time. If they do so in such a way as to be occupied always by the same set of substantial points, they are substantial lines. A field of such a kind has *substantial vector lines*. A substantial line that once is a vector line of $\mathbf{f}$ is then always a vector line of $\mathbf{f}$.

[1] A clear explanation of the idea is included in the paper by BOLDER cited on p. 104. The standard way to introduce the convected derivative begins from the *Lie derivative* $\pounds_{\mathbf{v}}$ based upon a vector field $\mathbf{v}$ and then sets $\mathbf{f}^{c} := \mathbf{f}' + \pounds_{\dot{\mathbf{x}}}\mathbf{f}$. For the Lie derivative a standard, old reference is Section 10 of J. A. SCHOUTEN, *Ricci-Calculus*, Berlin, Springer-Verlag, 1954.

To determine the fields $\mathbf{f}$ that have substantial vector lines, we let a curve $\mathscr{C}$ in $\boldsymbol{\kappa}(\mathscr{B})$ be given by the parametric representation $\mathbf{X} = \mathbf{H}(a)$ and proceed as we did above in considering substantial surfaces, and we assume that $\mathbf{f}$ does not vanish anywhere. By use of $(3)_4$ we obtain

$$(\mathbf{f} \wedge \partial_a \mathbf{x})^{\cdot} = \dot{\mathbf{f}} \wedge \partial_a \mathbf{x} + \mathbf{f} \wedge \mathbf{G}\, \partial_a \mathbf{x}. \qquad \text{(II.13-8)}$$

The material line generated by $\mathscr{C}$ is presently a vector line of $\mathbf{f}$ if and only if there is a scalar field A such that

$$\partial_a \mathbf{x} = A\mathbf{f}. \qquad \text{(II.13-9)}$$

The substantial line then remains always a vector line if and only if (9) implies for all t that

$$(\mathbf{f} \wedge \partial_a \mathbf{x})^{\cdot} = \mathbf{0}. \qquad \text{(II.13-10)}$$

Putting (9) into (8) yields

$$(\mathbf{f} \wedge \partial_a \mathbf{x})^{\cdot} = A(\dot{\mathbf{f}} - \mathbf{G}\mathbf{f}) \wedge \mathbf{f}. \qquad \text{(II.13-11)}$$

Comparison with (7) yields the ***Helmholtz-Zorawski criterion***: *The field* $\mathbf{f}$ *has substantial vector lines if and only if*

$$\mathbf{f} \wedge \mathbf{f}^{\mathrm{c}} = \mathbf{0}. \qquad \text{(II.13-12)}$$

We can express this statement equivalently in terms of the unit vector $\mathbf{e}$ in the direction of $\mathbf{f}$, that is, $\mathbf{e} := |\mathbf{f}|^{-1}\mathbf{f}$:

$$\mathbf{f}^{\mathrm{c}} = (\mathbf{e} \cdot \mathbf{f}^{\mathrm{c}})\mathbf{e}. \qquad \text{(II.13-13)}$$

In this formula we may, if we like, choose $\mathbf{f}$ to be a field $\mathbf{e}$ of unit magnitude. Then

$$\mathbf{e}^{\mathrm{c}} = (\mathbf{e} \cdot \mathbf{e}^{\mathrm{c}})\mathbf{e}. \qquad \text{(II.13-14)}$$

Of course $\mathbf{e} \cdot \dot{\mathbf{e}} = 0$, and from (7) we see that $\mathbf{e} \cdot \mathbf{e}^{\mathrm{c}} = -\mathbf{e} \cdot \mathbf{G}\mathbf{e}$, and so in general $\mathbf{e} \cdot \mathbf{e}^{\mathrm{c}} \neq 0$.

By use of (13) we may express the conclusions in Section II.12 in more general forms, without recourse either to homogeneous transplacements or to infinitesimals. The student should convince himself of this fact by solving the following exercise.

Exercise II.13.4 (STOKES, BOUSSINESQ, GOSIEWSKI, TRUESDELL & TOUPIN, WANG[1]). Let $\mathbf{e}$ be a field of unit vectors having substantial vector lines. Then

$$\dot{\mathbf{e}} = [\mathbf{D} + \mathbf{W} - (\mathbf{e}\cdot\mathbf{De})\mathbf{1}]\mathbf{e}. \tag{II.13-15}$$

If such an $\mathbf{e}$ presently lies in a principal axis of stretching, it is presently suffering a rigid motion with spin $\mathbf{W}$. If $\mathbf{m}$ and $\mathbf{n}$ are unit vector fields having substantial vector lines, then

$$\begin{aligned}(\mathbf{m}\cdot\mathbf{n})^{\cdot} &= 2\mathbf{m}\cdot\mathbf{Dn} - (\mathbf{m}\cdot\mathbf{Dm} + \mathbf{n}\cdot\mathbf{Dn})(\mathbf{m}\cdot\mathbf{n}),\\ \mathbf{m}\cdot\dot{\mathbf{n}} - \mathbf{n}\cdot\dot{\mathbf{m}} &= 2\mathbf{m}\cdot\mathbf{Wn} + (\mathbf{m}\cdot\mathbf{Dm} - \mathbf{n}\cdot\mathbf{Dn})(\mathbf{m}\cdot\mathbf{n}).\end{aligned} \tag{II.13-16}$$

Thus substantial lines orthogonal at one instant do not generally remain orthogonal. For example, the principal axes of stretching (*cf.* the Euler–Cauchy–Stokes Decomposition (II.11-8)) are not generally substantial. These conclusions may be related to those on homogeneous transplacements, given above in Section II.12.

The vector lines of a flow are called its *streamlines*. Generally these lines vary from one time to another; they are not generally the paths of the substantial points. It is plain that *the streamlines and the paths of the substantial points coincide if and only if both are steady.*[2] In order that a family of lines be steady, it is necessary and sufficient that any tangent field shall suffer change only in magnitude, not in direction. Therefore, in order that the streamlines of a non-vanishing flow $\dot{\mathbf{x}}$ be steady, it is necessary and sufficient that

$$\dot{\mathbf{x}} \wedge \dot{\mathbf{x}}' = \mathbf{0}. \tag{II.13-17}$$

This same formula should emerge also as a condition for the streamlines to be substantial, and it does. Indeed, if we apply $(7)_1$ to $\dot{\mathbf{x}}$, we find that $\dot{\mathbf{x}}^{\mathrm{c}} = \dot{\mathbf{x}}'$, and placing this conclusion in (12) yields (17). Of course (17) is satisfied by any steady flow.

Let $\mathbf{S}$ denote a field of skew tensors. A curve whose tangent at each $\mathbf{x}$ lies in the nullspace of $\mathbf{S}(\mathbf{x})$ is a *vector line* of $\mathbf{S}$. In discussing such vector lines we shall presume that $\dim \mathscr{E} = 3$. Then if $\mathbf{S} \neq \mathbf{0}$, the vector lines of the field $\mathbf{S}$ are the vector lines of the field of axes of $\mathbf{S}$.

[1] Further conclusions and interpretations are provided by C.-C. WANG, "On Gosiewski's theorem," *Archives of Mechanics* **24** (1972): 309–314.

[2] For analysis of streamlines, pathlines, and streaklines, with illustrations both graphic and analytic, see Sections 70–71 of CFT.

Exercise II.13.5. If $\dim \mathscr{E} = 3$, a substantial line that is once a vector line of the skew-tensor field **S** remains always a vector line of **S** if and only if

$$\mathbf{S}\mathbf{S}^c = \mathbf{S}^c\mathbf{S}. \qquad \text{(II.13-18)}$$

Hence in order for the vector lines of **S** to be substantial it is sufficient but not necessary that the flux of **S** through each substantial surface shall remain constant in time.

The vector lines of **W** are called *vortex lines*. If we take **W** for **S** in (18) and use (II.11-48), we obtain the condition

$$\mathbf{W}\mathbf{W}_{\mathfrak{a}} = \mathbf{W}_{\mathfrak{a}}\mathbf{W}. \qquad \text{(II.13-19)}$$

Since two non-null skew tensors commute if and only if they have the same axis, from (19) we read off a theorem due to POINCARÉ: *In a rotational flow, for a substantial line that is once a vortex line to remain always a vortex line, it is necessary and sufficient that either* $\mathbf{W}_{\mathfrak{a}} = \mathbf{0}$ *or the vector lines of* $\mathbf{W}_{\mathfrak{a}}$ *be the vortex lines.* The former alternative yields the celebrated ***Second Vorticity Theorem*** of HELMHOLTZ: *In a rotational flow that satisfies the D'Alembert-Euler condition* (II.11-43), *a substantial line that is once a vortex line is always a vortex line.*

In Section II.11 we have mentioned and used KELVIN's transformation when applied to the velocity field on a surface $\mathscr{S}$ whose border is the circuit $\mathscr{C}$:

$$C(\mathscr{C}) := \int_{\mathscr{C}} \dot{\mathbf{x}}\cdot d\mathbf{x} = \int_{\mathscr{S}} \mathbf{W}\cdot(\partial_a \mathbf{x} \wedge \partial_b \mathbf{x})\, da\, db. \qquad \text{(II.11-35)}_r$$

We now interpret the statement in general: *The circulation of a circuit equals the flux of the spin through any surface whose border is that circuit.* The usual convention of orientation is adopted here, and the fields and surfaces are presumed smooth enough to ensure the validity of the transformation.

A surface consisting entirely of vortex lines is a *vortex surface*. From (II.11-35) we see that at a given instant, *a surface is a vortex surface if and only if the circulation of every sufficiently small circuit on it is null.* In particular, *a flow in a region is irrotational if and only if the circulation of every sufficiently small circuit in that region is null.*

A flow such that the circulation of every substantial circuit is constant in time is said to *preserve circulation*. Because of (II.11-35) we may express HELMHOLTZ's Third Vorticity Theorem and its converse as follows: *The D'Alembert-Euler condition* (II.11-43) *is necessary and sufficient that the flow preserve circulation.* KELVIN's proof of this fact amounts to substitution of $\dot{\mathbf{x}}$ for $\mathbf{f}$ in (1) so as to obtain for a substantial circuit $\mathscr{C}$

$$\frac{d}{dt}\int_{\mathscr{C}} \dot{\mathbf{x}}\cdot d\mathbf{x} = \int_{\mathscr{C}} (\ddot{\mathbf{x}} + \operatorname{grad}(\tfrac{1}{2}\dot{x}^2))\cdot d\mathbf{x}, \\ = \int_{\mathscr{C}} \ddot{\mathbf{x}}\cdot d\mathbf{x}, \qquad \text{(II.13-20)}$$

the second step being a consequence of the fact that $\mathscr{C}$ is a circuit. In virtue of a standard theorem on lamellar fields, the integral on the right-hand side vanishes for all $\mathscr{C}$ if and only if grad $\ddot{\mathbf{x}}$ is symmetric. The conclusion then follows by (II.11-45).

There are several ways to see that *every irrotational flow preserves circulation*. One way has been indicated in Exercise II.11.13.

Exercise II.13.6. The simple vortex (II.11-18) preserves circulation for all choices of ω, and the circulation of the circle r = const., $0 \leqq \theta < 2\pi$, z = const., described counterclockwise, is $2\pi r^2\omega(r)$. Hence for the irrotational vortex the circulation of a curve which encircles the axis n times counterclockwise and m times clockwise is $2\pi(n-m)K$.

Exercise II.13.7 (Kelvin). The Helmholtz Theorems and the Lagrange-Cauchy Theorem (Exercise II.11.14) expressed for a substantial region follow directly from the concept of circulation.

Exercise II.13.8 (Appell, in principle). A motion with substantial vortex lines preserves circulation if and only if the vorticity satisfies the differential relation

$$(Jw)^{\cdot} = Jw\mathbf{n}\cdot\mathbf{D}\mathbf{n}, \qquad \text{(II.11-50)}_{\mathrm{r}}$$

$\mathbf{n}$ being either unit vector in the nullspace of $\mathbf{W}$.

We can now prove an important theorem of Appell: *A rotational motion with substantial vortex lines preserves circulation if and only if*

$$\int_{\mathscr{C}} \frac{ds}{Jw} = \text{const.} \qquad \text{(II.13-21)}$$

for every finite segment $\mathscr{C}$ of a substantial vortex line. Indeed, because of (2)

$$\frac{d}{dt}\int_{\mathscr{C}} \frac{ds}{Jw} = \int_{\mathscr{C}} \left[-\frac{(Jw)^{\cdot}}{(Jw)^2} + \frac{1}{Jw}\mathbf{t}\cdot\mathbf{D}\mathbf{t}\right] ds, \qquad \text{(II.13-22)}$$

$\mathbf{t}$ being either of the two continuous fields of unit vectors tangent to the present shape of $\mathscr{C}$. To conclude the proof, we apply the conclusion of the preceding exercise. The theorem shows that if the motion preserves circulation, the substantial vortex lines grow longer if w/ρ increases, shorter if w/ρ decreases. The same result may be inferred also from HELMHOLTZ's Second Theorem. For an isochoric motion that preserves circulation the statement is still simpler: the vortex lines stretch or shrink according as the spin at points upon them increases or decreases.

A motion whose streamlines are steady need not be a steady motion.[1]

Streamlines are often fairly easy to observe or trace in flows of water and other fluids. There is a vast literature concerning the kinematics of isochoric motions that preserve circulation. The streamlines of a steady, isochoric, potential flow essentially determine the quantities associated with that flow, but there are exceptions. In what follows now we shall consider the much broader class of steady, isochoric flows that preserve circulation. While such motions, like potential flows, are thought of mainly as pertaining to the solutions of the dynamical equation of an Eulerian fluid (defined below in Sections IV.4 and IV.7), they can be regarded and studied as purely kinematical developments of the purely kinematical postulate (II.11-45), which is called "the D'Alembert–Euler condition".[2]

A fascinating instance is provided by HAMEL's analysis[3] of isochoric, potential flows having constant speed on each streamline. He claimed to have proved that the streamlines of such a flow had to be parallel straight lines or circular helices mounted on concentric cylinders, but in the paper he published he did not give his formal proof, which, he wrote, filled a small notebook. During the war of the nineteen-forties the notebook disappeared. Of his exposition, MARRIS wrote[4] that he "essentially explained how he had achieved a proof of the theorem, rather than presenting an explicitly demonstrated proof." While various students attempted in vain to construct a demonstration, others took HAMEL's word for what he had done and used his statement as if it had been established. PRIM[5], mentioning some properties of flows as having been "known to HAMEL",

[1] *Cf.* A. W. MARRIS, "Unsteady motions with steady streamlines," *Archive for Rational Mechanics and Analysis* **109**(1990): 95–106.

[2] Among the studies adopting this approach are *The Kinematics of Vorticity*, cited at the end of this chapter, and the paper by A. W. MARRIS, "On steady three-dimensional motions," *Archive for Rational Mechanics and Analysis* **35**(1969): 122–168.

[3] G. HAMEL, "Potentialströmungen mit konstanter Geschwindigkeit," *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, physisch-mathematische Klasse (1937), pp. 5–20.

[4] A. W. MARRIS, "Hamel's theorem," *Archive for Rational Mechanics and Analysis* **51**(1973): 85–105.

[5] R. C. PRIM, "Steady rotational flow of ideal gases," *Journal of Rational Mechanics and Analysis* **1**(1952): 425–497. See Section Vb.

proved a cognate theorem: If there is a steady, isochoric, complex-lamellar flow which preserves circulation and has constant speed on each streamline, there is a steady isochoric, potential flow which has the same streamlines and constant speed on each of them. Finally MARRIS, through a difficult analysis,[1] achieved a formal proof. Later,[2] by a short and elegant argument, MARRIS proved that the only steady, isochoric, rotational flow that preserves circulation and has as its streamlines those of a potential flow is complex-lamellar and has constant speed on each streamline. Consequently the potential flow has constant magnitude on each streamline and therefore is covered by the HAMEL–MARRIS theorem.

A rotational motion whose vortex lines and streamlines coincide is a *screw motion*. STOKES once thought he had proved such motions to be impossible, but he later recognized his error. CRAIG was the first to study them.

In doing the following exercises the student might profit from the material in Section App. IIC.6, which introduces Ω, the *abnormality* of the vector lines.

Exercise II.13.9. In a screw motion $\Omega \neq 0$, and

$$\mathbf{w} = \Omega \dot{\mathbf{x}}. \tag{II.13-23}$$

Exercise II.13.10 (BELTRAMI). In a screw motion

$$\Omega = \frac{1}{w^2}\mathbf{w} \cdot \operatorname{curl} \mathbf{w}. \tag{II.13-24}$$

If $\dot{\mathbf{x}}$ or $\mathbf{w}$ is steady, so is Ω, and the motion is steady if and only if its spin is steady.

Exercise II.13.11 (BELTRAMI). In a screw motion

$$w^2 = \dot{\mathbf{x}} \cdot \operatorname{curl} \mathbf{w}. \tag{II.13-25}$$

Thus $\operatorname{curl} \mathbf{w}$ subtends upon $\dot{\mathbf{x}}$ an acute angle, possibly naught.

Exercise II.13.12 (GROMEKA, BELTRAMI). A screw motion preserves circulation if and only if it is steady. An acceleration-potential for it is $-\frac{1}{2}\dot{x}^2$.

[1]The paper cited in Footnote 4 on p. 141 rests essentially on the analysis of MARRIS & J.-F. SHIAU, "Hamel's theorem: the three polynomial integrals," *Rendiconti del Circolo Matematico di Palermo* (2)**22**(1973): 185–216.

[2]A. W. MARRIS, "Isochoric circulation-preserving motions with stream-lines of a potential motion," *Archive for Rational Mechamics and Analysis* **90**(1985): 213–218.

Exercise II.13.13 (GROMEKA, BELTRAMI). In a screw motion with steady density the surfaces

$$\Omega/\rho = \text{const.} \tag{II.13-26}$$

are stream surfaces; in particular Ω/ρ is constant on each streamline.

Exercise II.13.14 (GROMEKA, BELTRAMI). For a screw motion to preserve circulation, two equivalent conditions are necessary and sufficient: the motion is steady, or its spin is steady. Conversely, if $\mathbf{w}$ is steady, so is $\dot{\mathbf{x}}$, and (II.11-9) reduces to $\ddot{\mathbf{x}} = \operatorname{grad}(\frac{1}{2}\dot{x}^2)$.

Exercise II.13.15 (GROMEKA, BELTRAMI, NEMENYI & PRIM, TRUESDELL).[1] The curl of a screw motion is also a screw field if and only if the abnormality Ω of $\dot{\mathbf{x}}$ is constant in space. Then $\operatorname{curl}\dot{\mathbf{x}}$ has the same abnormality Ω as does $\dot{\mathbf{x}}$, and $\operatorname{div}\dot{\mathbf{x}} = 0$. Moreover, all successive curls of $\dot{\mathbf{x}}$ are solenoidal screw fields of abnormality Ω.

There has been much study of the kinematics of screw motions. BELTRAMI gave a simple example in cartesian components:

$$\dot{x}_1 = \sin(\Omega x_3), \qquad \dot{x}_2 = \cos(\Omega x_3), \qquad \dot{x}_3 = 0, \qquad \Omega = \text{const.}, \tag{II.13-27}$$

a uniplanar, isochoric flow of unit magnitude. NEMENYI & PRIM proved that the speed of a screw flow is spatially constant if and only if the streamlines are rectilinear. WANG noticed that if Ωx_3 is replaced by a function of x_3 in BELTRAMI's example (27), a uniplanar, isochoric flow of unit magnitude results. ERICKSEN proved that if Ω is spatially constant, BELTRAMI's uniplanar flow is the only possible steady, isochoric screw flow of unit magnitude. MARRIS & WANG[2] have proved a theorem that subsumes all the foregoing limitations upon screw flows.

14. Changes of Frame. Frame-Indifference

The concept of frame has been explained in Section I.6, and the transformations induced by a change of frame have been developed in Section I.9. The motion (I.7-7) of a body $\mathscr{B}$ is described with respect to a certain frame $\oint$; with

[1] For a fuller elaboration of this theorem see CFT, Section App. 34.

[2] A. W. MARRIS & C.-C. WANG, "Solenoidal screw fields of constant magnitude," *Archive for Rational Mechanics and Analysis* 39(1970): 227–244.

respect to another frame $\oint^*$, it is given by the mapping, say,

$$\mathbf{x}^* = \chi^*(X, t^*). \tag{II.14-1}$$

We regard a change of frame (I.9-5) as expressing the relation between the places and times, $(\mathbf{x}, t)$ and $(\mathbf{x}^*, t^*)$, of the same event as it appears to different observers. Thus, if (I.7-7) and (1) are to represent the same experiences of a body as apparent to observers in $\oint$ and $\oint^*$, respectively, the motions χ and χ^* must be related by (I.9-11), which we rewrite here:

$$\chi^*(X, t+a) = \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\chi(X, t) - \mathbf{x}_0), \tag{II.14-2}$$

$\mathbf{x}_0$ and a being the place and time with respect to $\oint$ of some assigned event, $\mathbf{x}_0^*$ being a function whose values are places, and $\mathbf{Q}$ being a function whose values are orthogonal tensors.

If we choose to describe the motion in terms of a reference placement $\boldsymbol{\kappa}$, the corresponding transplacements $\chi_{\boldsymbol{\kappa}}^*$ and $\chi_{\boldsymbol{\kappa}}$ are related in the same way:

$$\chi_{\boldsymbol{\kappa}}^*(\mathbf{X}, t+a) = \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\chi_{\boldsymbol{\kappa}}(\mathbf{X}, t) - \mathbf{x}_0). \tag{II.14-3}$$

As the notation indicates, we here use the same reference placement $\boldsymbol{\kappa}$ in forming from (2) the transplacements $\chi_{\boldsymbol{\kappa}}$ and $\chi_{\boldsymbol{\kappa}}^*$. We may change the reference placement also. To do so, we simply use (II.7-4).

Exercise II.14.1 (V. Bjerknes). Let a subscript $\oint$ denote the frame used; let ω be the angular speed at which $\oint$ is rotating with respect to $\oint^*$; let $\mathrm{C}(\mathscr{C})$ denote the circulation of a circuit $\mathscr{C}$. Then

$$\mathrm{C}_{\oint^*}(\mathscr{C}_{\oint^*}) - \mathrm{C}_{\oint}(\mathscr{C}_{\oint}) = 2\omega A_{\mathrm{eq}}(\mathscr{C}_{\oint}), \tag{II.14-4}$$

in which A_{eq} is a signed area of the region bounded by the projection of $\mathscr{C}_{\oint}$ onto a plane normal to the axis of spin of $\oint$ with respect to $\oint^*$.

In Section I.11 we have introduced the concept of frame-indifference. Briefly, a function of place and time whose values are scalar is frame-indifferent if it is in fact a function of events, independent of frame; one whose values are vectors, is frame-indifferent if its value in $\oint^*$ effects the same translation of the places of events in $\oint^*$ as its value in $\oint$ effects upon the places of these same events in $\oint$; one whose values are tensors, is frame-indifferent if it transforms each frame-indifferent vector into a frame-indifferent vector. Formally, as we have shown in Section I.11, these three conditions for frame-indifference are

$$\begin{aligned} F^* &= F && \text{for scalars,} \\ \mathbf{v}^* &= \mathbf{Q}\mathbf{v} && \text{for vectors,} \\ \mathbf{T}^* &= \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{T}} && \text{for tensors (of second order),} \end{aligned} \qquad \text{(II.14-5)}$$

the asterisks indicating quantities appropriate to the frame $\oint^*$, and $\mathbf{Q}$ being the orthogonal tensor that occurs in the change of frame (2).

When a quantity is defined by a prescription valid in all frames, conditions such as (5) may or may not be satisfied. In Section I.9 we have calculated the relation (I.9-14) connecting the velocities $\dot{\chi}$ and $\dot{\chi}^*$ as obtained in $\oint$ and $\oint^*$ whence we see that generally $\dot{\chi}^* \neq \mathbf{Q}\dot{\chi}$, and so the velocity is not frame-indifferent. Indeed, (I.9-14) shows that the spin $\mathbf{A}$ of $\oint^*$ with respect to $\oint$ gives rise to a velocity in $\oint^*$, which is in fact the velocity corresponding to a rigid motion for which $\oint^*$ is a rest frame (*cf.* Section I.10). Likewise, the relation (I.9-21) connecting the accelerations $\ddot{\chi}$ and $\ddot{\chi}^*$ in $\oint$ and $\oint^*$ shows that the acceleration is not frame-indifferent.

Now we shall consider the effect of change of frame upon quantities for whose definition not only a frame of reference but also a reference placement is employed. We begin with the transplacement gradient. Since the definition (II.5-1) applies both in $\oint$ and in $\oint^*$, we have

$$\mathbf{F} := \nabla\chi_\kappa(\mathbf{X}, t), \qquad \mathbf{F}^* := \nabla\chi_\kappa^*(\mathbf{X}, t). \qquad \text{(II.14-6)}$$

Taking the gradient of (3) shows that

$$\mathbf{F}^* = \mathbf{Q}\mathbf{F}. \qquad \text{(II.14-7)}$$

Thus the transplacement gradient is not frame-indifferent.

By applying to (7) the polar decomposition (II.9-1), we see that

$$\mathbf{R}^*\mathbf{U}^* = \mathbf{Q}\mathbf{R}\mathbf{U}. \qquad \text{(II.14-8)}$$

Because $\mathbf{QR}$ is orthogonal and because the polar decomposition of an invertible tensor is unique,

$$\mathbf{R}^* = \mathbf{Q}\mathbf{R}, \qquad \text{and} \qquad \mathbf{U}^* = \mathbf{U}. \qquad \text{(II.14-9)}$$

Hence

$$\mathbf{V}^* = \mathbf{R}^*\mathbf{U}^*\mathbf{R}^{*\mathsf{T}} = \mathbf{QRU}(\mathbf{QR})^{\mathsf{T}},$$
$$= \mathbf{QVQ}^{\mathsf{T}}. \tag{II.14-10}$$

Thus we have shown that $\mathbf{V}$ is frame-indifferent, while $\mathbf{F}$, $\mathbf{R}$, and $\mathbf{U}$ are not. Of course, $\mathbf{C}^* = \mathbf{C}$ and $\mathbf{B}^* = \mathbf{QBQ}^{\mathsf{T}}$, as is immediate by applying $(9)_2$ and $(10)_3$ to the definitions (II.9-5).

If we differentiate (7) with respect to time, we find that

$$\dot{\mathbf{F}}^* = \mathbf{Q}\dot{\mathbf{F}} + \dot{\mathbf{Q}}\mathbf{F}, \tag{II.14-11}$$

but by (II.11-5) $\dot{\mathbf{F}} = \mathbf{GF}$ and $\dot{\mathbf{F}}^* = \mathbf{G}^*\mathbf{F}^*$, and so

$$\mathbf{G}^*\mathbf{F}^* = \mathbf{QGF} + \dot{\mathbf{Q}}\mathbf{F},$$
$$= \mathbf{QGQ}^{\mathsf{T}}\mathbf{F}^* + \dot{\mathbf{Q}}\mathbf{Q}^{\mathsf{T}}\mathbf{F}^*. \tag{II.14-12}$$

Because $\mathbf{F}^*$ is invertible, it may be cancelled from this equation, which by use of the Euler–Cauchy–Stokes Decomposition (II.11-8) becomes

$$\mathbf{D}^* + \mathbf{W}^* = \mathbf{Q}(\mathbf{D} + \mathbf{W})\mathbf{Q}^{\mathsf{T}} + \mathbf{A}, \tag{II.14-13}$$

$\mathbf{A}$ being the spin (I.9-15) of $\oint$ with respect to $\oint^*$:

$$\mathbf{A} := \dot{\mathbf{Q}}\mathbf{Q}^{\mathsf{T}} = -\mathbf{A}^{\mathsf{T}}. \tag{II.14-14}$$

Since a decomposition into symmetric and skew parts is unique,

$$\mathbf{D}^* = \mathbf{QDQ}^{\mathsf{T}}, \qquad \mathbf{W}^* = \mathbf{QWQ}^{\mathsf{T}} + \mathbf{A}. \tag{II.14-15}$$

These formulae embody the ***Theorem of Zaremba and Zorawski***: *The stretching is frame-indifferent, while the spin in $\oint^*$ is the sum of the spin in $\oint$ and the spin of $\oint$ with respect to $\oint^*$.* The assertion is intuitively plain, since a change of frame in effect superimposes a rigid motion, possibly followed by a reflection, neither of which alters the stretchings of elements though the former does rotate the directions in which those stretchings seem to occur. A conclusion in Exercise I.11.3 makes the principal stretchings and the principal axes of stretching likewise frame-indifferent.

If we differentiate (II.9-10) n times with respect to τ and then put $\tau = t$, by appeal to the definition (II.11-31) we conclude that

$$\overset{(n)}{\mathbf{C}} = \mathbf{F}^{\mathsf{T}}\mathbf{A}_n\mathbf{F}, \tag{II.14-16}$$

$\mathbf{A}_n$ being the n^{th} Rivlin–Ericksen tensor. Applying the polar decomposition theorem (II.9-1) to (16), we obtain

$$\mathbf{U}^{-1}\overset{(n)}{\mathbf{C}}\mathbf{U}^{-1} = \mathbf{R}^{\mathsf{T}}\mathbf{A}_n\mathbf{R}. \tag{II.14-17}$$

Likewise

$$\mathbf{U}^{*-1}\overset{(n)}{\mathbf{C}^*}\mathbf{U}^{*-1} = \mathbf{R}^{*\mathsf{T}}\mathbf{A}_n^*\mathbf{R}^*. \tag{II.14-18}$$

We have shown a little above that $\mathbf{U}^* = \mathbf{U}$ and $\mathbf{C}^* = \mathbf{C}$; hence also $\overset{(n)}{\mathbf{C}^*} = \overset{(n)}{\mathbf{C}}$, and so the left-hand sides of (17) and (18) are equal. Therefore

$$\mathbf{R}^{*\mathsf{T}}\mathbf{A}_n^*\mathbf{R}^* = \mathbf{R}^{\mathsf{T}}\mathbf{A}_n\mathbf{R}. \tag{II.14-19}$$

By $(9)_1$ we conclude that

$$\mathbf{A}_n^* = \mathbf{Q}\mathbf{A}_n\mathbf{Q}^{\mathsf{T}}. \tag{II.14-20}$$

Thus *the Rivlin–Ericksen tensors are frame indifferent*. This statement generalizes the first assertion in the Zaremba–Zorawski Theorem. The second is equally easy to generalize, but the generalization is not so easy to interpret.

Exercise II.14.2. $\mathbf{U}_t$ is frame-indifferent, and

$$\mathbf{R}_t^*(\tau) = \mathbf{Q}(\tau)\mathbf{R}_t(\tau)\mathbf{Q}(t)^{\mathsf{T}}. \tag{II.14-21}$$

General References

Sections 15–25 of NFTM ("The Non-Linear Field Theories of Mechanics," *Handbuch der Physik* 3_3, Berlin, Heidelberg, and New York, Springer-Verlag, 1965).

Sections 13–171 (exhaustive treatment of kinematics in component notation) of CFT ("The Classical Field Theories," *Handbuch der Physik* 3_1, Berlin, Göttingen, and Heidelberg, Springer-Verlag, 1960).

C. Truesdell, *The Kinematics of Vorticity*, Bloomington, Indiana University Press, 1954.

This Page Intentionally Left Blank

Chapter III

The Stress Tensor

Although I envisage here very great generality both in the nature of the fluid and in the forces that act upon each of its particles, I have no fear of those reproaches often levelled with good reason at them who have undertaken to generalize the researches of others. I agree that often an excessive generality obscures rather than enlightens, and that sometimes it leads to calculations so messy as to make it extremely hard to draw any conclusions from them for the simplest cases. When generalizations are subject to this drawback, most certainly we ought abstain from them altogether and limit our studies to particular cases.

But in the subject I intend to explain, just the opposite happens: The generality that I embrace, far from dazzling our lights, will reveal to us rather the veritable laws of Nature in all their brilliance, and in them we shall find even stronger reason to admire her beauty and her simplicity. It will be an important lesson to learn that some principles till now believed bound to some special case are of greater breadth. Finally, these researches will demand calculations scarcely any more troublesome, and it will be easy to apply them to all special cases we might set up.

EULER
General principles of the state of
equilibrium of fluids
Mémoires de l'Académie des Sciences
de Berlin **11** (1757): 217-273

The geometers who have investigated the equations of equilibrium or motion of thin plates or of surfaces, either elastic or inelastic, have distinguished two kinds of forces, the one produced by dilatation or contraction, the other by the bending of these surfaces .... It has seemed to me that these two kinds of forces could be reduced to a single one, which ought to be called always *tension* or *pressure*, a force which acts upon each element of a section chosen at will, not only in a flexible surface but also in a solid, whether elastic or inelastic, and which is of the same kind as the hydrostatic pressure exerted by a fluid at rest upon the exterior surface of a body, except that the new pressure does not always remain perpendicular to the faces subject to it, nor is it the same in all directions at a given point.

CAUCHY
On the pressure or tension in a solid body
Exercices de Mathématiques, Seconde Année (1827)

One way of introducing the notion of stress into an abstract conceptual scheme of Rational Mechanics is to accept it as a fundamental notion derived from experience. The notion is simply that of mutual action between two bodies in contact, or between two parts of the same body separated by an imagined surface; and the physical reality of such modes of action is, in this view, admitted as part of the conceptual scheme .... This was the method followed by Euler in his formulation of the principles of Hydrostatics and Hydrodynamics, and by Cauchy in his earliest writings on Elasticity. When this method is followed, a distinction is established between the two types of forces which we have called "body forces" and "surface tractions," the former being conceived as due to direct action at a distance, and the latter to contact action.

LOVE
Note B, *A Treatise on the Mathematical Theory of Elasticity*, 2nd ed. (1906)

In many otherwise good textbooks a standing confusion reigns between three groups of forces: 1. Internal and external forces. 2. Volume and surface forces—a distinction which the mechanics of points is altogether incapable of perceiving. 3. Applied forces and forces of reaction.

HAMEL
On the foundations of mechanics
Mathematische Annalen **66** (1909): 350-397

1. Forces and Torques. The Laws of Dynamics. Body Forces and Contact Forces

Forces and torques, like bodies, motions, and masses, are primitive elements of mechanics. They are mathematical quantities introduced *a priori*, represented by symbols, and subjected to mathematical axioms that delimit their properties and render them clear and useful for the description of mechanical phenomena in nature. Axioms for a system of forces in general have been presented in Section I.5; torques have been defined as the moments of forces in Section I.8; general axioms of dynamics, which relate forces and torques to the motion they effect upon a given body, have been given in Sections I.12 and I.13. In the remainder of this book, except in passages where we discuss frame-indifference, we shall suppose that the frame $\oint$ is an inertial one, and we shall base dynamics on ***Euler's Laws of Motion***:

$$\mathbf{f}^a = \dot{\mathbf{m}}, \qquad \mathbf{F}^a = \dot{\mathbf{M}}. \tag{I.13-11$_r$}$$

That is, the rate of increase of the linear momentum of any body equals the applied force $\mathbf{f}^a$ upon it, and the rate of increase of the rotational momentum with respect to $\mathbf{x}_0$ equals the applied torque $\mathbf{F}^a$ upon it, the place $\mathbf{x}_0$ being stationary in the inertial frame.

We begin by restating these laws in more explicit forms, referred to a part $\mathscr{P}$ of $\mathscr{B}$ and to its shape $\chi(\mathscr{P})$ in the inertial frame $\oint$. These forms, which follow at once from (I.8-5) and (I.13-10), are

$$\begin{aligned} \int_{\chi(\mathscr{P})} \rho\ddot{\mathbf{x}}\,dV &= \mathbf{f}^a(\mathscr{P}), \\ \int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\ddot{\mathbf{x}}\,dV &= \mathbf{F}^a(\mathscr{P})_{\mathbf{x}_0}; \end{aligned} \tag{III.1-1}$$

we recall that the applied force $\mathbf{f}^a$ and the applied torque $\mathbf{F}^a$ may depend upon the time t, as does the shape $\chi(\mathscr{P})$, though we do not so indicate in the notation. Thus the applied force and torque upon $\mathscr{P}$ are expressed in terms of integrals over the actual shape of $\mathscr{P}$. As always, $\ddot{\mathbf{x}}$ is the acceleration field on $\chi(\mathscr{B})$, and we assume that it is essentially bounded.

In continuum mechanics two different systems of forces are introduced: *body forces* $\mathbf{f}_B$, which may be exerted mutually by bodies, whether or not they be in contact, and which are presumed related to the masses of the bodies, and *contact forces* $\mathbf{f}_C$, which are exerted by one body on another through their common surface of contact and are presumed related to that surface, distributed over it, and independent of the masses of the bodies on either side.

In this section we prepare the way toward deriving equations of motion expressed in terms of these special kinds of forces. We do so twice. First, imitating the great treatises of the preceding century, we follow a line of argument deriving from EULER, CAUCHY, and others; versions of this route are common in modern textbooks. Analytic precision wants; seeking only to make the desired conclusions clear and easily comprehensible, we bring in tacitly whatever assumptions of smoothness will do to get from one step to the next.

The second presentation aims to maintain the level of modern analysis. It is a sequel to the treatment of shapes in Section II.1, upon which we build.

In the traditional presentation the force $\mathbf{f}^{\mathrm{a}}$ applied to the part $\mathscr{P}$ in its shape $\chi(\mathscr{P})$ at the time t is assumed to be the sum of resultant forces of two different kinds:

$$\mathbf{f}^{\mathrm{a}} = \mathbf{f}^{\mathrm{r}}_{\mathrm{B}} + \mathbf{f}^{\mathrm{r}}_{\mathrm{C}}\,, \tag{III.1-2}$$

both of them obtained from densities,

$$\begin{aligned} \mathbf{f}^{\mathrm{r}}_{\mathrm{B}}(\mathscr{P}) &= \int_{\chi(\mathscr{P})} \rho \mathbf{b}_{\chi(\mathscr{P})}\, dV, \\ \mathbf{f}^{\mathrm{r}}_{\mathrm{C}}(\mathscr{P}) &= \int_{\partial\chi(\mathscr{P})} \mathbf{t}_{\partial\chi(\mathscr{P})}\, dA, \end{aligned} \tag{III.1-3}$$

$\mathscr{P}$ being any part of the body $\mathscr{B}$. The corresponding torques are given by

$$\begin{aligned} \mathbf{F}^{\mathrm{r}}_{\mathrm{B}}(\mathscr{P})_{\mathbf{x}_0} &= \int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0) \wedge \rho \mathbf{b}_{\chi(\mathscr{P})}\, dV, \\ \mathbf{F}^{\mathrm{r}}_{\mathrm{C}}(\mathscr{P})_{\mathbf{x}_0} &= \int_{\partial\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0) \wedge \mathbf{t}_{\partial\chi(\mathscr{P})}\, dA, \end{aligned} \tag{III.1-4}$$

and

$$\mathbf{F}^{\mathrm{r}}_{\mathbf{x}_0} = \mathbf{F}^{\mathrm{r}}_{\mathrm{B}\mathbf{x}_0} + \mathbf{F}^{\mathrm{r}}_{\mathrm{C}\mathbf{x}_0}\,. \tag{III.1-5}$$

As is clear from $(3)_1$, the *resultant applied body force* $\mathbf{f}^{\mathrm{r}}_{\mathrm{B}}$ is an absolutely continuous function of volume. For brevity, its density $\mathbf{b}_{\chi(\mathscr{P})}$ with respect to mass will be called henceforth the *body-force field* or even simply the *body force*. Moreover, we shall limit attention in this book to the case in which $\mathbf{b}_{\chi(\mathscr{P})}$

itself is an assigned function of place and time and hence independent of $\chi(\mathscr{P})$:

$$\mathbf{b}_{\chi(\mathscr{P})} = \mathbf{b}(\mathbf{x}, t) \qquad \forall\chi. \tag{III.1-6}$$

Such fields of body force are called *external*.[1] Commonly the external body force is assumed to be *lamellar*:

$$\mathbf{b} = -\operatorname{grad}\varpi; \tag{III.1-7}$$

the scalar function[2] ϖ is a *potential* of $\mathbf{b}$. If ϖ is a potential of $\mathbf{b}$, so is $\varpi + h$ if h is a function of time alone; the student should always recall this fact, and so we shall leave h unwritten henceforth, not only for the potential ϖ but also for other potentials. The student shall remember that *potentials are determinable* (apart from boundary conditions) *only to within a function of time alone*.

A steady lamellar body force is called *conservative*. A conservative body force has steady potentials, and of course in dealing with such a body force we always choose one of these. If $\mathbf{b}$ is constant in space and time, as is appropriate to heavy bodies near the surface of the earth, it is called the field of *uniform gravity*.[3] For such a $\mathbf{b}$ the potentials are affine functions of the distance $h(\mathbf{x})$ of the place $\mathbf{x}$ from some fixed plane; in the case of uniform gravity a convenient choice is $\varpi = gh$, the constant g being the gravitational acceleration and $h(\mathbf{x})$ being the height of $\mathbf{x}$ above the surface of the earth.

Exercise III.1.1. Two systems of forces applied to the shape of a body are said to be *equipollent* if they give rise to the same resultant force and resultant torque on that shape. The field of uniform gravity is equipollent to a single force acting at the center of mass of the body, directed parallel to $\mathbf{b}$ and in the same sense ("downward"), and equal in magnitude to the weight of the body.

[1] Not all external body forces are included in (6). For example, the density of force exerted by a magnetic or electric field is a function of $\dot{\mathbf{x}}$ and of constitutive properties of the body on which it acts. For the purposes of this book (6) is sufficient.

[2] In this book we always use the term "function" for a mapping, called in the older literature a "single-valued" function. "Cyclic" or "many-valued" potentials are important in many problems concerning multiply connected regions. Since this book is concerned mainly with local aspects of mechanics, and since "cyclic functions" are locally functions in the ordinary modern sense, we shall not take up the complications that may result from use of body forces with cyclic potentials. The reader already familiar with cyclic potentials can easily state for himself the generalizations to which they give rise in the few theorems in this book where they might be introduced. An example is EULER's corollary in Section IV.8.

A clear, elementary discussion of cyclic potentials may be found in Sections 49–54 of H. LAMB, *Hydrodynamics*, 2nd–6th eds., Cambridge, Cambridge University Press, 1895/1932, variously reprinted. A good example of a body force with cyclic potential is discussed in Section 6 of A. SOMMERFELD's *Mechanics of Deformable Bodies*, New York, Academic Press, 1950.

[3] The force of universal gravitation is a mutual body force, not an external one, and hence is not treated in this book except for a summary remark at the very end of this section.

Body forces are of secondary interest in continuum mechanics, which concerns mainly the effects of contact forces, to which we now address ourselves.

According to $(3)_2$ the *resultant contact force* $\mathbf{f}_C^r$ is an absolutely continuous function of the area of the bounding surface $\partial\chi(\mathscr{P})$ on which it acts. The surface density $\mathbf{t}_{\partial\chi(\mathscr{P})}$ is called the *traction field* on $\partial\chi(\mathscr{P})$. If that field be known, the resultant contact force is determined and is independent of whatever may be occurring at places not lying upon $\partial\chi(\mathscr{P})$. In this sense, the traction field is equipollent to the action upon $\mathscr{P}$ of the bodies outside $\mathscr{P}$ and adjacent to it. The assumption that the contact force is of this kind is the ***cut principle*** of EULER and CAUCHY: *Within the shape of a body at any given time, conceive a smooth, closed diaphragm; then the action of the part of the body outside that diaphragm and adjacent to the part inside is equipollent to that of a field of vectors defined on the diaphragm.*

Of course, the diaphragm may be chosen as the shape of the boundary of a body, the exterior of which we prefer not to specify, and in this case the cut principle does not furnish an interpretation for the traction $\mathbf{t}_{\partial\chi(\mathscr{B})}$. Rather, tractions upon the boundary of the largest body entering the statement of the problem at hand are regarded as prescribed by other considerations. For example, so as to represent the application of given forces upon the surface of a given body, without including in the theory such other bodies as may bring those forces to bear, we impose a *boundary condition of traction* by assigning $\mathbf{t}_{\partial\chi(\mathscr{B})}$, or a field closely related to it, on a given boundary surface such as $\partial\chi(\mathscr{B})$ or $\partial\boldsymbol{\kappa}(\mathscr{B})$. Examples of such conditions are given and discussed below in Sections III.2 and III.8–III.9. In other cases we may leave $\mathbf{t}_{\partial\chi(\mathscr{B})}$ to be determined on such surfaces by imposing a *boundary condition of place*, typically by prescribing on $\partial\chi(\mathscr{B})$ the transplacement $\chi_{\kappa}(\mathbf{X})$ or some quantity derived from it.

We assume that $\partial\chi(\mathscr{P})$ is orientable, and we write $\mathbf{n}$ for its outer unit normal. If $\mathbf{t}_{\partial\chi(\mathscr{P})}\cdot\mathbf{n} > 0$, the traction is said to be a *tension*; if $\mathbf{t}_{\partial\chi(\mathscr{P})}\cdot\mathbf{n} < 0$, it is a *pressure*.

If we substitute (2)–(5) into EULER's laws (1), we obtain the ***Basic Laws of Motion*** of continuum mechanics, as far as this book is concerned:[1]

$$\int_{\chi(\mathscr{P})} \rho\ddot{\mathbf{x}}\,dV = \int_{\partial\chi(\mathscr{P})} \mathbf{t}_{\partial\chi(\mathscr{P})}\,dA + \int_{\chi(\mathscr{P})} \rho\mathbf{b}\,dV,$$

$$\int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\ddot{\mathbf{x}}\,dV = \int_{\partial\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\mathbf{t}_{\partial\chi(\mathscr{P})}\,dA + \int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\mathbf{b}\,dV, \qquad \text{(III.1-8)}$$

for all parts $\mathscr{P}$ of all bodies in the universe.

[1]More general formulations relax the assumption that $\ddot{\mathbf{x}}$ exist everywhere at all times and take account of body couples, couple stresses, multipolar stresses, spin momentum, director stresses, *etc.*, as well as counterparts for diffusion and chemically reacting mixtures.

As we have stated in Section I.13, all forces are frame-indifferent. Thus, in particular, the contact forces and applied body forces are frame-indifferent. Consequently their densities are frame-indifferent vector fields:

$$\mathbf{b}^* = \mathbf{Q}\mathbf{b}, \qquad \mathbf{t}^*_{\partial\chi^*(\mathscr{P})} = \mathbf{Q}\mathbf{t}_{\partial\chi(\mathscr{P})}, \tag{III.1-9}$$

$\mathbf{Q}$ being the orthogonal tensor occurring in the change of frame (II.14-2).

Of course the forms (8) expressing the principles of linear and rotational momentum are valid only in an inertial frame. To obtain corresponding forms in a general frame, we need only replace the acceleration field $\ddot{\mathbf{x}}$ by the frame-indifferent vector field $\mathbf{a}$ that reduces to $\ddot{\mathbf{x}}$ when the frame is inertial. That frame-indifferent vector field we have calculated already and recorded as (II.4-7). With this replacement, the integrals on the left-hand sides of (8) become frame-indifferent, as are all four integrals on the right-hand sides.

The reader who is content to accept these equations, supplemented by axioms endowing the densities $\mathbf{b}$ and $\mathbf{t}_{\partial\chi(\mathscr{P})}$ with some smoothness, may pass straight on to the next section.

The more critical reader will see two objections. First, the resultant contact force $\mathbf{f}^r_C$ does not define the traction uniquely, since to any $\mathbf{t}_{\partial\chi(\mathscr{P})}$ that satisfies $(3)_2$ we may add $\mathbf{S}\mathbf{n}$ if for $\mathbf{S}$ we take any tensor field such that $\operatorname{div}\mathbf{S} = \mathbf{0}$, and $\mathbf{f}^r_C$ will be the same. Second, the resultant body force $\mathbf{f}_B$ and resultant contact force $\mathbf{f}^r_C$ are not clearly related to the general concept of a system of forces, which is a function defined on pairs of separate bodies rather than on single bodies. For such a reader this section concludes with an analysis which delivers the classical assumptions (3) and (4) as theorems[1] proved from assumptions of continuity phrased in terms of the modern theory of bodies and systems of forces, which appears above in Sections I.2–I.5.

For the text following now through the end of this section I am indebted both to W. O. WILLIAMS, as I was for its predecessor in the first edition, and to E. VIRGA, who provided the following formulation and arguments in terms of reduced boundaries, fit regions, and contacts.

Considering a fixed time t and not indicating it in the notation, we assume that $\mathbf{f}$ is a system of forces defined on $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$; accordingly, it satisfies Axioms

[1]The main sources of the material presented here are the papers by NOLL & VIRGA and by GURTIN, ŠILHAVÝ, WILLIAMS, and ZIEMER, cited above in Footnote 1 on p. 90 and Footnote 1 on p. 88. Those works were influenced by earlier researches, especially the paper of M. E. GURTIN & W. O. WILLIAMS, "An axiomatic foundation for continuum thermodynamics," *Archive for Rational Mechanics and Analysis* **26** (1967): 83–117. In that paper they are phrased in terms of scalar-valued functions having a thermomechanical rather than purely mechanical interpretation, but the mathematics is essentially the same.

F1, F2, and F3 in Section I.5, though we do not assume that it satisfies Axiom F4. Rather, fulfilling the promise made in Section I.5, we shall sketch a proof that under a physically natural axiom of bounds $\mathbf{f}$ in continuum mechanics must obey Axiom F4. Likewise, we shall outline an argument that delivers the systems of forces $\mathbf{f}_B$ and $\mathbf{f}_C$ whose resultants $\mathbf{f}_B^r$ and $\mathbf{f}_C^r$ appear as posited entities in (3).

To that end we introduce the idea that in continuum mechanics forces are exerted upon pairs of bodies in virtue of their masses and the areas of contact of their shapes, and that these forces diminish at least linearly with those masses and areas when both are sufficiently small. To express this idea within the general framework built in Section II.1, where the shapes of bodies are taken as fit regions in a three-dimensional Euclidean space $\mathscr{E}$, we have to specify what "area of contact" is to mean. We call the *contact* of two disjoint fit regions the intersection of their reduced boundaries, and we call *area of contact* the two-dimensional Hausdorff measure of the contact.[1] The student will recall from the end of Section II.1 that the shapes of separate bodies are disjoint fit regions.

Axiom on Forces in Continuum Mechanics. *Let $\mathscr{A}$ and $\mathscr{C}$ be separate bodies, the area of contact of whose shapes is sufficiently small, and let the mass of $\mathscr{A}$ be sufficiently small. Then*

$$|\mathbf{f}(\mathscr{A}, \mathscr{C})| \leq KA(\partial^*\chi(\mathscr{A}) \cap \partial^*\chi(\mathscr{C})) + K_{\mathscr{C}}M(\mathscr{A}), \qquad \text{(III.1-10)}$$

K being a positive constant and $K_{\mathscr{C}}$ being a positive, bounded function of $\mathscr{C}$ such that

$$\lim_{M(\mathscr{C})\to 0} K_{\mathscr{C}} = 0. \qquad \text{(III.1-11)}$$

This axiom seems broader and more natural than the classical assumptions (3).

The second addend in (10) is independent of the motion χ; it refers to the bodies $\mathscr{A}$ and $\mathscr{C}$ alone, independently of their shapes. The first addend depends upon χ not only as the notation indicates but also through K. The motion χ as it proceeds affects the values of the functions on the right-hand side of (10), but the property that (10) asserts remains unaffected.

The proof that these assumptions do lead to applied forces conform with the statements (2) and (3) is not easy. The argument does not require the system

[1]This definition comes from the paper by Noll & Virga cited above in Footnote 1 on p. 90. It follows from (II.1-5) that the area of contact of two fit regions is always finite.

of forces $\mathbf{f}$ to be balanced; at the beginning, it does not even require the forces to be pairwise equilibrated.

Theorem. *There are systems of forces* $\mathbf{f}_B$ *and* $\mathbf{f}_C$ *such as to satisfy the bounds*

$$|\mathbf{f}_B(\mathscr{A}, \mathscr{C})| \leqq K_{\mathscr{C}} M(\mathscr{A}),$$
$$|\mathbf{f}_C(\mathscr{A}, \mathscr{C})| \leqq K A(\partial^* \boldsymbol{\chi}(\mathscr{A}) \cap \partial^* \boldsymbol{\chi}(\mathscr{C})), \tag{III.1-12}$$

for all $\mathscr{A}$ *and* $\mathscr{C}$ *in* $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$, *and*

$$\mathbf{f} = \mathbf{f}_B + \mathbf{f}_C . \tag{III.1-13}$$

This major theorem decomposes the system of forces $\mathbf{f}$ uniquely into the system of body forces $\mathbf{f}_B$ and the system of contact forces $\mathbf{f}_C$. The bound it provides for $\mathbf{f}_B(\mathscr{A}, \mathscr{C})$ depends only upon $\mathscr{A}$ and $\mathscr{C}$, independently of the motions those bodies may undergo.

Because of (II.2-9), in $(12)_1$ $M(\mathscr{A})$ may be replaced by $V(\boldsymbol{\chi}(\mathscr{A}))$, but then the multiplier $K_{\mathscr{C}}$ will depend in general upon $\boldsymbol{\chi}$. The contact force $\mathbf{f}_C$ depends upon the shapes $\boldsymbol{\chi}(\mathscr{A})$ and $\boldsymbol{\chi}(\mathscr{C})$, but the existence of its stated bound is unaffected by whatever motion takes place. The body force $\mathbf{f}_B(\mathscr{A}, \mathscr{C}) \to 0$ as $V(\boldsymbol{\chi}(\mathscr{A})) \to 0$; the contact force $\mathbf{f}_C(\mathscr{A}, \mathscr{C}) \to 0$ as $A(\partial^* \boldsymbol{\chi}(\mathscr{A}) \cap \partial^* \boldsymbol{\chi}(\mathscr{C})) \to 0$.

The traditional treatment starting from (3) in effect posits the bounds (12) when $\mathscr{C} = \mathscr{A}^e$; here they are established for general $\mathscr{C}$.

Proof of the theorem. For given $\mathscr{A}$ and $\mathscr{C}$,

$$\Gamma_{\mathscr{A}}^{(\mathscr{C})} := \{\mathscr{B} \in \mathbf{\Omega}\colon\ \mathscr{B} \prec \mathscr{A} \ \&\ A((\partial^* \boldsymbol{\chi}(\mathscr{A}) \cap \partial^* \boldsymbol{\chi}(\mathscr{C})) \backslash (\partial^* \boldsymbol{\chi}(\mathscr{B}) \cap \partial^* \boldsymbol{\chi}(\mathscr{C})) = 0\}. \tag{III.1-14}$$

(Here the student might well draw a sketch.) The subbody $\mathscr{B}$ of $\mathscr{A}$ is an element of $\Gamma_{\mathscr{A}}^{(\mathscr{C})}$ if and only if the contact of $\boldsymbol{\chi}(\mathscr{B})$ and $\boldsymbol{\chi}(\mathscr{C})$ differs from the contact of $\boldsymbol{\chi}(\mathscr{A})$ and $\boldsymbol{\chi}(\mathscr{C})$ by a set of null area. If $\mathscr{D}$ and $\mathscr{F}$ are in $\Gamma_{\mathscr{A}}^{(\mathscr{C})}$, then, because $\mathscr{D} = (\mathscr{D} \wedge \mathscr{F}) \vee (\mathscr{D} \backslash \mathscr{F})$, $\mathscr{F} = (\mathscr{D} \wedge \mathscr{F}) \vee (\mathscr{F} \backslash \mathscr{D})$, and $\mathbf{f}$ is bi-additive,

$$\begin{aligned} |\mathbf{f}(\mathscr{D}, \mathscr{C}) - \mathbf{f}(\mathscr{F}, \mathscr{C})| &= |\mathbf{f}(\mathscr{D} \backslash \mathscr{F}, \mathscr{C}) - \mathbf{f}(\mathscr{F} \backslash \mathscr{D}, \mathscr{C})|, \\ &\leqq K_{\mathscr{C}} \max(M(\mathscr{D}), M(\mathscr{F})). \end{aligned} \tag{III.1-15}$$

The latter bound follows because $M(\mathscr{D}\backslash\mathscr{F}) \leqq M(\mathscr{D})$, $M(\mathscr{F}\backslash\mathscr{D}) \leqq M(\mathscr{F})$, and because the contact of the shapes of $\mathscr{D}\backslash\mathscr{F}$ and $\mathscr{C}$ has the same area as the contact of the shapes of $\mathscr{F}\backslash\mathscr{D}$ and $\mathscr{C}$. Because (II.2-9) implies that bodies of arbitrarily small volume have arbitrarily small mass, the definition

$$\mathbf{f}_{\mathrm{C}}(\mathscr{A}, \mathscr{C}) := \lim_{M(\mathscr{D})\to 0} \mathbf{f}(\mathscr{D}, \mathscr{C}), \qquad \mathscr{D} \in \Gamma_{\mathscr{A}}^{(\mathscr{C})}, \tag{III.1-16}$$

makes sense, and

$$|\mathbf{f}_{\mathrm{C}}(\mathscr{A}, \mathscr{C})| \leqq KA(\partial^*\chi(\mathscr{A}) \cap \partial^*\chi(\mathscr{C})) \tag{III.1-17}$$

since $A(\partial^*\chi(\mathscr{D}) \cap \partial^*\chi(\mathscr{C})) = A(\partial^*\chi(\mathscr{A}) \cap \partial^*\chi(\mathscr{C}))$ for each $\mathscr{D} \in \Gamma_{\mathscr{A}}^{(\mathscr{C})}$. Thus $\mathbf{f}_{\mathrm{C}}$ has the bound specified by $(12)_2$. Straightforward computations show that $\mathbf{f}_{\mathrm{C}}$ is bi-additive.

Defining $\mathbf{f}_{\mathrm{B}}$ thus:

$$\mathbf{f}_{\mathrm{B}} := \mathbf{f} - \mathbf{f}_{\mathrm{C}}, \tag{III.1-18}$$

we make $\mathbf{f}_{\mathrm{B}}$ bi-additive, and because of (16)

$$\mathbf{f}_{\mathrm{B}}(\mathscr{A}, \mathscr{C}) = \lim_{M(\mathscr{D})\to 0} (\mathbf{f}(\mathscr{A}, \mathscr{C}) - \mathbf{f}(\mathscr{D}, \mathscr{C})), \qquad \mathscr{D} \in \Gamma_{\mathscr{A}}^{(\mathscr{C})}. \tag{III.1-19}$$

We now show that $\mathbf{f}_{\mathrm{B}}$ has the bound specified by $(12)_1$. For every $\mathscr{D} \in \Gamma_{\mathscr{A}}^{(\mathscr{C})}$, let $\tilde{\mathscr{D}} := \mathscr{A} \wedge \mathscr{D}^{\mathrm{e}}$; thus

$$\mathscr{A} = \mathscr{D} \vee \tilde{\mathscr{D}}. \tag{III.1-20}$$

Since $\mathbf{f}$ is bi-additive, it follows from (19) and (20) that

$$\mathbf{f}_{\mathrm{B}}(\mathscr{A}, \mathscr{C}) = \lim_{M(\mathscr{D})\to 0} \mathbf{f}(\tilde{\mathscr{D}}, \mathscr{C}). \tag{III.1-21}$$

It is geometrically plain and will be proved soon (see (III.1-28), below), that

$$A(\partial^*\tilde{\mathscr{D}} \cap \partial^*\mathscr{C}) = 0; \tag{III.1-22}$$

for every $\mathscr{D} \in \Gamma_{\mathscr{A}}^{(\mathscr{C})}$. Thus, by (10) and (21),

$$|\mathbf{f}_{\mathrm{B}}(\mathscr{A}, \mathscr{C})| \leqq \lim_{M(\mathscr{D})\to 0} K_{\mathscr{C}} M(\tilde{\mathscr{D}}) \tag{III.1-23}$$

and the desired conclusion follows, since (20) implies that

$$\lim_{M(\mathscr{D})\to 0} M(\tilde{\mathscr{D}}) = M(\mathscr{A}). \quad \triangle \tag{III.1-24}$$

Taking up first the system of contact forces $\mathbf{f}_C$, we shall prove that because of $(12)_2$, $\mathbf{f}_C(\mathscr{A}, \mathscr{C})$ depends upon $\mathscr{A}$ and $\mathscr{C}$ only through the contact of their present shapes.

Lemma (GURTIN & WILLIAMS). *Let* $(\mathscr{A}, \mathscr{C})$ *and* $(\hat{\mathscr{A}}, \hat{\mathscr{C}})$ *be pairs of separate bodies; suppose that* $\hat{\mathscr{A}} \prec \mathscr{A}$ *and* $\hat{\mathscr{C}} \prec \mathscr{C}$; *suppose further that the shapes of* $\mathscr{A}$ *and* $\mathscr{C}$ *share the same area of contact as the shapes of* $\hat{\mathscr{A}}$ *and* $\hat{\mathscr{C}}$:

$$A(\partial^*\chi(\mathscr{A}) \cap \partial^*\chi(\mathscr{C})) = A(\partial^*\chi(\hat{\mathscr{A}}) \cap \partial^*\chi(\hat{\mathscr{C}})). \tag{III.1-25}$$

Then

$$\mathbf{f}_C(\mathscr{A}, \mathscr{C}) = \mathbf{f}_C(\hat{\mathscr{A}}, \hat{\mathscr{C}}). \tag{III.1-26}$$

Proof. We set

$$\tilde{\mathscr{A}} := \mathscr{A} \wedge \hat{\mathscr{A}}^{\mathrm{e}}, \qquad \tilde{\mathscr{C}} := \mathscr{C} \wedge \hat{\mathscr{C}}^{\mathrm{e}}, \tag{III.1-27}$$

and note that $(\tilde{\mathscr{A}}, \mathscr{C})$ and $(\hat{\mathscr{A}}, \tilde{\mathscr{C}})$ are also pairs of separate bodies. Then a sketch makes plausible the following statement, which was proved by GURTIN, WILLIAMS, & ZIEMER:

$$A(\partial^*\chi(\tilde{\mathscr{A}}) \cap \partial^*\chi(\mathscr{C})) = 0, \tag{III.1-28}$$

$$A(\partial^*\chi(\hat{\mathscr{A}}) \cap \partial^*\chi(\tilde{\mathscr{C}})) = 0. \tag{III.1-29}$$

An elegant proof follows now. To shorten the formulae, we introduce the temporary notation $\mathsf{A} := \chi(\mathscr{A})$, $\mathsf{C} := \chi(\mathscr{C})$, *etc.*

Let A, B, and C be mutually separate shapes. Then

$$\partial^*\mathsf{A} \cap \partial^*\mathsf{B} \cap \partial^*\mathsf{C} = \varnothing. \tag{III.1-30}$$

Suppose for contradiction that $\mathbf{x} \in \partial^*\mathsf{A} \cap \partial^*\mathsf{B} \cap \partial^*\mathsf{C}$. Since $\mathbf{x} \in \partial^*\mathsf{A} \cap \partial^*\mathsf{B}$,

$$\mathbf{n}_{\mathsf{A}} = -\mathbf{n}_{\mathsf{B}}, \tag{III.1-31}$$

where $\mathbf{n_A}$ and $\mathbf{n_B}$ are, respectively, the outer normal fields to $\mathbf{A}$ and $\mathbf{B}$ at $\mathbf{x}$ (see Section II.1). Similarly,

$$\mathbf{n_B} = -\mathbf{n_C}\,, \qquad \mathbf{n_C} = -\mathbf{n_A}\,. \tag{III.1-32}$$

It follows from (31) and (32) that

$$\mathbf{n_A} = -\mathbf{n_A}\,, \tag{III.1-33}$$

which is a contradiction.

Since $\tilde{\mathbf{A}} \subset \mathbf{A}$,

$$\partial^*\tilde{\mathbf{A}} \cap \partial^*\mathbf{C} = \partial^*\tilde{\mathbf{A}} \cap \partial^*\mathbf{A} \cap \partial^*\mathbf{C}. \tag{III.1-34}$$

Then, because $\tilde{\mathbf{A}}$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{C}}$ are mutually separate, from (25) we conclude that

$$A(\partial^*\tilde{\mathbf{A}} \cap \partial^*\mathbf{C}) = A(\partial^*\tilde{\mathbf{A}} \cap \partial^*\hat{\mathbf{A}} \cap \partial^*\hat{\mathbf{C}}) = 0. \tag{III.1-35}$$

Thus (28) is proved. Similarly, $\hat{\mathbf{A}} \subset \mathbf{A}$ and $\tilde{\mathbf{C}} \subset \mathbf{C}$ imply that

$$\partial^*\hat{\mathbf{A}} \cap \partial^*\tilde{\mathbf{C}} = \partial^*\hat{\mathbf{A}} \cap \partial^*\tilde{\mathbf{C}} \cap \partial^*\mathbf{C} = \partial^*\mathbf{A} \cap \partial^*\hat{\mathbf{A}} \cap \partial^*\tilde{\mathbf{C}} \cap \partial^*\mathbf{C}, \tag{III.1-36}$$

and so, because $\hat{\mathbf{A}}$, $\tilde{\mathbf{C}}$ and $\hat{\mathbf{C}}$ are separate,

$$A(\partial^*\hat{\mathbf{A}} \cap \partial^*\tilde{\mathbf{C}}) = A(\partial^*\hat{\mathbf{A}} \cap \partial^*\tilde{\mathbf{C}} \cap \partial^*\hat{\mathbf{C}}) = 0. \tag{III.1-37}$$

Thus also (29) is proved.

In view of $(12)_2$, the contact force exerted upon each other by bodies whose shapes have area of contact 0 is $\mathbf{0}$. Thus (28) and (29) show that

$$\mathbf{f_C}(\tilde{\mathscr{A}}, \mathscr{C}) = \mathbf{0}, \qquad \mathbf{f_C}(\hat{\mathscr{A}}, \tilde{\mathscr{C}}) = \mathbf{0}. \tag{III.1-38}$$

Now we see that since

$$\mathscr{A} = \hat{\mathscr{A}} \vee \tilde{\mathscr{A}} \qquad \text{and} \qquad \mathscr{C} = \hat{\mathscr{C}} \vee \tilde{\mathscr{C}}, \tag{III.1-39}$$

and since $\mathbf{f_C}$ is bi-additive on $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$,

$$\begin{aligned}\mathbf{f_C}(\mathscr{A}, \mathscr{C}) &= \mathbf{f_C}(\hat{\mathscr{A}}, \hat{\mathscr{C}} \vee \tilde{\mathscr{C}}) + \mathbf{f_C}(\tilde{\mathscr{A}}, \mathscr{C}),\\ &= \mathbf{f_C}(\hat{\mathscr{A}}, \hat{\mathscr{C}}) + \mathbf{f_C}(\hat{\mathscr{A}}, \tilde{\mathscr{C}}) + \mathbf{f_C}(\tilde{\mathscr{A}}, \mathscr{C}).\end{aligned} \tag{III.1-40}$$

In virtue of (38) this statement reduces to (26). $\triangle$

We use the lemma so as to render explicit the nature of the function $\mathbf{f}_C$. We shall call a subset of the underlying contact a *subcontact*. A contact is orientable by assigning it the orientation of the outer normal to one of the shapes in contact. Clearly, any contact may have one of two opposite orientations. We agree to assign a subcontact the same orientation as has the contact to which it belongs. Letting $\mathscr{S}$ be any subcontact of the contact of the shape of some body $\mathscr{C}$ with the shape of some other body, we define a new function on the subcontacts $\mathscr{Q}$ of $\mathscr{S}$. Noting that since $\mathscr{Q} \subset \partial^*\boldsymbol{\chi}(\mathscr{C})$, we see that there is some $\mathscr{A}$ such that $\mathscr{Q} = \partial^*\boldsymbol{\chi}(\mathscr{A}) \cap \partial^*\boldsymbol{\chi}(\mathscr{C})$, and we set

$$\mathbf{f}_C^S(\mathscr{Q}, \mathscr{S}) := \mathbf{f}_C(\mathscr{A}, \mathscr{C}). \tag{III.1-41}$$

The lemma guarantees that this definition is unambiguous. That is, if $\hat{\mathscr{C}}$ is another body such that $\mathscr{S} \subset \partial^*\boldsymbol{\chi}(\hat{\mathscr{C}})$, there is a body $\hat{\mathscr{A}}$ such that $\mathscr{Q} = \partial^*\boldsymbol{\chi}(\hat{\mathscr{A}}) \cap \partial^*\boldsymbol{\chi}(\hat{\mathscr{C}})$, and by (26) we conclude that $\mathbf{f}_C(\hat{\mathscr{A}}, \hat{\mathscr{C}})$ if used in (41) would yield the same function $\mathbf{f}_C^S$. Thus the new function $\mathbf{f}_C^S$, which is defined on contacts and their subcontacts by (41), completely determines the old function $\mathbf{f}_C$, defined on $(\bar{\boldsymbol{\Omega}} \times \bar{\boldsymbol{\Omega}})_0$. We have shown, then, that any system of contact forces is defined completely by an appropriate function whose arguments are contacts and their subcontacts.

Since, by assumption, $\mathbf{f}_C(\cdot, \mathscr{C})$ is defined and additive on the subbodies of $\mathscr{C}^e$, the function $\mathbf{f}_C^S(\cdot, \mathscr{S})$ is defined and additive on the subcontacts of $\mathscr{S}$. From (41) it follows also that if $\mathscr{S}'$ is a subcontact of $\mathscr{S}$ and if $\mathscr{U}$ is a subcontact of $\mathscr{S}'$, then

$$\mathbf{f}_C^S(\mathscr{U}, \mathscr{S}) = \mathbf{f}_C^S(\mathscr{U}, \mathscr{S}'). \tag{III.1-42}$$

The requirement $(12)_2$ now assumes the form

$$|\mathbf{f}_C^S(\mathscr{U}, \mathscr{S})| \leqq KA(\mathscr{U}), \tag{III.1-43}$$

if $\mathscr{U} \subset \mathscr{S}$ and if $A(\mathscr{U})$ is sufficiently small. Since $\mathbf{f}_C^S(\cdot, \mathscr{S})$ is additive and obeys (43) on a rich collection of subsets of $\mathscr{S}$, an exercise in measure theory shows that it has a countably additive extension to the Borel sets of $\mathscr{S}$. Then the Radón-Nikodym theorem provides it a representation as an integral:

$$\mathbf{f}_C^S(\mathscr{U}, \mathscr{S}) = \int_{\mathscr{U}} \mathbf{t}_{\mathscr{S}}\, dA. \tag{III.1-44}$$

Moreover, it follows from (42) that if the contacts $\mathscr{S}$ and $\mathscr{S}'$ have the same orientation and if $\mathscr{S}' \subset \mathscr{S}$, then

$$\mathbf{t}_{\mathscr{S}'} = \mathbf{t}_{\mathscr{S}} \tag{III.1-45}$$

at almost every point of $\mathscr{S}'$. Now going back to the definition (41) of $\mathbf{f}_C^S$, we may interpret (44) as demonstrating the following

Traction Theorem (GURTIN & WILLIAMS). *If a system of forces* $\mathbf{f}_C$ *on* $(\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0$ *satisfies* $(12)_2$, *there is an essentially bounded density* $\mathbf{t}_{\mathscr{S}}$ *such that*

$$\mathbf{f}_C(\mathscr{A}, \mathscr{C}) = \int_{\mathscr{S}} \mathbf{t}_{\mathscr{S}}\, dA, \qquad \mathscr{S} := \partial^* \boldsymbol{\chi}(\mathscr{A}) \cap \partial^* \boldsymbol{\chi}(\mathscr{C}); \tag{III.1-46}$$

moreover, $\mathbf{t}_{\mathscr{S}'} = \mathbf{t}_{\mathscr{S}}$ *if* $\mathscr{S}'$ *is a subcontact of* $\mathscr{S}$.

Taking $\mathscr{P}$ for $\mathscr{A}$ and $\mathscr{P}^e$ for $\mathscr{C}$ in this theorem yields $(3)_2$.

Now we take up the system of body forces $\mathbf{f}_B$. The details of the reduction of $\mathbf{f}_B$ to an integral representation are too involved to present here, but we may briefly sketch the argument used by GURTIN, WILLIAMS, & ZIEMER. First, for a fixed $\mathscr{B}$ the mapping $\mathscr{A} \mapsto \mathbf{f}_B(\mathscr{A}, \mathscr{B}^e)$ is additive for all $\mathscr{A} \prec \mathscr{B}$; then use of the bound $(12)_1$ ensures a representation of the form

$$\mathbf{f}_B(\mathscr{A}, \mathscr{B}^e) = \int_{\boldsymbol{\chi}(\mathscr{A})} \mathbf{b}^e\, dM - \int_{\boldsymbol{\chi}(\mathscr{A})} \ddot{\mathbf{x}}\, dM \tag{III.1-47}$$

Here we have invoked the Axioms of Inertia in Section I.13 to make $\ddot{\mathbf{x}}$ an identifiable part of the density of $\mathbf{f}_B$ with respect to mass. Since we have assumed that in an inertial frame $\ddot{\mathbf{x}}$ is bounded, it follows that $\mathbf{b}^e$ is essentially bounded. If

$$\mathbf{f}_B^e(\mathscr{A}, \mathscr{B}^e) := \int_{\boldsymbol{\chi}(\mathscr{A})} \mathbf{b}^e\, dM, \tag{III.1-48}$$

putting $\mathscr{P}$ for $\mathscr{A}$ and $\mathscr{B}$ in (48) leads to $(3)_1$.

Second, by circuitous arguments resting mainly on (11) and $(12)_1$, GURTIN, WILLIAMS, & ZIEMER show that if $\mathscr{C} \prec \mathscr{A}^e \wedge \mathscr{B}$, then

$$\mathbf{f}_B(\mathscr{A}, \mathscr{C}) = \int_{\boldsymbol{\chi}(\mathscr{A}) \times \boldsymbol{\chi}(\mathscr{C})} \mathbf{b}^m\, dM^2 \tag{III.1-49}$$

with $\mathbf{b}^m$ an integrable function defined on $\mathscr{B} \times \mathscr{B}$. Here $\mathbf{b}^m(\mathbf{x}, \mathbf{y})$ represents a *mutual body force* between body-points occupying the places $\mathbf{x}$ and $\mathbf{y}$ in $\boldsymbol{\chi}(\mathscr{B})$.

Since in this book we shall, as already remarked, deal with no body forces other than external ones, we take $\mathbf{b}^{\mathrm{m}}$ to be the null function, and henceforth we write $\mathbf{b}$ for $\mathbf{b}^{\mathrm{e}}$.

Now, for the first time in our treatment based upon the axiom (10), we suppose that *the system of forces* $\mathbf{f}$ *is pairwise equilibrated*:

$$\mathbf{f}(\mathscr{B}, \mathscr{C}) = -\mathbf{f}(\mathscr{C}, \mathscr{B}) \qquad \forall(\mathscr{B}, \mathscr{C}) \in (\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0 , \qquad \text{(I.5-6)}_{\mathrm{r}}$$

Theorem of Action and Reaction (NOLL). *If* $\mathbf{f}$ *is pairwise equilibrated, then both the system of body forces and the system of contact forces are pairwise equilibrated*:

$$\left.\begin{aligned} \mathbf{f}_{\mathrm{B}}(\mathscr{A}, \mathscr{C}) &= -\mathbf{f}_{\mathrm{B}}(\mathscr{C}, \mathscr{A}) \\ \mathbf{f}_{\mathrm{C}}(\mathscr{A}, \mathscr{C}) &= -\mathbf{f}_{\mathrm{C}}(\mathscr{C}, \mathscr{A}) \end{aligned}\right\} \qquad \forall(\mathscr{A}, \mathscr{C}) \in (\bar{\mathbf{\Omega}} \times \bar{\mathbf{\Omega}})_0 . \qquad \text{(III.1-50)}$$

We recall from Section I.5 that every balanced system of forces is pairwise equilibrated, and we note that neither of the systems $\mathbf{f}_{\mathrm{B}}$ and $\mathbf{f}_{\mathrm{C}}$ need be balanced.

Proof of the theorem. Again we let $\mathscr{S}$ denote the contact of the shapes of the separate bodies $\mathscr{A}$ and $\mathscr{C}$. We may choose sequences of parts $\mathscr{A}_n$ and $\mathscr{C}_n$ of $\mathscr{A}$ and $\mathscr{C}$, respectively, such that in the limit as $n \to \infty$ the volumes of their shapes vanish, yet they retain $A(\mathscr{S})$ as their area of contact. This statement, which is easy to demonstrate in elementary geometry, for bounded sets of finite perimeter has been proved by GURTIN, WILLIAMS, & ZIEMER. Formally,

$$\mathscr{A}_n \prec \mathscr{A}, \qquad \mathscr{C}_n \prec \mathscr{C}, \qquad A(\mathscr{S} \backslash \partial^* \chi(\mathscr{A}_n) \cap \partial^* \chi(\mathscr{C}_n)) = 0, \quad \text{(III.1-51)}$$

and, since in continuum mechanics mass is an absolutely continuous function of volume (Section II.2),

$$M(\mathscr{A}_n) \to 0, \qquad M(\mathscr{C}_n) \to 0. \qquad \text{(III.1-52)}$$

The lemma of GURTIN & WILLIAMS shows that

$$\begin{aligned} \mathbf{f}_{\mathrm{C}}(\mathscr{A}_n , \mathscr{C}_n) &= \mathbf{f}_{\mathrm{C}}(\mathscr{A}, \mathscr{C}), \\ \mathbf{f}_{\mathrm{C}}(\mathscr{C}_n , \mathscr{A}_n) &= \mathbf{f}_{\mathrm{C}}(\mathscr{C}, \mathscr{A}). \end{aligned} \qquad \text{(III.1-53)}$$

Of course $\mathscr{A}_n$ and $\mathscr{C}_n$, being parts of separate bodies, are separate, and so we may substitute $\mathscr{A}_n$ for $\mathscr{A}$ and $\mathscr{C}_n$ for $\mathscr{C}$ in (I.5-6) and by use of (53) obtain

$$\mathbf{f}_{\mathrm{B}}(\mathscr{A}_n, \mathscr{C}_n) + \mathbf{f}_{\mathrm{C}}(\mathscr{A}, \mathscr{C}) = -\mathbf{f}_{\mathrm{B}}(\mathscr{C}_n, \mathscr{A}_n) - \mathbf{f}_{\mathrm{C}}(\mathscr{C}, \mathscr{A}). \quad \text{(III.1-54)}$$

As $n \to \infty$, both body forces $\mathbf{f}_{\mathrm{B}}(\mathscr{A}_n, \mathscr{C}_n)$ and $\mathbf{f}_{\mathrm{B}}(\mathscr{C}_n, \mathscr{A}_n)$ vanish because of $(12)_1$, and so $(50)_2$ follows. By use of (8) and the assumption that $\mathbf{f}$ is pairwise equilibrated we deduce also $(50)_1$. $\triangle$

Corollary. *Let* $\mathbf{f}$ *be pairwise equilibrated; let* $-\mathscr{S}$ *denote the contact having the same underlying set as* $\mathscr{S}$ *but opposite orientation; then*

$$\mathbf{t}_{-\mathscr{S}} = -\mathbf{t}_{\mathscr{S}}. \quad \text{(III.1-55)}$$

Exercise III.1.2. Use of the Lebesgue Differentiation Theorem[1] proves that (55) follows from $(50)_1$ and (46).

Thus far we have not called upon the principle of linear momentum, though to prove (50) we have assumed (I.5-6), which is a corollary of that principle. Now, finally, we turn our attention to the resultant forces on a body $\mathscr{P}$ in $\bar{\mathbf{\Omega}}$. Because, as we see from (13),

$$\mathbf{f}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}) = \mathbf{f}_{\mathrm{B}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}) + \mathbf{f}_{\mathrm{C}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}), \quad \text{(III.1-56)}$$

the assumption that $\mathbf{f}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}) = \mathbf{0}$ for all $\mathscr{P}$ then delivers

$$\mathbf{f}_{\mathrm{C}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}) = -\mathbf{f}_{\mathrm{B}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}). \quad \text{(III.1-57)}$$

If we consider a sequence of parts $\mathscr{P}$ such that $V(\chi(\mathscr{P})) \to 0$, by (II.2-9) and $(12)_1$ (in which we recall that $K_{\mathscr{C}}$ is a bounded function of $\mathscr{C}$ and hence know that $K_{\mathscr{P}^{\mathrm{e}}}$ is a bounded function of $\mathscr{P}$) we conclude from (57) that there is a positive number K^{e} such that

$$|\mathbf{f}_{\mathrm{C}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}})| \leqq K^{\mathrm{e}} V(\chi(\mathscr{P})), \quad \text{(III.1-58)}$$

and so

$$\mathbf{f}_{\mathrm{C}}(\mathscr{P}, \mathscr{P}^{\mathrm{e}}) \to \mathbf{0}. \quad \text{(III.1-59)}$$

[1] *E.g.* Theorem 8.8 in the book by RUDIN cited above in Footnote 2 on p. 90.

This fact is stated formally by the following

Theorem (CAUCHY). *If the system of forces is balanced, then on a sequence of shapes whose volumes tend to* 0 *the resultant contact forces tend to* $\mathbf{0}$.

This theorem will play a major part in Section III.3, where we develop the nature of the contact forces in the interiors of the shapes of bodies.

Also the foregoing conclusions lead to the statements $(\mathrm{I}.13\text{-}11)_1$, (2), and (3), namely the assumptions with which the traditional treatment begins.

In the remainder of this book we follow largely the classical, local, and informal style of argument in continuum mechanics.

2. Reactions upon Containers and Submerged Obstacles

With little more than the concept of a system of contact forces and the theorem of action and reaction in the form (III.1-50) we can sometimes evaluate the force and the torque that a deforming body exerts upon a container or an object submerged in it. Analyses of this kind go back to the earliest days of mechanics; they remain of great utility to engineers because they require very little detailed knowledge of either the body or its motion; and for just the same reason they sometimes provide essential steps in the precise, mathematical treatment of qualitative problems. Here we shall consider only three examples, the simplest. We present them upon a general framework due, more or less, to v. MISES, CISOTTI, and BOGGIO.

First we substitute (II.6-12), (III.1-2), and (III.1-5) into Euler's Laws (I.13-11), thus obtaining expressions for the resultant contact force $\mathbf{f}_{\mathrm{C}}^{\mathrm{r}}$ and resultant contact torque $\mathbf{F}_{\mathrm{C}}^{\mathrm{r}}$ on the present shape $\chi(\mathscr{P})$ of the part $\mathscr{P}$ of a body:

$$-\mathbf{f}_{\mathrm{C}}^{\mathrm{r}} = -\int_{\chi(\mathscr{P})} (\rho\dot{\mathbf{x}})'\,dV - \int_{\partial\chi(\mathscr{P})} \rho\dot{\mathbf{x}}\otimes\dot{\mathbf{x}}\mathbf{n}\,dA + \mathbf{f}_{\mathrm{B}}^{\mathrm{r}},$$

$$-\mathbf{F}_{\mathrm{C}}^{\mathrm{r}} = -\int_{\chi(\mathscr{P})} \mathbf{p}\wedge(\rho\dot{\mathbf{x}})'\,dV - \int_{\partial\chi(\mathscr{P})} \mathbf{p}\wedge\rho\dot{\mathbf{x}}\otimes\dot{\mathbf{x}}\mathbf{n}\,dA + \mathbf{F}_{\mathrm{B}}^{\mathrm{r}}. \tag{III.2-1}$$

Here $\mathbf{F}_{\mathrm{C}}^{\mathrm{r}}$ is taken with respect to a fixed place $\mathbf{x}_0$ in an inertial frame, and $\mathbf{p} = \mathbf{x} - \mathbf{x}_0$ as in Sections I.13 and II.12. Because of (III.1-50), the left-hand sides of (1) are the contact force and torque, respectively, exerted by $\mathscr{P}$ upon its exterior $\mathscr{P}^e$. It is these quantities that we wish to evaluate in special problems. Together they are called the *reaction* of $\mathscr{P}$ upon its exterior.

In general terms (1) asserts that if the resultant applied force and torque upon $\mathscr{P}$ are known, *then the reaction of $\mathscr{P}$ upon its exterior is determined by the fields of density and velocity over the present shape of $\mathscr{P}$*. Kinematical data thus determine the reaction.

The essential field $\rho\dot{\mathbf{x}}$, which has the dimensions of momentum per unit volume, is called the *mass flow*. In the three examples we shall give now we shall assume that the mass flow is steady: $(\rho\dot{\mathbf{x}})' = \mathbf{0}$.

Example 1. *Flow in a Stationary Container.* Let $\mathscr{P}$ be confined by a bounded, stationary container. On the walls $\partial\chi(\mathscr{P})$ of the container the condition (II.6-17) is satisfied, and so the integral over $\partial\chi(\mathscr{P})$ vanishes. A little reflection enables us to derive the same conclusion even if $\mathscr{P}$ does not fill the container entirely. We have shown that *a body in motion with steady mass flow within a bounded, stationary container exerts upon that container just the reaction it would exert, were it at rest.*

Example 2. *Flow in a Pipe.* Suppose that a body is flowing through a stationary pipe of arbitrary form. We consider the part $\mathscr{P}$ contained in the portion of the pipe cut off by two surfaces, the inlet $\mathscr{S}_\mathrm{i}$ and the outlet $\mathscr{S}_\mathrm{o}$. Upon the walls of the pipe (II.6-17) is satisfied, and so, if $(\rho\dot{\mathbf{x}})' = \mathbf{0}$, (1) becomes

$$
\begin{aligned}
-\mathbf{f}_\mathrm{C}^\mathrm{r} &= \int_{\mathscr{S}_\mathrm{i}} \rho\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA - \int_{\mathscr{S}_\mathrm{o}} \rho\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA + \mathbf{f}_\mathrm{B}^\mathrm{r}, \\
-\mathbf{F}_\mathrm{C}^\mathrm{r} &= \int_{\mathscr{S}_\mathrm{i}} \mathbf{p} \wedge \rho\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA - \int_{\mathscr{S}_\mathrm{o}} \mathbf{p} \wedge \rho\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA + \mathbf{F}_\mathrm{B}^\mathrm{r}.
\end{aligned}
\tag{III.2-2}
$$

In writing the integral over $\mathscr{S}_\mathrm{i}$ we have taken the normal $\mathbf{n}$ as directed inward, so as to emphasise that a difference is being calculated.

We have not yet called upon the Traction Theorem. Doing so, we use (III.1-46) and (III.1-50)$_2$ to express the contact force and torque $\mathbf{f}_\mathrm{p}$ and $\mathbf{F}_\mathrm{p}$ on the pipe alone:

$$
\begin{aligned}
\mathbf{f}_\mathrm{C}^\mathrm{r} &= -\mathbf{f}_\mathrm{p} - \int_{\mathscr{S}_\mathrm{i}} \mathbf{t}_{\mathscr{S}_\mathrm{i}}\, dA + \int_{\mathscr{S}_\mathrm{o}} \mathbf{t}_{\mathscr{S}_\mathrm{o}}\, dA, \\
\mathbf{F}_\mathrm{C}^\mathrm{r} &= -\mathbf{F}_\mathrm{p} - \int_{\mathscr{S}_\mathrm{i}} \mathbf{p} \wedge \mathbf{t}_{\mathscr{S}_\mathrm{i}}\, dA + \int_{\mathscr{S}_\mathrm{o}} \mathbf{p} \wedge \mathbf{t}_{\mathscr{S}_\mathrm{o}}\, dA.
\end{aligned}
\tag{III.2-3}
$$

Eliminating the left-hand sides of (2) and (3), we find that

$$\mathbf{f}_{\mathrm{P}} = \int_{\mathscr{S}_{\mathrm{i}}} (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n} - \mathbf{t}_{\mathscr{S}_{\mathrm{i}}})\, dA - \int_{\mathscr{S}_{\mathrm{o}}} (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n} - \mathbf{t}_{\mathscr{S}_{\mathrm{o}}})\, dA + \mathbf{f}^{\mathrm{r}}_{\mathrm{B}},$$

$$\mathbf{F}_{\mathrm{P}} = \int_{\mathscr{S}_{\mathrm{i}}} \mathbf{p} \wedge (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n} - \mathbf{t}_{\mathscr{S}_{\mathrm{i}}})\, dA - \int_{\mathscr{S}_{\mathrm{o}}} \mathbf{p} \wedge (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n} - \mathbf{t}_{\mathscr{S}_{\mathrm{o}}})\, dA + \mathbf{F}^{\mathrm{r}}_{\mathrm{B}}. \tag{III.2-4}$$

From these formulae we see that *measurement of* ρ, $\dot{\mathbf{x}}$, *and* $\mathbf{t}_{\mathscr{S}}$ *at the inlet and the outlet suffice to determine the reaction exerted on a stationary pipe by a body moving through it with steady mass flow,* provided the applied force and torque on the body be known.

Various simplifying assumptions reduce the general expressions (4) to examples of great use in hydraulics. Instances of $(4)_1$ are called "Bernoulli's theorem", "the flow energy theorem", "the impulse theorem", *etc*. For the truth of the result it is not necessary that the body fill the pipe or that the fields ρ and $\dot{\mathbf{x}}$ be smooth within it; of course the cross-sectional area of the body must not vanish at any cross-section of the pipe.

Example 3. *Reaction upon a Submerged Object.* In this example we shall suppose that $\rho' = 0$, $\dot{\mathbf{x}}' = \mathbf{0}$, $\mathbf{f}^{\mathrm{r}}_{\mathrm{B}} = \mathbf{0}$, and $\mathbf{F}^{\mathrm{r}}_{\mathrm{B}} = \mathbf{0}$. We consider a body filling all of space except for a stationary, rigid, bounded object. The shape of the object need not be specified in the present context, for upon it the boundary condition (II.6-17) is satisfied, and so the integrals of integrands proportional to $\dot{\mathbf{x}}$ over the boundary of the obstacle vanish. We consider the region $\mathscr{R}_{\mathrm{c}}$ between the obstacle and a closed *control surface* $\mathscr{S}_{\mathrm{c}}$ so large as to contain the obstacle entirely. We denote by $\mathscr{P}$ the part of the body whose shape is the region $\mathscr{R}_{\mathrm{c}}$. We may take for $\mathscr{S}_{\mathrm{c}}$ the surface of a sphere if we wish to. From (1) we conclude that

$$-\mathbf{f}^{\mathrm{r}}_{\mathrm{C}} = -\int_{\mathscr{S}_{\mathrm{c}}} \rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA,$$

$$-\mathbf{F}^{\mathrm{r}}_{\mathrm{C}} = -\int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n})\, dA. \tag{III.2-5}$$

Exercise III.2.1. If $\mathbf{v}$ is any constant vector field,

$$\int_{\mathscr{S}_{\mathrm{c}}} \rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n}\, dA = \int_{\mathscr{S}_{\mathrm{c}}} \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes (\dot{\mathbf{x}} - \mathbf{v})\mathbf{n}\, dA + \left[\int_{\mathscr{S}_{\mathrm{c}}} \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes \mathbf{n}\, dA\right] \mathbf{v},$$

$$\begin{aligned}\int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge (\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}}\mathbf{n})\, dA = &\int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge [\rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes (\dot{\mathbf{x}} - \mathbf{v})\mathbf{n}]\, dA \\ &+ \int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge [\rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes \mathbf{v}\mathbf{n}]\, dA + \mathbf{m} \wedge \mathbf{v},\end{aligned} \tag{III.2-6}$$

$\mathbf{m}$ being the momentum of $\mathscr{P}$.

We now call upon the Traction Theorem. If we write $\mathbf{f}_{\mathrm{obs}}$ and $\mathbf{F}_{\mathrm{obs}}$ for the force and torque exerted by $\mathscr{P}$ upon the obstacle, then (III.1-46) and (III.1-50)$_2$ show that

$$\begin{aligned} \mathbf{f}^{\mathrm{r}}_{\mathrm{C}} &= -\mathbf{f}_{\mathrm{obs}} + \int_{\mathscr{S}_{\mathrm{c}}} \mathbf{t}_{\mathscr{S}_{\mathrm{c}}}\, dA, \\ \mathbf{F}^{\mathrm{r}}_{\mathrm{C}} &= -\mathbf{F}_{\mathrm{obs}} + \int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge \mathbf{t}_{\mathscr{S}_{\mathrm{c}}}\, dA. \end{aligned} \tag{III.2-7}$$

Putting (5), (6), and (7) together, we obtain finally

$$\begin{aligned} \mathbf{f}_{\mathrm{obs}} &= -\int_{\mathscr{S}_{\mathrm{c}}} \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes (\dot{\mathbf{x}} - \mathbf{v})\mathbf{n}\, dA - \left[\int_{\mathscr{S}_{\mathrm{c}}} \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes \mathbf{n}\, dA\right]\mathbf{v} \\ &\quad + \int_{\mathscr{S}_{\mathrm{c}}} (\mathbf{t}_{\mathscr{S}_{\mathrm{c}}} + P\mathbf{n})\, dA, \\ \mathbf{F}_{\mathrm{obs}} &= -\int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes (\dot{\mathbf{x}} - \mathbf{v})\mathbf{n}\, dA \\ &\quad - \int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge \rho(\dot{\mathbf{x}} - \mathbf{v}) \otimes \mathbf{v}\mathbf{n}\, dA - \mathbf{m} \wedge \mathbf{v} + \int_{\mathscr{S}_{\mathrm{c}}} \mathbf{p} \wedge (\mathbf{t}_{\mathscr{S}_{\mathrm{c}}} + P\mathbf{n})\, dA; \end{aligned} \tag{III.2-8}$$

here we have added to each right-hand side the resultant force and resultant torque of a constant scalar pressure P, both of these resultants being null.

These formulae serve to evaluate the reaction exerted by the motion of $\mathscr{P}$ upon a stationary, rigid obstacle immersed in it. The obstacle itself seems not to enter the final results. All we need know is the steady kinematical fields ρ and $\dot{\mathbf{x}}$ and the traction field $\mathbf{t}_{\mathscr{S}_{\mathrm{c}}}$ upon the control surface $\mathscr{S}_{\mathrm{c}}$. The choice of $\mathscr{S}_{\mathrm{c}}$ is ours. Generally it is advantageous to choose it very large; we expect then that the effect of the obstacle upon the fields of ρ and $\dot{\mathbf{x}}$ be lessened. To evaluate the integrals we may adjust as we like the form of $\mathscr{S}_{\mathrm{c}}$ and the values of the arbitrary constants $\mathbf{v}$ and P.

The most celebrated example is provided by the steady, uniform flow of a body past an obstacle. Then $\dot{\mathbf{x}} \to \mathbf{v}$, say, at ∞. If also the traction field at great distances from the obstacle is approximately a uniform hydrostatic pressure p_∞, then $\mathbf{t}_{\mathscr{S}_{\mathrm{c}}} \to -p_\infty \mathbf{n}$ at ∞. So as to model this condition, we consider the body $\mathscr{P}_{\mathrm{r}}$ that presently occupies the space between the obstacle and the surface $\mathscr{S}_{\mathrm{r}}$ of the sphere of radius r, centered upon some fixed point. To $\mathscr{P}_{\mathrm{r}}$ for any large enough r we may apply (8). If the integrals in (8) converge as $r \to \infty$, we obtain definite expressions for $\mathbf{f}_{\text{obs}}$ and $\mathbf{F}_{\text{obs}}$. For example, if as $r \to \infty$

$$\dot{\mathbf{x}} - \mathbf{v} = \mathbf{o}(r^{-2}), \qquad \rho = O(1), \qquad \mathbf{t}_{\mathscr{S}_r} + p_\infty \mathbf{n} = \mathbf{o}(r^{-2}), \quad \text{(III.2-9)}$$

then all the integrals in $(8)_1$ converge to naught as $r \to \infty$, and so

$$\mathbf{f}_{\text{obs}} = \mathbf{0}. \qquad \text{(III.2-10)}$$

That is, *the conditions* (9) *are sufficient that the infinite body with steady density in steady flow past the obstacle exert no resultant force on the obstacle.*

The conditions (9) are of the essence for the proof. Without some conditions of this kind, no such conclusion follows. They assert that the disturbance due to the presence of the obstacle falls off rapidly at great distances from it; indeed, they specify the rate at which it falls off. In Volume 2 we shall show that they are satisfied by an irrotational flow of a homogeneous, incompressible, Eulerian fluid body filling all of space outside an obstacle. For other bodies they are not. In general, $\mathbf{f}_{\text{obs}} \neq \mathbf{0}$.

EULER, treating a very special instance, was the first to obtain (10). His reasoning provides a primitive example of that which we have given in general terms. D'ALEMBERT, much later, rediscovered or appropriated the assertion; his unduly special reasoning applies only to obstacles of great symmetry. The fact itself he announced as a paradox. The name has stuck: "the d'Alembert paradox". Both the name and the fact have given rise to perennial confusion.

If as $r \to \infty$

$$\dot{\mathbf{x}} - \mathbf{v} = \mathbf{o}(r^{-3}), \qquad \rho = O(1), \qquad \mathbf{t}_{\partial\mathscr{S}_r} + p_\infty \mathbf{n} = \mathbf{o}(r^{-3}), \quad \text{(III.2-11)}$$

then of course (10) holds, and also the integrals in $(8)_2$ converge. The infinite body occupying the region $\mathscr{S}_\infty$ outside the obstacle has finite *relative momentum* $\mathbf{m}_\infty$, given by

$$\mathbf{m}_\infty := \int_{\mathscr{S}_\infty} \rho(\dot{\mathbf{x}} - \mathbf{v})\, dV, \qquad \text{(III.2-12)}$$

and from $(8)_2$ we obtain

$$\mathbf{F}_{\text{obs}} = \mathbf{v} \wedge \mathbf{m}_\infty. \qquad \text{(III.2-13)}$$

Here, not in the proof of (10), belong appeals to symmetry, for they suffice to show that $\mathbf{F}_{\text{obs}} = \mathbf{0}$ for some obstacles, though not for others.

3. The Traction Field. The Cauchy Postulate and the Hamel–Noll Theorem

In Section III.1 we have expressed the principles of balance of linear and rotational momentum in terms of the traction field $\mathbf{t}_{\partial\chi(\mathscr{P})}$ on the boundary $\partial\chi(\mathscr{P})$ of the shape of each part $\mathscr{P}$ of $\mathscr{B}$:

$$\int_{\chi(\mathscr{P})} \rho\ddot{\mathbf{x}}\,dV = \int_{\partial\chi(\mathscr{P})} \mathbf{t}_{\partial\chi(\mathscr{P})}\,dA + \int_{\chi(\mathscr{P})} \rho\mathbf{b}\,dV,$$

$$\text{(III.1-8)}_r$$

$$\int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\ddot{\mathbf{x}}\,dV = \int_{\partial\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\mathbf{t}_{\partial\chi(\mathscr{P})}\,dA + \int_{\chi(\mathscr{P})} (\mathbf{x}-\mathbf{x}_0)\wedge\rho\mathbf{b}\,dV;$$

here we continue to consider a particular time t but do not indicate it in the notation. To reduce these integral equations to equivalent field equations, we must express the traction field $\mathbf{t}_{\partial\chi(\mathscr{P})}$, which is defined only upon $\partial\chi(\mathscr{P})$, in terms of fields defined in an open set containing $\partial\chi(\mathscr{P})$. To an extension of this kind we now address ourselves.

A place $\mathbf{x}$ on $\partial\chi(\mathscr{P})$ obviously lies also upon the boundaries $\partial\chi(\mathscr{Q})$ of infinitely many parts $\mathscr{Q}$ of $\mathscr{P}$. The traction $\mathbf{t}$ for these various boundaries having the point $\mathbf{x}$ in common depends, in general, upon $\partial\chi(\mathscr{Q})$. For the reader who skipped the developments in Section III.1 following (III.1-9), we here repeat that we call contact of the shapes of two separate bodies the surface that their boundaries have in common. Here a contact is a smooth surface, oriented by assigning it one of its two normal fields. A subcontact is any subset of a contact that is assigned the same orientation. In Section III.1 we have shown that if $\mathscr{S}'$ is a subcontact of $\mathscr{S}$, then $\mathbf{t}_{\mathscr{S}'} = \mathbf{t}_{\mathscr{S}}$, but we have not established any relation between $\mathbf{t}_{\mathscr{S}'}$ and $\mathbf{t}_{\mathscr{S}}$ for more general pairs of contacts $\mathscr{S}$ and $\mathscr{S}'$, for example if $\mathscr{S}'$ and $\mathscr{S}$ have in common only the one place $\mathbf{x}$ we are considering. Classical continuum mechanics as developed by Cauchy and his successors assumes that *the tractions on all like-oriented contacts with a common tangent plane at* $\mathbf{x}$ *are the same at* $\mathbf{x}$. That is, $\mathbf{t}_{\mathscr{S}}$ at $\mathbf{x}$ is assumed to depend upon $\mathscr{S}$ only through the normal $\mathbf{n}$ of $\mathscr{S}$ at $\mathbf{x}$:

$$\mathbf{t}_{\mathscr{S}} = \mathbf{t}(\mathbf{x}, \mathbf{n}). \qquad \text{(III.3-1)}$$

This statement may be called ***the Cauchy Postulate***. $\mathscr{S}$ is oriented so that its normal $\mathbf{n}$ points out of $\chi(\mathscr{B})$ if $\mathscr{S}$ is a part of $\partial\chi(\mathscr{B})$. Thus $\mathbf{t}(\mathbf{x}, -\mathbf{n})$ is the traction at $\mathbf{x}$ on all surfaces $\mathscr{S}$ tangent to $\partial\chi(\mathscr{B})$ and forming parts of the boundaries of bodies in the exterior $\chi(\mathscr{B}^e)$ of $\chi(\mathscr{B})$. In this sense $\mathbf{t}(\mathbf{x}, \mathbf{n})$ is the traction exerted upon $\mathscr{B}$ at $\mathbf{x}$ by the contiguous bodies outside $\mathscr{B}$, while $\mathbf{t}(\mathbf{x}, -\mathbf{n})$ is the traction exerted there by $\mathscr{B}$ on the contiguous bodies outside it.

As a trivial corollary of (III.1-55) follows ***Cauchy's Fundamental Lemma***:

$$\mathbf{t}(\mathbf{x}, -\mathbf{n}) = -\mathbf{t}(\mathbf{x}, \mathbf{n}). \tag{III.3-2}$$

For those readers who have not stopped to follow the theorems given in the latter parts of Section III.1 we include here a sketch of CAUCHY's own argument to prove (2) as a consequence of (1) and the balance of linear momentum, of course without use of (III.1-55).

Proof. In view of (1), it suffices to consider an oriented disk $\mathscr{S}$ of sufficiently small radius, centered at $\mathbf{x}$; then $-\mathscr{S}$ is the oppositely oriented disk. Assuming that the universe of shapes is rich enough in sets that every right-circular cylinder of sufficiently small base and altitude is the shape of some body, for $\chi(\mathscr{B})$ we take a circular cylinder which is normal to $\mathscr{S}$ and is bisected transversely by $\mathscr{S}$. If ϵ denotes the altitude of this cylinder, then $V(\chi(\mathscr{B})) = \epsilon A(\mathscr{S})$. We assume that $\mathbf{b} - \ddot{\mathbf{x}}$ is essentially bounded. *Cf.* the statement following (III.1-47). We apply the balance of linear momentum as expressed by (III.1-8) to $\chi(\mathscr{B})$ and then take the limit as $\epsilon \to 0$, the disk $\mathscr{S}$ being kept fixed, so that its area remains constant. The limit of the difference of the volume integrals is $\mathbf{0}$, and so

$$\lim_{\epsilon \to 0} \int_{\partial\chi(\mathscr{B})} \mathbf{t}(\mathbf{x}, \mathbf{n})\, dA = \mathbf{0}. \tag{III.3-3}$$

(This statement is a special case of (III.1-59), but the present proof is intended for the reader who skipped the part of Section III.1 that follows (III.1-9).) In the passage to the limit $\mathbf{n}$ does not vary, but the set of $\mathbf{x}$ over which the integral is taken shrinks down to $\mathscr{S}$, twice over. If we assume that $\mathbf{t}(\cdot, \mathbf{n})$ is an essentially bounded function of $\mathbf{x}$, the limit of the integral over the mantle of the cylinder is $\mathbf{0}$, since the area of that mantle tends to 0. Thus only the limits of the two integrals over $\mathscr{S}$ remain to be considered. If we assume further that $\mathbf{t}(\cdot, \mathbf{n})$ is a continuous function of $\mathbf{x}$, then the limits of these integrals equal the integrals of the limit functions in the two cases:

$$\int_{\mathscr{S}} [\mathbf{t}(\mathbf{x}, \mathbf{n}) + \mathbf{t}(\mathbf{x}, -\mathbf{n})]\, dA = \mathbf{0}. \tag{III.3-4}$$

Since $\mathscr{S}$ is any sufficiently small disk normal to $\mathbf{n}$ at $\mathbf{x}$, (2) follows. $\triangle$

The reader who is content to lay down the Cauchy Postulate (1) and to make the assumptions of smoothness concerning $\mathbf{b} - \ddot{\mathbf{x}}$ and $\mathbf{t}$ that we have stated in the course of proving CAUCHY's Fundamental Lemma should now pass on to the next section. On the other hand, the reader who has followed the development in Section III.1 will have noted that one of the assumptions of smoothness made just now is unnecessarily strong, while another has already been proved to hold in the mathematical theory based on (III.1-10). In fact, as we shall see now, the Cauchy Postulate can be proved true as a consequence of the principle of linear momentum and very mild assumptions of smoothness. In the proof we shall appeal to some of the conclusions demonstrated in the part of Section III.1 that begins a little before (III.1-10).

Theorem (HAMEL (imperfectly), NOLL). *Suppose that the contact force* $\mathbf{f}_C(\mathscr{A}, \mathscr{C})$ *exerted upon any part* $\mathscr{A}$ *of* $\mathscr{B}$ *by the separate body* $\mathscr{C}$ *be determined by a traction field* $\mathbf{t}_{\mathscr{S}}$ *through* (III.1-46). *Then the Cauchy Postulate* (1) *holds almost everywhere on every surface* $\mathscr{S}$ *interior to* $\chi(\mathscr{B})$.

The reader should recall that (III.1-46) has been proved to hold as a consequence of (III.1-10).

Proof of the Hamel-Noll theorem.[1] By (III.1-43), (III.1-44), and the Lebesgue differentiation theorem we know that at almost all points of the surface $\mathscr{S}$

$$\mathbf{t}_{\mathscr{S}}(\mathbf{x}) = \lim_{m\to\infty} \frac{\int_{\mathscr{U}_m} \mathbf{t}_{\mathscr{S}}\, dA}{A(\mathscr{U}_m)} \tag{III.3-5}$$

if $\mathscr{U}_m$ is a suitably selected sequence of sets on $\mathscr{S}$ shrinking down to $\mathbf{x}$. We are to show that if $\mathscr{T}$ and $\mathscr{S}$ have a common oriented normal $\mathbf{n}$ at $\mathbf{x}$, and if both $\mathbf{t}_{\mathscr{T}}(\mathbf{x})$ and $\mathbf{t}_{\mathscr{S}}(\mathbf{x})$ exist, then

$$\mathbf{t}_{\mathscr{T}}(\mathbf{x}) = \mathbf{t}_{\mathscr{S}}(\mathbf{x}). \tag{III.3-6}$$

The common value of the two functions $\mathbf{t}_{\mathscr{T}}$ and $\mathbf{t}_{\mathscr{S}}$ at $\mathbf{x}$ is then a function of $\mathbf{n}$ only and may be denoted by $\mathbf{t}(\mathbf{x}, \mathbf{n})$; if so, the HAMEL–NOLL theorem will have been proved.

If $\mathscr{S}$ and $\mathscr{T}$ coincide near $\mathbf{x}$, the claim is trivial. Otherwise, at the regular point $\mathbf{x}$ common to $\mathscr{S}$ and $\mathscr{T}$ we describe a circular cylinder of small radius Δr

[1] A good presentation of this proof with appropriate figures may be found in Section II.5 of IRE, cited at the end of this chapter.

about the common normal $\mathbf{n}$ and denote the parts of its interior lying between $\mathscr{S}$ and $\mathscr{T}$ by $\Delta\mathscr{D}$. We denote the cylindrical part of $\partial\Delta\mathscr{D}$ by $\Delta\mathscr{A}^*$ and the parts of $\partial\Delta\mathscr{D}$ common to $\mathscr{S}$ and $\mathscr{T}$, respectively, by $\Delta\mathscr{A}$ and $\Delta\mathscr{A}'$.

First we suppose that $\mathscr{T}$ is the tangent plane to $\mathscr{S}$ at $\mathbf{x}$, that $\mathbf{x}$ is an elliptic point for $\mathscr{S}$, and that (5) holds at $\mathbf{x}$ for $\mathscr{S}$. Then not only does $\mathscr{S}$ lie entirely on one side of $\mathscr{T}$ near $\mathbf{x}$, but also we may construct a paraboloid of revolution $\mathscr{S}^*$ with vertex $\mathbf{x}$, with $\mathscr{T}$ as its tangent plane, and such as to include between itself and $\mathscr{T}$ all of $\mathscr{S}$, for sufficiently small Δr. Specifically, if $z = f(x, y)$ is a cartesian equation of $\mathscr{S}$ near $\mathbf{x}$, the coordinates x and y being in $\mathscr{T}$ and z being distance along the normal to $\mathscr{T}$, then $\partial_x^2 f \geqq 0$, and $\partial_y^2 f \geqq 0$ at $\mathbf{x}$, and if

$$K := \max(\partial_x^2 f, \partial_x\partial_y f, \partial_y^2 f) \qquad \text{when} \quad x^2 + y^2 \leqq \Delta r^2, \tag{III.3-7}$$

a paraboloid of the kind desired is given by

$$z = 2K(x^2 + y^2). \tag{III.3-8}$$

The area of the cylindrical part $\Delta\mathscr{A}^*$ of $\partial\Delta\mathscr{D}$ is not greater than that of the part of the cylinder between the plane and the paraboloid. Thus

$$\begin{aligned} A(\Delta\mathscr{A}^*) &\leqq (2\pi\Delta r)\cdot 2K(\Delta r)^2, \\ &= o(\Delta r^2) \end{aligned} \tag{III.3-9}$$

as $\Delta r \to 0$. Likewise, the volume of $\Delta\mathscr{D}$ is bounded by that of the region between $\mathscr{S}^*$, the cylinder, and the plane. Thus

$$\begin{aligned} V(\Delta\mathscr{D}) &\leqq \pi K \Delta r^4, \\ &= o(\Delta r^3). \end{aligned} \tag{III.3-10}$$

Of course

$$A(\Delta\mathscr{A}') = \pi\Delta r^2. \tag{III.3-11}$$

Exercise III.3.1.

$$A(\Delta\mathscr{A}) = \pi\Delta r^2 + o(\Delta r^2). \tag{III.3-12}$$

Hence

$$\begin{aligned} A(\partial\Delta\mathscr{D}) &= 2\pi\Delta r^2 + o(\Delta r^2) \qquad \text{as} \quad \Delta r \to 0, \\ V(\Delta\mathscr{D}) &= o(A(\partial\Delta\mathscr{D})) \qquad \text{as} \quad \Delta r \to 0. \end{aligned} \tag{III.3-13}$$

We assume that the universe of shapes is rich enough in sets that $\Delta\mathscr{D}$, no matter what be Δr, is the shape of some body $\mathscr{P}_{\Delta r}$. Because of $(13)_2$ and (III.1-58)

$$\lim_{\Delta r\to 0}\frac{\mathbf{f}_C(\mathscr{P}_{\Delta r},\mathscr{P}^e_{\Delta r})}{A(\partial\Delta\mathscr{D})}=0. \tag{III.3-14}$$

Orienting the tangent plane $\mathscr{T}$ so that its normal points into $\mathscr{S}$ at $\mathbf{x}$, we conclude that for almost all $\mathbf{x}$ on $\mathscr{S}$

$$\lim_{\Delta r\to 0}\frac{1}{A(\partial\Delta\mathscr{D})}\left(\int_{\Delta\mathscr{A}}\mathbf{t}_{\mathscr{S}}\,dA+\int_{\Delta\mathscr{A}'}\mathbf{t}_{\mathscr{T}}\,dA+\int_{\Delta\mathscr{A}^*}\mathbf{t}_{\Delta\mathscr{A}^*}\,dA\right)=\mathbf{0}. \tag{III.3-15}$$

In view of (9) and $(13)_1$ the limit of the third term vanishes, and so by (11) and (12) we see that

$$\lim_{\Delta r\to 0}\frac{1}{A(\Delta\mathscr{A})}\int_{\Delta\mathscr{A}}\mathbf{t}_{\mathscr{S}}\,dA=-\lim_{\Delta r\to 0}\frac{1}{A(\Delta\mathscr{A}')}\int_{\Delta\mathscr{A}'}\mathbf{t}_{\mathscr{T}}\,dA, \tag{III.3-16}$$

provided either limit exist. Now the value of the limit on the left-hand side is $\mathbf{t}_{\mathscr{S}}(\mathbf{x})$ by assumption, and so the limit on the right-hand side is proved to exist by the argument given. The limit on the right-hand side is independent of the choice of $\mathscr{S}$. Hence $\mathbf{t}_{\mathscr{S}}(\mathbf{x})$ is the same for all surfaces $\mathscr{S}$ that are elliptic at $\mathbf{x}$, provided $\mathbf{x}$ be a place where (5) holds for $\mathscr{S}$.

Exercise III.3.2. The proof of the HAMEL–NOLL theorem is completed by letting $\mathbf{x}$ be a saddle point for $\mathscr{S}$ or for $\mathscr{T}$ or for both. $\triangle$

4. Cauchy's Fundamental Theorem: Existence of the Stress Tensor

The Cauchy Postulate (III.3-1) and its consequence, CAUCHY's Fundamental Lemma (III.3-2), enable us to determine the way the traction field $\mathbf{t}$ depends upon $\mathbf{n}$. Indeed, it is a linear function of $\mathbf{n}$, as shown by

Cauchy's Fundamental Theorem. *If* $\mathbf{t}(\cdot,\mathbf{n})$ *is a continuous function, there is a tensor field* $\mathbf{T}$ *such that*

$$\mathbf{t}(\mathbf{x},\mathbf{n})=\mathbf{T}(\mathbf{x})\mathbf{n}. \tag{III.4-1}$$

CAUCHY himself interpreted this theorem as stating that the tractions on any three linearly independent planes at a point determine the traction on any and every surface at that point. Indeed, let $\{\mathbf{e}_k\}$ be a basis in $\mathscr{V}$, so that $\mathbf{n} = n^k\mathbf{e}_k$. Then (1) asserts that

$$\mathbf{t} = \mathbf{T}(n^k\mathbf{e}_k) = n^k(\mathbf{T}\mathbf{e}_k) = n^k\mathbf{t}_k\,, \tag{III.4-2}$$

$\mathbf{t}_k$ being the traction on a plane whose outer unit normal is $\mathbf{e}_k$. The value $\mathbf{T}(\mathbf{x})$ of the tensor field $\mathbf{T}$ at $\mathbf{x}$ is called the *stress tensor*, and CAUCHY's Fundamental Theorem asserts *the existence of the stress-tensor field.* The letter $\mathbf{T}$ should recall "tension," since $\mathbf{n}\cdot\mathbf{T}(\mathbf{x})\mathbf{n} > 0$ if and only if the traction $\mathbf{t}(\mathbf{x})$ is a tension. Sometimes, accordingly, $-\mathbf{T}$ is called the *pressure tensor*.

CAUCHY himself proved his theorem by applying (III.1-58) to a tetrahedron, three of whose four faces were mutually perpendicular. In this way he concluded by use of an orthonormal basis $\{\mathbf{e}_k\}$ at $\mathbf{x}$ that

$$\mathbf{T} = \sum_{k=1}^{3}\mathbf{t}(\mathbf{x}, \mathbf{e}_k)\otimes\mathbf{e}_k\,, \tag{III.4-3}$$

a statement equivalent to (2). CAUCHY's proof, which suggests a method for discovering the theorem, has been reproduced again and again in the textbooks. Here we shall give a proof due to NOLL[1], which is similar to CAUCHY's in resting essentially upon (III.1-58) but differs in that it uses a construction employing any two linearly independent vectors rather than an orthogonal triad.

Proof. The function $\mathbf{t}(\mathbf{x}, \cdot)$ is defined by (III.3-1) only for arguments which are unit vectors. We may extend it as follows to all of $\mathscr{V}$:

$$\mathbf{t}(\mathbf{x}, \mathbf{v}) := \begin{cases} |\mathbf{v}|\,\mathbf{t}\left(\mathbf{x}, \dfrac{\mathbf{v}}{|\mathbf{v}|}\right) & \text{if} \quad \mathbf{v} \neq \mathbf{0} \\ \mathbf{0} & \text{if} \quad \mathbf{v} = \mathbf{0}. \end{cases} \tag{III.4-4}$$

If $A > 0$ and $\mathbf{v} \neq \mathbf{0}$, then by $(4)_1$

$$\mathbf{t}(\mathbf{x}, A\mathbf{v}) = |A\mathbf{v}|\,\mathbf{t}\left(\mathbf{x}, \frac{A\mathbf{v}}{|A\mathbf{v}|}\right) = A\mathbf{t}(\mathbf{x}, \mathbf{v}), \tag{III.4-5}$$

[1] We follow the presentation by M. E. GURTIN in Section 15 of "The linear theory of elasticity," FLÜGGE's *Handbuch der Physik* **VIa/2**, ed. C. TRUESDELL, Berlin and New York, Springer-Verlag, 1972. In the same section GURTIN gives a rigorous and efficient version of CAUCHY's original proof. Rigorous presentations more or less close to CAUCHY's path of discovery may be found in CFT, Section 203; in Section 16 of C.-C. WANG's *Mathematical Principles of Mechanics and Electromagnetism*, N.Y. and London, Plenum, 1979; and Section 14 of M. E. GURTIN's *Introduction to Continuum Mechanics*, N.Y. *etc.*, Academic Press, 1981.

and the same conclusion follows trivially if $\mathbf{v} = \mathbf{0}$ or $A = 0$. If $A < 0$, then (5) and CAUCHY's Fundamental Lemma (III.3-2) show that

$$\mathbf{t}(\mathbf{x}, A\mathbf{v}) = \mathbf{t}(\mathbf{x}, -|A|\mathbf{v}) = |A|\mathbf{t}(\mathbf{x}, -\mathbf{v}) = A\mathbf{t}(\mathbf{x}, \mathbf{v}). \tag{III.4-6}$$

Thus $\mathbf{t}(\mathbf{x}, \cdot)$ is a homogeneous function of vectors.

We wish now to show that $\mathbf{t}(\mathbf{x}, \cdot)$ is additive:

$$\mathbf{t}(\mathbf{x}, \mathbf{v}_1 + \mathbf{v}_2) = \mathbf{t}(\mathbf{x}, \mathbf{v}_1) + \mathbf{t}(\mathbf{x}, \mathbf{v}_2). \tag{III.4-7}$$

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent, (7) follows at once from the homogeneity of $\mathbf{t}(\mathbf{x}, \cdot)$. We suppose, then, that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. At a given place $\mathbf{x}_0$ the planes $\mathbb{P}_1$ and $\mathbb{P}_2$ normal to $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively, are distinct. We set

$$\mathbf{v}_3 := -(\mathbf{v}_1 + \mathbf{v}_2) \tag{III.4-8}$$

and consider the wedge A that is bounded by these two planes, the plane $\mathbb{P}_3$ normal to $\mathbf{v}_3$ at the place $\mathbf{x}_0 + \epsilon\mathbf{v}_3$, the planes $\mathbb{P}_4$ and $\mathbb{P}_5$ distant δ from $\mathbf{x}_0$ and parallel to the plane of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. We suppose both ϵ and δ small enough that A be the shape of some part of $\mathscr{B}$, and we denote by $\partial_i\mathsf{A}$ the portion of the plane $\mathbb{P}_i$ that makes a part of the boundary of A. We shall hold δ fixed and let ϵ approach 0. $\partial_5\mathsf{A}$ is a triangle in the plane of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. If the lengths of its sides normal to these vectors are, respectively, l_1, l_2, and l_3, then consideration of similar triangles shows that $l_1/l_3 = |\mathbf{v}_1|/|\mathbf{v}_3|$ and $l_2/l_3 = |\mathbf{v}_2|/|\mathbf{v}_3|$. Also $l_3 = O(\epsilon)$. Hence if we write A_i for the area of $\partial_i\mathsf{A}$, we see that

$$A_1 = \frac{|\mathbf{v}_1|}{|\mathbf{v}_3|}A_3, \qquad A_2 = \frac{|\mathbf{v}_2|}{|\mathbf{v}_3|}A_3,$$

$$A_3 = O(\epsilon) \qquad \text{as} \quad \epsilon \to 0, \tag{III.4-9}$$

$$V(\mathsf{A}) = \tfrac{1}{2}\epsilon|\mathbf{v}_3|A_3 = 2\delta A_4 = 2\delta A_5 .$$

If

$$\mathbf{c} := \frac{|\mathbf{v}_3|}{A_3}\int_{\partial\mathsf{A}} \mathbf{t}(\mathbf{x}, \mathbf{n})\, dA, \tag{III.4-10}$$

from (9) and the assumption that $\mathbf{t}(\cdot, \mathbf{n})$ is continuous we see that

$$\mathbf{c} = \sum_{i=1}^{3} \frac{|\mathbf{v}_i|}{A_i} \int_{\partial_i \mathbf{A}} \mathbf{t}\left(\mathbf{x}, \frac{\mathbf{v}_i}{|\mathbf{v}_i|}\right) dA + O(\epsilon) \quad \text{as} \quad \epsilon \to 0, \qquad \text{(III.4-11)}$$

where we have used CAUCHY's Fundamental Lemma (III.3-2) so as to incorporate into the remainder the integrals over $\partial_4 \mathbf{A}$ and $\partial_5 \mathbf{A}$. Since $\mathbf{t}$ is a homogeneous function of its second argument and a continuous function of its first argument,

$$\mathbf{c} \to \sum_{k=1}^{3} \mathbf{t}(\mathbf{x}_0, \mathbf{v}_k) \quad \text{as} \quad \epsilon \to 0. \qquad \text{(III.4-12)}$$

On the other hand, by $(9)_4$ and (III.1-59), which is a consequence of the balance of momentum, we see that $\mathbf{c} \to \mathbf{0}$ as $\epsilon \to 0$. Therefore, since the sum in (12) is independent of ϵ, it must vanish:

$$\sum_{k=1}^{3} \mathbf{t}(\mathbf{x}_0, \mathbf{v}_k) = \mathbf{0}. \qquad \text{(III.4-13)}$$

By (7), $\mathbf{t}(\mathbf{x}_0, \cdot)$ is additive. Because every homogeneous additive function is linear, we have shown that $\mathbf{t}(\mathbf{x}, \mathbf{v}) = \mathbf{T}(\mathbf{x})\mathbf{v}$, and restriction of this statement to unit vectors yields CAUCHY's Fundamental Theorem (1). $\triangle$

If, as we have assumed in demonstrating CAUCHY's Fundamental Theorem, $\mathbf{t}(\cdot, \mathbf{n})$ is a continuous function, then so is $\mathbf{T}$.

CAUCHY's Fundamental Theorem must not be confused with any standard theorem in measure theory. The proof rests essentially upon (III.1-59), which reflects the balance of linear momentum. Of course the theorem can be phrased more abstractly, without verbal reference to mechanics,[1] but if we do not impose (III.1-59), then $\mathbf{t}$ need not be a linear function of $\mathbf{n}$. In some important theories of continuum mechanics there are, indeed, traction fields that are not delivered by a stress tensor,[2] but we shall not take up such theories in this book.

Exercise III.4.1 (GURTIN). Let $\mathbf{k}$ be a vector other than $\mathbf{0}$. If $\mathbf{t}(\mathbf{x}, \mathbf{n}) = (\mathbf{k} \cdot \mathbf{n})\mathbf{n}$, then (III.1-59) holds, but the important bound (III.1-58) and CAUCHY's Fundamental Lemma do not.

[1] *Cf.* the papers cited above in Footnote 2 on p. 88.

[2] *Cf.* R. A. TOUPIN, "Elastic materials with couple stresses," *Archive for Rational Mechanics and Analysis* **11** (1962), 387–414. TOUPIN calls the tensor whose contravariant components he denotes by t^{ij} the stress tensor, but according to his Equation $(7.19)_2$ the traction vector on a boundary surface is given only in part by $t^{ij}n_j$. TOUPIN's traction is not essentially bounded in the sense expressed by $(\text{III.1-12})_2$.

In the proof of CAUCHY's Fundamental Theorem the assumption that $\mathbf{t}(\cdot, \mathbf{n})$ be *continuous* is crucial. Then there is a sequence of sets $\mathscr{A}_m$ in the plane through $\mathbf{x}$ and normal to $\mathbf{n}$ such that

$$\mathbf{t}(\mathbf{x}, \mathbf{n}) = \lim_{m\to\infty} \frac{\displaystyle\int_{\mathscr{A}_m} \mathbf{t}_{\mathscr{A}_m}\, dA}{A(\mathscr{A}_m)}, \tag{III.4-14}$$

in which $\mathbf{t}_{\mathscr{A}_m}$ is the traction as defined by (III.1-46). This fact suggests that we define as follows the *average traction* $\mathbf{t}_r(\mathbf{x}, \mathbf{n})$ over a closed disk $\mathscr{D}_r$ of positive radius r, centered at $\mathbf{x}$ and normal to $\mathbf{n}$:

$$\mathbf{t}_r(\mathbf{x}, \mathbf{n}) := \frac{\displaystyle\int_{\mathscr{D}_r} \mathbf{t}_{\mathscr{D}_r}\, dA}{A(\mathscr{D}_r)}. \tag{III.4-15}$$

We say that $\mathbf{f}_C$ has *uniform average traction* in the actual shape $\chi(\mathscr{B})$ if for each fixed $\mathbf{n}$ the one-parameter family of functions $\{\mathbf{t}_r(\mathbf{x}, \mathbf{n})\}$ is, as $r \to 0$, uniformly Cauchy-convergent in the set of $\mathbf{x}$ belonging to any compact subset $\chi(\mathscr{B})$. Using $(\text{III.1-8})_1$ and (III.3-2), one can then show that the map $\mathbf{x} \mapsto \mathbf{t}_r(\mathbf{x}, \mathbf{n})$ is continuous.

Theorem (GURTIN & MARTINS).[1] *Let* $\mathbf{f}_C$ *satisfy* (III.1-58). *Then the following two statements are equivalent*:

(i) $\mathbf{f}_C$ *has uniform average traction.*

(ii) *for every* $\mathbf{x}$ *and every* $\mathbf{n}$, *the average traction* $\mathbf{t}_r(\mathbf{x}, \mathbf{n})$ *tends to a limit, say*

$$\lim_{r\to 0} \mathbf{t}_r(\mathbf{x}, \mathbf{n}) =: \mathbf{t}(\mathbf{x}, \mathbf{n}), \tag{III.4-16}$$

and $\mathbf{t}(\cdot, \mathbf{n})$ *is continuous.*

Thus the continuity of $\mathbf{t}(\cdot, \mathbf{n})$ is characterized directly in terms of the nature of the system of contact forces $\mathbf{f}_C$. Can the linearity expressed by CAUCHY's Fundamental Theorem be demonstrated on the basis of the bound (III.1-58) without further assumptions? We state without proof the following *generalization* of CAUCHY's Theorem, which, altogether dispensing with the assumption of continuity, refers directly to the existence of the average traction.

Theorem (GURTIN & MARTINS). *Again let* $\mathbf{f}_C$ *satisfy* (III.1-58). *At all places where the limit exists, let* $\mathbf{t}$ *be defined by* (16). *Then on* $\chi(\mathscr{B})$, *to within a set of*

[1] M. E. GURTIN & L. C. MARTINS, "Cauchy's theorem in classical physics," *Archive for Rational Mechanics and Analysis* **60** (1976): 305–324.

null volume, there is a tensor field $\mathbf{T}$ *such that*

$$\mathbf{t}(\mathbf{x}, \mathbf{n}) = \mathbf{T}(\mathbf{x})\mathbf{n} \qquad \text{(III.4-1)}_r$$

for every unit vector $\mathbf{n}$.

The normal component $\mathbf{n}\cdot\mathbf{t}(\mathbf{x}, \mathbf{n})$ of the traction corresponding to $\mathbf{n}$ at $\mathbf{x}$ is called the *normal traction* at $\mathbf{x}$ on the surfaces that have the outer unit normal $\mathbf{n}$ at $\mathbf{x}$, while the tangential component $\mathbf{t} - (\mathbf{n}\cdot\mathbf{t})\mathbf{n}$ is called the *shear traction* on those surfaces. From CAUCHY's Fundamental Theorem (1) we see that the normal traction is the normal component of $\mathbf{T}$:

$$\mathbf{n}\cdot\mathbf{t} = \mathbf{n}\cdot\mathbf{T}\mathbf{n}. \qquad \text{(III.4-17)}$$

If $\mathbf{n}$, $\mathbf{e}$, $\mathbf{f}$ is an orthonormal basis at $\mathbf{x}$, then

$$\mathbf{t} - (\mathbf{n}\cdot\mathbf{t})\mathbf{n} = (\mathbf{e}\cdot\mathbf{T}\mathbf{n})\mathbf{e} + (\mathbf{f}\cdot\mathbf{T}\mathbf{n})\mathbf{f}. \qquad \text{(III.4-18)}$$

The components $\mathbf{e}\cdot\mathbf{T}\mathbf{n}$ and $\mathbf{f}\cdot\mathbf{T}\mathbf{n}$ of the shear traction are called the *shear stresses* in the directions of $\mathbf{e}$ and $\mathbf{f}$, respectively, on a surface normal to $\mathbf{n}$ at $\mathbf{x}$.

Exercise III.4.2. Cauchy's Reciprocal Theorem:

$$\mathbf{n}\cdot\mathbf{t}(\cdot, \mathbf{m}) = \mathbf{m}\cdot\mathbf{t}(\cdot, \mathbf{n}) \quad \forall \mathbf{m}, \mathbf{n} \qquad \Leftrightarrow \qquad \mathbf{T} = \mathbf{T}^{\mathsf{T}}. \qquad \text{(III.4-19)}$$

In view of CAUCHY's Fundamental Theorem (1) we may express the principles of balance of linear and rotational momentum $(\text{III.1-8})_{1,2}$ as follows in terms of the stress tensor $\mathbf{T}$:

$$\int_{\chi(\mathscr{P},t)} \rho\ddot{\mathbf{x}}\, dV = \int_{\partial\chi(\mathscr{P},t)} \mathbf{T}\mathbf{n}\, dA + \int_{\chi(\mathscr{P},t)} \rho\mathbf{b}\, dV,$$

$$\text{(III.4-20)}$$

$$\int_{\chi(\mathscr{P},t)} (\mathbf{x} - \mathbf{x}_0) \wedge \rho\ddot{\mathbf{x}}\, dV = \int_{\partial\chi(\mathscr{P},t)} (\mathbf{x} - \mathbf{x}_0) \wedge (\mathbf{T}\mathbf{n})\, dA + \int_{\chi(\mathscr{P},t)} (\mathbf{x} - \mathbf{x}_0) \wedge \rho\mathbf{b}\, dV,$$

in which now we have restored the time t in the notation. These are the forms in which the two principles of momentum are commonly stated in continuum mechanics.

Since all forces are frame-indifferent, under change of frame from $\oint$ to $\oint^*$

$$\mathbf{t}^* = \mathbf{Q}\mathbf{t}, \qquad \mathbf{b}^* = \mathbf{Q}\mathbf{b}, \tag{III.4-21}$$

$\mathbf{Q}(t)$ being the orientation of $\oint^*$ relative to $\oint$ at the time t. As we have seen in Exercise I.11.4, the unit normal $\mathbf{n}$ to a surface in $\mathscr{E}$ is frame-indifferent:

$$\mathbf{n}^* = \mathbf{Q}\mathbf{n}. \tag{III.4-22}$$

CAUCHY's Fundamental Theorem (1) applies in $\oint^*$ as well as in $\oint$. Hence $\mathbf{T}$ transforms frame-indifferent vectors into frame-indifferent vectors. From the analysis leading to (I.9-10) we conclude that *the stress tensor is frame-indifferent*:

$$\mathbf{T}^* = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{T}}. \tag{III.4-23}$$

The traction $\mathbf{t}$ is called the *Cauchy traction*; the stress tensor $\mathbf{T}$, the *Cauchy stress*. Because it determines the contact force per unit area acting upon any surface in the present placement of a body, the CAUCHY stress is fundamental to our understanding of the effects of systems of forces in and on continuous bodies. For the solution of boundary-value problems and the discussion of constitutive properties of materials, the tractions referred to unit area in a reference placement are usually more convenient; in modern applications and studies of fundamental aspects of particular theories of materials, consequently, referential tractions occur much more frequently than the Cauchy traction $\mathbf{t}$. The referential tractions are determined by the first and second *Piola stresses*, which we shall introduce at our first need for them, namely in Volume 3.

The traction fields on the surfaces $\partial\chi(\mathscr{P}, t)$, which the traditional approach through (III.1-3) assumes to exist, and which in the approach through the modern theory of systems of forces GURTIN & WILLIAMS' Traction Theorem (III.1-46) proves to exist, CAUCHY's Fundamental Theorem (1) replaces by values of a linear function of $\mathbf{n}$, namely $\mathbf{T}(\mathbf{x}, t)\mathbf{n}$. Because the traction field $\mathbf{t}_{\partial\chi(\mathscr{P}, t)}$ is defined only upon $\partial\chi(\mathscr{P}, t)$, it does not lend itself easily to further development. In replacing it by the action of a tensor field $\mathbf{T}(\mathbf{x}, t)$ defined on all of $\chi(\mathscr{P}, t)$, CAUCHY's theorem suggests use of the divergence theorem. That use, as we shall see below in Section III.6, leads to the partial differential equation expressing locally the balance of linear momentum in continuum mechanics and requires $\mathbf{T}$ to be symmetric. Because the technique for using the divergence theorem applies to various sciences, we proceed to present it in fairly general terms.

5. The General Balance

The integral equation (III.4-20)$_1$, like other fundamental equations of classical physics, has the form of an *equation of balance*:

$$\frac{d}{dt}\int_{\mathscr{P}} \Psi \, dM = \int_{\chi(\mathscr{P},t)} \rho \dot{\Psi} \, dV = \int_{\partial\chi(\mathscr{P},t)} \mathcal{U}\mathbf{n} \, dA + \int_{\chi(\mathscr{P},t)} \rho \sigma \, dV. \tag{III.5-1}$$

Here $\dot{\Psi}$ and σ are tensor fields of the same order, defined over $\mathscr{P}$ and $\chi(\mathscr{P}, t)$, and $\mathcal{U}$ is a tensor field of order greater by 1 than that of $\dot{\Psi}$ and σ. One instance, trivial, has been encountered earlier, for the principle of conservation of mass (Section I.4) may be expressed in this form by the choices $\Psi = 1$, $\mathcal{U} = 0$, $\sigma = 0$. More generally, we interpret the equation of balance (1) as asserting that the rate of increase of the total Ψ in a part $\mathscr{P}$ of a body may be expressed as the sum of two effects: inflow through the boundary of the shape of $\mathscr{P}$ and growth at places within that shape. If a statement of the form (1) holds, $\mathcal{U}$ is called an *efflux* of Ψ and σ is called a *supply* of Ψ.

Of course, neither $\mathcal{U}$ nor σ is determined uniquely by (1). For example, any divergenceless tensor of the same order as $\mathcal{U}$ may be added to $\mathcal{U}$ without effecting any alteration of the other two terms in (1).

Equations of balance have two common applications: in regions where the fields occurring in them are smooth, and at certain kinds of discontinuities. We consider the former application here. When we come to treat elasticity (Volume 3), we shall illustrate the latter.

$\dot{\Psi}$, $\mathcal{U}$, and σ are functions of t and $\mathbf{x}$, integrable in the latter. We assume also that $\mathcal{U}$ is continuously differentiable in $\chi(\mathscr{P}, t)$ and continuous on $\partial\chi(\mathscr{P}, t)$. Then Green's transformation may be applied to regions with smooth boundaries:

$$\int_{\partial\chi(\mathscr{P},t)} \mathcal{U}\mathbf{n} \, dA = \int_{\chi(\mathscr{P},t)} \operatorname{div} \mathcal{U} \, dV, \tag{III.5-2}$$

and so the general balance (1) becomes

$$\int_{\chi(\mathscr{P},t)} (\rho\dot{\Psi} - \operatorname{div} \mathcal{U} - \rho\sigma) \, dV = 0. \tag{III.5-3}$$

Now a principle of balance is asserted to hold *for all bodies* and hence for all shapes of all parts of a given body. In particular, then, (3) follows for all parts whose shapes at the time t are sufficiently small spheres about the place $\mathbf{x}$ in the interior of $\chi(\mathscr{B}, t)$. If the integral of a continuous function f over every small sphere about $\mathbf{x}$ vanishes, then f itself vanishes at $\mathbf{x}$. Conversely, of course, if f vanishes at all points, so does its integral over every region contained in its domain. Thus we obtain the following

Theorem. *If the equation of balance* (1) *holds for all sufficiently small spheres in the fit region* $\chi(\mathscr{B}, t)$, *then at all interior points of* $\chi(\mathscr{B}, t)$ *where* $\rho\dot{\Psi} - \operatorname{div}\mathcal{U} - \rho\sigma$ *is continuous in its argument* $\mathbf{x}$ *the following differential equation holds*:

$$\rho\dot{\Psi} = \operatorname{div}\mathcal{U} + \rho\sigma. \tag{III.5-4}$$

Conversely, if (4) *holds at all interior points of a region, and if* $\mathcal{U}$ *is continuous in its argument* $\mathbf{x}$ *on the boundary of that region, the general balance* (1) *holds at the time t for the body occupying that region.*

The differential equation (4) is called the ***General Field Equation***. In the spatial description the substantial derivative $\dot{\Psi}$ is to be calculated by (II.6-3) or one of its generalizations.

6. Cauchy's Laws of Motion

We now obtain local forms of the principles of balance of linear and rotational momentum. First, with the choices $\Psi = \dot{\mathbf{x}}$, $\mathcal{U} = \mathbf{T}$, and $\sigma = \mathbf{b}$ we reduce (III.5-1) to $(\text{III.4-20})_1$, and so by (III.5-4) we obtain ***Cauchy's First Law of Motion***:

$$\rho\ddot{\mathbf{x}} = \operatorname{div}\mathbf{T} + \rho\mathbf{b}, \tag{III.6-1}$$

as a necessary and sufficient condition that linear momentum be balanced for all subbodies in the interior of a region where $\rho\ddot{\mathbf{x}}$, $\rho\mathbf{b}$, $\mathbf{T}$, and $\operatorname{div}\mathbf{T}$ are continuous.

The treatment of $(\text{III.4-20})_2$ is not quite so immediate.

Exercise III.6.1. For any tensor field $\mathbf{S}$ continuously differentiable in the fit region $\mathscr{R}$

$$\int_{\partial\mathscr{R}} (\mathbf{x} - \mathbf{x}_0) \wedge \mathbf{S}\mathbf{n}\, dA = \int_{\mathscr{R}} [(\mathbf{x} - \mathbf{x}_0) \wedge \operatorname{div}\mathbf{S} - 2\operatorname{skw}\mathbf{S}]\, dV. \tag{III.6-2}$$

If we substitute (2) into $(\text{III.4-20})_2$ and suppose that CAUCHY's First Law (1) holds, as a necessary and sufficient condition for the balance of rotational momentum we obtain

$$\int_{\chi(\mathscr{P},t)} \operatorname{skw} \mathbf{T}\, dV = \mathbf{0}. \tag{III.6-3}$$

Since $\mathbf{T}$ is continuous, $\operatorname{skw}\mathbf{T} = \mathbf{0}$; that is,

$$\mathbf{T}^{\mathsf{T}} = \mathbf{T}. \tag{III.6-4}$$

This is ***Cauchy's Second Law of Motion:*** *Under the hypotheses leading to the First Law, and on the assumption that the First Law holds, balance of rotational momentum and symmetry of the stress tensor are equivalent.*

CAUCHY's Laws assert that $\mathbf{T}^{\mathsf{T}} - \mathbf{T}$ and $\rho\ddot{\mathbf{x}} - \operatorname{div}\mathbf{T} - \rho\mathbf{b}$ vanish in an inertial frame. Since these quantities are frame-indifferent in the galilean class of any given frame, they vanish in one inertial frame if and only if they vanish in all inertial frames. *Cf.* Section I.13.

Since $\mathbf{T}$ is frame-indifferent, so is $\mathbf{T}^{\mathsf{T}} - \mathbf{T}$. Thus CAUCHY's Second Law is a frame-indifferent statement. That is, it holds for one frame if and only if it holds for all frames.

CAUCHY's First Law as stated is not frame-indifferent, but of course it can be modified so as to become so. Indeed, both $\operatorname{div}\mathbf{T}$ and $\rho\mathbf{b}$ are frame-indifferent, reflecting the fact that all forces and masses are frame-indifferent. The acceleration $\ddot{\mathbf{x}}$, in contradistinction, is not frame-indifferent (*cf.* Section I.9 and I.11). In accord with the frame-indifferent statement of the Axioms of Inertia in Section I.13, CAUCHY's First Law in a general rigid frame $\oint^*$ assumes the form

$$\rho\mathbf{a}_{\oint} = \operatorname{div}\mathbf{T}^* + \rho\mathbf{b}^*; \tag{III.6-5}$$

here

$$\mathbf{a}_{\oint} = \ddot{\mathbf{x}}^* - \ddot{\mathbf{x}}_0^* - 2\mathbf{A}(\dot{\mathbf{x}}^* - \dot{\mathbf{x}}_0^*) - (\dot{\mathbf{A}} - \mathbf{A}^2)(\mathbf{x}^* - \mathbf{x}_0^*), \qquad \text{(II.4-7)}_{\mathrm{r}}$$

$$\mathbf{T}^* = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{T}}, \qquad \text{(III.4-23)}_{\mathrm{r}}$$

$$\mathbf{b}^* = \mathbf{Q}\mathbf{b}, \qquad \text{(III.4-21)}_{2\mathrm{r}}$$

in which $\mathbf{Q}$, $\mathbf{x}_0$, and $\mathbf{x}_0^*$ are the quantities defining the change of frame (II.14-2) when applied to an inertial frame $\oint$, while $\dot{\mathbf{x}}^*$ and $\ddot{\mathbf{x}}^*$ are the fields of velocity and acceleration relative to $\oint^*$, and $\mathbf{A}$ is the spin of $\oint$ with respect to $\oint^*$. The student may refer to (I.9-15), (I.13-9), (I.11-3), and (I.13-13).

Some authors prefer to transfer the terms following the minus signs in (II.4-7) to the right-hand side of CAUCHY's First Law and call them "forces" or "apparent forces." According to their point of view, CAUCHY's First Law in the form (1) is valid in all frames, but the body force $\mathbf{b}$ must be augmented for forces "due to the motion" of the observer's frame with respect to an inertial frame. Since just the same equations result, this point of view is legitimate, but it is scarcely felicitous, since it obscures the basic nature of frames and of systems of forces.

In this book, except in passages where requirements of frame-indifference are developed, we shall always presume that the frame used is inertial.

CAUCHY's two Laws of Motion have been derived by arguments that apply only to interior points. On the boundary of a body not in contact with any other body, those arguments have no force. We shall assume that the stress field $\mathbf{T}(\mathbf{x}, t)$ is continuous on $\mathrm{clo}\,\chi(\mathscr{B}, t)$ at the time t. Thus the second law (4) holds on the boundary. At boundary points we shall assume that the First Law (1) holds in the sense of an interior limit.

In some recent theories of continuum mechanics stress tensors that are not symmetric appear. In these theories either there are torques that are not moments of forces, or the density of rotational momentum is not simply the moment of the density of linear momentum, or both. We shall not have need of these more general ideas in this course, for which the classical laws of CAUCHY will suffice as local statements of the principles of linear and rotational momentum.

Because the stress tensor is symmetric, it has a spectral decomposition:

$$\mathbf{T} = \sum_{k=1}^{3} t_k \mathbf{e}_k \otimes \mathbf{e}_k \,, \tag{III.6-6}$$

$\{\mathbf{e}_k\}$ being a suitably selected orthonormal basis at $(\mathbf{x}, t)$. The numbers t_k are called the *principal stresses*, and the directions of the $\mathbf{e}_k$ are called the *principal axes of stress*. If the three principal stresses are distinct, the principal axes of stress are unique; otherwise, $\mathbf{T}$ has infinitely many triads of principal axes. In general, there are no surfaces everywhere normal to the fields $\mathbf{e}_k$ (*cf.* Section App. IIC.5), but of course there are infinitely many surfaces normal to each $\mathbf{e}_k$ at any one given place $\mathbf{x}$ at the time t. On such a surface at the place $\mathbf{x}$ and the time t, the principal stress t_k is the normal traction, and all shear stresses vanish.

If $\mathbf{e}_1$ is a unit vector, a stress of the form $t_1 \mathbf{e}_1 \otimes \mathbf{e}_1$ is called a *uniaxial tension* of *amount* t_1 in the direction of $\mathbf{e}_1$. The spectral decomposition (6) asserts that the stress at $(\mathbf{x}, t)$ may be regarded as the sum of uniaxial tensions along the principal axes of stress.

Exercise III.6.2. The stress is said to be *hydrostatic* at $\mathbf{x}$ if, for all $\mathbf{n}$

$$|\mathbf{t}(\mathbf{x}, \mathbf{n})| \text{ is independent of } \mathbf{n}, \qquad \text{and} \qquad \mathbf{n}\cdot\mathbf{t}(\mathbf{x}, \mathbf{n}) \text{ is of one sign,} \tag{III.6-7}$$

and

$$\mathbf{n}\cdot\mathbf{t}(\mathbf{x},\mathbf{n}) \text{ is independent of } \mathbf{n}, \tag{III.6-8}$$

and

$$\mathbf{t}(\mathbf{x},\mathbf{n}) \text{ is parallel to } \mathbf{n}. \tag{III.6-9}$$

In view of CAUCHY's Second Law, any one of these statements implies the others and also

$$\mathbf{T} = -p(\mathbf{x})\mathbf{1}. \tag{III.6-10}$$

What relations can be established among these statements if $\mathbf{T}$ is not symmetric? A particular time t, not indicated in the notation, is understood throughout the argument.

The working W of a system of forces has been defined by (I.8-7). In an inertial frame W is expressed in terms of the kinetic energy K and the power P by (I.14-1). The power (I.14-2) in continuum mechanics is the rate of working of the contact force and the body force:

$$P = \int_{\partial\chi(\mathscr{P},t)} \dot{\mathbf{x}}\cdot\mathbf{t}\,dA + \int_{\chi(\mathscr{P},t)} \rho\dot{\mathbf{x}}\cdot\mathbf{b}\,dV. \tag{III.6-11}$$

Exercise III.6.3 (PIOLA's power theorem). If $\mathbf{t}$ and $\mathbf{b}$ are regarded as given fields in (11), then in order that $P = 0$ in every rigid motion of $\mathscr{P}$, it is necessary and sufficient that the resultant force and resultant torque applied to $\mathscr{P}$ shall vanish. Equivalently, in an inertial frame the linear momentum and rotational momentum of $\mathscr{P}$ are both constant.

Exercise III.6.4 (STOKES's power formula).

$$W = \int_{\chi(\mathscr{P},t)} w\,dV, \tag{III.6-12}$$

in which w, which is called the *stress power*, is given by

$$w = \mathbf{T}\cdot\mathbf{G} = \mathbf{T}\cdot\mathbf{D}. \tag{III.6-13}$$

Hence $w = 0$ in a rigid motion, and also in an isochoric motion subject to hydrostatic stress. Also w is a frame-indifferent scalar.

Exercise III.6.5 (Balance of mechanical energy). If

$$P_C := \int_{\partial\chi(\mathscr{P},t)} \dot{\mathbf{x}}\cdot\mathbf{t}\,dA, \qquad \text{(III.6-14)}$$

and if the body force is lamellar (*cf.* (III.1-7)), the corresponding *potential energy* $U(\mathscr{P}, t)$ is defined as follows:

$$U := \int_{\chi(\mathscr{P},t)} \rho\varpi\,dV. \qquad \text{(III.6-15)}$$

Then

$$P = P_C - \dot{U} + \int_{\chi(\mathscr{P},t)} \rho\varpi'\,dV, \qquad \text{(III.6-16)}$$

and hence

$$W = P_C - (\dot{K} + \dot{U}) + \int_{\chi(\mathscr{P},t)} \rho\varpi'\,dV. \qquad \text{(III.6-17)}$$

The definition of "mechanically perfect" in Section I.14 shows that in a mechanically perfect motion of a body subject to conservative body force and to boundary tractions normal to the velocities at the points where they act,

$$K + U = \text{const.}, \qquad \text{(III.6-18)}$$

provided, of course, that a steady ϖ be selected.

The foregoing exercise asserts a theorem of *conservation of purely mechanical energy*. It provides motivation of our having called "conservative" a body force that is steady as well as lamellar. Its hypotheses hold in some cases governed by some classical theories of continua, but not very generally.

Exercise III.6.6 (Balance of internal energy and working) (Fourier, Stokes, Maxwell, Kirchhoff, C. Neumann). Looking back at the Balance of Energy, expressed by (I.15-4), suppose that E has a density ϵ with respect to mass, that w as given by (III.16-13) is the density of the net working W with respect to volume, and that Q has both a superficial density q and a density s with respect to mass:

$$Q = \int_{\partial\chi(\mathscr{P},t)} q\,dA + \int_{\chi(\mathscr{P},t)} \rho s\,dV. \qquad \text{(III.6-19)}$$

Then there is a vector field $\mathbf{h}$ such that

$$q = \mathbf{h} \cdot \mathbf{n}, \tag{III.6-20}$$

and

$$\rho\dot{\epsilon} = w + \operatorname{div} \mathbf{h} + \rho s. \tag{III.6-21}$$

It is bruited that the differential equations expressing balance of linear and rotational momenta follow generally from (21) by applying a superposed rigid motion. That belief seems to have grown from a paper by A. E. GREEN & R. S. RIVLIN, "On Cauchy's equations of motion," *Zeitschrift für Angewandte Mathematik und Physik* **15**(1964): 290–291. If it seems strange that theorems about momenta alone can emerge from an assumption about internal energy, it is strange. An expert in continuum mechanics has written of this, "There are so many assumptions that the main argument becomes trivial." In fact, internal energy is introduced only so as to subtract it out at the first step, and no use of (21) is necessary to the end desired, which is the outcome of PIOLA's power theorem (Exercise III.6.3).

7. Mean Values and Lower Bounds for the Stress Field

SIGNORINI remarked that since

$$\operatorname{div}(\Psi \mathbf{T}) = \mathbf{T} \operatorname{grad} \Psi + \Psi \operatorname{div} \mathbf{T}, \tag{III.7-1}$$

CAUCHY's First Law (III.6-1) yields

$$\mathbf{T} \operatorname{grad} \Psi = \operatorname{div}(\Psi \mathbf{T}) + \rho \Psi (\mathbf{b} - \ddot{\mathbf{x}}), \tag{III.7-2}$$

and so if we integrate this identity over the present shape $\chi(\mathscr{B})$ of a body and then use the divergence theorem, we obtain

$$\int_{\chi(\mathscr{B})} \mathbf{T} \operatorname{grad} \Psi \, dV = \int_{\partial\chi(\mathscr{B})} \Psi \mathbf{T}\mathbf{n} \, dA + \int_{\chi(\mathscr{B})} \rho \Psi (\mathbf{b} - \ddot{\mathbf{x}}) \, dV. \tag{III.7-3}$$

The left-hand side is proportional to a certain weighted mean of the stress field over the shape $\chi(\mathscr{B})$. It is determined by the value of $\mathbf{T}$ upon the boundary $\partial\chi(\mathscr{B})$ and by the corresponding mean of $\rho \Psi (\mathbf{b} - \ddot{\mathbf{x}})$. The conclusion seems to be of interest mainly for a body at rest, and so $\ddot{\mathbf{x}} = \mathbf{0}$. Then it expresses the *mean values of the stress field in terms of the load alone*: $\mathbf{T}\mathbf{n}$ upon the boundary, $\rho\mathbf{b}$ in the interior.

An example due to CHREE and FINGER makes the point clear: If we take for Ψ the position vector $\mathbf{p}$, and if

$$\mathbf{L} := \frac{1}{V(\chi(\mathscr{B}))}\left[\int_{\partial\chi(\mathscr{B})} \mathbf{p}\otimes\mathbf{Tn}\,dA + \int_{\chi(\mathscr{B})} \rho\mathbf{p}\otimes\mathbf{b}\,dV\right], \qquad \text{(III.7-4)}$$

then the mean value $\bar{\mathbf{T}}$ of the stress field in $\chi(\mathscr{B})$ is given by

$$\bar{\mathbf{T}} = \mathbf{L}. \qquad \text{(III.7-5)}$$

The skew part of this equation merely reaffirms Cauchy's Second Law, but the symmetric part has some interesting applications.

First let us suppose that $\chi(\mathscr{B})$ is the region bounded without by a closed surface $\mathscr{S}_o$ and within by a closed surface $\mathscr{S}_i$, which lies wholly inside the region bounded by $\mathscr{S}_o$, so that the region interior to $\mathscr{S}_i$ bounds a cavity $\mathscr{C}$ of positive volume, and let $\mathscr{S}_o$ and $\mathscr{S}_i$ be subject to uniform hydrostatic pressures p_o and p_i (Exercise III.6.2). We suppose also that $\mathbf{b} = \mathbf{0}$, and we write $V(\mathscr{C})$ for the volume of the cavity. Then $\mathbf{L}$ is easy to evaluate, and (5) yields

$$-\bar{\mathbf{T}} = \left[p_o + \frac{V(\mathscr{C})}{V(\chi(\mathscr{B}))}(p_o - p_i)\right]\mathbf{1}. \qquad \text{(III.7-6)}$$

Thus *hydrostatic loading gives rise to a stress field that is hydrostatic in mean*. If $p_o = p_i$, the mean stress is the applied pressure. If $p_o > p_i$, the mean pressure always exceeds p_o.

Next we consider a body in a shape $\chi(\mathscr{B})$ subject to surface tractions alone, all of which are parallel to a certain vector $\mathbf{e}$. If $\mathbf{f}$ is any vector normal to $\mathbf{e}$, then $\mathbf{Tf} = \mathbf{0}$ on $\partial\chi(\mathscr{B})$, and so from (4) we obtain

$$\mathbf{Lf} = \frac{1}{V(\chi(\mathscr{B}))}\int_{\partial\chi(\mathscr{B})} (\mathbf{Tn}\cdot\mathbf{f})\mathbf{p}\,dA = \frac{1}{V(\chi(\mathscr{B}))}\int_{\partial\chi(\mathscr{B})} (\mathbf{n}\cdot\mathbf{Tf})\mathbf{p}\,dA = \mathbf{0}. \qquad \text{(III.7-7)}$$

Thus (5) yields

$$\bar{\mathbf{T}}\mathbf{f} = \mathbf{0}: \qquad \text{(III.7-8)}$$

The stress field corresponding to uniaxial surface load is uniaxial in mean.

Exercise III.7.1. Let $\mathscr{A}$ be a portion of a plane normal to $\mathbf{e}$, and upon $\mathscr{A}$ let $\mathbf{Te}$ be a constant multiple of $\mathbf{e}$. If $F\mathbf{e}$ denotes the resultant contact load upon $\mathscr{A}$, and if the centroid of $\mathscr{A}$ is at $\mathbf{p}_0(\mathscr{A})$, then $(\int_{\mathscr{A}} \mathbf{p} \otimes \mathbf{Te}\,dA)\mathbf{e} = F\mathbf{p}_0(\mathscr{A})$. Hence if the shape $\chi(\mathscr{B})$ of a body at rest has two plane, parallel faces normal to $\mathbf{e}$, upon each of which a uniform tensile load is applied, and if the body is otherwise free, then the mean tensile stress is given by

$$\mathbf{e}\cdot\bar{\mathbf{T}}\mathbf{e} = \frac{Fd}{V(\chi(\mathscr{B}))}, \tag{III.7-9}$$

F being the resultant tensile force applied to either face, and d being the distance between the plane faces.

Numerous other relations of this kind were obtained by SIGNORINI and his school. They studied also *moments of stress* $\overline{\mathbf{p}\otimes\mathbf{p}\otimes\cdots\otimes\mathbf{p}\otimes\mathbf{T}}$ and showed that many components of those moments can be determined from the moments of the load.

Exercise III. 7.2 (SIGNORINI). Let L be the third-order tensor whose components L_{kmq} are defined as follows in terms of the components p_s of the position vector:

$$L_{kpq} = \tfrac{1}{2}(M_{qkp} + M_{pkq} - M_{pqk}),$$
$$M_{rst} := \frac{1}{V(\mathscr{S})}\left[\int_{\partial\mathscr{S}} p_r p_s T_{tq} n_q\,dA + \int_{\mathscr{S}} \rho p_r p_s (b_t - \ddot{x}_t)\,dV\right]. \tag{III.7-10}$$

Then

$$\overline{\mathbf{p}\otimes\mathbf{T}} = \mathsf{L}. \tag{III.7-11}$$

SIGNORINI showed how to obtain lower bounds for the mean stresses in terms of other, more accessible means such as $\mathbf{L}$ and L. His method was extended by GRIOLI. Their conclusions are most easily expressed if we regard $\mathbf{T}$ as a 6-dimensional vector field, which we shall do for the remainder of this section.

Let the functions F_{a}, $\mathrm{a} = 1, 2, \ldots, m$, be orthonormal in mean over the present shape $\chi(\mathscr{B})$ of a body:

$$\overline{F_{\mathrm{a}}F_{\mathrm{b}}} = \delta_{\mathrm{ab}}\,. \tag{III.7-12}$$

Let $\mathbf{K}$ be any symmetric, non-negative tensor over the space of 6-dimensional vectors, and let $\mathbf{C}_{\mathrm{a}}$, $\mathrm{a} = 1, 2, \ldots, m$, be vectors in that space. Then

$$0 \leqq \mathbf{K}\cdot\overline{\left(\mathbf{T} - \sum_{\mathrm{a}=1}^{m} F_{\mathrm{a}}\mathbf{C}_{\mathrm{a}}\right)\otimes\left(\mathbf{T} - \sum_{\mathrm{b}=1}^{m} F_{\mathrm{b}}\mathbf{C}_{\mathrm{b}}\right)}\,. \tag{III.7-13}$$

Calculating the mean value of this expression, we obtain

$$0 \leqq \mathbf{K}\cdot\overline{\mathbf{T}\otimes\mathbf{T}} + \mathbf{K}\cdot\sum_{\mathrm{a}=1}^{m}\mathbf{C}_{\mathrm{a}}\otimes(\mathbf{C}_{\mathrm{a}} - 2\overline{F_{\mathrm{a}}\mathbf{T}}). \tag{III.7-14}$$

The vectors $\mathbf{C}_{\mathrm{a}}$ have been arbitrary so far. We now choose them as follows:

$$\mathbf{C}_{\mathrm{a}} = \overline{F_{\mathrm{a}}\mathbf{T}}. \tag{III.7-15}$$

Then (14) reduces to

$$\mathbf{K}\cdot\overline{\mathbf{T}\otimes\mathbf{T}} \geqq \mathbf{K}\cdot\sum_{\mathrm{a}=1}^{m}\overline{F_{\mathrm{a}}\mathbf{T}}\otimes\overline{F_{\mathrm{a}}\mathbf{T}}. \tag{III.7-16}$$

The non-negative tensor $\mathbf{K}$ and the orthonormal functions F_{a} remain ours to choose. Thus (16) provides infinitely many lower bounds for the components of $\overline{\mathbf{T}\otimes\mathbf{T}}$, bounds which depend upon the shape of the body and the loads applied to it. One general conclusion is worth noting before we descend to particular applications: If $\mathbf{K}$ is positive rather than merely non-negative, (13) and (15) show that equality is achieved in (16) if and only if

$$\mathbf{T} = \sum_{\mathrm{a}=1}^{m} F_{\mathrm{a}}\overline{F_{\mathrm{a}}\mathbf{T}}. \tag{III.7-17}$$

Therefore, among all stress fields that have in common the m means $\overline{F_{\mathrm{a}}\mathbf{T}}$ for a given set of orthonormal functions F_{a}, such fields as satisfy (17) give a minimum value to $\mathbf{K}\cdot\overline{\mathbf{T}\otimes\mathbf{T}}$ for every choice of the positive tensor $\mathbf{K}$.

For example, we may take for $\mathbf{K}$ the tensor whose components with respect to a particular basis are all 0 except for K_{kk}, which has the value 1. Since $\overline{T_k^2} \leqq \max T_k^2$, from (16) we see that

$$\max T_k^2 \geqq \overline{T_k^2} \geqq \sum_{\mathrm{a}=1}^{m}(\overline{F_{\mathrm{a}}T_k})^2, \qquad k = 1, 2, \ldots, 6. \tag{III.7-18}$$

Thus lower bounds for the magnitude of every component of $\mathbf{T}$ with respect to a constant basis field have been obtained.

The bounds we have demonstrated are expressed in terms of the means $\overline{F_{\mathrm{a}}\mathbf{T}}$ and thus might seem more difficult to calculate than such direct means as $\overline{\mathbf{T}\otimes\mathbf{T}}$. On the contrary, the conclusions reached earlier in this chapter show that for suitable choices of the functions F_a the means $\overline{F_a\mathbf{T}}$ can be evaluated in terms of the applied loads. We have obtained two theorems of this kind, namely, (5) and (11). We shall see now that they may be used so as to evaluate the right-hand side of (16) in terms of the shape of the body and the loads acting upon it.

To do so, we are guided by the properties of the center of mass and the Euler tensor of a body (Sections I.8, I.10). These assure us that by choice of a system of cartesian co-ordinates we can satisfy the relations $\int_{\chi(\mathscr{B})} z_k\, dV = 0$ and $\int_{\chi(\mathscr{B})} z_p z_q\, dV = 0$ if $p \neq q$. We could describe these co-ordinates as having their origin at the centroid of $\chi(\mathscr{B})$ and their axes parallel to principal axes of inertia of a body of uniform density having the shape $\chi(\mathscr{B})$. To express the outcome, it is convenient to write A_k for the reciprocal of the square root of the k^{th} axial momentoid of inertia:

$$A_k := 1/\sqrt{\overline{z_k^2}}, \qquad k = 1, 2, 3. \tag{III.7-19}$$

Then the following 4 functions F_a are orthonormal in mean over $\mathscr{S}$:

$$F_0 := 1, \qquad F_k := A_k z_k\,, \qquad k = 1, 2, 3. \tag{III.7-20}$$

The relation (5) may be expressed as $\overline{F_0\mathbf{T}} = \mathbf{L}$. In the present notation, which regards $\mathbf{T}$ as a 6-dimensional vector field, the third-order tensor L defined in terms of the applied loads by (10) becomes a triple of vectors $\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$, and we may express (11) in the form $\overline{z_k\mathbf{T}} = \mathbf{L}_k$, $k = 1, 2, 3$. Thus, finally, if

$$N_a := \begin{cases} \mathbf{L} & \text{if} \quad a = 0, \\ A_k\mathbf{L}_k & \text{if} \quad a = 1, 2, 3, \end{cases} \tag{III.7-21}$$

by using in (16) the particular set of orthonormal functions (20) we obtain an elegant inequality discovered by SIGNORINI:

$$\mathbf{K}\cdot\overline{\mathbf{T}\otimes\mathbf{T}} \geqq \sum_{a=0}^{3} \mathbf{N}_a\cdot\mathbf{K}\mathbf{N}_a\,. \tag{III.7-22}$$

The 4 vectors $\mathbf{N}_a$ on the right-hand side are determined by (5) and (11) from the shape of the body and the loads acting upon it.

The estimate (22) may be rendered more specific by considering special loadings upon special shapes. Perhaps the most interesting application is the most trivial. Various older theories of plasticity lay down an axiom that for an appropriate choice of $\mathbf{K}$ there is a constant C such that

$$\mathbf{T}\cdot\mathbf{K}\mathbf{T} \leqq C. \tag{III.7-23}$$

SIGNORINI's inequality (22) shows that such an axiom cannot hold unless the loads and the shape are such that

$$\sum_{a=0}^{4} \mathbf{N}_a \cdot \mathbf{K}\mathbf{N}_a \leqq C. \tag{III.7-24}$$

Thus (23) cannot serve as a general assumption in any theory designed to represent the behavior of bodies of arbitrary shape subject to arbitrary loads.

8. Load. Boundary Condition of Traction

The applied force $\mathbf{f}^a$ and applied torque $\mathbf{F}^a$ acting upon the shape $\chi(\mathscr{P}, t)$ appear on the right-hand sides of (III.4-20). When ρ is regarded as given, both of them are determined by the field $\mathbf{b}$ upon $\chi(\mathscr{P}, t)$ and the field $\mathbf{Tn}$ upon $\partial\chi(\mathscr{P}, t)$. These two fields by integration deliver the *load* on $\mathscr{P}$ in its shape $\chi(\mathscr{P}, t)$. In many particular problems of continuum mechanics the load on some given shape is prescribed. The condition

$$\mathbf{Tn} = \mathbf{t} \qquad \text{upon} \quad \partial\chi(\mathscr{B}, t), \tag{III.8-1}$$

$\mathbf{t}$ being a given function of $\mathbf{x}$ and t, is the boundary condition of *prescribed traction*. Such a condition supplements the kinematical boundary conditions mentioned in Section II.6.

For example, if p is a prescribed scalar field, and if

$$\mathbf{Tn} = -p\mathbf{n} \qquad \text{upon} \quad \partial\chi(\mathscr{B}, t), \tag{III.8-2}$$

the body is subject to the *pressure* p upon its boundary. Of course this condition does not require the stress field throughout $\chi(\mathscr{B}, t)$ to be hydrostatic [*cf.* (III.7-6)].

When the field p in (2) has a constant value on $\partial\chi(\mathscr{B}, t)$, the body is subject to *uniform pressure*. This boundary condition often is regarded as a model for the contact load exerted by a quiet body of gas upon a body submerged in it. A field of pressure proportional to the distance from a fixed plane provides a common model for the contact load exerted by a quiet body of *heavy liquid* of uniform density upon a body partly or wholly submerged in it. The fixed plane represents the horizontal upper surface of the body of liquid.

Exercise III.8.1 (ARCHIMEDES, STEVIN, EULER). Let a body whose shape is a fit region be submerged partly or wholly in a heavy liquid of uniform density. The centroid of the part of the shape below the horizontal upper surface of the liquid is called the *center of buoyancy*. The line connecting the center of buoyancy to the center of mass is called the *hydrostatic axis*. If the part of the boundary above the upper surface is free of traction, then *the resultant contact load on the body is equipollent to a force applied at the center of buoyancy directed upwards, of magnitude equal to the weight of the fluid displaced* by the body. If the body force applied to the body results from the same gravitational field as that which acts on the liquid, then the body is isolated, as the term is defined by (I.13-12), if and only if

1. the hydrostatic axis is vertical, and
2. the weight of the displaced fluid equals the weight of the body.

An important special kind of pressure is that exerted by *surface tension*. In contrast with the other special cases just presented, this one reflects the nature of the body as well as its shape and the nature of its surroundings:

$$\begin{aligned} p &= 2\sigma k, \\ k &= \text{mean curvature of } \partial\chi(\mathscr{B}, t), \\ \sigma &= \text{const.} \end{aligned} \tag{III.8-3}$$

The constant σ, called the *coefficient of surface tension*, is adjustable so as to model, more or less, the nature of the parts of $\mathscr{B}$ adjacent to the inside of $\partial\chi(\mathscr{B}, t)$ and the parts of the surroundings adjacent to the outside of $\partial\chi(\mathscr{B}, t)$. It is the first example of a *constitutive modulus* in this book. Others, referring to the material of which $\mathscr{B}$ is composed, will appear in later chapters.

Exercise III.8.2. The contact load of surface tension upon a shape whose boundary has a continuous unit normal field is null. Thus *a body loaded by surface tension alone is isolated.*

A body $\mathscr{B}$ subjected to null loads, that is,

$$\mathbf{b} \equiv \mathbf{0} \quad \text{in } \chi(\mathscr{B}, t), \qquad \mathbf{Tn} \equiv \mathbf{0} \quad \text{upon } \partial\chi(\mathscr{B}, t), \tag{III.8-4}$$

is *free*. Of course a free body is isolated, but an isolated body need not be free. *Cf.* Section I.13.

9. Motion of a Free Body

The theory presented thus far is so general as to impose little restriction upon the motions of a body. Theories of particular materials, developed in the succeeding chapter and applied throughout the remainder of this book, impose systematic restrictions upon bodies by requiring that they consist of particular *materials*. In addition to these *constitutive* restrictions, or in some cases instead of them, kinematic assumptions are sometimes imposed directly, and these may severely limit the kind of motion possible. For example, rigid motions have been discussed in Section I.10, motions that preserve circulation in Section II.13, and some effects of kinematic boundary conditions have been demonstrated in Section II.11. The dynamic boundary conditions discussed in the preceding section also have their effects, as we shall see now in what appears to be the simplest case, namely, the motion of a free body.

In the mechanics of systems of mass-points the motion of a free body is trivial. In contrast, a free rigid body may rotate about its center of mass in a most complicated way. When we come to deformable bodies, the problem of free motion becomes indeterminate. Nevertheless, something definite can be learned about it. For example, by putting a position vector $\mathbf{p}$ for Ψ in (III.7-3) we see that

$$\bar{\mathbf{T}} = -\overline{\rho \mathbf{p} \otimes \ddot{\mathbf{x}}}, \qquad \text{(III.9-1)}$$

by which the mean stress at each time restricts, or is restricted by, the acceleration field. In particular, if we denote by p the arithmetic mean of the normal tractions, $p := -\frac{1}{3}\operatorname{tr}\mathbf{T}$, then from (1) it follows that

$$\bar{p} = \frac{1}{3}\overline{\rho \mathbf{p} \cdot \ddot{\mathbf{x}}}. \qquad \text{(III.9-2)}$$

Day, acknowledging influence of Sundman's classic work on the three-body problem of analytical dynamics, has exploited (2) so as to prove a theorem relating $\bar{p}$ to the diameter of the shape of a body $\mathscr{B}$ supposed free when $t \geqq 0$; in particular, the rotational momentum $\mathbf{M}$ is an assigned constant. By the diameter $d(t)$ of $\chi(\mathscr{B}, t)$ is meant the supremum of the distances between its points. The position vector $\mathbf{p}(t)$ will henceforth be taken with respect to the center of mass $\bar{p}(t)$ of $\chi(\mathscr{B}, t)$; thus $\mathbf{p} := \mathbf{x} - \bar{\mathbf{p}}$, $\dot{\mathbf{p}} = \dot{\mathbf{x}} - \dot{\bar{\mathbf{p}}}$, $\ddot{\bar{\mathbf{p}}} = \mathbf{0}$.

Theorem. *Let* $\mathbf{M}$ *be the rotational momentum of* $\mathscr{B}$ *with respect to its center of mass. If* $\mathbf{M} \neq \mathbf{0}$, *then either* $d(t) \to \infty$ *or there is a positive time* t^* *such that* $\bar{p}(t^*) < 0$.

In particular, in the interior of a free body that remains within a bounded part of space for all time, a region in which at least one of the principal stresses is a tension must develop unless the rotational momentum of the body is null. The proof follows now.

Exercise III.9.1. If

$$P(t) := \int_{\chi(\mathscr{B},t)} \rho|\mathbf{p}|^2\,dV, \tag{III.9-3}$$

then

$$P \leqq \tfrac{1}{2}Md^2. \tag{III.9-4}$$

Also

$$\begin{aligned}|\mathbf{M}|^2 &= \left|\int_{\chi(\mathscr{B},t)} \rho\mathbf{p}\wedge\dot{\mathbf{p}}\,dV\right|^2 \leqq \left(\int_{\chi(\mathscr{B},t)} \rho|\mathbf{p}\wedge\dot{\mathbf{p}}|\,dV\right)^2, \\ &\leqq P\int_{\chi(\mathscr{B},t)} \rho\frac{|\mathbf{p}\wedge\dot{\mathbf{p}}|^2}{|\mathbf{p}|^2}\,dV;\end{aligned} \tag{III.9-5}$$

the last step is a consequence of the Cauchy-Schwarz inequality. But for any vectors **a** and **b**

$$|\mathbf{a}\wedge\mathbf{b}|^2 = 2|\mathbf{a}|^2|\mathbf{b}|^2 - 2(\mathbf{a}\cdot\mathbf{b})^2 \leqq 2|\mathbf{a}|^2|\mathbf{b}|^2. \tag{III.9-6}$$

Thus it follows from (5) that

$$|\mathbf{M}|^2 \leqq 4PK, \tag{III.9-7}$$

in which K denotes the kinetic energy of $\chi(\mathscr{B}, t)$ with respect to the body's center of mass.

We now take the substantial derivative of (3) and use (2) to obtain

$$\ddot{P} = 4K + 6\bar{p}V, \tag{III.9-8}$$

V being the volume of $\chi(\mathscr{B}, t)$. If $\bar{p}(t) \geqq 0$ when $t \geqq 0$, then from (8) it follows that $\ddot{P} \geqq 4K$. Thus (7) yields

$$2P\ddot{P} \geqq 2|\mathbf{M}|^2. \tag{III.9-9}$$

Hence

$$(P^2)^{\cdot\cdot} = 2P\ddot{P} + 2\dot{P}^2 \geqq 2|\mathbf{M}|^2 > 0. \tag{III.9-10}$$

Integrating this inequality twice shows that $P(t) \to \infty$. A glance at (4) suffices to prove that $d(t) \to \infty$. $\triangle$

Some materials are regarded as being unable to support tension. DAY's theorem shows that a freely spinning body of such material will ultimately fly asunder unless its diameter tends to ∞. For further development of this idea, see below, Section IV.19.

General References

Sections 199–238 of CFT. The treatment of stress is exhaustive for 1960 but now partly superseded. Various difficulties result from an attempt to confine attention to resultant forces without bringing into the open the systems of forces giving rise to those resultants.

W. NOLL, "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* (1957), Amsterdam, North-Holland Publ., 1959, pp. 266–281. Reprinted in NOLL's *The Foundations of Mechanics and Thermodynamics*, New York, Heidelberg, and Berlin, Springer-Verlag, 1974.

Chapter II of IRE (C.-C. WANG & C. TRUESDELL, "Introduction to Rational Elasticity," Leyden, Noordhoff, 1973).

Chapter IV

Constitutive Relations

If geometry is to serve as a model for the treatment of axioms of physics, we shall try first to cover with a few axioms as large a class of physical phenomena as possible, and then by adjoining further axioms, one after another, to arrive at the more special theories .... Also the mathematician will have to take account not only of those theories that come near to reality but also, as in geometry, of all logically possible theories, and he must always be careful to obtain a complete survey of the consequences implied by the system of axioms laid down.

Further, it is the task of the mathematician, complementing the physicist's way of looking at things, in each instance to examine exactly whether the further axioms be compatible with the foregoing ones. The physicist regards himself often as being compelled by the results of his experiments to make new hypotheses from time to time *during* the development of his theory; ... he calls only upon those experiments or a certain physical intuition, a practice which in the rigorously logical erection of a theory is not admissible.

HILBERT
in regard to his Sixth Problem,
"Mathematical Treatment of the
Axioms of Physics" (1900)
Mathematische Probleme
Archiv für Mathematik und Physik
(3)**1** (1901), 44–63, 213–237

1. Dynamic Processes

A motion χ assigns to a body $\mathscr{B}$ a shape $\chi(\mathscr{B}, t)$ at the time t. At a point $\mathbf{x}$ of that shape, the stress tensor $\mathbf{T}(\mathbf{x}, t)$ determines the traction on every surface that is the boundary of the shape of some interior part $\mathscr{P}$ of $\mathscr{B}$. In this sense the stress field $\mathbf{T}$ is assigned to the body in its motion. The ordered pair $(\chi, \mathbf{T})$ is called a *dynamic process* for $\mathscr{B}$ if χ and $\mathbf{T}$ are related in such a way as to satisfy the principles of balance of linear and rotational momentum.

At interior points of regions where χ and $\mathbf{T}$ are sufficiently smooth, the principles of linear and rotational momentum are expressed by CAUCHY's Laws of Motion. The second law (III.6-4) requires that the stress be symmetric. The first law (III.6-1) relates the stress field to the acceleration $\ddot{\mathbf{x}}$ in an inertial frame, provided the body force field $\mathbf{b}$ be known. We regard $\mathbf{b}$, which represents the action on $\mathscr{B}$ of unspecified bodies exterior to $\mathscr{B}$, as assignable. While in practice only a few special body forces like that of gravity are available in laboratories or daily life—indeed, typically in specific problems of continuum mechanics we suppose that $\mathbf{b} = \mathbf{0}$—in principle we have no way of delimiting the class of all *possible* fields of body force. Therefore, in arguments concerning the totality of all possible motions of a body, we shall necessarily think of $\mathbf{b}$ as being unrestricted. Whatever be χ and $\mathbf{T}$, a field $\mathbf{b}$ such as to satisfy the principle of balance of linear momentum is determined by (III.6-1), or, if the frame of reference is not inertial, by (III.6-5). Thus CAUCHY's First Law imposes no restriction at all upon χ and $\mathbf{T}$.

A dynamic process is defined in terms of a frame $\oint$. Thus, properly, we should refer to $\{\chi, \mathbf{T}\}$ as being a *dynamic process in* $\oint$. Suppose now we consider another frame $\oint^*$. We have reason to regard the motion χ^* obtained from χ by the transformation (I.9-11) as being the very same motion as apparent in $\oint^*$:

$$\mathbf{x}^* = \chi^*(X, t^*) = \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\chi(X, t) - \mathbf{x}_0),$$
$$t^* = t + a, \qquad \text{(I.9-11)}_r$$

a, $\mathbf{x}_0$, $\mathbf{x}_0^*(t)$, and $\mathbf{Q}(t)$ being prescribed. In contrast, as we have seen in Section III.4, the stress $\mathbf{T}$ is frame-indifferent:

$$\mathbf{T}^*(\mathbf{x}^*, t^*) = \mathbf{Q}(t)\mathbf{T}(\mathbf{x}, t)\mathbf{Q}(t)^{\mathsf{T}}; \qquad \text{(IV.1-1)}$$

here $\mathbf{x}^*$ and t^* are determined from $\mathbf{x}$ and t by (I.9-11). Finally, if the dynamic process $\{\chi, \mathbf{T}\}$ determines a body force $\mathbf{b}$ in $\oint$, then $\mathbf{b}^*$ as given by

$$\mathbf{b}^*(\mathbf{x}^*, t^*) = \mathbf{Q}(t)\mathbf{b}(\mathbf{x}, t) \tag{IV.1-2}$$

serves to balance $\{\chi^*, \mathbf{T}^*\}$ in $\oint^*$, and, of course, CAUCHY's First Law is understood to hold in the frame-indifferent form (III.6-5). Not only is $\{\chi^*, \mathbf{T}^*\}$ a dynamic process defined in terms of $\oint^*$, but also the body force $\mathbf{b}^*$ corresponding with it is the same, in the sense of frame-indifference, as the body force required to equilibrate $\{\chi, \mathbf{T}\}$ in $\oint$. Thus the definition of a dynamic process is frame-indifferent, and the process $\{\chi^*, \mathbf{T}^*\}$ in $\oint^*$ may be regarded as describing the same phenomena of nature as does $\{\chi, \mathbf{T}\}$ in $\oint$. We shall say formally that $\{\chi^*, \mathbf{T}^*\}$ is the process in $\oint^*$ that is *equivalent* to the process $\{\chi, \mathbf{T}\}$ in $\oint$ if the two are related through (I.9-11) and (1).

The foregoing statements enable us to substitute (I.7-7) and $(\text{I.9-11})_1$ into (1), so obtaining

$$\mathbf{T}^*(\chi^*(X, t^*), t^*) = \mathbf{Q}(t)\mathbf{T}(\chi(X, t), t)\mathbf{Q}(t)^{\mathsf{T}}. \tag{IV.1-3}$$

2. Constitutive Relations. Noll's Axioms

The principles and definitions so far presented express properties *common* to all bodies and motions. The *diversity* of natural bodies, which arises from the differences among the materials that make up those bodies, we represent in the theory by *constitutive relations*. In mechanics, a constitutive relation is a restriction upon the forces or the motions or both. In popular terms, forces applied to a body "cause" it to undergo a motion, and the motion "caused" differs according to the nature of the body.

In this regard some constitutive relations are trivial, in the sense that a constant function is a trivial special case of a function. External body forces are of this kind. The assumption that the body force is external, since it restricts the class of body forces to those unaffected by the motions of such bodies as may occupy the part of space in which they act, is a constitutive relation, but it is not the kind subjected to study in continuum mechanics, in which, in typical problems, we simply assume that $\mathbf{b} = \text{const.}$ or even $\mathbf{0}$ and go on to analyse in detail the different responses to these trivial body forces effected upon bodies in which there are different kinds of contact forces.

Indeed, the only forces of much interest in continuum mechanics are contact forces. As we have seen in Section III.4, these are determined from the stress tensor field $\mathbf{T}$. Just as different figures defined in geometry idealize certain important natural objects, in continuum mechanics *ideal materials* are defined by particular relations between the stress tensor and the motion of a body. Some special materials, like some special figures, are important in themselves, but it is more efficient to study infinite classes of geometric objects and infinite classes of materials, distinguished by properties of symmetry and invariance. A *general theory of constitutive relations* lays down overriding restrictions that all constitutive relations must satisfy in order to represent mathematically the kinds of behavior observed in materials in nature. In the class of all such constitutive relations a rational scheme of classification is then introduced, and theorems characterizing or describing the members of this class are then proved.

The approach is like that of Euclidean geometry, in which, after statement of the axioms satisfied by all geometric objects, theorems characterizing and relating classes of figures are proved. Since mechanics is a discipline vastly more subtle and sensitive than geometry, the parallel stops here and does not extend to the theorems themselves or even to the methods of constructing proofs.

Continuum mechanics, like any other branch of mathematics, has its own characteristic concepts and methods. These were created in large part by JAMES BERNOULLI, EULER, CAUCHY, GREEN, STOKES, KELVIN, MAXWELL, and HUGONIOT, but only in recent years have they been subjected to general and collective scrutiny and forged into a unified doctrine.

The further development of continuum mechanics in this book will fall within the axioms laid down by NOLL in 1958. These, which we now state, while they are by no means the most general considered today, are nevertheless far more general than necessary for our purpose in this introductory book, but they are so clear and easy to grasp that more special statement here would only blunt them.

Axiom N1. Principle of Determinism. *The stress at the place occupied by the substantial point X at the time t is determined by the history χ^t of the transplacement of $\mathscr{B}$ up to the time t:*

$$\mathbf{T}(\chi(X, t), t) = \mathfrak{F}(\chi^t; X, t). \tag{IV.2-1}$$

Here $\mathfrak{F}$ denotes a mapping of histories χ^t, substantial points X, and times t onto symmetric tensors. It defines the *material* composing $\mathscr{B}$. The domain of the first argument χ^t is the set of possible motions of $\mathscr{B}$ (and not merely their restrictions to the substantial point X). The range of $\mathfrak{F}$, for each X, is some subset of the set of histories of symmetric tensors. The mapping $\mathfrak{F}$ is the *constitutive mapping* of the substantial point X; and the substantial point X itself is now called a *material point* of $\mathscr{B}$. The relation (1) is the *constitutive relation* of the material defined by $\mathfrak{F}$. The mapping $\mathfrak{F}$ is neither more nor less than a rule which, for each material point and at each time, assigns to the history up to the time t of each conceivable motion of $\mathscr{B}$ a unique stress tensor $\mathbf{T}(\chi(X, t), t)$ at the place $\mathbf{x}$ occupied by X at the time t. As X ranges over $\mathscr{B}$, the value of $\mathfrak{F}$ at the time t delivers the stress field $\mathbf{T}(\mathbf{x}, t)$ acting upon $\chi(\mathscr{B}, t)$. In rough terms, *the past and present placements given by the motion of $\mathscr{B}$ to the material points it comprises determine the stress field over its present shape $\chi(\mathscr{B}, t)$.*

The concept of material here defined represents the common observation that many natural bodies exhibit *memory* of their past experiences, sometimes continuing to respond to the effects of change of form long after the change itself took place. For this reason $\mathfrak{F}$ is often called a *memory functional.* Of course, those special $\mathfrak{F}$ that depend on χ only through its present value, which model materials without memory, or through the present values of its time derivatives, which model materials with short-range memory, are not excluded.

Only frames preserving the sense of time are allowed in mechanics, as has been stated in Section I.6. In view of this fact and the definition (II.10-1) of the history χ^t, the constitutive relation (1) respects the sense of time. While past and present motion determine present stress, it by no means follows that future and present motion do the same. In the materials of nature the past of a specimen cannot generally be reconstructed from its present and future conditions, and irreversibility of this kind is allowed for by the mathematical theory from the start. Indeed, irreversibility is the rule, not the exception, in continuum mechanics, and the study of various precise senses of that word is one of the main aims of the theory. In this study, continuum mechanics quits the tradition of analytical dynamics, in which, in typical cases, such as that presented above in Section I.14, past and future are interchangeable.

It is possible that (1) be invertible in the sense that the motion χ of a body is determined, conversely, from the history $\mathbf{T}^t$ of the stress field defined over it. However, such cannot be the case in general, since in Eulerian hydrodynamics, defined by the special constitutive relation (IV.4-4), below, a knowledge of the pressure field for all times and at all places determines nothing more about the placement $\chi(\cdot, t)$ than its mass density ρ. Thus an inverted relation giving χ in terms of $\mathbf{T}^t$ *cannot possibly be general.*

Axiom N2. Principle of Local Action. The principle of determinism allows the motions of material points Z that lie far away from X to affect the stress at X. The notion of contact force makes it natural to exclude action at a distance as a material property. Accordingly, we assume a second constitutive axiom: *The motion of material points at a finite distance from X in some shape of $\mathscr{B}$ may be disregarded in calculating the stress at X from the past and present shapes of $\mathscr{B}$.* (Of course, by the smoothness assumed for χ, material points once a finite distance apart are always a finite distance apart.)

Formally, if χ and χ are motions such that for some neighborhood $\mathcal{N}(X)$

$$\chi^t(Z, s) = \chi^t(Z, s) \qquad \forall s \geqq 0, \quad \forall Z \in \mathcal{N}(X), \tag{IV.2-2}$$

then

$$\mathfrak{F}(\chi^t; X, t) = \mathfrak{F}(\chi^t; X, t). \tag{IV.2-3}$$

If we apply Axiom 1 to the transformation rule (IV.1-3), we obtain

$$\mathfrak{F}^*(\chi^{*t^*}; X, t^*) = \mathbf{Q}(t)\mathfrak{F}(\chi^t; X, t)\mathbf{Q}(t)^{\mathsf{T}}. \tag{IV.2-4}$$

Here the constitutive mappings $\mathfrak{F}$ and $\mathfrak{F}^*$ may differ. In conformity with general experience gained from observations of the behavior of materials, we assume that the properties of a given material as represented by $\mathfrak{F}$ do not differ for different rigid frames. *If the constitutive relation* (1) *holds for the dynamic process* $\{\chi, \mathbf{T}\}$, *it holds also for every equivalent dynamic process* $\{\chi^*, \mathbf{T}^*\}$ as defined in Section IV.1. In other words, we cannot distinguish one rigid frame from another by measurement of the stresses in a body of given material. Formally, for any two rigid frames $\oint$ and $\oint^*$

$$\mathfrak{F}^* = \mathfrak{F}. \tag{IV.2-5}$$

The assumption (5) expresses

Axiom N3. Principle of Material Frame-Indifference.

$$\mathfrak{F}(\chi^{*t^*}; X, t^*) = \mathbf{Q}(t)\mathfrak{F}(\chi^t; X, t)\mathbf{Q}(t)^{\mathsf{T}}. \tag{IV.2-6}$$

In fact (6) should hold for two generalized frames $\oint$ and $\oint^*$ in the same rigid class. (Generalized frames are defined above in the passage in fine print preceding the discussion of units of length and physical distance in Section I.6.)

Now referring to a particular frame $\oint$ and a particular constitutive mapping $\mathfrak{F}$, let us consider two motions χ_1 and χ_2 that differ by a rigid motion, a time shift, and possibly a reflection:

$$\chi_2(X, t + a) = \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\chi_1(X, t) - \mathbf{x}_0) \tag{IV.2-7}$$

for a real number a, a fixed place $\mathbf{x}_0$, a vector $\mathbf{x}_0^*(t)$, and an orthogonal tensor $\mathbf{Q}(t)$. While χ_1 and χ_2 need not be given a kinematical interpretation here,

χ_2^{t+a} may be regarded as formally related to χ_1^t just as χ^{*t^*} is related to χ^t, and so by using (6) we obtain the following functional equation, which each constitutive mapping $\mathfrak{F}$ must satisfy:

$$\mathfrak{F}(\chi_2^{t+a}; X, t+a) = \mathbf{Q}(t)\mathfrak{F}(\chi_1^t; X, t)\mathbf{Q}(t)^{\mathsf{T}}. \tag{IV.2-8}$$

This statement asserts ***Invariance under Superposed Rigid Motions***: *In any given frame let the motion* χ_2 *be obtained from the motion* χ_1 *by superposing a rigid motion, a time-shift, and possibly a reflection. Under that superposition the value of a constitutive mapping transforms like the stress under a change of frame* (I.9-11) *with the orthogonal tensor* $\mathbf{Q}(t)$.

While the analysis just given refers to a single frame only, its conclusion is not restricted to that frame, because if (8) holds in one frame, it holds in all, for a change of frame if considered as a purely mathematical statement superposes a rigid motion perhaps combined with a reflection.

We have considered the behavior of $\mathfrak{F}$ for a class of motions. Material frame-indifference was used to effect the proof of Invariance under Superposed Rigid Motions. Now we shall show that, conversely, Invariance under Superposed Rigid Motions implies the Principle of Material Frame-Indifference. The assumption is now (8), with χ_2 derived from χ_1 by superposing a rigid motion, a time-shift, and possibly a reflection, expressed by (7).

For any two rigid frames $\not{f}$ and $\not{f}^*$, any two equivalent motions χ and χ^* are related by (I.9-11), and hence they satisfy (7) with $\chi_1 = \chi$, $\chi_2 = \chi^*$, and $t + a = t^*$. Thus our assumption (8) asserts that

$$\mathfrak{F}(\chi^{*t^*}; X, t) = \mathbf{Q}(t)\mathfrak{F}(\chi^t; X, t)\mathbf{Q}(t)^{\mathsf{T}}. \tag{IV.2-6$_r$}$$

From (6) and (4) we conclude that

$$\mathfrak{F}(\chi^{*t^*}; X, t^*) = \mathfrak{F}^*(\chi^{*t^*}; X, t). \tag{IV.2-9}$$

Because χ^* in (9) is arbitrary, (5) follows.

Using the abbreviations MFI and SRM to denote the two principles, we may abbreviate as follows the theorem just proved:

$$\underset{\text{(all frames)}}{\text{MFI}} \Leftrightarrow \underset{\text{(some one frame)}}{\text{SRM}} . \tag{IV.2-10}$$

In other words, *the two principles are equivalent.*

For the foregoing analysis I am indebted to R. SEGEV.

From physically oriented circles come claims that SRM is the "correct" principle, being more "physical", while MFI is not correct. Some use the term "objectivity" to denote SRM. Tendentious terms are dangerous; in ordinary life, use of a tendentious term often indicates deceit, and those who claim to be "objective" in their judgments sometimes mislead themselves while trying to pull the wool over the eyes of others.

Neither MFI nor SRM is objective in any sense of that adjective found in dictionaries. Both are assumptions. On the basis of Axiom N1, each implies the other.

The real difference of opinion comes with the sign of $\det \mathbf{Q}$. Rotation of a body does not turn it inside out. Therefore, $\det \mathbf{Q} = +1$ in transformations intended to represent motions of natural bodies. Anyone who so wishes may impose that requirement. To impose it for SRM automatically imposes it for MFI, and conversely. For all purposes in this book, it makes no difference. In the proofs given in this section, $\mathbf{Q}$ may always be replaced by $-\mathbf{Q}$ at pleasure.

Constitutive mappings whose values are tensors of odd order appear in theories of heat conduction, optics, *etc.* They are not invariant under change of the sign of $\mathbf{Q}$.

While some steps may be taken to delimit the class of all constitutive mappings that satisfy Axioms N1–N3, in this book we shall consider only *simple materials.* To this special class, which is still general enough to include all the older theories of continua and many of the more recent ones, we now address ourselves.

3. Simple Materials

The constitutive axioms N1 and N2 state that the history of the motion of an arbitrarily small neighborhood of a material point determines the stress at the place presently occupied by that point. The first approximation to the transplacement χ_κ at $\mathbf{X}$ is provided, at each t, by the transplacement gradient $\mathbf{F}_\kappa(\mathbf{X}, t)$, the properties of which we have discussed in Section II.5 and thereafter. Thus the history of $\mathbf{F}_\kappa$, which we shall denote by $\mathbf{F}_\kappa^t$, provides a first approximation near $\mathbf{X}$ to the history χ_κ^t of the transplacement χ_κ of $\mathscr{B}$. If a knowledge of this first approximation suffices to determine the stress at $\mathbf{X}$, the corresponding material point X is called *simple*. Formally, (IV.2-1) becomes in this instance

$$\mathbf{T}(\mathbf{x}, t) = \mathbf{T}(\chi_\kappa(\mathbf{X}, t), t) = \mathfrak{G}_\kappa(\mathbf{F}_\kappa^t(\mathbf{X}), \mathbf{X}). \qquad \text{(IV.3-1)}$$

Clearly the principles of determinism and local action, Axioms N1 and N2, are satisfied. We shall consider Axiom N3 presently.

The mapping $\mathfrak{G}_\kappa$ is called the *response with respect to* κ. If it is such as to satisfy Axiom N3, the Principle of Material Frame-Indifference, it defines a particular *simple material*; otherwise, it does not. The domain of its first argument, for fixed $\mathbf{X}$, is a suitable class of histories of invertible tensors. Its range is some subset of the set of all symmetric tensor fields over the present shape $\chi(\mathscr{B}, t)$ of $\mathscr{B}$ in $\mathscr{E}$. In other words, at a given time t and at a fixed place $\mathbf{X}$ in the reference shape it maps the history of an invertible tensor function of time onto a symmetric tensor at the place $\mathbf{x}$ presently occupied by X.

As in Section II.7, let $\boldsymbol{\lambda}$ map the reference shape $\kappa_1(\mathscr{B})$ onto another one, $\kappa_2(\mathscr{B})$, and let $\mathbf{P} := \nabla\boldsymbol{\lambda}$. Thus $\mathbf{P}$ is a given function of $\mathbf{X}$. Substituting (II.7-5) into (1) yields

$$\mathfrak{G}_{\kappa_1}(\mathbf{F}^t_{\kappa_1}) = \mathfrak{G}_{\kappa_1}(\mathbf{F}^t_{\kappa_2}\mathbf{P}), \tag{IV.3-2}$$

in a notation which omits the place $\mathbf{X}$ in $\kappa_1(\mathscr{B})$ that the material point X occupies. Thus if for any invertible history $\mathbf{F}^t$

$$\mathfrak{G}_{\kappa_2}(\mathbf{F}^t) := \mathfrak{G}_{\kappa_1}(\mathbf{F}^t\mathbf{P}), \tag{IV.3-3}$$

the constitutive relation (1) assumes the form

$$\mathbf{T}(\mathbf{x}, t) = \mathfrak{G}_{\kappa_2}(\mathbf{F}^t(\mathbf{X}), \mathbf{X}), \tag{IV.3-4}$$

provided now $\mathbf{F}$ be interpreted as the gradient of χ_{κ_2} at $\mathbf{X}$, and $\mathbf{x} = \chi_{\kappa_2}(\mathbf{X}, t)$. Thus $\mathbf{T}$ is determined just as well by the history of the transplacement gradient from $\kappa_2(\mathscr{B})$ as by the history of the transplacement gradient from $\kappa_1(\mathscr{B})$. Although the response $\mathfrak{G}_{\kappa_2}$ is not generally the same mapping as is the response $\mathfrak{G}_{\kappa_1}$, the existence of such a mapping is a fact independent of the choice of reference placement. Therefore, the definition of a simple material, while it mentions a reference placement, does not depend upon that placement and hence could be expressed without any use of reference placements.

Homogeneous transplacements were defined and analysed in Section II.12. A history $\mathbf{F}^t$ of the gradient of a homogeneous transplacement can be constructed from any invertible tensor function $\mathbf{F}$. By exhausting the class of histories of such transplacements, we exhaust the domain of the response $\mathfrak{G}_\kappa(\cdot, \mathbf{X})$. Thus *the response of a simple material point is determined for all histories by its restriction to the histories of gradients of homogeneous transplacements.*

In laboratories of experimental mechanics great weight is laid upon homogeneous transplacements, and the results of more complicated transplacements are commonly explained in terms of them. In this sense, though unconsciously, experimenters tend to presume that a material found in the laboratory may be modelled sufficiently well in the mathematical theory by some simple material. In this book we shall sometimes use the term "experiment" in an ideal sense. We shall imagine an experiment in which a subbody containing X is subjected to a particular transplacement history χ^t_κ with respect to the reference placement κ, and we shall suppose that the resulting stress $\mathbf{T}$ is then measured. As χ_κ^t ranges over various transplacement histories, various values of $\mathbf{T}$ result. We shall describe the constitutive relation itself as expressing the outcome of these experiments. In this sense we may say that *the material point X of the body $\mathscr{B}$ is* simple *if the outcomes of all experiments at X are determined by the outcomes of all experiments on homogeneous transplacements of parts of $\mathscr{B}$ near X*. In Section IV.9, so as to delimit the ideal experimental program suggested by this fact, we shall determine the homogeneous transplacements that can be produced by the action of uniform body forces.

The definition of a simple material and its interpretations are independent of the choice of reference placement κ. The response $\mathfrak{G}_\kappa$ with respect to κ is not, nor are the experiments just mentioned. A homogeneous transplacement of $\kappa_1(\mathscr{B})$ is not a homogeneous transplacement of $\kappa_2(\mathscr{B})$ unless $\kappa_2 \circ \kappa_1^{-1}$ is an affine mapping. The responses $\mathfrak{G}_{\kappa_1}$ and $\mathfrak{G}_{\kappa_2}$ are in general different mappings, each being determined uniquely from the other by (3).

Henceforth in this book *we shall consider only simple materials*. When, as will usually be the case, a reference placement κ is selected once and for all, we shall write the constitutive relation (1) of a simple material in the abbreviated form

$$\mathbf{T}(t) = \mathfrak{G}(\mathbf{F}^t). \tag{IV.3-5}$$

The theory of simple materials includes most of the common purely mechanical theories of continua studied in works on mechanics, engineering, physics, applied mathematics, *etc*. While modern studies of continuum physics include microstructure, electromagnetism, chemical reactions, diffusion, and relativistic phenomena, we shall not consider those in this book. Modern continuum thermomechanics incorporates the effects of heating and change of temperature, but those, too, this book will not go into.

4. Some Classical Instances. Specimens of the Effect of the Principle of Material Frame-Indifference

In this section we shall define some of the special materials, the theory of which furnished the main subject of study in continuum mechanics in former times, and we shall use them to illustrate the power of the Principle of Material Frame-Indifference to reduce the apparent generality of a class of hypothetical constitutive relations. The reader who is already familiar with classical theories or who desires only a consecutive, systematic development of continuum mechanics should skip this section and pass to the next.

An *elastic material* is defined by the instance in which the mapping $\mathfrak{G}$ in (IV.3-5) reduces to a function $\mathfrak{g}$ of the present transplacement gradient $\mathbf{F}(\mathbf{X}, t)$, irrespective of the values $\mathbf{F}^t(\mathbf{X}, s)$ of the history $\mathbf{F}^t$ in the past; *i.e.*, when $s > 0$,

$$\mathbf{T} = \mathfrak{g}(\mathbf{F}, \mathbf{X}), \tag{IV.4-1}$$

$\mathfrak{g}(\cdot, \mathbf{X})$ being a function which maps invertible tensors $\mathbf{F}$ onto symmetric tensors $\mathbf{T}$. Not all such functions define elastic materials, however, since the Principle of Material Frame-Indifference, stated in Section IV.2 as Axiom N3, is not satisfied unless $\mathfrak{g}$ is of a special kind, as delimited in the following

Theorem (CELLERIER, RICHTER). *Let the polar decomposition of the transplacement gradient be* $\mathbf{F} = \mathbf{RU}$. *Then the constitutive relation of an elastic material is of the form*

$$\mathbf{T} = \mathbf{R}\mathfrak{g}(\mathbf{U}, \mathbf{X})\mathbf{R}^{\mathsf{T}}. \tag{IV.4-2}$$

in which $\mathfrak{g}(\cdot, \mathbf{X})$ *maps positive, symmetric tensors onto symmetric tensors. Conversely, any such* $\mathfrak{g}$ *serves by means of* (2) *to define a particular elastic material.*

Proof. We invoke Axiom N3 only in a weakened form. Indeed, since (1) involves $\mathbf{F}^t$ only through $\mathbf{F}$, which is $\mathbf{F}^t(0)$, we need specify nothing about the orthogonal tensor history $\mathbf{Q}^t$ mentioned in Axiom N3 except its present value $\mathbf{Q}^t(0)$, which we shall denote by $\mathbf{Q}$. Under a change of frame $\mathbf{F}$ obeys the transformation rule (II.14-7). Thus, according to (IV.2-6),

$$\mathfrak{g}(\mathbf{QRU}) = \mathbf{Q}\mathfrak{g}(\mathbf{RU})\mathbf{Q}^{\mathsf{T}}, \tag{IV.4-3}$$

$\mathbf{X}$ being omitted from the notation since it is held fixed in this proof. The functional equation (3) must hold for all orthogonal $\mathbf{Q}$, all orthogonal $\mathbf{R}$, and all positive and symmetric $\mathbf{U}$. In particular, (3) must hold if we choose $\mathbf{Q} = \mathbf{R}^{\mathsf{T}}$. Hence (2) follows as a necessary condition. That it is also sufficient, is trivial to verify. $\triangle$

The elastic material is of intrinsic interest because it is the simplest example of a simple material that springs to mind, as natural to mechanics as is the circle to geometry. The constitutive relation (1) provides a precise, generalized formulation of HOOKE's "ut tensio sic vis." Moreover, many real materials conform with it roughly when $|\mathbf{U} - \mathbf{1}|$ is sufficiently small—for glass, very small indeed at room temperature, and for rubber, many times larger. Furthermore, if we apply (IV.3-5) to any constant history $\mathbf{F}^t = \mathbf{F}_0$, say, it reduces to (1). Thus, in problems of statics every simple material behaves like a corresponding elastic material: *The statics of simple materials is elastostatics.* This obvious fact has many important consequences in the theory of simple materials.

If $\mathfrak{g}(\mathbf{U}) = g(\det \mathbf{U})\mathbf{1}$, then (2) reduces because of (II.9-8) to

$$\mathbf{T} = -p(\rho)\mathbf{1}; \tag{IV.4-4}$$

p is the *pressure function*, which, as the notation indicates, determines the pressure field from the density field. The material so defined is called an *elastic fluid* or *ideal fluid* or *Eulerian fluid*; it provides the basis for much of the classical theory of compressible fluids.

As we have remarked before, the velocity field of a motion is called a flow. The term "flow" also has a popular or physical meaning, and so as to reconcile common language with kinematics, sometimes the capacity of a fluid to flow is attributed to its failure to sustain shear stress when at rest in any placement whatever. We shall see below in Section IV.17 that this property, while common to all simple fluids, does not suffice to define them. The Eulerian fluid satisfies it *a fortiori*, since it never sustains shear stress, whether it be at rest or in motion.

A class of materials more general than the Eulerian fluids and also not subsumed under elastic materials may be defined by the functional relation

$$\mathbf{T} = \mathfrak{r}(\mathbf{G}, \rho, \dot{\mathbf{x}}, \mathbf{x}, t), \tag{IV.4-5}$$

the first argument, $\mathbf{G}$, being the velocity gradient (II.11-7). We shall see now that the Principle of Material Frame-Indifference forces the last three arguments to drop out and imposes further restrictions upon the function $\mathfrak{r}$. Indeed, Axiom N3 requires that for an arbitrary orthogonal tensor function of time $\mathbf{Q}$, an arbitrary place-valued function of time $\mathbf{x}_0^*$, an arbitrary place $\mathbf{x}_0$, and an arbitrary constant a, the function $\mathfrak{r}$ shall satisfy for all arguments $\mathbf{G}$, ρ, $\dot{\mathbf{x}}$, $\mathbf{x}$, t the identity

$$\begin{aligned} &\mathfrak{r}(\mathbf{G}, \rho, \dot{\mathbf{x}}, \mathbf{x}, t) \\ &\quad = \mathbf{Q}^{\mathsf{T}}\mathfrak{r}(\mathbf{D}^* + \mathbf{W}^*, \rho^*, \dot{\mathbf{x}}^*, \mathbf{x}^*, t^*)\mathbf{Q}, \\ &\quad = \mathbf{Q}^{\mathsf{T}}\mathfrak{r}(\mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}} + \mathbf{Q}\mathbf{W}\mathbf{Q}^{\mathsf{T}} + \mathbf{A}, \rho, \\ &\qquad \mathbf{Q}\dot{\mathbf{x}} + \dot{\mathbf{x}}_0^* + \mathbf{A}(\mathbf{x}^* - \mathbf{x}_0^*), \mathbf{x}_0^* + \mathbf{Q}(\mathbf{x} - \mathbf{x}_0), t + a)\mathbf{Q}, \end{aligned} \tag{IV.4-6}$$

to explicate which we have used (II.11-8), (II.14-13), (I.9-11), and (I.9-14). Let us consider particular, fixed arguments $\mathbf{G}$, ρ, $\dot{\mathbf{x}}$, $\mathbf{x}$, t and choose a function

$\mathbf{Q}$ such that $\mathbf{Q}(t) = \mathbf{1}$, $\mathbf{A} = \dot{\mathbf{Q}}(t) = -\mathbf{W}$; a function $\mathbf{x}_0^*$ such that $\dot{\mathbf{x}}_0^*(t) = -\dot{\mathbf{x}} - \mathbf{A}(\mathbf{x}^* - \mathbf{x}_0^*(t))$, $\mathbf{x}_0^*(t) = \mathbf{x}_0 - (\mathbf{x} - \mathbf{x}_0)$; and a constant a such that $a = -t$. Then (6) yields the following necessary condition at each argument of $\mathfrak{r}$:

$$\mathfrak{r}(\mathbf{G}, \rho, \dot{\mathbf{x}}, \mathbf{x}, t) = \mathfrak{r}(\mathbf{D}, \rho, \mathbf{0}, \mathbf{x}_0, 0), \tag{IV.4-7}$$

in which $\mathbf{x}_0$ is any fixed place. Thus $\mathfrak{r}$ reduces to a function $\mathfrak{h}$ of $\mathbf{D}$ and ρ alone:

$$\mathfrak{r}(\mathbf{G}, \rho, \dot{\mathbf{x}}, \mathbf{x}, t) = \mathfrak{h}(\mathbf{D}, \rho). \tag{IV.4-8}$$

Roughly, we may describe the formal reasoning just given as showing that since the spin and the velocity may be transformed away by a suitable change of frame, and since any place and time may similarly be transformed into any other, these four arguments cannot enter a frame-indifferent constitutive relation of the presumptive class asserted by (5). But that is not all. If we substitute (8) back into (6), we obtain the relation

$$\mathfrak{h}(\mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}}, \rho) = \mathbf{Q}\mathfrak{h}(\mathbf{D}, \rho)\mathbf{Q}^{\mathsf{T}}. \tag{IV.4-9}$$

This identity must be satisfied by all symmetric tensors $\mathbf{D}$ and all orthogonal tensors $\mathbf{Q}$. Conversely, if it is satisfied, so also is (6). Thus we have the following

Theorem (NOLL). *In order that the relation* (5) *satisfy the Principle of Material Frame-Indifference, it is necessary and sufficient that* $\mathfrak{r}$ *reduce to a function* $\mathfrak{h}$ *of* $\mathbf{D}$ *and* ρ *alone and also satisfy* (9) *as an identity in* $\mathbf{Q}$ *and* $\mathbf{D}$.

A function $\mathfrak{h}$ that maps tensors onto tensors and satisfies the functional equation (9) is called *isotropic*. In a sense which we shall make precise in Section IV.14, NOLL's theorem asserts that *all materials whose constitutive relations are subsumed under* (5) *are isotropic materials*.

A material having a constitutive relation in the class defined by (5) when $\mathfrak{r}$ is made to be an affine function of its first argument is called a *linearly viscous fluid*. By NOLL's theorem, such a fluid must have a constitutive relation of the form inferred by STOKES:

$$\mathbf{T} = \mathfrak{h}(\mathbf{D}, \rho), \tag{IV.4-10}$$

in which $\mathfrak{h}$ is an affine, isotropic mapping of the set of symmetric tensors into itself.

We shall now determine the most general function of that kind. For later use we shall at first leave aside the condition that $\mathfrak{h}$ be affine. The dimension of the vector space considered, so long as it be finite, plays no part in the conclusions or the analysis.

Transfer Theorem (RIVLIN & ERICKSEN, SERRIN, NOLL, GUO). *Let* $\mathfrak{h}$ *map tensors onto tensors. If* $\mathbf{A}$ *and* $\mathbf{A}^{\mathsf{T}}$ *have a common proper vector* $\mathbf{e}$, *and if for all orthogonal* $\mathbf{Q}$

$$\mathfrak{h}(\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathsf{T}}) = \mathbf{Q}\mathfrak{h}(\mathbf{A})\mathbf{Q}^{\mathsf{T}}, \tag{IV.4-11}$$

then $\mathbf{e}$ *is a proper vector of both* $\mathfrak{h}(\mathbf{A})$ *and* $\mathfrak{h}(\mathbf{A}^{\mathsf{T}})$.

Proof. Let $\mathbf{e}$ be a unit proper vector of $\mathbf{A}$, and let $\mathbf{R}_{\mathbf{e}}$ be the reflection across the plane normal to $\mathbf{e}$:

$$\mathbf{R}_{\mathbf{e}} = \mathbf{1} - 2\mathbf{e}\otimes\mathbf{e}. \tag{IV.4-12}$$

Then for any $\mathbf{A}$

$$\mathbf{R}_{\mathbf{e}}\mathbf{A}\mathbf{R}_{\mathbf{e}}^{\mathsf{T}} = \mathbf{A} - 2\mathbf{e}\otimes\mathbf{A}^{\mathsf{T}}\mathbf{e} - 2\mathbf{A}\mathbf{e}\otimes\mathbf{e} + 4(\mathbf{e}\cdot\mathbf{A}\mathbf{e})(\mathbf{e}\otimes\mathbf{e}). \tag{IV.4-13}$$

If $\mathbf{A}$ and $\mathbf{A}^{\mathsf{T}}$ have a common proper vector $\mathbf{e}$, their corresponding proper numbers are the same, and so it follows that

$$\mathbf{R}_{\mathbf{e}}\mathbf{A}\mathbf{R}_{\mathbf{e}}^{\mathsf{T}} = \mathbf{A}. \tag{IV.4-14}$$

Since $\mathbf{R}_{\mathbf{e}}$ is orthogonal, (11) requires that

$$\begin{aligned}\mathbf{R}_{\mathbf{e}}\mathfrak{h}(\mathbf{A})\mathbf{R}_{\mathbf{e}}^{\mathsf{T}} &= \mathfrak{h}(\mathbf{R}_{\mathbf{e}}\mathbf{A}\mathbf{R}_{\mathbf{e}}^{\mathsf{T}}),\\ &= \mathfrak{h}(\mathbf{A}),\end{aligned} \tag{IV.4-15}$$

the second step being a consequence of (14). Thus $\mathbf{R}_{\mathbf{e}}$ commutes with $\mathfrak{h}(\mathbf{A})$, and so

$$\mathbf{R}_{\mathbf{e}}\mathfrak{h}(\mathbf{A})\mathbf{e} = \mathfrak{h}(\mathbf{A})\mathbf{R}_{\mathbf{e}}\mathbf{e} = -\mathfrak{h}(\mathbf{A})\mathbf{e}. \tag{IV.4-16}$$

That is, $\mathbf{R}_{\mathbf{e}}$ transforms $\mathfrak{h}(\mathbf{A})\mathbf{e}$ into its negative. Hence $\mathfrak{h}(\mathbf{A})\mathbf{e}$ is parallel to $\mathbf{e}$. Similarly, $\mathbf{e}$ is also a proper vector of $\mathfrak{h}(\mathbf{A}^{\mathsf{T}})$. $\triangle$

Among the tensors to which the transfer theorem applies are those that are symmetric, skew, or orthogonal.

The transfer theorem is widely useful. In this volume we make only one application of it, which follows now.

Theorem (CAUCHY). *In order that a function $\mathfrak{h}$ mapping symmetric tensors onto symmetric tensors be both isotropic and affine, it must have the representation*

$$\mathfrak{h}(\mathbf{A}) = (\alpha + \beta \operatorname{tr} \mathbf{A})\mathbf{1} + \gamma \mathbf{A}, \tag{IV.4-17}$$

in which α, β, and γ are constants. Conversely, if (17) *holds, $\mathfrak{h}$ is isotropic and affine.*

Proof (GURTIN). The projection $\mathbf{P}_\mathbf{e}$ has as proper vectors $\mathbf{e}$ itself and all vectors normal to $\mathbf{e}$. By the transfer theorem, these are proper vectors of $\mathfrak{h}(\mathbf{P}_\mathbf{e})$. The spectral representation of $\mathfrak{h}(\mathbf{P}_\mathbf{e})$ is therefore

$$\mathfrak{h}(\mathbf{P}_\mathbf{e}) = \beta(\mathbf{e})\mathbf{1} + \gamma(\mathbf{e})\mathbf{P}_\mathbf{e}, \tag{IV.4-18}$$

$\beta(\mathbf{e})$ being the proper number that corresponds to the vectors normal to $\mathbf{e}$, and $\beta(\mathbf{e}) + \gamma(\mathbf{e})$ being the proper number that corresponds to $\mathbf{e}$. It is sufficient to restrict the argument of β and γ to unit vectors. If $\mathbf{f}$ is any unit vector, there is an orthogonal tensor $\mathbf{Q}$ such that $\mathbf{Qe} = \mathbf{f}$. Then $\mathbf{P}_\mathbf{f} = \mathbf{QP}_\mathbf{e}\mathbf{Q}^\mathsf{T}$. Using this fact and (11) and then appealing to (18) twice, we show that

$$\begin{aligned} \mathbf{0} &= \mathbf{Q}\mathfrak{h}(\mathbf{P}_\mathbf{e})\mathbf{Q}^\mathsf{T} - \mathfrak{h}(\mathbf{QP}_\mathbf{e}\mathbf{Q}^\mathsf{T}), \\ &= \mathbf{Q}\mathfrak{h}(\mathbf{P}_\mathbf{e})\mathbf{Q}^\mathsf{T} - \mathfrak{h}(\mathbf{P}_\mathbf{f}), \\ &= [\beta(\mathbf{e}) - \beta(\mathbf{f})]\mathbf{1} + [\gamma(\mathbf{e}) - \gamma(\mathbf{f})]\mathbf{P}_\mathbf{f}. \end{aligned} \tag{IV.4-19}$$

Because $\mathbf{1}$ and $\mathbf{P}_\mathbf{f}$ are linearly independent,

$$\beta(\mathbf{e}) = \beta(\mathbf{f}) \qquad \text{and} \qquad \gamma(\mathbf{e}) = \gamma(\mathbf{f}). \tag{IV.4-20}$$

Since $\mathbf{e}$ and $\mathbf{f}$ are any two unit vectors, β and γ are constants.

Suppose now that $\mathfrak{h}$ is an affine function. Then there is a constant symmetric tensor $\mathbf{K}$ and a linear function $\mathfrak{l}$ such that

$$\mathfrak{h} = \mathbf{K} + \mathfrak{l}, \tag{IV.4-21}$$

and, by (11),

$$\mathbf{K} + \mathfrak{l}(\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathsf{T}}) = \mathbf{Q}\mathbf{K}\mathbf{Q}^{\mathsf{T}} + \mathbf{Q}\mathfrak{l}(\mathbf{A})\mathbf{Q}^{\mathsf{T}} \tag{IV.4-22}$$

for all symmetric $\mathbf{A}$ and for all orthogonal $\mathbf{Q}$. Since $\mathfrak{l}$ is linear, $\mathfrak{l}(\mathbf{0}) = \mathbf{0}$. According to (22), then, the constant symmetric tensor $\mathbf{K}$ commutes with every orthogonal tensor.

Exercise IV.4.1. If $\mathbf{K}$ commutes with every orthogonal tensor, then

$$\mathbf{K} = \alpha\mathbf{1}. \tag{IV.4-23}$$

Thus the linear function $\mathfrak{l}$ in (21) must satisfy (11). The fact that it is linear allows us to conclude from (18) that

$$\begin{aligned}
\mathfrak{l}(\mathbf{A}) &= \mathfrak{l}\left(\sum_{k=1}^{n} a_k \mathbf{P}_{\mathbf{e}_k}\right), \\
&= \sum_{k=1}^{n} a_k \mathfrak{l}(\mathbf{P}_{\mathbf{e}_k}), \\
&= \sum_{k=1}^{n} a_k (\beta\mathbf{1} + \gamma\mathbf{P}_{\mathbf{e}_k}), \\
&= \left(\beta\sum_{k=1}^{n} a_k\right)\mathbf{1} + \gamma\left(\sum_{k=1}^{n} a_k \mathbf{P}_{\mathbf{e}_k}\right), \\
&= \beta(\operatorname{tr}\mathbf{A})\mathbf{1} + \gamma\mathbf{A},
\end{aligned} \tag{IV.4-24}$$

the numbers $a_1, a_2, \ldots, a_n$ being the latent roots of $\mathbf{A}$. Putting (23) and (24) into (21) shows that $\mathfrak{h}$ must have the form (17).

Conversely, it is plain that (17) is an isotropic affine function for every choice of α, β, and γ. $\triangle$

By combining the theorems of CAUCHY and NOLL we obtain the following

Theorem (STOKES). *The constitutive relation of a linearly viscous fluid is*

$$\mathbf{T} = (-p + \lambda \operatorname{tr}\mathbf{D})\mathbf{1} + 2\mu\mathbf{D}, \tag{IV.4-25}$$

in which p, λ, and μ are functions of ρ. Every such relation defines a linearly viscous fluid.

The theory based on (25) is called the *Navier–Stokes Theory of Fluids*; under various hypotheses, (25) or major special cases of it were derived by NAVIER, CAUCHY, ST. VENANT, and STOKES. The coefficients λ and μ are called the *viscosities* of the fluid. In rigid motions the Navier–Stokes Theory reduces to Eulerian hydrodynamics, and so the fluid it defines exhibits the phenomenon of "flow" in the sense described above, namely, in a state of rest it can sustain only hydrostatic stress. If $\lambda = \mu = 0$, the linearly viscous fluid reduces to an elastic fluid, and for this reason elastic fluids are sometimes called "inviscid" or "perfect".

A material slightly more general than any of those introduced so far in this section is defined by reducing the mapping $\mathfrak{G}$ in (IV.3-5) to a function of $\mathbf{F}(\mathbf{X}, t)$ and $\dot{\mathbf{F}}(\mathbf{X}, t)$ which is affine in $\dot{\mathbf{F}}$:

$$\mathbf{T} = \mathsf{K}(\mathbf{F}, \mathbf{X})[\dot{\mathbf{F}}] = \mathsf{L}(\mathbf{F}, \mathbf{X})[\mathbf{G}]; \qquad \text{(IV.4-26)}$$

the second form follows from the first by (II.11-5), and the domain of the affine operator L, indicated by the brackets, is the space of tensors over $\mathscr{V}$. Such a material is called *linearly viscous.*

Exercise IV.4.2. The relation (26) satisfies the Principle of Material Frame-Indifference if and only if

$$\mathbf{R}^{\mathsf{T}}\mathbf{T}\mathbf{R} = \mathsf{M}(\mathbf{C}, \mathbf{X})[\mathbf{R}^{\mathsf{T}}\mathbf{D}\mathbf{R}], \qquad \text{(IV.4-27)}$$

$\mathsf{M}(\mathbf{C}, \mathbf{X})$ being an affine operator on the space of symmetric tensors over $\mathscr{V}$.

BOLTZMANN's accumulative theory of visco-elasticity is obtained if we suppose the mapping $\mathfrak{G}$ in (IV.3-5) to be expressible as an integral from $s = 0$ to $s = \infty$. In this case, too, the Principle of Material Frame-Indifference imposes a restriction upon the class of putative constitutive mappings, but we defer to Volume 3 the appropriate reduction.

In the Boltzmann theory, as in the theory of elasticity, a further simplification is often attained at the cost of supposing that $|\mathbf{F} - \mathbf{1}|$ or some measure of the magnitude of $\mathbf{F}^t - \mathbf{1}$ be small in some sense. Approximations of this kind make it easier to solve some special problems but are more confusing than helpful in analysis of the general theory.

Exercise IV.4.3. Other than a constant multiple of $\mathbf{1}$, there is no affine function $\mathfrak{g}$ in (1) that satisfies the Principle of Material Frame-Indifference. (Do not confuse this condition with that of taking the restriction of $\mathfrak{g}$ to positive symmetric arguments as affine.) This fact may be interpreted in terms of the theory of elasticity.

In this section we have defined and named some of the most important special materials of old. Also, we have illustrated the force of the Principle of Material Frame-Indifference by showing how it serves to *delimit those mappings that may enter a putative class of constitutive relations*. In the next section we shall encounter a more general argument of the same kind, an argument which applies to all simple materials.

5. Material Frame-Indifference. Reduced Constitutive Relations

According to Axiom N3, the response $\mathfrak{G}$ must be such as to make the constitutive relation (IV.3-5) satisfy the Principle of Material Frame-Indifference. Under the change of frame (II.14-3) the transplacement gradient $\mathbf{F}$ obeys the transformation rule (II.14-7), and hence its history $\mathbf{F}^t$ obeys the rule

$$\mathbf{F}^{*t^*} = \mathbf{Q}^t\mathbf{F}^t, \tag{IV.5-1}$$

$\mathbf{Q}^t$ being the history of the orthogonal tensor function $\mathbf{Q}$ occurring in (II.14-3), while the stress tensor $\mathbf{T}$ satisfies (IV.1-1). Hence in order that Axiom N3 hold, $\mathfrak{G}$ must be such that

$$\mathfrak{G}(\mathbf{Q}^t\mathbf{F}^t) = \mathbf{Q}(t)\mathfrak{G}(\mathbf{F}^t)\mathbf{Q}(t)^{\mathsf{T}} \tag{IV.5-2}$$

for *every* orthogonal tensor history $\mathbf{Q}^t$ and for *every* invertible tensor history $\mathbf{F}^t$ in a suitable class. Here $\mathbf{Q}(t)$ is the present value of the function $\mathbf{Q}$, so that $\mathbf{Q}^t(0) = \mathbf{Q}(t)$. Conversely, if (2) is satisfied, so is Axiom N3.

Following an analysis first given by NOLL, we can solve the equation (2) for $\mathfrak{G}$, once and for all. Indeed, by the polar decomposition theorem (II.9-1) we see that $\mathbf{F}^t = \mathbf{R}^t\mathbf{U}^t$, so that (2) becomes

$$\mathbf{Q}(t)^{\mathsf{T}}\mathfrak{G}(\mathbf{Q}^t\mathbf{R}^t\mathbf{U}^t)\mathbf{Q}(t) = \mathfrak{G}(\mathbf{F}^t). \tag{IV.5-3}$$

We may now choose the orthogonal tensor history $\mathbf{Q}^t$ in such a way that $\mathbf{Q}^t(s) = \mathbf{R}^t(s)^{\mathsf{T}}$, $0 \leqq s < \infty$. Hence $\mathbf{Q}(t) = \mathbf{R}(t)^{\mathsf{T}}$, and (3) becomes

$$\mathfrak{G}(\mathbf{F}^t) = \mathbf{R}(t)\mathfrak{G}(\mathbf{U}^t)\mathbf{R}(t)^{\mathsf{T}}. \tag{IV.5-4}$$

Conversely, if (4) holds, it is easy to show that Axiom N3 is satisfied. We have proved the following

Reduction Theorem (NOLL). *Let* $\mathfrak{G}$ *denote a mapping of positive symmetric tensor histories onto symmetric tensors. Then every constitutive relation for a simple material is of the form*

$$\mathbf{T}(t) = \mathbf{R}(t)\mathfrak{G}(\mathbf{U}^t)\mathbf{R}(t)^{\mathsf{T}}, \tag{IV.5-5}$$

and conversely, any such mapping defines a simple material.

A constitutive equation of this kind, in which the mappings or functions occurring are not subject to any further restriction upon the class of putative responses set down for study, is called a *reduced form*.

The reduction (5) shows that while the stretch history $\mathbf{U}^t$ of a simple material may influence its present stress in any way whatever, past rotations have no influence at all. The present rotation $\mathbf{R}$ enters (5) explicitly. Thus the reduced form enables us to dispense with considering rotation in determining the response to a motion. If we like, we may regard (4) as effecting an extension of $\mathfrak{G}$ from a domain of positive, symmetric tensor histories to the full domain of invertible tensor histories. In writing it and similar formulae henceforth we shall usually omit the argument t of $\mathbf{T}$, $\mathbf{U}$, $\mathbf{R}$, *etc.*, although of course t must still appear in the notation for histories $\mathbf{U}^t$, *etc.*

The reduced form enables us also, in principle, to reduce the number of tests needed to determine the response $\mathfrak{G}$ by experiment. Indeed, consider pure stretch histories: $\mathbf{R}^t = \mathbf{1}$. If we know the stress $\mathbf{T}$ corresponding to an arbitrary, homogeneous, pure stretch history $\mathbf{U}^t$, we have a relation of the form $\mathbf{T} = \mathfrak{G}(\mathbf{U}^t)$. By (5) we then know $\mathbf{T}$ for all deformation histories. Alternatively, consider irrotational histories: $\mathbf{W} = \mathbf{0}$. Given any $\mathbf{U}^t$, we can determine $\mathbf{R}^t$ by integrating $(\text{II}.11\text{-}26)_2$ with $\mathbf{W}$ set equal to $\mathbf{0}$. If we know the stress corresponding to an arbitrary irrotational history, by putting the corresponding values of $\mathbf{R}$ into (5) we can again determine $\mathfrak{G}$. Thus we may characterize simple materials in either of two more economical ways: *A material is simple if and only if its response in general is determined by its restriction to homogeneous, pure stretch histories, or to homogeneous, irrotational histories.*

In the polar decomposition (II.9-1) two measures of stretch, $\mathbf{U}$ and $\mathbf{V}$, are introduced. Kinematically, there is no reason to prefer one to the other. From (4) we see that use of $\mathbf{U}$ as a measure of stretch history leads to a simple reduced form for the constitutive equations of simple materials. If we like, of course we may use $\mathbf{V}$ instead. Since $\mathbf{U}^t = (\mathbf{R}^t)^{\mathsf{T}}\mathbf{V}^t\mathbf{R}^t$, substitution into (4) shows that by using $\mathbf{V}$ we do not generally eliminate the rotation history $\mathbf{R}$. Consequently, use of $\mathbf{V}$ does not lead to a simple statement.

There are many other tensors that measure stretch just as well as $\mathbf{U}$ and $\mathbf{V}$. In the older literature one or another of these is called a "strain" tensor, but the term "strain" has led to such confusion that we are better advised to avoid it altogether.

Exercise IV.5.1. Had we started from a relation of the form

$$\mathbf{T}(X, t) = \mathfrak{G}(\mathbf{F}^t, \mathbf{X}, \dot{\mathbf{x}}, \mathbf{x}, t) \tag{IV.5-6}$$

as the definition of a simple material point, the Principle of Material Frame-Indifference would have reduced it to our actual starting point (IV.3-1). (*Cf.* the analysis of the assumption (IV.4-5) in Section IV.4.)

Exercise IV.5.2. All the reductions obtained in Section IV.4 are in fact instances of the reduction indicated in the preceding exercise, followed by the reduction of (IV.3-5) to (5).

There are infinitely many other reduced forms for the constitutive relation of a simple material. Since $\mathbf{C}^t = (\mathbf{U}^t)^2$, one such form is

$$\begin{aligned} \mathbf{T} &= \mathbf{R}\mathbf{U}\mathbf{U}^{-1}\mathfrak{G}(\sqrt{\mathbf{C}^t})\mathbf{U}^{-1}\mathbf{U}\mathbf{R}^{\mathsf{T}}, \\ &= \mathbf{F}\mathfrak{K}(\mathbf{C}^t)\mathbf{F}^{\mathsf{T}}, \end{aligned} \tag{IV.5-7}$$

$\mathfrak{K}$ being defined as follows:

$$\mathfrak{K}(\mathbf{C}^t) := \sqrt{\mathbf{C}}^{-1}\mathfrak{G}(\sqrt{\mathbf{C}^t})\sqrt{\mathbf{C}}^{-1}. \tag{IV.5-8}$$

In Section II.8 we constructed the kinematical apparatus for using the actual placement as the reference placement. It is natural to ask if the response of a simple material can be described entirely in terms of this apparatus. Of course the answer is no, but an analysis due to NOLL shows just how far we can go toward expressing the constitutive relation in terms of $\mathbf{F}_t^t$ rather than $\mathbf{F}^t$. To do so, we note from (II.8-7) and (II.9-1) that for given $\mathbf{X}$

$$\mathbf{F}(\tau) = \mathbf{R}_t(\tau)\mathbf{U}_t(\tau)\mathbf{R}(t)\mathbf{U}(t). \tag{IV.5-9}$$

Thus

$$\mathbf{F}(\tau) = \mathbf{R}_t(\tau)\mathbf{R}(t)[\mathbf{R}(t)^{\mathsf{T}}\mathbf{U}_t(\tau)\mathbf{R}(t)]\mathbf{U}(t). \tag{IV.5-10}$$

In the notation (II.10-1) for histories, (10) reads as

$$\mathbf{F}^t = \mathbf{R}_t^t\mathbf{R}(t)[\mathbf{R}(t)^{\mathsf{T}}\mathbf{U}_t^t\mathbf{R}(t)]\mathbf{U}(t), \tag{IV.5-10A}$$

and so if

$$\mathbf{Q}^t(s) := (\mathbf{R}_t^t(s)\mathbf{R}(t))^{\mathsf{T}}, \tag{IV.5-11}$$

we obtain

$$\mathbf{Q}^t\mathbf{F}^t = \mathbf{R}(t)^{\mathsf{T}}\mathbf{U}_t^t\mathbf{R}(t)\mathbf{U}(t). \tag{IV.5-12}$$

We may write the requirement of frame-indifference (2) in the equivalent form

$$\mathfrak{G}(\mathbf{F}^t) = \mathbf{Q}(t)^{\mathsf{T}}\mathfrak{G}(\mathbf{Q}^t\mathbf{F}^t)\mathbf{Q}(t) \qquad \forall\mathbf{Q}^t. \tag{IV.5-13}$$

With the choice of $\mathbf{Q}^t$ given by (11) we have $\mathbf{Q}(t) = \mathbf{R}(t)^{\mathsf{T}}$, and so (12) and (13) yield

$$\mathbf{R}^{\mathsf{T}}\mathbf{T}\mathbf{R} = \mathfrak{G}(\mathbf{R}^{\mathsf{T}}\mathbf{U}_t^t\mathbf{R}\mathbf{U}). \tag{IV.5-14}$$

Because the right-hand side of this equation may be thought of as the value of a mapping of two arguments, $\mathbf{R}^{\mathsf{T}}\mathbf{U}_t^t\mathbf{R}$ and $\mathbf{U}$, a more useful equivalent expression in terms of $\mathbf{C}_t^t$ and $\mathbf{C}(t)$ is

$$\mathbf{R}^{\mathsf{T}}\mathbf{T}\mathbf{R} = \mathfrak{R}(\mathbf{R}^{\mathsf{T}}\mathbf{C}_t^t\mathbf{R}, \mathbf{C}). \tag{IV.5-15}$$

NOLL's reduced forms (14) and (15) show that it is not possible to express the effect of the transplacement history in determining the stress entirely by reference to the present shape. While the effect of all the *past* history, $0 < s < \infty$, is accounted for in this way, a fixed reference placement is required, in general, to allow for the effect of the deformation and rotation at the present instant, as shown by the appearance of $\mathbf{R}$ and $\mathbf{C}$ in (15). The relative rotation $\mathbf{R}_t$ has no effect at all.

We conclude this section by remarking upon an important instance. A material point is said to have a *placement at ease* $\boldsymbol{\kappa}_0$ if the stress vanishes when a neighborhood of that point has been held at rest in $\boldsymbol{\kappa}_0$ at all times, past and present. In general, of course, a material point need not have any such place ment, as is shown by the case of an Eulerian fluid, defined by (IV.4-4), since usually the pressure function p is assumed to be such that $p(\rho) > 0$ unless $\rho = 0$, the exception $\rho = 0$ being excluded because it violates the condition (II.2-5). When a placement at ease $\boldsymbol{\kappa}_0$ exists, if we choose it as the reference placement $\boldsymbol{\kappa}$ we obtain

$$\mathfrak{G}(\mathbf{1}^t) = \mathbf{0}, \tag{IV.5-16}$$

$\mathbf{1}^t$ being the history up to the time t of the tensor function $\mathbf{F}$ such that $\mathbf{F}(t) = \mathbf{1}$ for all times t. By (2) we see that

$$\mathfrak{G}(\mathbf{Q}^t) = \mathbf{0}. \tag{IV.5-17}$$

Thus any rotation, constant or varying in time in any way, carries one placement at ease into another. The converse is not true, for a material point may have two distinct placements at ease that are not obtained from one another by a rotation.

In this book we shall not assume in general that any material point has a placement at ease.

6. Internal Constraints

So far, we have been assuming that the material is capable, if subjected to appropriate forces, of undergoing any smooth deformation. Such a material is said to be *unconstrained*. An *a priori* restriction of possible transplacements at interior points of $\chi(\mathscr{B}, t)$ for all t in the domain of $\chi(\mathscr{B}, \cdot)$ is called an *internal constraint*; a material subject to one or more internal constraints is said to be *constrained*. An *elastic constraint* is expressed by requiring the transplacement gradient $\mathbf{F}$ to satisfy an equation of the form

$$\gamma(\mathbf{F}) = 0, \tag{IV.6-1}$$

where γ is a scalar function. The set of transplacement gradients satisfying (1) is called the *constraint set*. This set must be frame-indifferent in the sense that if it contains a particular transplacement gradient $\mathbf{F}$, then it contains also $\mathbf{QF}$ for all orthogonal $\mathbf{Q}$. This condition is satisfied if γ is a frame-indifferent function.

More generally, a *simple constraint*[1] is expressed by a relation like (1) except that the argument $\mathbf{F}$ is replaced by $\mathbf{F}^t$. In this book the only constraints we shall study are elastic constraints.

Exercise IV.6.1. γ is frame-indifferent if and only if

$$\gamma(\mathbf{F}) = \gamma(\mathbf{U}). \tag{IV.6-2}$$

[1] Simple constraints were introduced and studied by S. ANTMAN, "Material constraints in continuum mechanics," *Atti della Accademia Nazionale dei Lincei, Rendiconti, Classe di Scienze fisiche, matematiche e naturali* (8)**70**(1981): 256–264.

Hence an elastic constraint may be written in the form

$$\lambda(\mathbf{C}) = 0, \tag{IV.6-3}$$

where λ is a scalar function. Let λ have been determined, and let f be any real function that vanishes at 0 only. Then $\mathbf{F}$ satisfies the frame-indifferent elastic constraint (1) if and only if it satisfies $f(\lambda(\mathbf{C})) = 0$.

If we differentiate (3) with respect to time at a given material point, we obtain

$$\dot{\lambda} = \partial_{\mathbf{C}}\lambda(\mathbf{C})\cdot\dot{\mathbf{C}} = 0. \tag{IV.6-4}$$

That is, in view of $(\text{II}.11\text{-}26)_1$,

$$(\mathbf{F}\partial_{\mathbf{C}}\lambda(\mathbf{C})\mathbf{F}^{\mathsf{T}})\cdot\mathbf{D} = 0 \tag{IV.6-5}$$

for $\mathbf{F}$ compatible with the constraint and for all $\mathbf{D}$ corresponding with such $\mathbf{F}$. Conversely, if (5) holds at each instant for the material point in question, by integration we conclude that $\lambda(\mathbf{C}) = \text{const.}$; therefore, (5) asserts that if (3) holds at one instant, it holds always. Thus (5) may be used alternatively as a general expression for a frame-indifferent elastic constraint.

In all examples so far found to be of interest, for every positive $\mathbf{C}$ satisfying (3)

$$\partial_{\mathbf{C}}\lambda(\mathbf{C}) \neq \mathbf{0}, \tag{IV.6-6}$$

and we shall consider only constraints of this kind. Because $\mathbf{F}\partial_{\mathbf{C}}\lambda(\mathbf{C})\mathbf{F}^{\mathsf{T}}$ is a symmetric tensor, we may interpret (5) as requiring all $\mathbf{D}$ corresponding to $\mathbf{F}$ to lie in a certain five-dimensional plane determined by $\mathbf{C}$.

7. Principle of Determinism for Constrained Simple Materials

Constraints, since they assert that some kinds of deformation cannot occur, must be maintained by forces. Since the constraints, by definition, are immutable, the forces maintaining them cannot be determined by the motion itself or its history. Internal constraints must be maintained by appropriate stresses, and the constitutive equation of a constrained material must be such as to allow these stresses to operate.

For constrained materials, accordingly, the principle of determinism must be relaxed. *A fortiori*, the necessary modification of that principle cannot be deduced from the principle itself but must be brought in through a more general axiom.

There are, presumably, many systems of forces which could effect any given constraint. The simplest are those whose power vanishes in any motion compatible with the constraint. In a constrained material stresses that do no work will therefore be assumed to remain arbitrary in the sense that *they generally will be conditioned by the transplacement history but not entirely determined by it.*

Thus we have given reasons for laying down the following

Axiom N1$_C$ (Principle of Determinism for Simple Materials Subject to Constraints). *The stress is determined by the history of the transplacement gradient only to within an arbitrary tensor that does no work in any motion compatible with the constraint. That is,*

$$\mathbf{T} = \mathbf{N} + \mathfrak{G}(\mathbf{F}^t), \qquad \text{(IV.7-1)}$$

$\mathbf{N}$ *being a stress for which the stress-power vanishes in any motion satisfying the constraint. The determinate response* $\mathfrak{G}$ *need be defined only for arguments* $\mathbf{F}^t$ *such as to satisfy the constraint.*

It is understood here that $\mathfrak{G}$ is not unique but that for each $\mathbf{X}$ in the reference placement $\boldsymbol{\kappa}$ a particular $\mathfrak{G}$ may be selected. Thus far, $\mathfrak{G}$ depends upon $\boldsymbol{\kappa}$.

The definition

$$\mathbf{S} := \mathbf{T} - \mathbf{N} \qquad \text{(IV.7-2)}$$

gives the *determinate stress* $\mathbf{S}$: It is the value of the response $\mathfrak{G}$, which appears in (1), for the history $\mathbf{F}^t$ of the transplacement gradient under consideration. Axiom N1$_C$ generalizes Axiom N1 of Section IV.2 and reduces to it when no internal constraint is imposed, for then the only stress that never does work is $\mathbf{0}$.

As has been stated already, in this book we consider only elastic constraints.

The problem now, given an internal constraint λ, is to find the $\mathbf{N}$ that corresponds with it. The stress-power w of a symmetric stress tensor $\mathbf{T}$ in a motion with stretching $\mathbf{D}$ is given by (III.6-13). Accordingly, we are to find the general solution $\mathbf{N}$ of the equation

$$\mathbf{N} \cdot \mathbf{D} = 0 \qquad \text{(IV.7-3)}$$

if $\mathbf{D}$ is any symmetric tensor that satisfies (IV.6-5). Hence the symmetric tensor $\mathbf{N}$ must be perpendicular to every vector $\mathbf{D}$ that is perpendicular to $\mathbf{F}\partial_{\mathbf{C}}\lambda(\mathbf{C})\mathbf{F}^{\mathsf{T}}$. Thus $\mathbf{N}$ is parallel to this latter vector:

$$\mathbf{N} = q\mathbf{F}\partial_{\mathbf{C}}\lambda(\mathbf{C})\mathbf{F}^{\mathsf{T}}, \tag{IV.7-4}$$

q being an arbitrary scalar field. This formula provides the general solution of (3).

If there are k constraints $\lambda^m(\mathbf{C}) = 0$, $m = 1, 2, \ldots, k$, then

$$\mathbf{N} = \sum_{m=1}^{k} q_m \mathbf{F}\partial_{\mathbf{C}}\lambda^m(\mathbf{C})\mathbf{F}^{\mathsf{T}}. \tag{IV.7-5}$$

That is, the symmetric tensor $\mathbf{F}^{-1}\mathbf{N}(\mathbf{F}^{-1})^{\mathsf{T}}$ must lie in the span of the k symmetric tensors $\partial_{\mathbf{C}}\lambda^m(\mathbf{C})$, $m = 1, 2, \ldots, k$. If the k tensors $\partial_{\mathbf{C}}\lambda^m(\mathbf{C})$ are linearly independent, their span is a k-dimensional plane. If $k \geqq 6$, no restriction upon $\mathbf{N}$ results. Thus in a material subject to 6 or more constraints with linearly independent gradients, the stress is altogether arbitrary.

The argument given here applies at a single material point. Usually the same constraints will be laid down for all points of a body. In that case (5) will result for each, but the theory does not require that the quantities q_m in (5) for the several points be related to one another in any particular way. In order to obtain a constitutive relation leading to definite solutions of specific problems it is customary to assume that each multiplier q_m is a smooth field $q_m(\mathbf{x}, t)$ on the present shape of $\mathscr{B}$. Substitution into (1) yields the general constitutive equation for simple material subject to k simple, frame-indifferent, elastic constraints.

The determinate response $\mathfrak{G}(\mathbf{F}^t)$ may be expressed in reduced forms like those found in Section IV.4 for unconstrained materials.

We now consider some examples of constraints.

1. *Incompressibility.* A material is said to be *incompressible* if it can experience only isochoric motions. By (II.5-10) and (II.9-7)$_9$, an appropriate constraint function for an incompressible material is

$$\lambda(\mathbf{C}) = \det \mathbf{C} - 1. \tag{IV.7-6}$$

Because

$$\mathbf{F}\partial_{\mathbf{C}}\lambda(\mathbf{C})\mathbf{F}^{\mathsf{T}} = \mathbf{F}\mathbf{C}^{-1}\mathbf{F}^{\mathsf{T}} \det \mathbf{C} = \mathbf{1}, \tag{IV.7-7}$$

(4) yields

$$\mathbf{N} = -p\mathbf{1}, \tag{IV.7-8}$$

where p is an arbitrary scalar. Thus we have verified a statement due in effect to POINCARÉ: *In an incompressible material the stress is determined by the history of the transplacement gradient only to within an arbitrary hydrostatic pressure.*

2. *Inextensibility.* If $\mathbf{e}_\kappa$ is a unit position vector in the reference shape $\kappa(\mathscr{B})$, $\mathbf{F}\mathbf{e}_\kappa$ is the vector $\mathbf{e}$ into which it is carried in a homogeneous transplacement with gradient $\mathbf{F}$, as we have seen in Section II.12. Accordingly, for a material *inextensible* in the actual direction $\mathbf{e}$ an appropriate constraint function is

$$\lambda(\mathbf{C}) = |\mathbf{F}\mathbf{e}_\kappa|^2 - 1 = \mathbf{e}_\kappa \cdot \mathbf{C}\mathbf{e}_\kappa - 1. \tag{IV.7-9}$$

Because

$$\partial_{\mathbf{C}}\lambda(\mathbf{C}) = \mathbf{e}_\kappa \otimes \mathbf{e}_\kappa\,, \tag{IV.7-10}$$

(4) yields

$$\mathbf{N} = q\mathbf{F}(\mathbf{e}_\kappa \otimes \mathbf{e}_\kappa)\mathbf{F}^{\mathsf{T}} = q\mathbf{e} \otimes \mathbf{e}. \tag{IV.7-11}$$

Since $\mathbf{N}$ is an arbitrary uniaxial tension in the direction of $\mathbf{e}$, we recover a conclusion due to ADKINS & RIVLIN: *In a material inextensible in a certain direction, the stress is determined by the history of the transplacement gradient only to within a uniaxial tension in that direction.*

3. *Rigidity.* A material is *rigid* if it is inextensible in every direction. By the theorem just established, the stress in a rigid material is determinate only to within an arbitrary tension in every direction. Therefore, the stress in a rigid material is altogether unaffected by the motion. That is to be expected in view of the fact, demonstrated in Section I.13, that the rigid motion of any body is determinable without knowledge of what the stress may be.

A body of rigid material is a rigid body in the sense defined at the end of Section I.10.

Exercise IV.7.1. For incompressible materials there are counterparts of the reduced forms (IV.5-5), (IV.5-7), (IV.5-14), and (IV.5-15). The constitutive relation of an incompressible elastic material is of the form

$$\mathbf{T} = -p\mathbf{1} + \mathbf{R}\mathfrak{g}(\mathbf{U})\mathbf{R}^{\mathsf{T}}, \qquad |\det \mathbf{U}| = 1; \tag{IV.7-12}$$

of an incompressible, elastic fluid,

$$\mathbf{T} = -p\mathbf{1}; \tag{IV.7-13}$$

of an incompressible, linearly viscous fluid,

$$\mathbf{T} = -p\mathbf{1} + 2\mu\mathbf{D}, \qquad \operatorname{tr}\mathbf{D} = 0. \tag{IV.7-14}$$

In all three cases the hydrostatic pressure p is indeterminate in the sense that it may be assigned independently of the history of the motion. Conversely, any relation having one of the above three forms defines, respectively, an incompressible elastic material, an incompressible elastic fluid, and an incompressible linearly viscous fluid.

For future reference we note that the constitutive relation of an *incompressible material* is

$$\mathbf{T} = -p\mathbf{1} + \mathbf{S}, \qquad \mathbf{S} = \mathfrak{G}(\mathbf{F}^t); \tag{IV.7-15}$$

the response $\mathfrak{G}$ need not be defined except for arguments such that $|\det \mathbf{F}^t| = 1$.

Exercise IV.7.2 (Energy theorem for incompressible fluids in classical hydrodynamics). All motions of an incompressible, elastic fluid body $\mathscr{B}$ are mechanically perfect. When $\mathscr{B}$ is subject to conservative body force, the conclusions of Exercise III.6.5 show that

$$K + U = \text{const.} \tag{III.6-18$_r$}$$

if on $\partial\chi(\mathscr{B}, t)$ the pressure is everywhere constant or the velocity is everywhere tangential (a more general statement of this kind is given below in Volume 3).

8. Simple Bodies. Equations of Motion. Homogeneous Universal Transplacements, Motions, and Flows

A body $\mathscr{B}$ all of whose points are of a single, simple material is a *simple body*.

A simple body, be it unconstrained or constrained, is *homogeneous* if there is a reference placement $\boldsymbol{\kappa}$ such as to render the response $\mathfrak{G}$ in (IV.3-5) or (IV.7-15) independent of the reference position $\mathbf{X}$. Such a $\boldsymbol{\kappa}$ is a *homogeneous reference* placement. For an unconstrained body, the constitutive relation is (IV.3-5), and the equation of motion is

$$\rho\ddot{\mathbf{x}} = \operatorname{div}\mathfrak{G}(\mathbf{F}^t) + \rho\mathbf{b}. \tag{IV.8-1}$$

We think of **b** as given—typically, as being a constant vector or even **0**—and then (1) becomes *a condition on the transplacement* χ_κ. In the older theories this condition is a differential equation of second order in the time and the co-ordinates, separately or jointly. In general, it is a differential-functional equation which in view of the reduced form (IV.5-5) *is never linear in the derivatives with respect to spatial co-ordinates.* The resources of analysis are far from sufficient today to approach the general solution of initial-value or boundary-value problems stated through such equations except for a few particular kinds of body. Nevertheless, a great deal is known about particular solutions for particular classes of responses $\mathfrak{G}$, and the rest of this book is devoted to proof and explanation of some of these now known theorems of rational mechanics.

We have just made it plain that a constrained body is by no means a special case of an unconstrained one. Rather, the reverse holds, and the unconstrained body emerges as special. The behavior of a constrained body is not the same as that of any corresponding unconstrained one which happens to experience a motion satisfying the constraint. For example, if an unconstrained body happens to have been subjected to an isochoric transplacement history, the stress field on its present shape is determined by that history. An incompressible body, by definition, can never be subjected to anything but isochoric transplacement histories, but its stress field is never completely determined by them, being always indeterminate to the extent of an arbitrary hydrostatic pressure field. We shall see below in this section an example to show how this hydrostatic pressure field, to within a function of t only, can be determined by the principle of balance of linear momentum. More to this effect will be found near the end of Section IV.10.

Some recent writers on hydrodynamics are guilty of propagating bad English and hence confusion when they refer to "incompressible flows". No such carelessness occurs in the classic treatises of LAMB and MILNE-THOMSON. A flow, in any sense of the term, cannot be compressed. A flow may or may not be isochoric, and a fluid may or may not be incompressible; the behavior of an incompressible fluid in a certain, necessarily isochoric flow is generally not at all the same as that of any compressible fluid undergoing the same isochoric flow.

A constrained body is susceptible of a smaller class of deformations than is an unconstrained one. Corresponding to this restriction are certain arbitrary stresses, arbitrary in the sense that they are not determined by the deformation history. When we seek to determine whether or not a given deformation history of a constrained body be compatible with the axioms of mechanics and an assigned body force, the presence of these arbitrary stresses gives us greater freedom than for an unconstrained body undergoing the same transplacement history subject to the same body force. In this sense a single transplacement history satisfying a certain internal constraint will correspond with infinitely many different stress fields, provided it correspond with any at all. Roughly, we may say that while a constrained body is susceptible, by definition, to a restricted class of transplacement histories, it is easier to solve problems concerning those histories for a constrained body than for a corresponding unconstrained one. We shall frequently illustrate this evident but important fact.

The extreme case is furnished by the rigid body, whose allowed transplacements reduce to so special a class that the stress, whatever it may be, has no effect at all on the motion of the body, which can be determined by solving ordinary differential equations expressing no more than the principles of linear and rotational momentum for the whole body, with no reference to what the actions of its subbodies upon one another may or may not be.

The most useful constrained body is the incompressible one. To obtain the equation of motion for it, we substitute (IV.7-15) into CAUCHY's First Law and so obtain

$$\rho(\ddot{\mathbf{x}} - \mathbf{b}) = -\operatorname{grad} p + \operatorname{div} \mathfrak{G}(\mathbf{F}^t). \qquad \text{(IV.8-2)}$$

If a field p satisfies this equation, so also does $p + h$ for an arbitrary function h of t alone. This arbitrariness must arise because any uniform pressure applied to the boundary of an incompressible body exerts no resultant force or torque on that body.

If we suppose $\mathbf{b}$ given, χ_κ must satisfy (1) for an unconstrained body, (2) for an incompressible one. In the former all fields χ_κ are eligible to compete, and few will be found successful. For the latter, only those fields χ_κ such that $\det \mathbf{F}^t = 1$ are allowed, but the scalar field p may be adjusted to aid in finding a solution. The condition upon the motion alone is now

$$\operatorname{skw} \operatorname{grad}[\operatorname{div} \mathfrak{G}(\mathbf{F}^t) - \rho(\ddot{\mathbf{x}} - \mathbf{b})] = \mathbf{0}, \qquad \text{(IV.8-3)}$$

a differential-functional equation of order higher than that of (1). If this condition is satisfied, then locally $\rho(\ddot{\mathbf{x}} - \mathbf{b}) - \operatorname{div} \mathfrak{G}(\mathbf{F}^t)$ has a potential, from which the pressure p required to complete the solution of (2) is easy to obtain.

To see the effect of this difference, we restrict attention to *homogeneous, incompressible bodies*. The term *homogeneous* applied to an incompressible body will be taken to mean not only that the response $\mathfrak{G}_\kappa$ does not depend upon $\mathbf{X}$ but also that ρ_κ is an assigned, positive constant. The reference placement $\boldsymbol{\kappa}$ will then be called *homogeneous* for $\mathscr{B}$, and it is motions homogeneous with respect to such a $\boldsymbol{\kappa}$ that we shall consider. We shall write ρ for ρ_κ and v for $1/\rho_\kappa$. Furthermore, we shall assume the body force lamellar with potential ϖ.

With the definition

$$\phi := pv + \varpi, \tag{IV.8-4}$$

we reduce (2) to

$$\ddot{\mathbf{x}} = -\operatorname{grad}\phi + v \operatorname{div}\mathfrak{G}(\mathbf{F}^t). \tag{IV.8-5}$$

Because ϖ and v are given, p is determined from ϕ through (4). Referring back to (IV.7-15), we thus determine the entire stress field:

$$\mathbf{T} = -\rho(\phi - \varpi)\mathbf{1} + \mathfrak{G}(\mathbf{F}^t). \tag{IV.8-6}$$

We note that p and ϖ enter the equation of motion only through the combination denoted by ϕ. Suppose, now, that for a given incompressible body, that is, for a given response $\mathfrak{G}$ and a given density ρ, a certain isochoric transplacement history satisfy (5) with a certain pressure field p_1 and a certain field of body force having the potential ϖ_1. Let p_2 and ϖ_2 be any scalar fields such that

$$p_2 + \rho\varpi_2 = p_1 + \rho\varpi_1 . \tag{IV.8-7}$$

A glance at (4) shows that the equation of motion (5) is satisfied when p_1 and ϖ_1 are replaced by p_2 and ϖ_2. Thus we have the following theorem, patterned upon a conclusion and argument given by EULER for ideal fluids: *Let a homogeneous incompressible body of density ρ undergo a flow subject to pressure p_1 and body force having the potential ϖ_1. Then that body may undergo the same flow subject to pressure p_2 and body force having the potential ϖ_2 if*

$$p_2 = p_1 + \rho(\varpi_1 - \varpi_2). \tag{IV.8-8}$$

When $\mathbf{b} = 0$, the only forces applied to $\mathscr{B}$ from without are tractions upon the boundary $\partial\chi(\mathscr{B}, t)$. Thus we have the following corollary: *A flow of a homogeneous incompressible body is possible subject to some lamellar field of body force if and only if it is possible for that same body subject to surface tractions alone.* Indeed, for surface tractions to suffice it is necessary and sufficient, starting from the flow subject to p_1 and ϖ_1, that we put $\varpi_2 = 0$ and so obtain $p_2 = p_1 + \rho\varpi_1$.

Since no more than adjustment of the pressure field is needed to convert the solution of a problem in which there is no body force at all into one in which some lamellar body force is applied, there is little loss in generality in supposing that $\mathbf{b} = \mathbf{0}$ when we treat problems concerning homogeneous, incompressible bodies; also there is scant gain in doing so.

Exercise IV.8.1. Let two homogeneous, incompressible bodies have responses $\mathfrak{G}$ and $(v_1/v_2)\mathfrak{G}$ and specific volumes v_1 and v_2, respectively. Then if the former body can undergo a flow subject to the pressure p_1 and to the body force having the potential ϖ_1, the latter body can undergo the same flow with v_1, p_1, and ϖ_1 replaced by v_2, p_2, and ϖ_2 provided that

$$p_2 v_2 + \varpi_2 = p_1 v_1 + \varpi_1 . \tag{IV.8-9}$$

If, continuing to presume that $\mathbf{b}$ has a potential, we substitute (II.11-48) and (II.11-41) into (5), we obtain a useful form of the condition of integrability necessary and sufficient for ϕ to exist:

$$\rho \mathbf{W}_{\mathbf{a}} = \rho(\dot{\mathbf{W}} + \mathbf{D}\mathbf{W} + \mathbf{W}\mathbf{D}) = \operatorname{skw} \operatorname{grad} \operatorname{div} \mathfrak{G}(\mathbf{F}^t). \tag{IV.8-10}$$

When specialized to classical fluids this relation is often called "the vorticity equation". We shall use it much in Chapters VII and VIII.

Exercise IV.8.2. For a motion of the incompressible body whose response is $\mathfrak{G}$ to preserve circulation, it is necessary and sufficient that for the $\mathbf{F}^t$ giving rise to that motion

$$\operatorname{skw} \operatorname{grad} \operatorname{div} \mathfrak{G}(\mathbf{F}^t) = \mathbf{0} \tag{IV.8-11}$$

and hence that during that motion there be a scalar field λ such that

$$\operatorname{div} \mathfrak{G}(\mathbf{F}^t) = -\operatorname{grad} \lambda. \tag{IV.8-12}$$

Because of (5) and (II.11-47)

$$\lambda = \rho(P_{\mathbf{a}} - \phi), \qquad p = \rho(P_{\mathbf{a}} - \varpi) - \lambda, \tag{IV.8-13}$$

$P_{\mathbf{a}}$ being an acceleration-potential of the flow, and so use of (6) yields the following

Theorem (COLEMAN & TRUESDELL). *For the homogeneous, incompressible body whose response is $\mathfrak{G}$, let a certain flow that preserves circulation be possible, subject to null body force. Then that flow is possible also for* *arbitrary* ϖ, *and*

$$\mathbf{T} = -[\rho(P_{\mathfrak{a}} - \varpi) - \lambda]\mathbf{1} + \mathfrak{G}(\mathbf{F}^t). \qquad \text{(IV.8-14)}$$

In other words, if the incompressible body whose response is $\mathfrak{G}$ may undergo a flow that is possible for an Eulerian fluid, both $P_{\mathfrak{a}}$ and λ will exist for it in that flow, ϖ will be assigned, and the stress that the body will experience will be determined by (14) to within an arbitrary hydrostatic pressure dependent on time only.

A careless glance might suggest that (14) merely repeats (6). That is not so. The latter reflects a condition of integrability for (5), that is, for the scalar field ϕ, which is delivered by a theorem of existence and hence may be hard to determine simply. The former uses a theorem of existence for λ, a potential of div $\mathfrak{G}$ in the particular flow considered, while presuming the existence of a potential $P_{\mathfrak{a}}$ for $\ddot{\mathbf{x}}$. If an acceleration field is known, it is easy to determine whether $P_{\mathfrak{a}}$ exists, and if it exists, to calculate it is straightforward. Below in Sections IV.10 and IV.15 the student will see several examples of a potential $P_{\mathfrak{a}}$ calculated explicitly.

A transplacement or motion or flow is called *universal* for a given class of bodies subject to a given class of body forces $\mathbf{b}$ if it satisfies the corresponding equation of motion with such $\mathbf{b}$ for all those bodies. When $\mathbf{b}$ is assigned, any corresponding universal transplacement or motion or flow may be produced by bringing to bear *suitable surface tractions* upon the boundary of the body in question. These tractions will vary in general from one body of the given class to another. If they can be measured in an experiment, they will provide information about the material of which the body made to undergo the known transplacement consists.

Universal transplacements are centrally important because they suggest experiments in which the transplacement is known, at least approximately, from the outset. Then the analysis of experimental data is not complicated by the need to determine at the same time an unknown transplacement. Many of the particular solutions presented in textbooks of elasticity or fluid dynamics involve universal transplacements or motions. A famous example follows now.

A homogeneous, incompressible Eulerian fluid has an assigned, constant density and the constitutive relation (IV.7-13); thus for it $\mathfrak{G} = \mathbf{0}$; hence we may take $\lambda = 0$ in (12) through (14) and so obtain from the foregoing theorem the following celebrated

Corollary (EULER). *All flows of homogeneous incompressible Eulerian fluids subject to lamellar body force are universal; they are the isochoric flows that preserve circulation. The corresponding pressure fields* *are given by*

$$pv = P_{\mathfrak{a}} - \varpi. \tag{IV.8-15}$$

Simple logic shows that *if a motion is universal for a class of bodies and body forces, it must be universal for every subclass of that class*. For example, a flow that is universal for the class of homogeneous, incompressible bodies subject to lamellar body force is also universal for incompressible, homogeneous, Eulerian fluids subject to lamellar body force. Therefore, by the foregoing corollary, it must preserve circulation. Roughly speaking, the more general is the class of bodies and body forces, the fewer are its universal solutions. Finally, to show that a motion is not universal in some class, we need only exhibit one member of the class that does not satisfy the equation of motion defining that class.

9. Universal Homogeneous Transplacements of Unconstrained Bodies

The constitutive relation of an unconstrained body with respect to the reference placement κ is

$$\mathbf{T}(\mathbf{X}, t) = \mathfrak{G}_\kappa(\mathbf{F}^t_\kappa(\mathbf{X}), \mathbf{X}). \qquad \text{(IV.3-1)}_{2\text{r}}$$

Therefore, as we have explained in Section IV.3, the restriction of the constitutive mapping of a body to the histories of transplacements homogeneous with respect to κ determines its response to all transplacement histories altogether. Thus in an ideal program of experiment we should subject a body of given material to every transplacement of the form (*cf.* Section II.12)

$$\mathbf{x} = \chi_\kappa(\mathbf{X}, t) = \mathbf{x}_0(t) + \mathbf{F}(t)(\mathbf{X} - \mathbf{X}_0), \qquad \det \mathbf{F}(t) \neq 0, \quad \text{(II.12-1)}_\text{r}$$

and record the stresses obtained. The results would amount to a determination of the response $\mathfrak{G}_\kappa$. We now ask whether such a program be possible in principle.

Can the transplacement (II.12-1) be produced in a body of the material defined by (IV.3-1)? If the body force $\mathbf{b}$ in CAUCHY's First Law (III.6-1) is disposable, the answer is of course yes. In contrast, while in considering the totality of dynamical processes we saw no reason to exclude any $\mathbf{b}$, it is a different matter when we come to think about particular experiments, for only very special body forces are available in the laboratory. Practically speaking, a uniform field $\mathbf{b}$ = const. is all we are likely to be able to produce, unless we call upon electromagnetic forces, the effects of which are not taken up in this book. We then ask whether the homogeneous transplacement (II.12-1) can be produced in the body defined by $(\mathrm{IV}.3\text{-}1)_2$ if suitable surface tractions be supplied. We shall approach the problem only for homogeneous bodies.

In a homogeneous body undergoing a history of homogeneous transplacement, at each time the stress field $\mathbf{T}$ has the same value at every place, and so

$$\operatorname{div}\mathbf{T} = \mathbf{0}. \qquad \text{(IV.9-1)}$$

The question we now put is, if the value of $\mathbf{b}$ is constant, is it possible to supply boundary tractions such as to produce the homogeneous transplacement (II.12-1) of an unconstrained simple body? Substitution of (1) into the equation of motion (IV.8-1) yields the condition

$$\rho\mathbf{b} = \rho\ddot{\mathbf{x}}. \qquad \text{(IV.9-2)}$$

This requirement is compatible with the motion (II.12-1) if and only if

$$\ddot{\mathbf{F}} = \mathbf{0}, \qquad \ddot{\mathbf{x}}_0 = \mathbf{b}. \qquad \text{(IV.9-3)}$$

Hence

$$\mathbf{F}(t) = \mathbf{F}_0(\mathbf{1} + t\mathbf{F}_1), \qquad \mathbf{x}_0(t) = \tfrac{1}{2}t^2\mathbf{b} + t\mathbf{e} + \mathbf{f}, \qquad \text{(IV.9-4)}$$

$\mathbf{F}_0$ being an arbitrary, constant, invertible tensor, $\mathbf{F}_1$ an arbitrary, constant tensor, $\mathbf{e}$ an arbitrary vector, and $\mathbf{f}$ an arbitrary, fixed place.

Exercise IV.9.1. In an interval of t in which $\mathbf{1} + t\mathbf{F}_1$ is invertible

$$\mathbf{G} = \mathbf{F}_0\mathbf{F}_1(\mathbf{1} + t\mathbf{F}_1)^{-1}\mathbf{F}_0^{-1}. \qquad \text{(IV.9-5)}$$

The foregoing analysis shows that *for unconstrained homogeneous bodies subject to constant body force, the homogeneous transplacements* (II.12-1) *are possible if and only if they satisfy the restrictions* (4). *If they are possible, they are universal.*

Therefore, the ideal program of determining $\mathfrak{G}_\kappa$ by effecting all homogeneous transplacements from κ cannot be carried out. This conclusion does not mean that no method of determining $\mathfrak{G}_\kappa$ may be found but merely that the vista of homogeneous transplacements, used to interpret the *definition* of a simple material, is not feasible for finding $\mathfrak{G}_\kappa$ by experiment without use of artificial body forces.

Once an $\mathbf{F}$ of the form $(4)_1$ be selected, by substituting $\mathbf{F}^t$ into $(\text{IV}.3\text{-}1)_2$ we obtain the stress required to effect the resulting universal transplacement in the body whose response is $\mathfrak{G}_\kappa$. To make bodies of different materials undergo one and the same universal transplacement, generally different fields of stress must be produced in them, and different fields of stress will act in their interiors. Thus universal motions bring into relief the effects of different constitutive relations.

Motions satisfying (4) generally exist only for a finite interval of time. By assumption, $\det \mathbf{F}(0) = \det \mathbf{F}_0 \neq 0$, and so (4) makes $\mathbf{F}$ invertible only so long as

$$\det(\mathbf{1} + t\mathbf{F}_1) \neq 0, \tag{IV.9-6}$$

that is, in an interval of time $]t_-, t_+[$ containing 0 and such that $-1/t$ never equals a proper number of $\mathbf{F}_1$. Since $\mathbf{F}_1$ is an arbitrary tensor, perhaps singular, nothing can be said in general about its proper numbers. The possibilities that $t_- = -\infty$ or $t_+ = +\infty$ are not excluded. *E.g.*, in an isochoric motion of this class,

$$|\det \mathbf{F}_0| = 1, \qquad \det(\mathbf{1} + t\mathbf{F}_1) = 1, \tag{IV.9-7}$$

and the interval in which the motion exists is $]-\infty, +\infty[$.

For some particular materials, limitation of t to a finite interval would not matter. Examples are the elastic materials and the linearly viscous fluids, defined above in Section IV.4. In contrast, the general constitutive relation (IV.3-5) for simple materials refers to the entire history $\mathbf{F}^t$, and so for general considerations $\mathbf{F}_1$ must be such that $t_- = -\infty$.

Exercise IV.9.2. The motions defined by (4) are isochoric if and only if

$$|\det \mathbf{F}_0| = 1, \qquad \operatorname{tr} \mathbf{F}_1 = 0, \qquad \operatorname{tr} \mathbf{F}_1^2 = 0, \qquad \det \mathbf{F}_1 = 0. \tag{IV.9-8}$$

A counterexample shows that $\mathbf{F}_1^2$ need not equal $\mathbf{0}$.

If $\mathbf{b} = \mathbf{0}$, we may rephrase the main conclusion from (3) and $(\text{II}.12\text{-}17)_2$ as follows: *A homogeneous transplacement is universal for unconstrained, homogeneous bodies if and only if it is accelerationless.*

An important example is furnished by *steady simple shearing*, which has been used traditionally to illustrate various special theories in continuum mechanics. Cartesian components of the flow are given by (II.11-17). In a suitable pair of cartesian systems, one on $\boldsymbol{\kappa}(\mathscr{B})$ and one on $\boldsymbol{\chi}(\mathscr{B}, t)$, the components of the transplacement are

$$\begin{aligned} x_1 &= X_1, \\ x_2 &= X_2 + \kappa t X_1, \\ x_3 &= X_3. \end{aligned} \tag{IV.9-9}$$

Thus

$$[\mathbf{F}] = \begin{Vmatrix} 1 & 0 & 0 \\ \kappa t & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \qquad \mathbf{b} = \mathbf{0}. \tag{IV.9-10}$$

and so

$$\mathbf{F}_0 = \mathbf{1}, \qquad [\mathbf{F}_1] = \kappa \begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \tag{IV.9-11}$$

and (8) is satisfied. In fact, $\mathbf{F}_1^2 = \mathbf{0}$. Therefore, steady, simple shearing is a universal flow for homogeneous, unconstrained bodies; it arises from a universal transplacement with respect to a homogeneous reference placement.

Another example is furnished by a homogeneous, irrotational, pure stretch: $\mathbf{R} = \mathbf{1}$, $\mathbf{W} = \mathbf{0}$, $\mathbf{U} = \mathbf{U}(t)$. From (II.11-26)$_2$ we see that $\mathbf{U}$ must satisfy the differential equation

$$\dot{\mathbf{U}}\mathbf{U} = \mathbf{U}\dot{\mathbf{U}}. \tag{IV.9-12}$$

Exercise IV.9.3. The condition (12) holds if $\mathbf{U}$ has an orthogonal triad of proper vectors $\mathbf{e}_k$ which are constant in time. Then

$$\mathbf{U} = \sum_{k=1}^{3} u_k(t)\mathbf{e}_k \otimes \mathbf{e}_k. \tag{IV.9-13}$$

The corresponding homogeneous, pure stretch has constant acceleration if and only if $\ddot{\mathbf{x}}_0 = \text{const.}$ and the u_k are positive, affine functions of t. A rectangular block with faces normal to the $\mathbf{e}_i$ is transformed by this motion into another such block at any time within the interval for which the motion exists. This motion is isochoric if and only if it reduces to a translation.

The two families of motions just exhibited are interesting members of the class of universal motions for homogeneous bodies subject to constant body force.

The class of body forces for which some homogeneous transplacements are universal for homogeneous, unconstrained bodies is very limited, as the following exercise shows.

Exercise IV.9.4. Use of (II.12-17) shows that for the homogeneous transplacement (II.12-1) to be possible in *some one* homogeneous, unconstrained body, $\mathbf{b}$ must satisfy the following condition:

$$\mathbf{b}(\mathbf{x}, t) = \mathbf{B}(t)(\mathbf{x} - \mathbf{x}_0) + \mathbf{b}_0(t). \qquad \text{(IV.9-14)}$$

For such a $\mathbf{b}$ the homogeneous transplacements (II.12-1) *are possible and hence universal for homogeneous, unconstrained bodies* if and only if

$$\ddot{\mathbf{F}} = \mathbf{BF}, \qquad \operatorname{skw} \mathbf{B} = \mathbf{0}, \qquad \ddot{\mathbf{x}}_0 = \mathbf{b}_0 . \qquad \text{(IV.9-15)}$$

As we shall see presently, the class of body forces compatible with universal motions for homogeneous, incompressible bodies is much greater, and the class of universal motions, while of course comprising only isochoric ones, is otherwise much broader.

10. Universal Homogeneous Transplacements of Incompressible Bodies

We now determine all homogeneous transplacements that are possible, and hence universal, for homogeneous, incompressible bodies subject to lamellar body force. The apparatus for obtaining universal solutions for incompressible bodies has been provided above in the statement of Exercise IV.8.2. In any homogeneous transplacement of an incompressible body, the determinate stress $\mathbf{S}$ is a function of t alone, and so

$$\operatorname{div} \mathbf{S} = \mathbf{0}; \qquad \text{(IV.10-1)}$$

therefore we take λ as 0 in (IV.8-12) and (IV.8-14). All that remains is to ensure that the flow preserves circulation. Therefore a *homogeneous, isochoric transplacement is possible, subject to boundary tractions alone, in every homogeneous, incompressible body if and only if* $\mathbf{F}$ *satisfies the condition*

$$\operatorname{skw}(\ddot{\mathbf{F}}\mathbf{F}^{-1}) = \mathbf{0}. \qquad \text{(II.11-46)}_{1r}$$

If $\mathbf{F}$ satisfies this differential equation, inspection of (II.12-17)$_{3,4}$ delivers the acceleration-potential $P_{\mathbf{a}}$:

$$\begin{aligned} -P_{\mathbf{a}} &= (\mathbf{x}-\mathbf{x}_0)\cdot[\tfrac{1}{2}\ddot{\mathbf{F}}\mathbf{F}^{-1}(\mathbf{x}-\mathbf{x}_0)+\ddot{\mathbf{x}}_0], \\ &= (\mathbf{x}-\mathbf{x}_0)\cdot[\tfrac{1}{2}(\dot{\mathbf{G}}+\mathbf{G}^2)(\mathbf{x}-\mathbf{x}_0)+\ddot{\mathbf{x}}_0], \end{aligned} \qquad \text{(IV.10-2)}$$

and (IV.8-14) yields the stress:

$$\begin{aligned} \mathbf{T} &= \rho[(\mathbf{x}-\mathbf{x}_0)\cdot(\tfrac{1}{2}\ddot{\mathbf{F}}\mathbf{F}^{-1}(\mathbf{x}-\mathbf{x}_0)+\ddot{\mathbf{x}}_0)+\varpi]\mathbf{1}+\mathfrak{G}(\mathbf{F}^t), \\ &= \rho[(\mathbf{x}-\mathbf{x}_0)\cdot[(\tfrac{1}{2}(\dot{\mathbf{G}}+\mathbf{G}^2)(\mathbf{x}-\mathbf{x}_0)+\ddot{\mathbf{x}}_0)+\varpi]\mathbf{1}+\mathfrak{G}(\mathbf{F}^t). \end{aligned} \qquad \text{(IV.10-3)}$$

Any unimodular solution $\mathbf{F}$ of (II.11-46)$_1$ and any point-valued function $\mathbf{x}_0$ if put into (II.12-1) yield a possible transplacement, and then substitution into (3) delivers the stress required to produce it, subject to the action of the body force with potential ϖ. The student will recall that for these motions the value of $\mathfrak{G}$ is a function of t only; *cf.* (1).

The solution of Exercise II.9.1 shows that all rigid motions are homogeneous, while Exercise II.11.17 implies that a rigid motion satisfies (II.11-46) if and only if its spin is steady.

Exercise IV.10.1. For a rotation with steady spin the expression (2) for $P_{\mathbf{a}}$ reduces to (II.11-52).

We proceed to determine the homogeneous motions that are irrotational. If we are given a homogeneous stretch history $\mathbf{U}^t$, we may set $\mathbf{W}=\mathbf{0}$ in (II.11-26) and integrate the resulting ordinary differential equation for $\mathbf{R}$. In this way we can determine a rotation history $\mathbf{R}^t$ such that the flow corresponding with $\mathbf{R}^t\mathbf{U}^t$ is irrotational. As we remarked in Section II.13, every irrotational flow preserves circulation. If $\det\mathbf{U}(t)=1$, the result demonstrated above shows that the motion just determined in principle can be produced in any homogeneous, incompressible body by applying suitable boundary tractions. Consequently, the ideal experimental program proposed initially can be achieved, for homogeneous, incompressible bodies, without calling upon artificial body forces, in fact without use of any body force at all, and by considering only irrotational histories.

More generally, starting from any homogeneous, pure stretch, we can construct homogeneous flows that preserve circulation. To do so, we substitute CAUCHY's criterion (II.11-44) into (II.11-26)$_2$ and so find that

$$\dot{\mathbf{R}}=\mathbf{R}\mathbf{Y}, \qquad \text{(IV.10-4)}$$

$\mathbf{Y}$ being defined as follows:

$$\mathbf{Y} := \tfrac{1}{2}(\mathbf{U}^{-1}\dot{\mathbf{U}} - \dot{\mathbf{U}}\mathbf{U}^{-1}) + \mathbf{U}^{-1}\mathbf{W}_\kappa\mathbf{U}^{-1}. \qquad \text{(IV.10-5)}$$

Suppose now a homogeneous stretch $\mathbf{U}$ and an arbitrary spin $\mathbf{W}_\kappa$ in the reference shape be given. Then $\mathbf{Y}$ is a known function of t. If $\mathbf{U}$ and $\mathbf{W}_\kappa$ are such that $\mathbf{Y}$ is continuous, the first-order linear differential equation (4) determines a unique rotation $\mathbf{R}(t)$ corresponding with any assigned initial rotation $\mathbf{R}(0)$. Therefore, the homogeneous motion whose deformation gradient is $\mathbf{R}(t)\mathbf{U}(t)$ preserves circulation. The theorem established near the end of Section IV.8 rests upon assuming that a particular flow preserves circulation. Because we have now exhibited the entire class of homogeneous motions that do so, we may apply the theorem to each of those motions and so obtain the following

Theorem (COLEMAN & TRUESDELL). *By applying suitable boundary tractions alone, it is possible to cause any homogeneous, incompressible body to undergo any desired isochoric, homogeneous stretch history* $\mathbf{U}^t$. *The corresponding rotation history* $\mathbf{R}^t$, *which is independent of the material, is obtained from the unique solution of* (4) *corresponding with assigned initial values* $\mathbf{R}(0)$ *and* $\mathbf{W}_\kappa$. *Conversely, the only homogeneous transplacements that can be effected in all homogeneous, incompressible bodies by the application of boundary tractions and lamellar body force are those in which* $\mathbf{R}$ *is determined from* $\mathbf{U}$, $\mathbf{R}(0)$, *and* $\mathbf{W}_\kappa$ *by* (4).

Putting $\mathbf{W}_\kappa = \mathbf{0}$ in the foregoing theorem, we recover the statement about irrotational histories proved just after Exercise IV.10.1. Clearly pure stretch histories do not suffice to achieve the ideal experimental program since $\mathbf{R} = \mathbf{1}$ is not generally a solution of (4).

Exercise IV.10.2. A pure stretch preserves circulation if and only if

$$\dot{\mathbf{U}}\mathbf{U} - \mathbf{U}\dot{\mathbf{U}} = \text{const.}, \qquad \text{(IV.10-6)}$$

a condition more general than (IV.9-12). Hence, in general, a homogeneous, isochoric, pure stretch cannot be produced in an arbitrary homogeneous, incompressible simple body by the effect of boundary tractions alone. Among those special homogeneous, isochoric, pure stretches that can be so produced are the irrotational ones. The class of homogeneous, isochoric, irrotational, pure stretch histories includes[1] those given by (IV.9-13) with the added restriction $u_1(t)u_2(t)u_3(t) = 1$.

[1]The wording here clarifies that following Equation (30.37) in NFTM and in the paper of COLEMAN & TRUESDELL cited on p. 73 of NFTM.

Exercise IV.10.2 shows that any homogeneous, isochoric, irrotational pure stretch history can be produced by the action of surface tractions alone or of any lamellar body force, in any homogeneous, incompressible body. Exemplifying the general discourse shortly after the beginning of Section IV.8, this conclusion illustrates the difference between the stress system in a compressible body that just happens to undergo an isochoric motion and that in a corresponding incompressible body undergoing a motion with the same transplacement gradient and the same response. For the unconstrained body, change of volume is avoided because the stresses are selected in just the right way, and that way is specified uniquely by the response $\mathfrak{G}$. For the incompressible body, no system of stresses can produce any motion but an isochoric one, and corresponding with that fact there is a hydrostatic pressure which is arbitrary in the sense that it is not determined by the history of the transplacement gradient but is determined, to within a time-dependent hydrostatic pressure, by the balance of linear momentum and is exhibited in (3).

When given body forces are applied, CAUCHY's First Law restricts that arbitrary pressure but does not determine it uniquely. In this sense a single isochoric transplacement history if possible at all for a given incompressible body is possible subject to infinitely many different body forces and surface tractions.

Exercise IV.10.3. Suppose that $\mathfrak{G}$ be the response of a certain unconstrained simple body $\mathscr{B}$, and that the restriction of $\mathfrak{G}$ to isochoric transplacement histories be the response of a certain incompressible simple body $\mathscr{B}_0$. How does the stress system required to effect a certain simple shearing in $\mathscr{B}$ differ from that required to effect just the same simple shearing in $\mathscr{B}_0$?

Internal constraints such as incompressibility reduce the class of possible motions but otherwise expand the class of stresses compatible with such motions as may take place. The theory of a constrained body is therefore essentially easier to work out. The far-reaching simplification that results from assuming the material to be incompressible was seen and exploited by RIVLIN in his pioneering researches on non-linear continuum theories in 1946–1955. Most of the explicit solutions known today concern incompressible bodies; several were discovered by RIVLIN and his associates.

11. Material Isomorphisms

Up to now we have considered the constitutive relation of a single material point, or a single homogeneous body made up of material points all having the same response relative to a given reference placement $\boldsymbol{\kappa}$. When can we say that two body-points X_1 and X_2 of $\mathscr{B}$ are of the *same* material? When it is possible to bring small portions $\mathscr{P}_1$ and $\mathscr{P}_2$ of $\mathscr{B}$ containing X_1 and X_2 into reference shapes $\boldsymbol{\kappa}_1(\mathscr{P}_1)$ and $\boldsymbol{\kappa}_2(\mathscr{P}_2)$ such that any subsequent history of transplacement gives rise to exactly the same stress at the places $\mathbf{x}_1$ and $\mathbf{x}_2$ occupied by X_1 and X_2. Thus no experimental measurement of stress as determined by transplacement histories can detect whether we started with the part $\mathscr{P}_1$ containing $\mathbf{X}_1$ in $\boldsymbol{\kappa}_1(\mathscr{B})$ or the part $\mathscr{P}_2$ containing $\mathbf{X}_2$ in $\boldsymbol{\kappa}_2(\mathscr{B})$, it being understood that $\mathbf{X}_1 = \boldsymbol{\kappa}_1(X_1)$, $\mathbf{X}_2 = \boldsymbol{\kappa}_2(X_2)$. This interpretation suggests also that we should require the densities $\rho_{\boldsymbol{\kappa}_1}$ and $\rho_{\boldsymbol{\kappa}_2}$ to be equal and uniform near X_1 and X_2, as we shall.

To render this idea formal, we erect the following

Definition (NOLL). *Let* $\mathfrak{G}_{\boldsymbol{\kappa}}$ *be the response of a simple material with respect to the reference placement* $\boldsymbol{\kappa}$. *The points* X_1 *and* X_2 *of* $\mathscr{B}$ *are* materially isomorphic *if there are reference placements* $\boldsymbol{\kappa}_1$ *and* $\boldsymbol{\kappa}_2$ *such that* $\rho_{\boldsymbol{\kappa}_1} = \rho_{\boldsymbol{\kappa}_2} = \text{const.}$ *near* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *and*

$$\mathfrak{G}_{\boldsymbol{\kappa}_1}(\mathbf{F}^t, \mathbf{X}_1) = \mathfrak{G}_{\boldsymbol{\kappa}_2}(\mathbf{F}^t, \mathbf{X}_2) \tag{IV.11-1}$$

for every transplacement history $\mathbf{F}^t$ *in the domains of* $\mathfrak{G}_{\boldsymbol{\kappa}_1}$ *and* $\mathfrak{G}_{\boldsymbol{\kappa}_2}$, *respectively.*

This definition embodies the idea just stated informally, for the value of the left-hand side is the stress at the place occupied by X_1 when the material points constituting $\mathscr{B}$ have been subjected to a history of transplacement $\mathbf{F}^t$ with respect to $\boldsymbol{\kappa}_1(\mathscr{B})$, while the right-hand side is the stress at the place occupied by X_2 when the material points constituting $\mathscr{B}$ have been subjected to just the same history of transplacement with respect to $\boldsymbol{\kappa}_2(\mathscr{B})$. Since (1) must hold for all $\mathbf{F}^t$ in the domains of the two responses considered, we can bring the parts $\mathscr{P}_1$ and $\mathscr{P}_2$ of the body that contain X_1 and X_2, respectively, into shapes indistinguishable by any measurement of stress.

If each body-point is materially isomorphic to every other one, then every sufficiently small part of $\mathscr{B}$ has just the same properties as every other sufficiently small part, and we say that the body is *uniform*. Now this quality requires that the responses of $\mathscr{B}$ at X_1 and X_2 be the same with respect to suitable reference placements $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$, that the responses of $\mathscr{B}$ at X_2 and X_3 be the same with respect to suitable reference placements $\boldsymbol{\kappa}_2'$ and $\boldsymbol{\kappa}_3'$, *etc*. There need be no single reference placement $\boldsymbol{\kappa}$ such that all the material points making up $\mathscr{B}$ have one and the same response: $\mathfrak{G}_{\boldsymbol{\kappa}}(\cdot, \mathbf{X}) = \mathfrak{G}_{\boldsymbol{\kappa}}(\cdot, \mathbf{Y})\ \forall \mathbf{X}, \mathbf{Y} \in \boldsymbol{\kappa}(\mathscr{B})$. In order to demonstrate the isomorphism of each pair of body-points it may be necessary (in imagination, of course) to cut the body into small pieces and bring each piece separately into an appropriate shape before beginning the experiment. These small pieces need not fit together to form a shape of all of $\mathscr{B}$.

If the isomorphism of all the points of a uniform body may be demonstrated by use of a single reference placement κ, the body is *homogeneous*. The response $\mathfrak{G}_\kappa$ with respect to this particular κ is independent of $\mathbf{X}$, and ρ_κ is likewise independent of $\mathbf{X}$, so that the definition of "homogeneous" in terms of the concept of material isomorphism is equivalent to the one we have introduced already at the beginning of Section IV.8.

While every homogeneous body is uniform, the converse is false. Uniform but inhomogeneous bodies seem to correspond in some cases with what in physics are called bodies with "defects" and "dislocations". In this book, henceforth, we shall consider only homogeneous bodies.[1]

The concept of material isomorphism is of far greater use than merely to define homogeneity, as we shall now see.

12. The Peer Group

Trivially, every point X of $\mathscr{B}$ is materially isomorphic to itself, but there may be also non-trivial isomorphisms of X with itself. We shall analyse this possibility by the aid of an arbitrarily selected reference placement κ_1, and since we shall consider now a single body-point X, we shall drop X from the notation. Thus (IV.11-1) yields the condition

$$\mathfrak{G}_{\kappa_1}(\mathbf{F}^t) = \mathfrak{G}_{\kappa_2}(\mathbf{F}^t). \qquad \text{(IV.12-1)}$$

If we can find a κ_2 distinct from κ_1 such that (1) holds for all $\mathbf{F}^t$ in the respective domains, we shall have shown that the response of the given body-point X is just the same in deformations with respect to two distinct reference placements. That is, in terms of the ideal experiments we sometimes invoke so as to visualize the assertions of the theory, no measurement of stress on the part of $\mathscr{B}$ near X can distinguish κ_2 from κ_1. Thus the reference placements κ_1 and κ_2 are *peers* at X.

If we choose a different reference placement, say κ^*, then $\mathfrak{G}_{\kappa^*}$ will generally determine a different set of peers.

The set of gradients at $\kappa(X)$ of transplacements carrying κ into its peers

[1] The general theory and solutions of particular problems concerning inhomogeneous, uniform, simple bodies are presented in the book republishing memoirs by W. Noll, R. A. Toupin, and C.-C. Wang, *Continuum Theory of Inhomogeneities in Simple Bodies*, Berlin and New York, Springer-Verlag, 1968, and also in Chapters V and VI of IRE.

forms a group called the *peer group*[1] $\mathscr{g}_\kappa$. Because material isomorphisms leave the mass density assigned by $\boldsymbol{\kappa}$ to X unchanged, the gradient $\mathbf{P}$ of a transplacement delivering peers is unimodular: $\det \mathbf{P} = \pm 1$. Thus the peer group $\mathscr{g}_\kappa$ of $\boldsymbol{\kappa}$ at X is a subgroup of the unimodular group $\mathscr{u}$:

$$\mathscr{g}_\kappa \subset \mathscr{u}. \tag{IV.12-2}$$

It is the group of gradients of *all maps that carry* $\boldsymbol{\kappa}$ *into its peers*, namely, the reference placements indistinguishable from $\boldsymbol{\kappa}$ by measurements of stress arising from deformation of parts of $\boldsymbol{\kappa}(\mathscr{B})$.

By substituting (IV.3-3) into (1), we find that the elements of the peer group $\mathscr{g}_\kappa$ are unimodular tensors $\mathbf{H}$ such that for all histories $\mathbf{F}^t$ in the domain of $\mathfrak{G}_\kappa$

$$\mathfrak{G}_\kappa(\mathbf{F}^t\mathbf{H}) = \mathfrak{G}_\kappa(\mathbf{F}^t), \tag{IV.12-3}$$

and conversely, any such $\mathbf{H}$ is an element of $\mathscr{g}_\kappa$. Here we assume that if the domain of $\mathfrak{G}_\kappa$ includes $\mathbf{F}^t$, it is large enough to include also the products $\mathbf{F}^t\mathbf{H}$ for every unimodular $\mathbf{H}$.

We have called the set of peers a group, but we have not yet shown that it deserves that appellation.

Exercise IV.12.1. The collection of solutions $\mathbf{H}$ of (3) forms a group.

As a part of the definition of the peer group we have required that its members be unimodular. We have done so in favor of the intended application rather than for any mathematical block against more general isomorphisms. By considering in (3) the case of the rest history $\mathbf{F}(t) := \mathbf{1}$, we see that if $\mathbf{H} \in \mathscr{g}_\kappa$ and $n = 1, 2, 3, \ldots$, then $\mathfrak{G}_\kappa((\mathbf{H}^n)^t) = \mathfrak{G}_\kappa(\mathbf{1}^t)$; here $(\mathbf{H}^n)^t$ denotes the history of the constant tensor $\mathbf{H}^n$, and $\mathbf{1}^t$ denotes the history of $\mathbf{1}$. If $|\det \mathbf{H}| < 1$, this conclusion and (II.5-4) show that we can find a placement which has arbitrarily large density and in which a part of the body can be held at rest indefinitely under just the same stress as that required for equilibrium in $\boldsymbol{\kappa}$. If $|\det \mathbf{H}| > 1$, the same can be said for a placement with arbitrarily small density. Such a material would be a strange one. In particular, no Eulerian fluid with invertible pressure function (*cf.* IV.4-4) is of this kind. In this book we merely leave out of account any $\mathbf{H}$ that satisfies (3) and is not unimodular, but the foregoing remarks would lend support

[1] The term "isotropy group", used by Noll in introducing these groups, is misleading here because it derives from the concept of turning, while the elements of the peer group need not all be rotations; "symmetry", while closer to the popular speech of physicists, would be equally misleading because it derives from the concept of distance, which is irrelevant in material response. The term "peer" is intended to suggest its root meaning, which is "equal in status before the law", the "law" being here the constitutive relation of the material.

to *requiring*, as part of the definition of simple material, that $\mathfrak{G}_\kappa$ allow no solutions $\mathbf{H}$ of (3) that are not unimodular.[1]

The members $\mathbf{H}$ of $\mathscr{g}_\kappa$ need not be orthogonal, but they may be. Since $\mathbf{1} \in \mathscr{g}_\kappa$ for every material point and every reference placement κ, at least one member of $\mathscr{g}_\kappa$ is orthogonal. If an orthogonal tensor $\mathbf{Q} \in \mathscr{g}_\kappa$, then also $\mathbf{Q}^\mathsf{T} \in \mathscr{g}_\kappa$ since $\mathscr{g}_\kappa$ is a group; also, if $\mathbf{F}^t$ runs over all invertible tensor histories, so does $\mathbf{QF}^t$. Thus, when $\mathbf{H} = \mathbf{Q}^\mathsf{T}$, (3) is equivalent to

$$\mathfrak{G}_\kappa(\mathbf{QF}^t\mathbf{Q}^\mathsf{T}) = \mathfrak{G}_\kappa(\mathbf{QF}^t). \tag{IV.12-4}$$

In the condition (IV.5-2), which expresses the Principle of Material Frame-Indifference, we select the particular history $\mathbf{Q}^t = \mathbf{Q}(t) = \mathbf{Q}$ and obtain

$$\mathfrak{G}_\kappa(\mathbf{QF}^t) = \mathbf{Q}\mathfrak{G}_\kappa(\mathbf{F}^t)\mathbf{Q}^\mathsf{T}. \tag{IV.12-5}$$

This relation holds for all $\mathbf{F}^t$ and for all orthogonal tensors $\mathbf{Q}$, while (4) holds only for those $\mathbf{Q}$ that belong to $\mathscr{g}_\kappa$. Combining the two relations yields

$$\mathfrak{G}_\kappa(\mathbf{QF}^t\mathbf{Q}^\mathsf{T}) = \mathbf{Q}\mathfrak{G}_\kappa(\mathbf{F}^t)\mathbf{Q}^\mathsf{T} \tag{IV.12-6}$$

as a necessary condition to be satisfied by all orthogonal members of $\mathscr{g}_\kappa$.

Exercise IV.12.2. Conversely, if $\mathbf{Q}$ satisfies (6), then $\mathbf{Q} \in \mathscr{g}_\kappa$.

Thus (6) is *a necessary and sufficient condition for the orthogonal tensor* $\mathbf{Q}$ *to belong to the peer group.*

From (6) we see that $-\mathbf{1} \in \mathscr{g}_\kappa$ for all materials and all κ. Since $-\mathbf{1}$ is a central inversion, it does not correspond with any deformation that could be effected physically but merely expresses the invariance of material properties under reflections of the reference placement.[2] Since $-\mathbf{1} \in \mathscr{g}_\kappa$ and $\mathscr{g}_\kappa$ is a group, $-\mathbf{H} \in \mathscr{g}_\kappa \Leftrightarrow \mathbf{H} \in \mathscr{g}_\kappa$. Thus $\mathscr{g}_\kappa$ can be expressed as the direct product of the trivial group consisting in $\mathbf{1}$ and $-\mathbf{1}$ alone and a group $\mathscr{g}_\kappa^+$ all of whose

[1] In a theory of thermomechanics it is possible to define peer groups and to prove that in order to satisfy certain reasonable requirements they must be subgroups of $\mathscr{u}$, as has been shown by M. E. Gurtin & W. O. Williams, "On the inclusion of the complete symmetry group in the unimodular group," *Archive for Rational Mechanics and Analysis* **23** (1966/7), 163–172 (1966).

[2] The reader should not extrapolate this statement to other theories such as those of heat conduction and electromagnetism; in them there is no such invariance, because the transplacement gradient $\mathbf{F}$ is not the only independent variable in the constitutive relations.

members have determinant $+\mathbf{1}$:

$$\mathscr{g}_\kappa = \{\mathbf{1}, -\mathbf{1}\} \otimes \mathscr{g}_\kappa^+, \tag{IV.12-7}$$

and only the elements of $\mathscr{g}_\kappa^+$ can be interpreted in terms not only of change of reference placement but also as gradients of transplacements that map one shape of a given body onto another. These are the transplacements that cannot be distinguished from one another by mental measurement,[1] but it is formally more convenient to retain the trivial central inversions and so operate with $\mathscr{g}_\kappa$ itself. We have shown, then, that $\{\mathbf{1}, -\mathbf{1}\}$ is the smallest possible peer group:

$$\{\mathbf{1}, -\mathbf{1}\} \subset \mathscr{g}_\kappa \subset \mathscr{u}. \tag{IV.12-8}$$

The foregoing constructions and conclusions in these precise, abstract forms were introduced by NOLL, generalizing earlier and more special notions.

Any subgroup of the unimodular group that includes $\{\mathbf{1}, -\mathbf{1}\}$ may be the peer group of a material point. Corresponding with any assigned unimodular subgroup $\mathscr{g}$, it is possible to construct infinitely many responses $\mathfrak{G}$; more specifically, it is possible to write $\mathfrak{G}$ in a reduced form such as to be frame-indifferent and to include automatically *all* materials having an assigned peer group, and these only.[2] In the following sections we shall consider only such $\mathscr{g}$ as are notable or lead to especially simple representations for $\mathfrak{G}$. In particular, we shall use the ideas and apparatus just given so as to define the concepts of "fluid", "solid", and "isotropic".

13. Comparison of Peer Groups with Respect to Different Reference Placements

The peer group $\mathscr{g}_\kappa$ at a material point depends, as does the response $\mathfrak{G}_\kappa$ of the material, upon the choice of reference placement κ. Since $\mathfrak{G}_{\kappa_1}$ determines $\mathfrak{G}_{\kappa_2}$ for all κ_2, the same should be true of $\mathscr{g}_{\kappa_1}$ and $\mathscr{g}_{\kappa_2}$. That is so, and either group determines the other through a rule found by NOLL:

$$\mathscr{g}_{\kappa_2} = \mathbf{P}\mathscr{g}_{\kappa_1}\mathbf{P}^{-1}. \tag{IV.13-1}$$

[1] Again the reader must be warned that while this fact expresses a proved theorem of the theory presented in this book, nothing of the sort holds for the peer groups that can be defined by parallel constructions in other theories such as optics.

[2] C.-C. WANG, "On a general representation theorem for constitutive relations," *Archive for Rational Mechanics and Analysis* **33** (1969), 1–25.

To prove this rule, we simply apply (IV.3-3) to each member of (IV.12-3) after replacing $\boldsymbol{\kappa}$ therein by $\boldsymbol{\kappa}_1$:

$$\mathfrak{G}_{\boldsymbol{\kappa}_2}(\mathbf{F}^t\mathbf{H}\mathbf{P}^{-1}) = \mathfrak{G}_{\boldsymbol{\kappa}_2}(\mathbf{F}^t\mathbf{P}^{-1}); \qquad \text{(IV.13-2)}$$

here $\mathbf{P} := \nabla\boldsymbol{\lambda}$ and $\boldsymbol{\lambda} := \boldsymbol{\kappa}_2 \circ \boldsymbol{\kappa}_1^{-1}$. As $\mathbf{F}^t$ runs over all invertible tensor histories, so does $\mathbf{F}^t\mathbf{P}^{-1}$ for any assigned invertible tensor $\mathbf{P}$. Hence (2) is equivalent to

$$\mathfrak{G}_{\boldsymbol{\kappa}_2}(\mathbf{F}^t\mathbf{P}\mathbf{H}\mathbf{P}^{-1}) = \mathfrak{G}_{\boldsymbol{\kappa}_2}(\mathbf{F}^t), \qquad \text{(IV.13-3)}$$

which is of the form (IV.12-3) with $\boldsymbol{\kappa}$ replaced by $\boldsymbol{\kappa}_2$. Since $\mathbf{PHP}^{-1}$ is unimodular if $\mathbf{H}$ is, every solution $\mathbf{H}$ of (IV.12-3) corresponds with a unimodular solution $\mathbf{PHP}^{-1}$ of (3), and conversely. NOLL's rule (1) is an abbreviated statement of this fact.

It is a trivial consequence of (IV.12-1) that *if $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ are peers, they have the same peer groups.*

While the members of $\mathscr{G}_{\boldsymbol{\kappa}_1}$ and $\mathscr{G}_{\boldsymbol{\kappa}_2}$ are unimodular tensors, the reference placements $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ themselves need not have the same density. In particular, if we let $\boldsymbol{\kappa}_2$ be obtained from $\boldsymbol{\kappa}_1$ by a dilatation, then $\mathbf{P} = K\mathbf{1}$ and $K \neq 0$, and so $\mathbf{P}^{-1} = K^{-1}\mathbf{1}$. Therefore (1) yields $\mathscr{G}_{\boldsymbol{\kappa}_2} = \mathscr{G}_{\boldsymbol{\kappa}_1}$. *Thus the peer group is unaltered by a dilatation.*

Whatever be $\mathscr{G}_{\boldsymbol{\kappa}_1}$, (1) shows that for some choice of $\boldsymbol{\kappa}_2$ we may expect to obtain a different peer group $\mathscr{G}_{\boldsymbol{\kappa}_2}$. Thus the concept of peerdom is a relative one, depending upon the choice of reference placement. It is possible, nonetheless, that $\mathscr{G}_{\boldsymbol{\kappa}_1} = \mathscr{G}_{\boldsymbol{\kappa}_2}$ for all choices of $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$. In that case we shall say that the material is *egalitarian*: No deformation can alter its peer group. A glance at (1) reveals two groups corresponding to egalitarian materials:

$$\mathscr{G} = \{\mathbf{1}, -\mathbf{1}\} \qquad \text{or} \qquad \mathscr{G} = \mathscr{U}. \qquad \text{(IV.13-4)}$$

According to a theorem of group theory,[1] the proper unimodular group $\mathscr{U}^+$ is "simple", which means that the equation

$$\mathscr{G} = \mathbf{P}\mathscr{G}\mathbf{P}^{-1} \qquad \forall \mathbf{P} \in \mathscr{U} \qquad \text{(IV.13-5)}$$

[1] My inquiries have not led to a simple, direct proof. The statement follows from more powerful theorems of group theory presented by J. J. ROTMAN, *The Theory of Groups*, 2^{nd} ed., Rockleigh, New Jersey, Allyn & Bacon, 1973. The projective special linear group $PSL(n, K) := SL(n, K)/Z_0$. Here K is an arbitrary field; $SL(n, K)$ is the multiplicative group of proper unimodular $n \times n$ matrices over K; and Z_0 is its center, that is, the group of all elements that commute with every element. ROTMAN's Theorem 8.25 asserts that $PSL(m, K)$ is simple if $m \geqq 3$; his Theorem 8.13, that the center of $SL(3, R)$ is the unit matrix. Thus $PSL(3, R) = SL(3, R)/Z_0 = SL(3, R)$, and so $SL(3, R)$ is simple.

has no solutions $\mathcal{g}$ that can be peer groups other than the trivial ones (4). *Thus the groups* (4) *correspond with the only possible egalitarian materials.*[1] In Section IV.16 we shall see an important consequence of this fact. Here we remark merely that the fact itself will not startle the student, since it asserts that only for the two extremes of response can no deformation either create new peers or unseat any of the old, as might well be expected from the definition of peers. At one extreme, all placements are peers; at the other, no placement has any peers but the two trivial ones.

The considerations of this section and the preceding apply equally to the determinate response of a constrained material, which is defined by (IV.7-2).

14. Isotropic Materials

A homogeneous body is isotropic if it can be brought into a shape, no rotations of which can be detected by measurement of stress. Isotropy is an example, and the most important one, of *material symmetry*. To consider material symmetries, we fix attention upon the peer groups of a single material point. In this section and the next two we shall use the phrase "a material is . . ." to abbreviate "a material point is . . . ." Since in the rest of the book we consider only homogeneous bodies, and so all the material points that make up a body must have the same material symmetry, we could just as well write in each case "a body is . . . ." The letter $\mathcal{o}$ will denote the full orthogonal group.

Definition (CAUCHY, NOLL). *A material is* isotropic *if there is a reference placement* $\boldsymbol{\kappa}$ *such that*

$$\mathcal{g}_{\boldsymbol{\kappa}} \supset \mathcal{o}. \tag{IV.14-1}$$

Such a placement $\boldsymbol{\kappa}$ is called *undistorted*; other placements, *distorted*. According to this definition, every orthogonal transplacement of an undistorted placement carries it into a peer. From NOLL's rule (IV.13-1) we see that for other placements $\boldsymbol{\kappa}'$ the peer groups $\mathcal{g}_{\boldsymbol{\kappa}'}$ need not contain $\mathcal{o}$. That is, rotations of $\boldsymbol{\kappa}'$ generally can be detected by experimental measurements of stress, though rotations of an undistorted placement $\boldsymbol{\kappa}$ cannot. Of course, that same rule shows us that *an orthogonal transplacement carries one undistorted placement of an isotropic material into another*, a fact which merely reflects the definition

[1]The reader should be warned not to expect that the statement proved here for the mechanics of simple materials can be extended to other theories in which a peer group may be defined. For example, in optics there are four groups that correspond to egalitarian materials: not only those given by (4) but also $\{1\}$ and $\mathcal{u}^+$.

of "isotropic material", besides showing that an isotropic material has infinitely many undistorted placements.

For an isotropic material, (IV.12-6) changes from an equation to be solved for certain $\mathbf{Q}$ into an identity satisfied by all $\mathbf{Q}$, and likewise (IV.12-3) is satisfied by all orthogonal $\mathbf{H}$. By this latter equation, then, the value of $\mathbf{T}$ is unchanged if we replace $\mathbf{F}^t$ by $\mathbf{F}^t\mathbf{Q}$, where $\mathbf{Q}$ is any constant orthogonal tensor. In particular, if we regard the present time t as a parameter which we may hold fixed, and if $\bar{\mathbf{F}}^t(s) := \mathbf{F}^t(s)\mathbf{R}(t)^{\mathsf{T}}$, then $\bar{\mathbf{F}}^t$ delivers the same stress as does $\mathbf{F}^t$ at the time t. $\bar{\mathbf{R}}$, the present rotation of $\bar{\mathbf{F}}$, equals $\mathbf{1}$. Thus $(\mathbf{F}_t^{\mathsf{T}}\mathbf{F}_t)^t = \bar{\mathbf{C}}_t^t = \mathbf{C}_t^t$ and $\bar{\mathbf{C}} = \mathbf{RCR}^{\mathsf{T}} = \mathbf{B}$. Putting $\bar{\mathbf{R}}$, $\bar{\mathbf{C}}_t^t$, and $\bar{\mathbf{C}}$ into (IV.5-15) delivers NOLL's *reduction of the constitutive relation for isotropic materials*:

$$\mathbf{T} = \mathfrak{K}(\mathbf{C}_t^t\,;\,\mathbf{B}), \tag{IV.14-2}$$

in which, as was to be expected, the rotation does not appear at all.

According to (IV.12-6), moreover, if $\mathbf{F}^t$ is replaced by $\mathbf{Q}\mathbf{F}^t\mathbf{Q}^{\mathsf{T}}$, for any $\mathbf{Q}$, the stress $\mathbf{T}$ is replaced by $\mathbf{QTQ}^{\mathsf{T}}$. In this replacement $\mathbf{C}_t^t$ and $\mathbf{B}$ are replaced by $\mathbf{Q}\mathbf{C}_t^t\mathbf{Q}^{\mathsf{T}}$ and $\mathbf{QBQ}^{\mathsf{T}}$, as is easily verified from (II.9-10) and (II.9-5). Thus the mapping $\mathfrak{K}$ in (2) must satisfy the condition

$$\mathfrak{K}(\mathbf{Q}\mathbf{C}_t^t\mathbf{Q}^{\mathsf{T}};\,\mathbf{QBQ}^{\mathsf{T}}) = \mathbf{Q}\mathfrak{K}(\mathbf{C}_t^t\,;\,\mathbf{B})\mathbf{Q}^{\mathsf{T}}, \tag{IV.14-3}$$

for every orthogonal tensor $\mathbf{Q}$, for every positive symmetric tensor history $\mathbf{C}_t^t$, and for every positive symmetric tensor $\mathbf{B}$.

A mapping satisfying this requirement for all $\mathbf{Q}$ is called *isotropic*. Thus, the concept of isotropic mapping generalizes that of isotropic function defined by (IV.4-9). Conversely, if (3) is satisfied by $\mathfrak{K}$, (2) gives the constitutive equation of an isotropic simple material, referred to an undistorted placement. If a distorted reference placement is used, the constitutive relation of an isotropic material cannot have the form (2) and generally shows no recognizable simplicity.

The solution of Exercise IV.7.1 enables us to reduce as follows the constitutive equation $(\text{IV.7-14})_2$ giving the determinate response of an *incompressible isotropic material* with respect to an undistorted placement:

$$\mathbf{S} = \mathfrak{K}(\mathbf{C}_t^t\,;\,\mathbf{B}); \tag{IV.14-4}$$

$\det \mathbf{C}_t^t = \det \mathbf{B} = 1$; and the mapping $\mathfrak{K}$ satisfies (3).

While (1) embodies a natural concept of isotropy, it seems more general than in fact it is. According to a theorem of group theory, the orthogonal group is maximal in the unimodular group.[1] That is, if $\mathscr{g}$ is a group such that $\mathscr{o} \subset \mathscr{g} \subset \mathscr{u}$, then

$$\text{either} \quad \mathscr{g} = \mathscr{o} \qquad \text{or} \qquad \mathscr{g} = \mathscr{u}. \tag{IV.14-5}$$

Thus the peer group of an isotropic material in an undistorted placement is either the orthogonal group or the unimodular group.

In either (2) or (4), let $\mathbf{C}_t^t$ be fixed. Then (3) reduces to (IV.4-9), and $\mathfrak{K}$ reduces to an isotropic mapping of symmetric tensors onto symmetric tensors.

A body of isotropic material is called an *isotropic body*.

15. Universal Transplacements of Isotropic Incompressible Bodies

In Section IV.8 we have defined universal motions and explained their great value for use in comparison of theory with experiment, and we have set out the scheme for finding universal transplacements for homogeneous, incompressible bodies. In Section IV.10 we have determined all homogeneous transplacements of homogeneous, incompressible bodies subject to lamellar body force. The constraint of incompressibility, which narrows the class of admissible transplacements to those that are isochoric, at the same time broadens it by allowing some universal transplacements that are not homogeneous. We shall now exhibit, again supposing the body force lamellar, five families of universal transplacements for homogeneous, isotropic, incompressible bodies, defined in the preceding section. We shall always presume that *the reference placement is homogeneous and undistorted.*

As for the homogeneous transplacements discussed in Sections IV.9 and IV.10, the analysis follows a *semi-inverse method.* Families of putative transplacements such as to model circumstances of interest in mechanics are set down. Each is written in terms of functions, at first arbitrary but later to be restricted in such a way as to deliver dynamically possible motions for every homogeneous, isotropic, incompressible body. The outcome of such analysis is a class (rather small, perhaps even empty) of solutions of the problem initially set. Comparison of the calculated solution with data from experiments on motions of real bodies idealized by members of the class laid down may then yield some information about the constitutive properties of the bodies used in the experiments.

The at first arbitrary functions denoted by capital letters in the following five examples denote twice differentiable functions of time only.

[1] *E.g.* W. NOLL, "Proof of the maximality of the orthogonal group in the unimodular group," *Archive for Rational Mechanics and Analysis* **18** (1965): 100–102, reprinted in NOLL's *Foundations of Mechanics and Thermodynamics*, Berlin and New York, Springer-Verlag, 1974.

Although our problem is essentially simple, its solution is achieved in several steps, some of them long. The constitutive and dynamical conclusions are due to CARROLL, who extended the work of TRUESDELL and others in elasticity to simple materials in general. We shall follow in the main the elegant presentation of FOSDICK,[1] which in outline goes as follows.

Step 0. We specify the five families. To this end we set out these putative universal transplacements in a referential description using two conveniently selected co-ordinate systems, one for the reference placement and one for the actual placement.

Step 1. For each of the five families we calculate $\mathbf{B}$ and $\mathbf{C}_t^t$.
This matter is purely kinematical, and the calculation is routine. We find that at least two of the shear components of $\mathbf{B}$ and $\mathbf{C}_t^t$ vanish, while the remaining components are the values of simple, explicit functions of one particular distance (labelled r or x), thus reducing the generality of the functions of time introduced at Step 0, and of the histories of those functions.

Step 2. We use the constitutive relation (IV.14-4) and the functional restriction (IV.14-3).
We show thereby that each null component of $\mathbf{B}$ and $\mathbf{C}_t^t$ corresponds with a null shear component of $\mathbf{S}$, that the remaining components of $\mathbf{S}$ are the values of functions of the variables occurring at Step 1 and of the histories of the functions of time occurring there, and that those functions are odd or even in certain of their arguments.

Step 3. Turning to the flows corresponding with the transplacements, we recall from Section IV.10 that every transplacement universal for homogeneous, incompressible bodies subject to lamellar body force must preserve circulation. Thus, on the assumption that the domain of flow is simply connected, an acceleration-potential $P_{\mathfrak{a}}$ stands at our disposition:

$$\ddot{\mathbf{x}} = -\operatorname{grad} P_{\mathfrak{a}}, \qquad \text{(II.11-33)}_{\text{r}}$$

whence (*cf.* (IV.8-12)) follows the existence of a scalar field λ such that

$$\operatorname{div}\mathbf{S} = -\operatorname{grad}\lambda. \qquad \text{(IV.15-1)}$$

The left-hand side of this relation, calculable from the conclusions of Step 2 for each of the five families, delivers λ for each.

[1] For further detail see the paper of R. L. FOSDICK, "Dynamically possible motions of incompressible, isotropic, simple materials," *Archive for Rational Mechanics and Analysis* **29**(1968): 272–288.

Step 4. For each of the five families in turn we apply the requirement that the flow preserve circulation, namely

$$\operatorname{skw}\operatorname{grad}\ddot{\mathbf{x}} = \mathbf{0}. \qquad \text{(II.11-45)}_{\mathrm{r}}$$

Restrictions upon the arbitrary functions of time appearing in the families set down at Step 0 result, and $P_{\mathfrak{a}}$ is determined for each family in turn. Because λ has been determined at Step 3, p for each family is now determined:

$$p = \rho(P_{\mathfrak{a}} - \varpi) - \lambda. \qquad \text{(IV.8-13)}_{2\mathrm{r}}$$

Step 4, although it is purely kinematical and requires only routine calculus, is the longest. For it we shall follow the analysis of WANG,[1] which includes and extends the contributions of several earlier authors.

We proceed now with the details.

Step 0. The five families of isochoric transplacements follow. In each the letters $A, B, \ldots$ stand for as yet arbitrary, twice-differentiable functions of t. To each putative universal transplacement we may add an arbitrary rotation, but to do so would add at Step 4 complications not worth the effort needed to take account of them. On the other hand, the student must be warned that two universal transplacements which are not identical may in fact differ only by some particular rigid motion. Moreover, the reference placement is arbitrary. One particular choice of it is made when a transplacement is specified. If, instead, as we shall do for some instances below in Section IV.18, we begin from a spatial velocity field, we must always remember that *infinitely many transplacements give rise to it*, one for each choice of reference placement.

Family 1 (Pure bending, stretching, and shearing of a rectangular block). X, Y, Z are cartesian co-ordinates in the reference placement; r, θ, z are cylindrical polar co-ordinates in the present placement.

$$r^2 = 2AX + B, \qquad \theta = CY + DZ + K, \qquad z = EY + FZ + L,$$
$$A(CF - DE) = 1. \qquad \text{(IV.15-2)}$$

From $(2)_4$ it follows that A is determined by C, D, E, and F, that $CF \neq DE$, and that $\operatorname{sgn} A = \operatorname{sgn}(CF - DE)$. Usually we shall neglect the arbitrary constants K and L.

[1] C.-C. WANG, "Universal solutions for incompressible laminated bodies," *Archive for Rational Mechanics and Analysis* **29**(1968): 161–192. For some further details consult IRE, especially Section 6 of Chapter 5.

Family 2 (Straightening, stretching, and shearing of a section of a hollow cylinder). R, Θ, Z are cylindrical polar co-ordinates in the reference placement, and x, y, z are cartesian co-ordinates in the present placement.

$$x = \tfrac{1}{2}AR^2, \qquad y = B\Theta + CZ, \qquad z = D\Theta + EZ,$$
$$A(BE - CD) = 1. \tag{IV.15-3}$$

From $(3)_4$ it follows that A is determined by B, C, D, and E, that $A \neq 0$ and $BE \neq CD$, and that $\operatorname{sgn} A = \operatorname{sgn}(BE - CD)$.

Family 3 (Inflation, eversion, bending, torsion, extension, and shearing of an annular wedge). Here R, Θ, Z and r, θ, z are cylindrical polar co-ordinates in the reference placement and the present placement, respectively.

$$r^2 = AR^2 + B, \qquad \theta = C\Theta + DZ + K, \qquad z = E\Theta + FZ + L,$$
$$A(CF - DE) = 1. \tag{IV.15-4}$$

Thus $A \neq 0$, $CF \neq DE$, $A = 1/(CF - DE)$, $\operatorname{sgn} A = \operatorname{sgn}(CF - DE)$. As we did for Family 1, here too we shall generally set aside K and L.

Family 4 (Inflation and eversion of a sector of a spherical shell). The co-ordinates R, Θ, Φ and r, θ, φ are spherical polar in the reference placement and present placement, respectively.

$$r^3 = \pm R^3 + A, \qquad \theta = \pm\Theta, \qquad \varphi = \Phi. \tag{IV.15-5}$$

Family 5 (Inflation, azimuthal bending and shearing, and extension of an annular wedge). The co-ordinate systems are the same as those used for Family 3.

$$r = AR, \qquad \theta = B \log R + C\Theta, \qquad z = DZ, \qquad A^2CD = 1. \tag{IV.15-6}$$

Thus $A \neq 0$ and $CD > 0$.

Exercise IV.15.1. The verbal descriptions of the five families are just.

Step 1. We shall refer vectors and tensors to their physical components, for which see Section App. IIC.9. Because (2), (3), (4), (5), and (6) are linear in two out of three co-ordinates, it is plain that the physical components of **B** and $\mathbf{C}_t^t$ will be functions of the one co-ordinate in which the transplacement is not linear.

Family 1. We show from $(\text{II}.9\text{-}6)_2$ that

$$[\mathbf{B}] = \begin{Vmatrix} A(r^2 - B)/r^2 & 0 & 0 \\ \cdot & [AC/(r^2 - B) + D^2]r^2 & [CEA/(r^2 - B) + DF]r \\ \cdot & \cdot & E^2A/(r^2 - B) + F^2 \end{Vmatrix}. \tag{IV.15-7}$$

Likewise by use of $(\text{II}.9\text{-}6)_1$ and (II.9-10) we conclude that (in physical components)

$$\begin{aligned} C_t^t{}_{rr} &= (A^t)^2 r^2/[A(A^t r^2 - AB^t - A^t B)], \\ C_t^t{}_{\theta\theta} &= A(C^tF - D^tE)^2(A^t r^2 + AB^t - A^tB)/r^2 + A^t(E^tF - EF^t)^2/r^2, \\ C_t^t{}_{zz} &= A(D^tC - DC^t)(A^t r^2 + AB^t - A^tB) + A^2(F^tC - E^tD)^2, \\ C_t^t{}_{\theta z} &= A(D^tC - DC^t)(C^tF - D^tE)(A^t r^2 + AB^t - A^tB)/r \\ &\quad + A^2(E^tF - EF^t)(F^tC - E^tD)/r, \qquad C_t^t{}_{r\theta} = C_t^t{}_{rz} = 0. \end{aligned} \tag{IV.15-8}$$

Exercise IV.15.2. The statements (7) and (8) are correct.

Family 2. $\mathbf{B}$ depends upon x alone, $\mathbf{C}_t^t$ is independent of x, and $B_{xy} = B_{xz} = C_t^t{}_{xy} = C_t^t{}_{xz} = 0$.

Exercise IV.15.3. $\mathbf{B}$ and $\mathbf{C}_t^t$ for Family 2 are to be calculated.

Exercise IV.15.4 (BHARATHA). The flow delivered by Family 2 is homogeneous.

Family 3. $\mathbf{B}$ and $\mathbf{C}_t^t$ for Family 3 are given by (7) and (8).

Exercise IV.15.5 (WANG). The flows delivered by Families 1 and 3 are the same. This fact is explained by showing that composition of a static instance of Family 1 with the general Family 3 delivers the general Family 1. The motions defining Families 1 and 3 differ from each other only by a change of reference placement.

Family 4.

$$\begin{aligned} [\mathbf{B}] &= \operatorname{diag}[(r^3 - A)^{4/3}/r^4,\ r^2/(r^3 - A)^{2/3},\ r^2/(r^3 - A)^{2/3}], \\ [\mathbf{C}_t^t] &= \operatorname{diag}[r^4/(r^3 - A + A^t)^{4/3},\ (r^3 - A + A^t)^{2/3},\ (r^3 - A + A^t)^{2/3}]. \end{aligned} \tag{IV.15-9}$$

Family 5.

$$[\mathbf{B}] = \begin{Vmatrix} A^2 & BA^2 & 0 \\ \cdot & A^2(B^2+C^2) & 0 \\ \cdot & \cdot & D^2 \end{Vmatrix},$$

$$\begin{aligned} C_t^{t}{}^{rr} &= [(A^t)^2/A^2][1+(B^tC-BC^t)^2/C^2], \\ C_t^{t}{}^{\theta\theta} &= (C^t)^2(A^t)^2/(C^2A^2), \qquad C_t^{t}{}^{zz} = (D^t)^2/D^2, \\ C_t^{t}{}^{r\theta} &= (A^t)^2C^t(B^tC-BC^t)/(A^2C^2). \end{aligned} \tag{IV.15-10}$$

Both $\mathbf{B}$ and $\mathbf{C}_t^t$ are independent of place.

Step 2. The mapping $\mathfrak{R}$ must satisfy the relation (IV.14-3) for all orthogonal tensors $\mathbf{Q}$. We first consider the particular $\mathbf{Q}$ that represents a rotation through a straight angle about the x^1-axis in some orthogonal co-ordinate system:

$$[\mathbf{Q}] := \operatorname{diag}(1, -1, -1). \tag{IV.15-11}$$

Then for any tensor $\mathbf{Y}$

$$[\mathbf{Q}\mathbf{Y}\mathbf{Q}^{\mathsf{T}}] = \begin{Vmatrix} Y^{11} & -Y^{12} & -Y^{13} \\ -Y^{21} & Y^{22} & Y^{23} \\ -Y^{31} & Y^{32} & Y^{33} \end{Vmatrix}. \tag{IV.15-12}$$

If $Y^{12} = Y^{13} = Y^{21} = Y^{31} = 0$, then for the $\mathbf{Q}$ given by (11) it follows that $\mathbf{Q}\mathbf{Y}\mathbf{Q}^{\mathsf{T}} = \mathbf{Y}$. Applying this observation to (IV.14-3), we see that when $B^{12} = B^{13} = C_t^{t}{}^{12} = C_t^{t}{}^{13} = 0$, then

$$\mathfrak{R}(\mathbf{C}_t^t; \mathbf{B}) = \mathbf{Q}\mathfrak{R}(\mathbf{C}_t^t; \mathbf{B})\mathbf{Q}^{\mathsf{T}} \tag{IV.15-13}$$

for the $\mathbf{Q}$ defined by (11). Thus $\mathbf{S} = \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathsf{T}}$, and hence (12) requires that

$$S^{12} = S^{13} = 0: \tag{IV.15-14}$$

The shear stresses corresponding to vanishing shears are null.

Next we consider the effect of a rotation through a straight angle about the x^3-axis:

$$[\mathbf{Q}] = \operatorname{diag}(-1, -1, 1). \tag{IV.15-15}$$

For any tensor $\mathbf{Y}$

$$[\mathbf{QYQ}^\top] = \begin{Vmatrix} Y_{11} & Y_{12} & -Y_{13} \\ Y_{21} & Y_{22} & -Y_{23} \\ -Y_{31} & -Y_{32} & Y_{33} \end{Vmatrix}. \tag{IV.15-16}$$

If, supposing that (14) is satisfied, we use in (IV.14-3) the $\mathbf{Q}$ given by (15), we see that when $C_t^{\langle 23 \rangle}$ and $B_{\langle 23 \rangle}$ are reversed in sign, also $S_{\langle 23 \rangle}$ is reversed in sign, while $S_{\langle 11 \rangle}$, $S_{\langle 22 \rangle}$, and $S_{\langle 33 \rangle}$ remain unchanged.

Now we apply the two foregoing statements to the five families in turn, with conclusions as follows.

Family 1. $S_{\langle r\theta \rangle} = S_{\langle rz \rangle} = 0$, and there are scalar-valued mappings $\tau^{(1)}$, $\sigma_1^{(1)}$, $\sigma_2^{(1)}$ such that

$$\begin{aligned} S_{\langle \theta z \rangle} &= \tau^{(1)}(B^t, C^t, D^t, E^t, F^t; r), \\ S_{\langle rr \rangle} - S_{\langle zz \rangle} &= \sigma_1^{(1)}(B^t, C^t, D^t, E^t, F^t; r), \\ S_{\langle \theta\theta \rangle} - S_{\langle zz \rangle} &= \sigma_2^{(1)}(B^t, C^t, D^t, E^t, F^t; r). \end{aligned} \tag{IV.15-17}$$

From (8) we see that to change the sign of $S_{\langle \theta z \rangle}$ while leaving unchanged the diagonal components of $\mathbf{S}$ we may either replace C^t and F^t by $-C^t$ and $-F^t$ or replace D^t and E^t by $-D^t$ and $-E^t$. Therefore τ is odd under such changes:

$$\begin{aligned} \tau^{(1)}(B^t, -C^t, D^t, E^t, -F^t; r) &= \tau^{(1)}(B^t, C^t, -D^t, -E^t, F^t; r), \\ &= -\tau^{(1)}(B^t, C^t, D^t, E^t, F^t; r). \end{aligned} \tag{IV.15-18}$$

Hence if the shear stress function $\tau^{(1)}$ is continuous at $C^t = D^t = E^t = F^t = 0$, then

$$\tau^{(1)}(B^t, 0, D^t, E^t, 0; r) = \tau^{(1)}(B^t, C^t, 0, 0, F^t; r) = 0, \tag{IV.15-19}$$

Also, when $\alpha = 1$ or 2,

$$\sigma_\alpha^{(1)}(B^t, -C^t, D^t, E^t, -F^t; r) = \sigma_\alpha^{(1)}(B^t, C^t, -D^t, -E^t, F^t; r)$$
$$= \sigma_\alpha^{(1)}(B^t, C^t, D^t, E^t, F^t; r). \tag{IV.15-20}$$

Exercise IV.15.6. The conclusions (18), (19), and (20) suffice for $\mathfrak{K}$ to satisfy (IV.14-3) in all instances of Family 1.

Family 2. Similar analysis shows that $S_{xy} = S_{xz} = 0$ and

$$\begin{aligned} S_{yz} &= \tau^{(2)}(B^t, C^t, D^t, E^t; x), \\ S_{xx} - S_{zz} &= \sigma_1^{(2)}(B^t, C^t, D^t, E^t; x), \\ S_{yy} - S_{zz} &= \sigma_2^{(2)}(B^t, C^t, D^t, E^t; x); \end{aligned} \tag{IV.15-21}$$

that $\tau^{(2)}$ changes sign when B^t and E^t are replaced by $-B^t$ and $-E^t$ and when C^t and D^t are replaced by $-C^t$ and $-D^t$; and that $\sigma_1^{(2)}$ and $\sigma_2^{(2)}$ are unchanged in sign by those transformations.

Family 3. The conclusions are the same as for Family 1.

Family 4. To obtain necessary and sufficient conditions here, we use not only (15) but also another $\mathbf{Q}$:

$$[\mathbf{Q}] = \begin{Vmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix}, \tag{IV.15-22}$$

which represents a rotation about an axis in the plane normal to the direction of r. The outcome is that $[\mathbf{S}]$ is diagonal, that $S^{\theta\theta} = S^{\varphi\varphi}$, and that

$$S^{rr} - S^{\theta\theta} = S^{rr} - S^{\varphi\varphi} = \tau^{(4)}(A^t; r). \tag{IV.15-23}$$

Family 5. $S^{rz} = S^{\theta z} = 0$, and

$$\begin{aligned} S^{r\theta} &= \tau^{(5)}(A^t, B^t, C^t), \\ S^{rr} - S^{zz} &= \sigma_1^{(5)}(A^t, B^t, C^t), \\ S^{\theta\theta} - S^{zz} &= \sigma_2^{(5)}(A^t, B^t, C^t). \end{aligned} \tag{IV.15-24}$$

The shearing stress and the differences of the normal stresses, like the components of $\mathbf{B}$ and $\mathbf{C}_t^t$, are independent of position.

Step 3. We now show that the dynamical condition (1) is satisfied for each family, and in so doing we calculate for each the function λ that is defined by (1) and will be used to determine p through (IV.8-11).

Family 1. From Section App. II C.9 we see that if the physical components of $\mathbf{S}$ are given by functions of r alone at each time t, then

$$
\begin{aligned}
(\operatorname{div}\mathbf{S})^{r} &= \partial_r S^{rr} + \frac{S^{rr} - S^{\theta\theta}}{r}, \\
(\operatorname{div}\mathbf{S})^{\theta} &= \frac{1}{r}\partial_r (r^2 S^{r\theta}), \\
(\operatorname{div}\mathbf{S})^{z} &= \frac{1}{r}\partial_r (r^2 S^{rz}).
\end{aligned}
\tag{IV.15-25}
$$

For Family 1, because $S^{r\theta} = S^{rz} = 0$, we see that (1) is satisfied, and

$$
-\lambda = S^{rr} + \int \frac{S^{rr} - S^{\theta\theta}}{r}\, dr + f(t). \tag{IV.15-26}
$$

S^{zz} may be taken as the value of an arbitrary function of r alone at each t; then the right-hand side of (26) is determined by the constitutive functions $\sigma_1^{(1)}$ and $\sigma_2^{(1)}$, obtained at Step 2.

Family 2. If the components of $\mathbf{S}$ are functions of x alone at each time t, then

$$
(\operatorname{div}\mathbf{S})_x = \partial_x S_{xx}, \qquad (\operatorname{div}\mathbf{S})_y = \partial_y S_{xy}, \qquad (\operatorname{div}\mathbf{S})_z = \partial_z S_{xz}. \tag{IV.15-27}
$$

For Family 2, because $S_{xy} = S_{xz} = 0$,

$$
-\lambda = S_{xx}. \tag{IV.15-28}
$$

S_{zz} may be taken as the value of an arbitrary function of x and t, and so the right-hand side of (28) is determined by the constitutive function $\sigma_1^{(2)}$, obtained at Step 2.

Family 3. Refer to Family 1.

Family 4. From Section App. IIC.11 we see that if the physical components of S are given by functions of r alone at each t, and if $S^{r\theta} = S^{r\varphi} = S^{\theta\varphi} = 0$, then

$$(\operatorname{div} \mathbf{S})^r = \partial_r S^{rr} + \frac{2}{r}(S^{rr} - S^{\theta\theta}),$$
$$(\operatorname{div} \mathbf{S})^\theta = 0, \qquad (\operatorname{div} S)^\varphi = 0. \tag{IV.15-29}$$

Hence

$$-\lambda = S^{rr} + 2\int \frac{S^{rr} - S^{\theta\theta}}{r}\, dr. \tag{IV.15-30}$$

Because $S^{\varphi\varphi}$ is given by an arbitrary function of r at each t and $S^{\theta\theta} = S^{\varphi\varphi}$, the right-hand side of (30) is determined by the constitutive function $\tau^{(4)}$, obtained at Step 2.

Family 5. We appeal to (25) again. This time the physical components of S are given by function of t alone, and hence

$$-\lambda = (S^{rr} - S^{\theta\theta})\log r + 2\theta S^{r\theta}. \tag{IV.15-31}$$

The right-hand side of (31) is determined by the constitutive functions $\tau^{(5)}$, $\sigma_1^{(5)}$, and $\sigma_2^{(5)}$, obtained at Step 2.

Step 4. On the assumption that the arbitrary functions of t appearing in the putative universal transplacements (2), (3), (4), (5), and (6) can be so chosen as to satisfy (II.11-45), we have exhausted the requirements of material symmetry and dynamics except for determining the required pressure p. We shall now solve (II.11-45) for each family and so determine for each the acceleration-potential $P_{\mathbf{a}}$. Substitution of $P_{\mathbf{a}}$ into (IV.8-11) will then deliver p. Accordingly, for each family in turn we calculate the flow to which it gives rise and then the acceleration field of that flow.

Family 1. From (2) we find that

$$
\begin{aligned}
\dot r &= \frac{\dot A}{2A} r + \left(\frac{\dot B}{2} - \frac{B\dot A}{2A}\right)\frac{1}{r},\\
\dot\theta &= \theta A(F\dot C - E\dot D) + zA(C\dot D - D\dot C),\\
\dot z &= \theta A(F\dot E - E\dot F) + zA(C\dot F - D\dot E),\\
\ddot r &= \left[\frac{\ddot A}{2A} - \left(\frac{\dot A}{2A}\right)^2\right] r + \left[\frac{\dot A}{A}\left(\frac{B\dot A}{A} - \dot B\right) - \left(\frac{B\ddot A}{A} - \ddot B\right)\right]\frac{1}{2r}\\
&\quad - \left(\frac{B\dot A}{A} - \dot B\right)^2 \frac{1}{4r^3}, \qquad \text{(IV.15-32)}\\
\ddot\theta &= \theta A(F\ddot C - E\ddot D) + zA(C\ddot D - D\ddot C),\\
\ddot z &= \theta A(F\ddot E - E\ddot F) + zA(C\ddot F - D\ddot C).
\end{aligned}
$$

From these we can determine the covariant components of the acceleration field:

$$
\begin{aligned}
\ddot x_r &= \ddot r - r\dot\theta^2,\\
\ddot x_\theta &= r^2\ddot\theta + 2r\dot r\dot\theta, \qquad \text{(IV.15-33)}\\
\ddot x_z &= \ddot z.
\end{aligned}
$$

Substituting (32) and (33) into (II.11-45) yields the differential system

$$
\begin{aligned}
C\dot D - D\dot C &= 0,\\
F\ddot E - E\ddot F &= 0, \qquad \text{(IV.15-34)}
\end{aligned}
$$

$$
F\ddot C - E\ddot D + (F\dot C - E\dot D)\left(AF\dot C - AE\dot D + \frac{\dot A}{A}\right) = 0.
$$

We wish to solve this system of differential equations together with the condition $(7)_4$ for isochoric motion. In doing so we recall that

$$
A(CF - DE) = 1. \qquad \text{(IV.15-2)}_{4\mathrm{r}}
$$

First we notice that the function B does not appear in the governing equations (34) and (2); thus it is arbitrary. Second, C and D cannot vanish simultaneously, for that would violate (2). We shall prove that C vanishes always or never. If

C does not vanish for any t, the integral of $(34)_1$ is

$$D = kC, \qquad k = \text{const.} \tag{IV.15-35}$$

Because $\dot{C}$ appears in the differential system (34), C must be continuous. The set

$$\mathscr{U} := \{t : C(t) \neq 0\}, \tag{IV.15-36}$$

because it is the inverse image of the open set $\mathscr{R} \setminus \{0\}$, is open, and if C does vanish at some t but not always, then

$$\varnothing \subset \mathscr{U} \subset \mathscr{R}, \tag{IV.15-37}$$

and both inclusions are proper. Thus there is a boundary point t_0 of $\mathscr{U}$ that does not belong to $\mathscr{U}$, *i.e.*,

$$C(t) \neq 0 \tag{IV.15-38}$$

on an open interval with t_0 as an end-point. Thus t_0 is a limit point of $\mathscr{U}$. Since (35) holds on the interval just mentioned, continuity requires that $D(t_0) = 0$, but it is impossible for C and D to vanish simultaneously.

We have shown that there are only two possibilities:

I) $\mathscr{U} = \mathscr{R}$; equivalently, $C \neq 0$ for any t. In this case (35) holds for all t. Then $(2)_4$ reduces to

$$AC(F - kE) = 1, \tag{IV.15-39}$$

and $(34)_3$ reduces to

$$\ddot{C} + \dot{C}\left(\frac{\dot{C}}{C} + \frac{\dot{C}}{A}\right) = 0, \tag{IV.15-40}$$

an integral of which is

$$\dot{C} = \frac{k_1}{AC}, \qquad k_1 = \text{const.} \tag{IV.15-41}$$

The differential equation $(34)_2$ has an obvious integral also:

$$F\dot{E} - E\dot{F} = k_2, \qquad k_2 = \text{const.} \tag{IV.15-42}$$

Now we consider the following three possibilities:

Ia) $F = 0$. In this subcase $k_2 = 0$, and (42) is satisfied for all E. Further, (39) reduces to $-kACE = 1$. Thus the constant k cannot be 0. The complete solution for this subcase is

$$
\begin{aligned}
B, E &= \text{arbitrary functions of } t, \text{ but } E(t) \neq 0 \text{ for any } t, \\
F &= 0, \\
C &= k'\left(\int_0^t E\,dt + k''\right), \quad C(t) \neq 0 \text{ for any } t, \\
D &= -k'''\left(\int_0^t E\,dt + k''\right) \neq 0, \\
A &= \frac{1}{k'''E\left(\int_0^t E\,dt + k''\right)},
\end{aligned}
\tag{IV.15-43}
$$

in which k', k'', and k''' are arbitrary, non-null constants.

If $F \neq 0$, we can still get local solutions near points where F does not vanish or does vanish, as follows.

Ib) $F(t_0) \neq 0$ for some t_0. We consider the solution for t near t_0. Integrating (42) yields

$$E = F\left(k' + k_2\int_0^t \frac{1}{F^2}\,dt\right). \tag{IV.15-44}$$

We put

$$\xi = \frac{1}{AC} = F\left[1 - k\left(k' + k_2\int_0^t \frac{1}{F^2}\,dt\right)\right]. \tag{IV.15-45}$$

Then the complete solution for t near t_0 is

$$
\begin{aligned}
B, F &= \text{arbitrary functions of } t, \\
E &= F\left(k' + k_2\int_{t_0}^t \frac{1}{F^2}\,dt\right) \neq 0, \\
C &= k''\left(\int_{t_0}^t \xi\,dt + k'''\right) \neq 0, \\
D &= kC = kk''\left(\int_{t_0}^t \xi\,dt + k'''\right), \\
A &= \frac{1}{C\xi},
\end{aligned}
\tag{IV.15-46}
$$

k, k', k'', k''', and k_2 being arbitrary constants.

Ic) $F(t_0) = 0$ for some t_0, but $F \neq 0$. Then t_0 must be an isolated root of F. To see that, in parallel with (36) we put

$$\mathscr{U} := \{t : F(t) \neq 0\}, \tag{IV.15-47}$$

and again $\mathscr{U}$ is open, and both inclusions in (37) are proper. It follows that the boundary of $\mathscr{U}$ is not empty. Thus $F(t_0) = 0$ if t_0 lies on the boundary of $\mathscr{U}$. Now suppose that $k_2 = 0$. Then we can integrate (42) as before and so show that $E = kF$ on $\mathscr{U}$. By continuity, $E(t_0) = F(t_0) = 0$, contradicting (39). Hence the assumption $k_2 = 0$ is false. Having proved that $k_2 \neq 0$, we conclude from (42) and the assumption $F(t_0) = 0$ that $\dot{F}(t_0) \neq 0$. Therefore t_0 is an isolated root of F.

Now we consider the solution for t near t_0. By (39), $E(t_0)$ and k do not vanish. Hence we can integrate (42) to get

$$F = -k_2 E \int_{t_0}^{t} \frac{dt}{E^2}. \tag{IV.15-48}$$

In parallel with (45) we put

$$\xi = \frac{1}{AC} = -E\left(k + k_2 \int_{t_0}^{t} \frac{dt}{E_2}\right). \tag{IV.15-49}$$

Then the complete solution is

$$\begin{gathered} B \text{ and } E \text{ are arbitrary functions of } t, \text{ but } E \neq 0, \\ F = -k_2 E \int_{t_0}^{t} \frac{dt}{E^2} \neq 0, \\ C = k_1 \left(\int_{t_0}^{t} \xi \, dt + k_0\right) \neq 0, \\ D = kC = kk_1 \left(\int_{t_0}^{t} \xi \, dt + k_0\right) \neq 0, \\ A = \frac{1}{C\xi}, \end{gathered} \tag{IV.15-50}$$

where k, k_0, k_1, and k_2 are arbitrary constants.

II) $C = 0$. In this case $(2)_4$ and (34) reduce to

$$ADE = -1,$$

$$F\dot{E} - E\dot{F} = k_2, \qquad \text{(IV.15-51)}$$

$$\ddot{D} + \dot{D}\left(\frac{\dot{A}}{A} + \frac{\dot{D}}{D}\right) = 0.$$

The last equation can be integrated at once, yielding $\dot{D} = k'''/AD$. Again there are three possibilities:

IIa) $F = 0$. The complete solution for this subcase is exactly the same as that of Subcase Ia) except that now $C = 0$.

IIb) $F(t_0) \neq 0$ for some t_0. The solutions for B, E, and F are the same as those of the subcase Ib) except that $E \neq 0$, and the solutions for A and D are the same as those of Subcase Ia).

IIc) $F(t_0) = 0$ for some t_0, but $F \neq 0$. Again t_0 must be an isolated root of F. The solutions for B, E, and F are the same as those of Subcase Ic), and the solutions for A and D are the same as those of Subcase Ia).

In general a solution may belong to different subcases in different intervals of time.

An acceleration-potential $P_{\mathbf{a}}$ for this family, in all cases, is given by

$$\begin{aligned} -P_{\mathbf{a}} = {} & \tfrac{1}{2}\left[\frac{\ddot{A}}{2A} - \left(\frac{\dot{A}}{2A}\right)^2\right] r^2 + \tfrac{1}{2}\left[\frac{\dot{A}}{A}\left(\frac{B\dot{A}}{A} - \dot{B}\right) - \left(\frac{B\ddot{A}}{A} - \ddot{B}\right)\right] \log r \\ & + \tfrac{1}{8}\left(\frac{B\dot{A}}{A} - \dot{B}\right)^2 \frac{1}{r^2} - \tfrac{1}{2}A^2(F\dot{C} - E\dot{D})^2 r^2\theta^2 \\ & + \tfrac{1}{2}(A\dot{B} - B\dot{A})(F\dot{C} - E\dot{D})\theta^2 + \tfrac{1}{2}A(C\ddot{F} - D\ddot{E})z^2. \end{aligned} \qquad \text{(IV.15-52)}$$

Family 2. While the conclusion of Exercise IV.15.4, above, reduces analysis of this family to an application of statements established in Section IV.8, also a direct attack is instructive. From $(3)_{1,2,3}$

$$\begin{aligned} \ddot{x} &= \frac{\ddot{A}}{A}x, \\ \ddot{y} &= yA(E\ddot{B} - D\ddot{C}) + zA(B\ddot{C} - C\ddot{B}), \\ \ddot{z} &= yA(E\ddot{D} - D\ddot{E}) + zA(B\ddot{E} - C\ddot{D}), \end{aligned} \qquad \text{(IV.15-53)}$$

along with $(3)_4$. Condition (II.11-45) now takes the form

$$B\ddot{C} - C\ddot{B} = E\ddot{D} - D\ddot{E}. \qquad \text{(IV.15-54)}$$

This differential equation can be integrated at once, yielding the integral

$$B\dot{C} - C\dot{B} = E\dot{D} - D\dot{E} + k, \qquad \text{(IV.15-55)}$$

where k is a constant. We consider the following two possibilities:

I) $B(t_0) \neq 0$ for some t_0. In this case the complete solution is

B, D, and E are arbitrary functions of t, but $B \neq 0$, and E and D do not vanish simultaneously,

$$C = B\left[k' + \int_{t_0}^{t} \frac{1}{B^2}(E\dot{D} - D\dot{E} + k)\,dt\right],$$

k' being a constant such that $E(t_0) - k'D(t_0) \neq 0$,

$$A = \frac{1}{BE - CD}. \qquad \text{(IV.15-56)}$$

II) $B(t_0) = 0$ for some t_0. In this case $C(t_0)D(t_0) \neq 0$. The complete solution near t_0 is

C, D, and E are arbitrary functions of t, except that $CD \neq 0$,

$$\begin{aligned} B &= -C\int_{t_0}^{t} \frac{1}{C^2}(E\dot{D} - D\dot{E} + k)\,dt, \\ A &= \frac{1}{BE - CD}. \end{aligned} \qquad \text{(IV.15-57)}$$

The root t_0 need not be isolated.

A solution for this family may belong to different cases on different intervals of time.

No matter how $A, B, \ldots, E$ are determined, an acceleration-potential is given by

$$\begin{aligned} -P_a = \tfrac{1}{2}\frac{\ddot{A}}{A}x^2 &+ \tfrac{1}{2}A(E\ddot{B} - D\ddot{C})y^2 \\ &+ A(B\ddot{C} - C\ddot{B})yz + \tfrac{1}{2}A(B\ddot{E} - C\ddot{D})z^2. \end{aligned} \qquad \text{(IV.15-58)}$$

Family 3. The conclusions for Family 1 hold.

Family 4. From (5) we find that

$$\dot{r} = \frac{\dot{A}}{3r^2}, \qquad \dot{\theta} = 0, \qquad \dot{\varphi} = 0,$$
$$\ddot{x}_r = \frac{\ddot{A}}{3r^2} - \frac{2\dot{A}}{9r^5}, \qquad \ddot{x}_\theta = \ddot{x}_\varphi = 0. \tag{IV.15-59}$$

The condition (II.11-45) does not restrict A; a velocity-potential and an acceleration-potential are given by

$$-P_{\mathrm{v}} = \frac{\dot{A}}{r^2}, \qquad -P_{\mathrm{a}} = -\frac{\ddot{A}}{3r} + \frac{\dot{A}^2}{18r^4}. \tag{IV.15-60}$$

Family 5. From (6) we find that

$$\begin{aligned}
\dot{r} &= \frac{\dot{A}}{A}r,\\
\dot{\theta} &= \left(\dot{B} - \frac{B\dot{C}}{C}\right)\log\frac{r}{A} + \frac{\dot{C}}{C}\theta,\\
\dot{z} &= \frac{\dot{D}}{D}z,\\
\ddot{x}_r &= \frac{\ddot{A}}{A}r - \left[\dot{B}\log\frac{r}{A} + \frac{\dot{C}}{C}\left(\theta - B\log\frac{r}{A}\right)\right]^2 r,\\
\ddot{x}_\theta &= \left[\ddot{B}\log\frac{r}{A} + \frac{\ddot{C}}{C}\left(\theta - B\log\frac{r}{A}\right)\right]r^2\\
&\quad + 2\frac{\dot{A}}{A}\left[\dot{B}\log\frac{r}{A} + \frac{\dot{C}}{C}\left(\theta - B\log\frac{r}{A}\right)\right]r^2,\\
\ddot{x}_z &= \frac{\ddot{D}}{D}z.
\end{aligned} \tag{IV.15-61}$$

Now the condition (II.11-45) implies that

$$\begin{aligned}
\ddot{B} + \dot{B}\left(\frac{\dot{C}}{C} + 2\frac{\dot{A}}{A}\right) &= 0,\\
\ddot{C} + \dot{C}\left(\frac{\dot{C}}{C} + 2\frac{\dot{A}}{A}\right) &= 0,\\
\ddot{B} - B\frac{\ddot{C}}{C} + 2\frac{\dot{A}\dot{B}}{A} - 2\frac{\dot{A}\dot{C}B}{AC} &= 0.
\end{aligned} \tag{IV.15-62}$$

We wish to solve this system of differential equations, subject to the condition $(6)_4$.

First, $(62)_2$ can be integrated at once, yielding

$$A^2 C\dot{C} = k_1 . \tag{IV.15-63}$$

Next subtracting $(62)_3$ from $(62)_1$, we obtain a differential equation and its integral:

$$\frac{\dot{B}\dot{C}}{C} + \frac{B\ddot{C}}{C} + 2\frac{\dot{A}\dot{C}B}{AC} = 0, \quad A^2 B\dot{C} = k_2. \tag{IV.15-64}$$

We consider the following two possibilities:

I) $k_1 = 0$. The complete solution for this case is

$$C = k \neq 0,$$

A is an arbitrary function except $A(t) \neq 0$ for all t,

$$\begin{aligned} B &= k' \int_0^t \frac{dt}{A^2} + k'', \\ D &= \frac{1}{kA^2}, \end{aligned} \tag{IV.15-65}$$

k, k', and k'', being arbitrary constants.

II) $k_1 \neq 0$. Now (63) makes $AC\dot{C}$ a non-null constant, and so neither $C(t)$ nor $\dot{C}(t)$ can vanish for any t. The quotient of (63) by $(64)_2$ makes B proportional to C, and the complete solution is

$$C = \text{an arbitrary non-vanishing function of } t, \text{ and } \dot{C}(t) \neq 0 \text{ for all } t,$$

$$A = \left(\frac{k_1}{C\dot{C}}\right)^{1/2}, \qquad B = k'C, \qquad D = \frac{1}{A^2 C}, \tag{IV.15-66}$$

k_1 and k' being arbitrary constants.

An acceleration-potential is given by

$$-P_a = \frac{r^2}{2}\left(\frac{\ddot{A}}{A} - \alpha^2 + \alpha\beta - \frac{\beta}{2}\right) + \frac{1}{2}\frac{\dot{D}}{D}z^2, \tag{IV.15-67A}$$

in which

$$\alpha = \left(\dot{B} - \frac{B\dot{C}}{C}\right)\log\frac{r}{A} + \frac{\ddot{C}}{C}\theta,$$

$$\beta = \dot{B} - \frac{B\dot{C}}{C}. \tag{IV.15-67}$$

For each instance we have determined $P_{\mathfrak{a}}$; by use of (IV.8-11) we thus obtain, for each family in turn, the pressure field p required to effect the motion.

Conclusion. We have now shown that the five families of putative universal motions for homogeneous, isotropic, incompressible bodies subject to lamellar body force are indeed solutions when the originally arbitrary functions of time occurring in them are suitably specialized. They remain universal motions when those functions of time are constant functions. Then $P_{\mathfrak{a}}$ = const., all components of **S** are constant, $\mathfrak{K}$ reduces to a function of **B** alone, and our solutions here reduce to universal transplacements in the *statics* of an *elastic* body, of course homogeneous, isotropic, and incompressible. As has been remarked in Section IV.4, the statics of simple bodies is elastostatics. Thus the placement at each time in one of these motions is a possible placement of rest for the body in question. Motions of this kind are called *quasi-equilibrated*.

In Volume 3 we shall show that *all universal motions of homogeneous, isotropic, incompressible, elastic bodies are quasi-equilibrated*. Because the class of universal motions cannot be greater for simple bodies than for elastic bodies, *all universal motions of homogeneous, isotropic, incompressible, simple bodies are quasi-equilibrated*.

Whether, for the same classes of bodies, there are universal solutions beyond those just exhibited, is presently unknown. To understand why that is so, the student may consult ERICKSEN's paper on universal solutions in the statics of homogeneous, incompressible, isotropic, elastic bodies.[1]

The logic used to obtain the five families of universal solutions shows that among the putatively arbitrary functions set out in (2), (3), (4), (5), and (6) *only* those satisfying the restrictions later derived give rise to universal solutions, and that *if* those restrictions are satisfied, they do deliver solutions of the differential equations locally. Two qualifications must be noted. First, the analysis takes no account of the fact that symbols such as r and θ represent co-ordinates and hence

[1] J. L. ERICKSEN, "Deformations possible in every isotropic, incompressible, perfectly elastic body," *Zeitschrift für angewandte Mathematik und Physik* 5(1954): 466–489, reprinted in *Problems of Non-Linear Elasticity* (edited by C. TRUESDELL), New York, *etc.*, Gordon & Breach, 1965.

are subject to restrictions such as $r \geqq 0$ and $0 \leqq \theta < 2\pi$. Second, the various integrals in terms of which the conclusions are stated are not tested or analysed. There are no theorems of existence corresponding with given initial conditions, for no such conditions are stated, and no attempt has been made to confront the conclusions with physical requirements. Steady, simple shearing, defined by (IV.9-12), indeed exists for all time, but the irrotational pure stretches defined by (IV.9-15) exist only so long as solutions of that differential equation exist. If we look back at (1) with the function C given by $(43)_3$ or $(46)_3$, we see that even if $C(0) = 1$, in general $C > 1$ or $C < 1$ at later times. In the former instance an angular wedge of a solid cylinder will be made to overlap itself, contrary to the requirement that $\chi(\cdot, t)$ be a homeomorphism. In the latter instance a cylinder with an angular wedge removed may be made to close up and fill the void, again violating the requirement of homeomorphism. Also cylinders $R =$ const. in the reference placement are generally expanded or contracted into spatial cylinders of greater or lesser radius r, and with A and B given by $(43)_{1,5}$ and $(46)_{1,5}$ it may happen that the radius of some cylinder $R =$ const. either becomes null at some time or originates at some time from a cylinder of null radius, and the transplacement fails to exist outside some interval of times. Specific examples of such phenomena are presented below in Section 18.

How important these considerations are, depends in part upon the materials of the bodies to which they are applied. For theories of the classical types, say elastic solids or Navier-Stokes fluids, existence in a tiny interval containing the present time suffices. For bodies of material with long memory, on the contrary, motions must exist in the time-interval $]-\infty, t]$, and all those that do not must be rejected. Also if motions do exist at a particular time, they may break down in one way or another shortly thereafter.

These observations sound a warning, but also they proffer encouragement. Some of the universal solutions may serve as models for explosions, implosions or tears and welds such as those that form from opening and closing cavities. Continuum mechanics is sometimes faulted for failure to include such phenomena. It is just, indeed, to state that they have not been much studied up to now in the serious literature on mathematical theory and will not be well understood until studied further.

16. Solids

In ordinary experience we commonly think of a body as being "solid" if after changing its form we can discern a difference in the way it responds to further deformation. A solid, then, has some placement, any non-rigid transplacement of which is detectable by some subsequent measurement of stress. Thus, still considering a particular material point, we lay down the following formal

Definition (NOLL). *A material is solid if there is a reference placement $\boldsymbol{\kappa}$ such that*

$$\mathscr{g}_{\boldsymbol{\kappa}} \subset \mathscr{o}. \qquad \text{(IV.16-1)}$$

Such a placement $\boldsymbol{\kappa}$ is called *undistorted.* According to this definition only orthogonal deformations belong to the peer group $\mathscr{g}_{\boldsymbol{\kappa}}$ corresponding with an undistorted $\boldsymbol{\kappa}$.

A material for which $\mathscr{g} = \{\mathbf{1}, -\mathbf{1}\}$ is solid. Such a material, which is called *triclinic*, furnishes an example of a *crystalline solid* in the classical sense. All the classical crystallographic groups, provided they be extended so as to include $-\mathbf{1}$, correspond with solids. So also do the groups defining "transversely isotropic" and "orthotropic" materials, and many others.

For solids, no particularly simple form of the constitutive relation has been found.

An *isotropic solid material,* of course, is a material that is both solid and isotropic. Both of these qualities have been defined in terms of the existence of special reference placements, both of which have been called "undistorted". Denoting by $\boldsymbol{\kappa}$ the one used to define "isotropic" and by $\bar{\boldsymbol{\kappa}}$ the one used to define "solid", for an isotropic solid

$$\mathscr{g}_{\boldsymbol{\kappa}} \supset \mathscr{o}, \qquad \mathscr{g}_{\bar{\boldsymbol{\kappa}}} \subset \mathscr{o}. \qquad \text{(IV.16-2)}$$

The relation between any such pair of reference placements is laid bare in the following

Theorem 1 (TRUESDELL & NOLL).

$$\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{g}_{\bar{\boldsymbol{\kappa}}} = \mathscr{o}. \qquad \text{(IV.16-3)}$$

Proof. According to the last statement in Section IV.14, either $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{o}$ or $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{u}$. If $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{u}$, then by (IV.13-1) $\mathscr{g}_{\bar{\boldsymbol{\kappa}}} = \mathscr{u}$; since $(2)_2$ contradicts this conclusion, we are left with the former alternative, $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{o}$. Thus $\boldsymbol{\kappa}$ is an undistorted placement of the solid.

If $\boldsymbol{\lambda} := \bar{\boldsymbol{\kappa}} \circ \boldsymbol{\kappa}^{-1}$ and $\mathbf{P} := \nabla\boldsymbol{\lambda}$, NOLL's rule (IV.13-1) ensures that there is a $\boldsymbol{\lambda}$ such that $\mathscr{g}_{\bar{\boldsymbol{\kappa}}} = \mathbf{P}\mathscr{o}\mathbf{P}^{-1}$. If we can find an orthogonal tensor $\mathbf{R}$ such that $\mathscr{g}_{\bar{\boldsymbol{\kappa}}} = \mathbf{R}\mathscr{o}\mathbf{R}^{-1}$, then we shall have proved the theorem, since the only orthogonal conjugate of $\mathscr{o}$ is $\mathscr{o}$ itself. That such an $\mathbf{R}$ exists, is a corollary of a more general theorem which is stated in the following exercise. $\triangle$

Exercise IV.16.1 (COLEMAN & NOLL). Let $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}^*$ be two undistorted placements of a solid, and let $\mathbf{P} = \nabla(\boldsymbol{\kappa}^* \circ \boldsymbol{\kappa}^{-1})$. If the polar decomposition of $\mathbf{P}$ is $\mathbf{P} = \mathbf{R}_0\mathbf{U}_0$, and if $\mathbf{Q}^*$ and $\mathbf{Q}$ are elements of $\mathscr{g}_{\boldsymbol{\kappa}^*}$ and $\mathscr{g}_{\boldsymbol{\kappa}}$ that correspond with one another through NOLL's rule, then

$$\mathbf{Q}^* = \mathbf{R}_0\mathbf{Q}\mathbf{R}_0^{\mathsf{T}}, \qquad \mathbf{U}_0 = \mathbf{Q}^{\mathsf{T}}\mathbf{U}_0\mathbf{Q}. \tag{IV.16-4}$$

Hence

$$\mathscr{g}_{\boldsymbol{\kappa}^*} = \mathbf{R}_0\mathscr{g}_{\boldsymbol{\kappa}}\mathbf{R}_0^{-1}. \tag{IV.16-5}$$

That is, $\mathscr{g}_{\boldsymbol{\kappa}^*}$ is an orthogonal conjugate of $\mathscr{g}_{\boldsymbol{\kappa}}$.

A body composed of solid material points is a *solid body*. A body of isotropic solid material is an *isotropic solid*. No confusion of "material" and "body" should ensue.

Returning to the consideration of a solid material in general, we remark that its peer group with respect to an undistorted placement may be any subgroup of the orthogonal group that contains $-\mathbf{1}$.

However, only certain particular kinds of anisotropy have attracted much notice until recently. These are the ones corresponding with the 32 crystal classes, which are defined by optical symmetries, and to two further types which correspond with some manufactured products. In order to define these particular symmetries, we let $\mathbf{R}_{\mathbf{a}}^{\varphi}$ denote a right-handed rotation of angle φ about an axis in the direction of the vector $\mathbf{a}$; we let $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ be an orthonormal basis, and we set $\mathbf{p} := \sqrt{\tfrac{1}{3}}(\mathbf{i}+\mathbf{j}+\mathbf{k})$. In view of (IV.12-7), it suffices to specify $\mathscr{g}^+$, which is a group of rotations.

A material such that $\mathscr{g}^+$ consists in $\mathbf{1}$ and all rotations $\mathbf{R}_{\mathbf{k}}^{\varphi}$ for a fixed $\mathbf{k}$ and all angles φ is called *transversely isotropic* with respect to $\mathbf{k}$.

The 32 crystal classes reduce to 11 in the context of the present, purely mechanical theory. Definitions of these, along with the standard crystallographic names, are given in the following table, summarizing conclusions derived by COLEMAN & NOLL. The directions of the particular unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are called the *crystallographic axes*.

Finally, a material is called *orthotropic* if $\mathscr{g}$ contains the reflections $-\mathbf{R}_{\mathbf{i}}^{\pi}$, $-\mathbf{R}_{\mathbf{j}}^{\pi}$, $-\mathbf{R}_{\mathbf{k}}^{\pi}$. Since $\mathbf{R}_{\mathbf{i}}^{\pi}\mathbf{R}_{\mathbf{j}}^{\pi} = \mathbf{R}_{\mathbf{k}}^{\pi}$ and $(\mathbf{R}_{\mathbf{i}}^{\pi/2})^2 = \mathbf{R}_{\mathbf{i}}^{\pi}$, the materials belonging to the classes numbered 3, 5, 6, and 7 in the table are orthotropic.

In this book we shall not have occasion to treat crystals or other materials of special symmetry, except, of course, isotropic materials. The definitions just given are included only so as to help the student understand the meanings of the terms, should he encounter them elsewhere.

	Crystal class	Generators[1] of $\mathscr{G}^+$	Order of $\mathscr{G}$
1.	*Triclinic system* all classes	$\mathbf{1}$	2
2.	*Monoclinic system* all classes	$\mathbf{R}_{\mathbf{k}}^{\pi}$	4
3.	*Rhombic system* all classes	$\mathbf{R}_{\mathbf{i}}^{\pi}, \mathbf{R}_{\mathbf{j}}^{\pi}$	8
4.	*Tetragonal system* tetragonal-disphenoidal tetragonal-pyramidal tetragonal-dipyramidal	$\mathbf{R}_{\mathbf{k}}^{\pi/2}$	8
5.	tetragonal-scalenohedral ditetragonal-pyramidal tetragonal-trapezohedral ditetragonal-dipyramidal	$\mathbf{R}_{\mathbf{k}}^{\pi/2}, \mathbf{R}_{\mathbf{i}}^{\pi}$	16
6.	*Cubic system* tetartoidal diploidal	$\mathbf{R}_{\mathbf{i}}^{\pi}, \mathbf{R}_{\mathbf{j}}^{\pi}, \mathbf{R}_{\mathbf{p}}^{2\pi/3}$	24
7.	hextetrahedral gyroidal hexoctahedral	$\mathbf{R}_{\mathbf{i}}^{\pi/2}, \mathbf{R}_{\mathbf{j}}^{\pi/2}, \mathbf{R}_{\mathbf{k}}^{\pi/2}$	48
8.	*Hexagonal system* trigonal-pyramidal rhombohedral	$\mathbf{R}_{\mathbf{k}}^{2\pi/3}$	6
9.	ditrigonal-pyramidal trigonal-trapezohedral hexagonal-scalenohedral	$\mathbf{R}_{\mathbf{i}}^{\pi}, \mathbf{R}_{\mathbf{k}}^{2\pi/3}$	12
10.	trigonal-dipyramidal hexagonal-pyramidal hexagonal-dipyramidal	$\mathbf{R}_{\mathbf{k}}^{\pi/3}$	12
11.	ditrigonal-dipyramidal dihexagonal-pyramidal hexagonal-trapezohedral dihexagonal-dipyramidal	$\mathbf{R}_{\mathbf{i}}^{\pi}, \mathbf{R}_{\mathbf{k}}^{\pi/3}$	24

[1] The members of a set of elements of a group $\mathscr{G}$ are called *the generators* of $\mathscr{G}$ if products of their powers exhaust $\mathscr{G}$.

Traditionally the use of these classical "point groups" is motivated by CAUCHY's theory of stress in a lattice of mass-points.[1] ERICKSEN[2] has pointed out that the arguments used apply only when $|\mathbf{F}-\mathbf{1}|$ is small. The theory of point lattices if taken seriously for transplacement gradients of great magnitude suggests that the peer groups of crystals should contain some non-orthogonal tensors, no matter what reference placement be used. Thus a crystal lattice does not serve as a model for a solid body in the sense defined by (1) and used throughout this book.

Returning to the consideration of solids as defined by (1), we note first that only certain particular placements will be undistorted. Indeed, if $\boldsymbol{\kappa}$ is undistorted, if $\boldsymbol{\kappa}^*$ is another reference placement, and if $\mathbf{P} := \nabla(\boldsymbol{\kappa}^* \circ \boldsymbol{\kappa}^{-1})$, by NOLL's rule we have

$$\mathcal{G}_{\boldsymbol{\kappa}^*} = \mathbf{P}\mathcal{G}_{\boldsymbol{\kappa}}\mathbf{P}^{-1}, \qquad \mathcal{G}_{\boldsymbol{\kappa}} \subset \mathcal{O}. \qquad \text{(IV.16-6)}$$

Now if $\mathbf{Q}$ is orthogonal, $\mathbf{PQP}^{-1}$ generally fails to be orthogonal. Thus *not all placements of a solid are undistorted.*

Exercise IV.16.2. Let $\mathcal{G}_{\boldsymbol{\kappa}}$ contain all rotations about $\mathbf{e}_3$, and let $\boldsymbol{\kappa}^*$ be obtained by the biaxial stretch such that $[\mathbf{P}] = \operatorname{diag}(\lambda, \lambda, \mu)$, $\lambda \neq \mu$. Then $\mathcal{G}_{\boldsymbol{\kappa}^*}$ contains all rotations about $\mathbf{e}_3$. Thus if $\boldsymbol{\kappa}$ is an undistorted placement of a material transversely isotropic with respect to $\mathbf{e}_3$, so also is $\boldsymbol{\kappa}^*$. For example, if $\boldsymbol{\kappa}$ is an undistorted placement of an isotropic solid (a special instance of a material transversely isotropic with respect to $\mathbf{e}_3$), then $\boldsymbol{\kappa}^*$ is an undistorted placement of a material transversely isotropic with respect to $\mathbf{e}_3$. In contrast, rotations of $\boldsymbol{\kappa}(\mathscr{B})$ about $\mathbf{e}_1$ are not carried into rotations of $\boldsymbol{\kappa}^*(\mathscr{B})$ about $\mathbf{e}_1$. Thus, even if $\boldsymbol{\kappa}$ is an undistorted placement of an isotropic solid $\mathscr{B}$, $\boldsymbol{\kappa}^*$ is a distorted placement of $\mathscr{B}$.

Exercise IV.16.3. Application of (5) shows that the peer groups corresponding with different undistorted placements of a particular solid are not generally the same, and that the undistorted placements of an anisotropic solid generally fail to be peers.

We may set ourselves the following task: to find the largest class of mappings $\boldsymbol{\lambda}$ that carry places in an undistorted placement $\boldsymbol{\kappa}$ defined by a given group $\mathcal{G}_{\boldsymbol{\kappa}}$ into places in another undistorted placement.

For the largest and smallest possible peer groups, the answer is easy to get. First, if $\mathcal{G}_{\boldsymbol{\kappa}} = \{\mathbf{1}, -\mathbf{1}\}$, then, as shown in Section IV.13, all placements are

[1]An outline of this theory is given by A. E. H. LOVE in Note B, "The notion of stress," in his *A Treatise on the Mathematical Theory of Elasticity*, Cambridge, Cambridge University Press, 2nd–4th editions, 1906/1927, variously reprinted.

[2]J. L. ERICKSEN, "Nonlinear elasticity of diatomic crystals," *International Journal of Solids and Structures* **6**(1970): 951–952, and Chapter IV of "Special topics in elastostatics," *Advances in Applied Mechanics*, **17**(1977): 179–244.

undistorted, and so any $\boldsymbol{\lambda}$ has the property sought. The second case is settled by

Theorem 2 (COLEMAN & NOLL). *A transplacement of an isotropic solid maps one undistorted shape onto another if and only if it is conformal.*[1]

The sufficiency of the condition follows trivially from a theorem established in Section IV.13. Necessity is a consequence, as we shall see presently, of the following more general

Theorem 3 (COLEMAN & NOLL). *Let $\boldsymbol{\kappa}$ be an undistorted placement of a solid body $\mathscr{B}$. If $\boldsymbol{\lambda} := \boldsymbol{\kappa}^* \circ \boldsymbol{\kappa}^{-1}$, so that $\boldsymbol{\lambda}$ maps $\boldsymbol{\kappa}(\mathscr{B})$ onto $\boldsymbol{\kappa}^*(\mathscr{B})$, then $\boldsymbol{\kappa}^*$ is undistorted if and only if the proper spaces of the right stretch tensor $\mathbf{U}_0$ of $\nabla\boldsymbol{\lambda}$ are invariant under all the rotations in the peer group $\mathscr{g}_{\boldsymbol{\kappa}}$.*

Proof. By $(4)_2$, every member $\mathbf{Q}$ of $\mathscr{g}_{\boldsymbol{\kappa}}$ commutes with $\mathbf{U}_0$. According to a theorem of algebra,[2] $\mathbf{Q}$ satisfies this condition if and only if it leaves the proper spaces of $\mathbf{U}_0$ invariant.

Exercise IV.16.4. The statement of sufficiency in Theorem 2 is a corollary of Theorem 3. $\triangle$

We turn now to the use of Theorem 3 so as to complete the proof of necessity in Theorem 2. By Theorem 1 we know that if $\boldsymbol{\kappa}$ is undistorted, $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{o}$. If $\boldsymbol{\lambda}$ carries $\boldsymbol{\kappa}$ into another undistorted placement $\boldsymbol{\kappa}^*$, then by Theorem 3 every orthogonal tensor must leave invariant the proper spaces of $\mathbf{U}_0$. Therefore, the proper space of $\mathbf{U}_0$ can be nothing but $\mathscr{V}$ itself. Hence $\mathbf{U}_0$ has only one proper number, so that $\mathbf{U}_0 = K\mathbf{1}$, and consequently $\nabla\boldsymbol{\lambda} = K\mathbf{R}$. $\triangle$

Theorem 3 itself may be used to determine the most general form of $\mathbf{U}_0$ compatible with a given $\mathscr{g}_{\boldsymbol{\kappa}}$ in cases other than the two already disposed of: isotropic or triclinic solids. The outcomes for the crystalline solids are shown in the following table, due to COLEMAN & NOLL. The numbers in parentheses

[1] A transplacement $\boldsymbol{\lambda}$ is conformal if it preserves the angles between material curves: equivalently, there is an orthogonal tensor $\mathbf{R}$ such that $\nabla\boldsymbol{\lambda} = K\mathbf{R}$ and $K \neq 0$.

[2] This theorem is a corollary of Theorem 2, Section 43, and Theorem 3, Section 79, of P. R. HALMOS, *Finite-Dimensional Vector Spaces*, 2nd ed., Princeton, Van Nostrand 1958.

refer to the definitions of the special kinds of aeolotropy in the table printed above on p. 267.

Type of aeolotropy	Restrictions on $\mathbf{U}_0$
Triclinic system (1)	no restriction
Monoclinic system (2)	$\mathbf{k}$ is a proper vector of $\mathbf{U}_0$
Rhombic system (3)	$\mathbf{i}, \mathbf{j}, \mathbf{k}$ are proper vectors of $\mathbf{U}_0$
Tetragonal system (4, 5)	
Hexagoanl system (8, 9, 10, 11)	$\mathbf{U}_0 = A\mathbf{1} + B\mathbf{k} \otimes \mathbf{k}$
Transverse isotropy	
Cubic system (6, 7)	$\mathbf{U}_0 = A\mathbf{1}$

17. Fluids

There are various physical notions concerned with fluids. One is that a fluid is a substance which can flow. "Flow" itself is a vague term. One meaning of "flow" is simply deformation under stress, which does not distinguish a fluid from any other material not rigid. Another is that steady velocity results from constant stress, which seems to be special and to apply only with difficulty and to particular flows. Another is inability to support shear stress when in equilibrium. Formally, within the theory of simple materials, such a definition would yield

$$\mathbf{T} = -p(\rho)\mathbf{1} + \mathfrak{F}(\mathbf{F}^t), \qquad \text{(IV.17-1)}$$

where $\mathfrak{F}(\mathbf{1}^t) = \mathbf{0}$, $\mathbf{1}^t$ being the history whose value is always $\mathbf{1}$. Since the material so defined may have any peer group whatever, including one of those already considered to define a solid, this definition does not lend itself to a criterion in terms of common response.

Exercise IV.17.1. The constitutive relation $\mathbf{T} = -K(3 - \operatorname{tr}\mathbf{U})\mathbf{1}$, $K = \text{const.} \neq 0$, defines an isotropic, elastic solid which has infinitely many placements at ease and never experiences non-vanishing shear stress, no matter how it be deformed.

Finally, a fluid is regarded as a material having "no preferred configuration". In terms of peer groups we may realize this somewhat vague idea by the following

Definition. *A fluid is an egalitarian material that is not solid.*

In Section IV.13 we have shown that for an egalitarian material either $\mathscr{g} = \{\mathbf{1}, -\mathbf{1}\}$ or $\mathscr{g} = \mathscr{u}$. The former case corresponds with a solid, according to the definition given in Section IV.16. Thus we have the following

Theorem. *A material is fluid if and only if for some* κ

$$\mathscr{g}_\kappa = \mathscr{u}. \tag{IV.17-2}$$

For a fluid material, (2) *holds for all* κ.

From this theorem, some preceding ones, and the definitions, we read off the following trivial but important corollaries:

1. *Every fluid is isotropic.*
2. *Every placement of a fluid is undistorted.*
3. *A material is egalitarian if and only if it is either a fluid or a triclinic solid.*
4. *The only isotropic materials are fluids and isotropic solids.*

The condition (2) was laid down as the definition of a fluid by NOLL, who derived thereupon the following

Fundamental Theorem on Fluids. *Every unconstrained fluid has a constitutive relation of the form*

$$\mathbf{T} = \mathfrak{K}(\mathbf{C}_t^t\,;\,\rho); \tag{IV.17-3}$$

also

$$\mathfrak{K}(\mathbf{Q}\mathbf{C}_t^t\mathbf{Q}^\mathsf{T};\,\rho) = \mathbf{Q}\mathfrak{K}(\mathbf{C}_t^t;\,\rho)\mathbf{Q}^\mathsf{T} \tag{IV.17-4}$$

for all orthogonal $\mathbf{Q}$ *and all arguments* $\mathbf{C}_t^t$ *and* ρ *that lie in the domain of* $\mathfrak{K}$. *Every such isotropic mapping of positive, symmetric tensor histories onto symmetric tensors defines a fluid. Furthermore,*

$$\mathfrak{K}(\mathbf{1}^t;\,\rho) = -p(\rho)\mathbf{1}. \tag{IV.17-5}$$

This last conclusion states that all fluids obey in rigid motion, such as a state of rest, the laws of Eulerian hydrostatics, according to which the stress is a hydrostatic pressure which depends on the density alone. In particular, a fluid exhibits the phenomenon of "flow" in one of the common senses, namely, it cannot support any shear stress when it has been at rest for all times, past and present, in any placement whatever. As we have shown at the beginning of this section, the converse is false: a material capable of "flow" in this sense may have any peer group.

Proof of Noll's theorem. Since a fluid is isotropic and every placement is undistorted, we may apply (IV.14-2) for any reference placement κ. Because the stress in a fluid cannot be changed by a static deformation from one placement to another with the same density, the dependence upon $\mathbf{B}(t)$ in (IV.14-2) must reduce to dependence on $\det \mathbf{B}(t)$, or, what is the same thing, dependence on ρ, and this fact establishes the necessity of (3). Furthermore, $\mathfrak{K}$ must satisfy (IV.14-3), which now reduces to (4). If $\mathbf{C}_t^t = \mathbf{1}^t$, then (4) yields

$$\mathbf{T} = \mathfrak{K}(\mathbf{1}^t; \rho) = \mathbf{Q}\mathfrak{G}(\mathbf{1}^t, \rho)\mathbf{Q}^{\mathsf{T}} = \mathbf{Q}\mathbf{T}\mathbf{Q}^{\mathsf{T}}. \qquad \text{(IV.17-6)}$$

Thus in a fluid which has always been at rest $\mathbf{T}$ commutes with every orthogonal tensor. The conclusion of Exercise IV.4.1 establishes the necessity of (5).

Exercise IV.17.2. The relations (3) and (4) imply that $\mathscr{g}_{\boldsymbol{\kappa}} = \mathscr{u}$ for every $\boldsymbol{\kappa}$. This exercise completes the proof of NOLL's theorem. △

We may express the foregoing theorem also as follows: The constitutive relation of a fluid is of the form

$$\mathbf{T} = -p(\rho)\mathbf{1} + \mathfrak{C}(\mathbf{C}_t^t - \mathbf{1}; \rho); \qquad \text{(IV.17-7)}$$

the mapping $\mathfrak{C}$ is isotropic, and its value is naught when its argument is the history $\mathbf{0}^t$ whose value is always $\mathbf{0}$. Conversely, every relation of this form defines a fluid.

A trivial corollary of the foregoing, which may be proved in several other ways, states that *any relation of the form* (IV.4-4) *defines an elastic fluid.* While in hydrodynamics it is customary to impose the condition that $p(\rho) > 0$ for all ρ, or at least the weaker requirement that $p(\rho) > 0$ for all but a discrete set of values of ρ, this condition does not follow from any general principle of mechanics. Because steady, hydrostatic tensions of some magnitude have been produced, with extreme pains, in very quiet laboratories, perhaps the condition $p(\rho) > 0$ should be regarded as expressing stability rather than a constitutive restriction.

From (IV.14.4) it is clear that the determinate stress of a *homogeneous incompressible fluid* has a constitutive relation of the form

$$\mathbf{S} = \mathfrak{K}(\mathbf{C}_t^t); \qquad \text{(IV.17-8)}$$

the mapping $\mathfrak{K}$ need not be defined except for arguments such that $\det \mathbf{C}_t^t = 1$; $\mathfrak{K}$ must satisfy the condition of isotropy expressed by (4); and there is no loss in generality if we require that $\mathfrak{K}(\mathbf{1}^t) = \mathbf{0}$. Conversely, every mapping of this kind defines a homogeneous, incompressible fluid.

Finally, since for a rigid motion $\mathbf{C}_t^t = \mathbf{1}^t$, the constitutive relation (3) reduces to (IV.4-4): *A body of unconstrained fluid in rigid rotation behaves like* *a body of Eulerian fluid.* For a homogeneous, incompressible fluid body a similar statement holds, and in steady rotation about a fixed axis

$$pv + \varpi = \tfrac{1}{2}\omega^2 r^2, \tag{IV.17-9}$$

r being the distance from the axis.

Exercise IV.17.3. The statement (9) follows by use of (II.11-53), (II.11-52), and (IV.10-2).

A body of fluid material is a *fluid body*. Often the noun "fluid" is used equally to refer to a material or a body. No confusion ought result.

A fluid may react to its entire transplacement history, yet its reaction cannot be different for different placements with the same density. A fluid reconciles these two seemingly contradictory qualities—ability to remember all its past and inability to regard one placement as different from another—by reacting to the past only insofar as it may differ from the present, which may be ever changing.

18. Universal Flows of Homogeneous Incompressible Fluids

Universal motions, transplacements, and flows have been defined in Section IV.9. Five families of universal transplacements for isotropic, incompressible, homogeneous bodies have been presented and discussed in Section IV.15. Since homogeneous, incompressible fluids constitute a proper subclass of homogeneous, incompressible, isotropic materials, the class of universal motions of fluids may be greater than that given by the five families presented and analysed in Section IV.15. Whether such is the case, is not presently known. Thus this section cannot include anything not a consequence of what has appeared in Section IV.15.

Nevertheless, we shall here remark upon some universal flows. The student will recall the main advantage of the spatial description: While a motion is defined in terms of a particular reference placement, the spatial velocity-field is unique, independent of reference placements. Infinitely many transplacements, one for each choice of reference placement, give rise to the same flow. We have seen examples in Section I.15: namely, the flow of Family 2 is homogeneous and hence corresponds also with a homogeneous motion, and Families 1 and 3 give rise to the same flow. In doing Exercises IV.15.4 and IV.15.5 the student will have confirmed this observation in two examples.

In researches on the dynamics of incompressible fluids of various kinds, members of two particular families of steady flows are often mentioned, usually because they satisfy the dynamical equations and provide exceptions, perhaps rather degenerate ones, to some otherwise general statements. Their definitions in terms of contravariant cylindrical components follow.

$$\dot{r} = -ar + b/r, \qquad \dot{\theta} = a\theta + \tfrac{1}{2}c, \qquad \dot{z} = az + g, \tag{IV.18-1}$$

$$\dot{r} = h/r, \qquad \dot{\theta} = k, \qquad \dot{z} = l\theta + m; \tag{IV.18-2}$$

a, b, c, g, h, j, l, and m are arbitrary constants. These flows generally represent expansion or contraction of concentric cylinders superimposed upon azimuthal and longitudinal stretches and shears; the most familiar instance has streamlines that are logarithmic spirals, which include as a limiting instance flow between concentric, co-axial cylinders. We note that the instance $a = 0$ in (1) and the instance $l = 0$ in (2) are identical.

Exercise IV.18.1 (MARRIS). These flows preserve circulation; they are irrotational if and only if, respectively, $a = c = 0$ and $k = l = 0$, and for their principal stretchings to be all constant, it is necessary and sufficient that $b = 0$ and $h = l = 0$, respectively. Also if $h = 0$, then $l \neq 0$ unless the motion is rigid.

Exercise IV.18.2. The flows (1) and (2) are steady instances of $(\text{IV}.15.32)_{1,2,3}$ to within an arbitrary, steady rotation about the z axis and a steady translation along it.

Following WANG & MARRIS,[1] we render the statement of the preceding exercise transparent by calculating relative transplacements corresponding with (1) and (2). A relative description of (1) is obtained by integrating the appropriate instance of (II.8-3). Using a prime to denote differentiation with respect to the time lapse s, we obtain

$$r' = -ar + b/r, \qquad \theta' = a\theta + \tfrac{1}{2}c, \qquad z' = az + g. \tag{IV.18-3}$$

Exercise IV.18.3. If $a \neq 0$,

$$\begin{aligned} r^2 - \frac{b}{a} &= \left(R^2 - \frac{b}{a}\right)e^{-2as}, \\ \theta + \frac{c}{2a} &= \left(\Theta + \frac{c}{2a}\right)e^{as}, \\ z + \frac{g}{a} &= \left(Z + \frac{g}{a}\right)e^{as}. \end{aligned} \tag{IV.18-4}$$

For (2)

$$r' = h/r, \qquad \theta' = k, \qquad z' = l\theta + m, \tag{IV.18-5}$$

[1] C.-C. WANG & A. W. MARRIS, "Proof that motions obtained in the preceding paper by Marris are universal for all incompressible isotropic simple materials," *Archive for Rational Mechanics and Analysis* **69**(1979): 381–390.

which includes as a special instance the result of taking $a = 0$ in (3). Hence

$$r^2 = R^2 + 2hs, \qquad \theta = \Theta + ks, \qquad z = Z + (l\theta + m)s + \tfrac{1}{2}kls. \quad \text{(IV.18-6)}$$

The minuscules $a, b, \ldots, m$ occurring on the right-hand sides of (4) and (6) are the very ones appearing in the flows (1) and (2).

In (4) and (6) the arbitrary constants of integration R, Θ, and Z may be taken as the polar co-ordinates of body points in a reference placement corresponding with $s = 0$, though no such placement need be occupied by the body, and that interpretation is not compelled.

Exercise IV.18.4 (WANG & MARRIS). The functions $A, B, \ldots, H$ in (IV.15-4) may be so chosen as to yield motions that deliver the flows (1) and (2).

Corresponding thus with instances of (IV.15-4), *the flows* (1) *and* (2) *are universal for homogeneous, incompressible, isotropic bodies* provided the conditions following from (IV.15-34) be satisfied; the corresponding stresses are then determined by (IV.15-17). The values $A(0), B(0), \ldots, H(0)$ deliver the position of a body-point in the chosen reference placement, which is presumed to be homogeneous and undistorted. That reference placement need not ever be occupied by the body-point in consideration, but of course it may be.

Exercise IV.18.5 (WANG & MARRIS). Examination of the effects of the signs of a and b in (4) and of h in (6) exemplifies the statements made above at the end of Section IV.15. In general the motions delivering (1) and (2) can be maintained with physically reasonable connotation at most for a semi-infinite interval of time.

The steady universal flows (1) and (2) have a special status. In performing Exercise IV.18.1 the student will have shown that with specified exceptions they are rotational and have fields of principal stretchings that are not constant in space. MARRIS[1] proved in a long and difficult analysis of a small class of fluids that *there are no other universal flows having these properties*. By a simple and direct calculation based on the flows to which the five families of universal transplacements discussed in Section IV.15 give rise, WANG & MARRIS[I] proved that among those flows all but (1) and (2) are either unsteady or irrotational or have three constant principal stretchings or do not preserve circulation. The conditions required to reduce the five families to their subclasses that do preserve circulation have been provided above in Section IV.15.

Perhaps there are universal flows not delivered by any of the five families of transplacements mentioned. Perhaps there are universal flows of incompressible fluids that are not universal for isotropic, incompressible solids. MARRIS's theor-

[1] A. W. MARRIS, "Steady universal motions of Rivlin-Ericksen fluids," *Archive for Rational Mechanics and Analysis* **69**(1979): 335–380.

em tells us that any further universal flows, be they of incompressible fluids only or of isotropic, incompressible materials in general, will necessarily have three constant principal stretchings or be unsteady or irrotational.

For incompressible fluids the class of all universal flows in which all three principal stretchings are constant and distinct seems to be abundant and very difficult to delimit.

19. Steady Rotation of a Homogeneous Body of Incompressible Fluid Loaded by Surface Tension

In Section III.9 we have presented the theorem of DAY which states that a free body whose rotational momentum is not null will ultimately develop negative pressures unless its diameter tends to ∞. In particular, such will be the case for a free body in rigid rotation. It is natural to ask if surface tension, as defined by (III.8-3) with a positive coefficient σ, can overcome the tendency of a spinning body of fluid to fly asunder. DAY, again acknowledging the influence of SUNDMAN's work, has found circumstances sufficient that such be true of a homogeneous body of incompressible simple fluid. For the generalized form in which his analysis is presented here I am indebted to R. BATRA.

Theorem. *Let a body of incompressible fluid of uniform density be loaded when $t \geqq 0$ by surface tension alone. If the rotational momentum H, mass M, and volume V are related as follows to the coefficient σ of surface tension*:

$$0 < H < 1.051\sigma^{1/2}M^{1/2}V^{2/3}, \qquad \text{(IV.19-1)}$$

then when $t \geqq 0$ the body may undergo a steady, rigid rotation in a shape such that $p > 0$ everywhere.

Preliminaries to the proof. The student will recall from Exercise III.8.2 that a body loaded by surface tension alone is isolated, and so its center of mass remains fixed and its rotational momentum constant. From the developments in Section IV.17 we know that any homogeneous, incompressible fluid body undergoing steady rotation obeys the constitutive relation of an Eulerian fluid; the pressure on such a fluid in steady rotation is given by (IV.17-9), which here reduces to

$$p = p_0 + \tfrac{1}{2}\rho\omega^2 r^2, \qquad \text{(IV.19-2)}$$

r being the distance from the axis of spin, ω the angular speed, and p_0 the pressure of the fluid at points on the axis. The theory presented in Section I.13

shows that that axis is parallel to the rotational momentum and is a principal axis of inertia of the body in its shape at the time 0 and thereafter.

We shall construct a solution for a body whose shape is bounded by a surface of revolution about the axis of spin. We shall call that axis the z-axis and assume that the body has an equatorial plane, and so we may take the generating curve as having the equation $z = \pm f(r)$, $0 \leqq r \leqq a$. We shall assume that the bounding surface of the body has a continuous normal field. Thus

$$f(a) = 0, \qquad f'(0) = 0, \qquad f'(r) \to -\infty \quad \text{as} \quad r \to a. \tag{IV.19-3}$$

Under these assumptions we shall show first that at most one choice of a and f exists. It will be easy then to see that that choice does indeed satisfy the conditions set down and so establishes the theorem.

Proof of Day's Theorem. Because of (2) the boundary condition (III.8-3) assumes the form

$$-\frac{\sigma}{r}\left[\frac{rf'}{(1+f'^2)^{1/2}}\right]' = p_0 + \frac{1}{2}\rho\omega^2 r^2, \qquad 0 \leqq r \leqq a. \tag{IV.19-4}$$

Integrating this differential equation and using the conditions (3), we find that

$$\begin{aligned} p_0 &= \frac{2\sigma}{a} - \frac{1}{4}\rho\omega^2 a^2, \\ -\frac{f'}{(1+f'^2)^{1/2}} &= \frac{r}{a}\left[1 - K\left(1 - \frac{r^2}{a^2}\right)\right], \end{aligned} \tag{IV.19-5}$$

in which

$$K = \frac{\rho\omega^2 a^3}{8\sigma}. \tag{IV.19-6}$$

Now using (2) and $(5)_1$, we calculate p explicitly:

$$p = \frac{2\sigma}{a}\left[1 - K\left(1 - \frac{2r^2}{a^2}\right)\right]. \tag{IV.19-7}$$

Thus in order that $p > 0$ throughout the shape of the body, it is necessary and sufficient that

$$K < 1. \tag{IV.19-8}$$

Integration of $(5)_2$ determines the shape of the body, but we are not directly interested in that.

Exercise IV.19.1.

$$V = 2\pi a^3 g_2(K), \qquad H = \frac{\pi\omega a^5 M}{V} g_4(K); \tag{IV.19-9}$$

the functions g_a are defined as follows:

$$g_a(K) = \int_0^1 \frac{x^{a+1}(1-K+Kx^2)}{[1-x^2(1-K+Kx^2)^2]^{1/2}}\,dx. \tag{IV.19-10}$$

From (9) and (6) we see that

$$\frac{H}{\sigma^{1/2}M^{1/2}V^{2/3}} = 2^{1/3}\pi^{-1/6}\frac{K^{1/2}g_4(K)}{[g_2(K)]^{7/6}} := h(K), \tag{IV.19-11}$$

say. Regarding V, H, M, and σ as given, we seek a value of K such as to satisfy (11). If we can find it, we can then determine a from $(9)_1$ and thereafter determine ω from $(9)_2$. We shall then have determined the shape and spin of the body. If, furthermore, we can satisfy (8), then it will follow that $p > 0$ everywhere in the shape of the body.

Thus it remains to solve (11). The definition (10) shows that g_2 and g_4 are continuous and positive on $[0, 1]$. Because, therefore, h as defined by (11) is continuous on $[0, 1]$, it assumes every value in the interval $[h(0), h(1)]$. Now $h(0) = 0$, and $h(1)$ can be estimated numerically. According to DAY,

$$h(1) = \frac{2^{1/3}\pi^{-1/6}g_4(1)}{[g_2(1)]^{7/6}} > 1.051. \tag{IV.19-12}$$

It follows, then, that if

$$\frac{H}{\sigma^{1/2}M^{1/2}V^{2/3}} < 1.051, \tag{IV.19-13}$$

there is a value of K in $[0, 1]$ such as to satisfy (11). Obviously 0 violates $(1)_1$, while (12) shows that 1 is too large. $\triangle$

20. Fluid Crystals

To exhaust the possible types of simple materials, any material that is not a solid we shall call a *fluid crystal*. For a fluid crystal, then $\mathscr{g}_{\boldsymbol{\kappa}} \not\subset \mathscr{o}$, no matter what be the reference placement $\boldsymbol{\kappa}$. Thus the peer group with respect to every placement has some elements which are not orthgonal. That is, there is always some change of shape that no experiment on the stress can detect. In this regard a fluid crystal is like a fluid, for which no change of shape without change of density is detectable by measurement of stress. Since it is impossible that $\mathscr{g}_\kappa \supset o$ unless the fluid crystal be in fact a fluid, for a fluid crystal not isotropic some rotations are detectable. In this property an anisotropic fluid resembles an anisotropic solid.

The definitions and theorems in the preceding section show that *a fluid crystal is a fluid if and only if it is isotropic.*

In this book we shall not go any further into the theory of fluid crystals.[1]

Exercise IV.20.1. A "Venn diagram" represents the exhaustive classification of peer groups.

21. Monotonous Motions

Continuum mechanics, even the mechanics of simple materials, covers so vast a range of possible behavior that little can be learnt from it without descending to instances. In this complexity continuum mechanics mirrors nature itself, for only by specifying particular features of a phenomenon can we so much as name it, let alone describe it. In the mechanics of simple materials two kinds of specialization are fruitful:

1. of the material,
2. of the motions a body is forced to undergo.

We have given examples of the former in the immediately preceding sections. The constitutive relations of fluids and isotropic solids are simpler than the general one, and we can expect the solution of problems for these two classes of bodies to be relatively easier than for anisotropic solids or fluid crystals.

[1] The peer groups of certain fluid crystals have been defined and interpreted by B. D. COLEMAN, "Simple liquid crystals," *Archive for Rational Mechanics and Analysis* **20**, 41–58 (1965), and C.-C. WANG, "A general theory of subfluids," *ibid.* **20**(1965): 1–40.

Fluid crystals as defined here are not to be confused with the "liquid crystals" occurring in physics; those liquid crystals do not fit into the framework established and studied in this book, although they are simple materials in the more general sense introduced in NOLL's paper of 1972, which is cited at the end of this chapter. Surveys of the vast literature on theories of liquid crystals are available:

1. *Static Theory.* J. L. ERICKSEN, "Equilibrium theory of liquid crystals," *Advances in Liquid Crystals*, Vol. 2, ed. G. BROWN, New York, Academic Press, 1976.
2. *Dynamic Theory.* F. M. LESLIE, "Theory of flow phenomena in liquid crystals," *ibid.* Vol. 4, New York, Academic Press, 1979.

The continuum mechanics of the last century carried this kind of specialization much further and restricted attention to materials specified by one or two constants. As a result, the solution of many boundary-value problems became easy—deceptively so, since only rarely can the properties of natural bodies be condensed adequately into one or two numbers fit to be tabulated in a manual.

We have given a specimen of the second simplification in Sections IV.9 and IV.10, where we have seen that we may determine, once and for all, all homogeneous transplacements that can be produced in an arbitrary homogeneous simple body by bringing to bear suitable tractions upon its boundary. In Section IV.15 we have derived and displayed the universal solutions made possible by two specializations: incompressibility as well as isotropy. We now define and analyse certain particular motions in which the effects of material memory, which for a simple material may indeed be various and complicated in a general motion, are given little chance to manifest themselves, because there is little to remember.

Consider, for example, the constitutive equation of a simple fluid:

$$\mathbf{T} = \mathfrak{R}(\mathbf{C}_t^t\,;\,\rho). \qquad \text{(IV.17-3)}_{\mathrm{r}}$$

In the particular case when ρ = const. and $\mathbf{C}_t^t(s)$ is the same function of s for all t, the stress becomes constant in time for a given material point. The fluid body may have undergone transplacements for all past time, but as each material point looks backward, so to speak, it sees the entire sequence of past transplacements referred to its present placement remain unchanged.

More generally, since the Principle of Material Frame-Indifference (Section IV.5) forbids past rotations to enter the constitutive relation and renders explicit the effect of present rotation, we should be able to simplify the constitutive relation almost as much in the more general circumstances when, for some orthogonal tensor $\mathbf{Q}(t)$,

$$\mathbf{C}_t^t(s) = \mathbf{Q}(t)\mathbf{C}_0^0(s)\mathbf{Q}(t)^{\mathsf{T}}, \qquad 0 \leqq s < \infty. \qquad \text{(IV.21-1)}$$

Here $\mathbf{C}_0^0$ denotes $\mathbf{C}_t^t$ when $t = 0$, and $\mathbf{Q}(0) = \mathbf{1}$. COLEMAN isolated motions of this kind as a class and called them *substantially stagnant*. In them, an observer situate upon the moving material point may choose his frame in such a way as to see behind him always the same transplacement history referred to the present placement. The proper numbers of $\mathbf{C}_t^t(s)$ are for a given s and any t the same as those of $\mathbf{C}_0^0(s)$, although the principal axes of the one tensor for a given s may rotate arbitrarily with respect to those of the other as t increases. Thus, while the principal relative stretches $v_{(t)k}$ generally vary with t, they do so in such a way that their histories up to the time t remain unchanged:

$$v^{t}_{(t)k} = v^{0}_{(0)k}, \qquad k = 1, 2, 3, \quad -\infty < t < \infty. \qquad \text{(IV.21-2)}$$

Thus a substantially stagnant motion is a motion having constant principal relative stretch histories. A simpler name is *monotonous motion*.

Since the definition of this property makes no use of a fixed reference placement, it pertains to the motion itself rather than to any of its embodiments as a transplacement. Moreover, in view of a conclusion derived in Exercise I.11.2, this property is a frame-indifferent one.

We turn now to the pure kinematics of monotonous motions. They are characterized by the following

Fundamental Theorem (NOLL). *A motion is monotonous if and only if there are an orthogonal tensor* $\mathbf{Q}(t)$, *a scalar* κ, *and a constant tensor* $\mathbf{N}_0$ *such that*

$$\begin{aligned} \mathbf{F}_0(\tau) &= \mathbf{Q}(\tau)e^{\tau\kappa\mathbf{N}_0}, \\ \mathbf{Q}(0) &= \mathbf{1}, \qquad |\mathbf{N}_0| = 1. \end{aligned} \qquad \text{(IV.21-3)}$$

Proof. We begin from the hypothesis (1) and set

$$\mathbf{H}(s) := \mathbf{C}_0(-s) = \mathbf{Q}(t)^{\mathsf{T}}\mathbf{C}_t(t-s)\mathbf{Q}(t). \qquad \text{(IV.21-4)}$$

By (II.8-8), $\mathbf{F}_t(\tau) = \mathbf{F}_0(\tau)\mathbf{F}_0(t)^{-1}$, and so

$$\begin{aligned} \mathbf{Q}(t)\mathbf{H}(s)\mathbf{Q}(t)^{\mathsf{T}} &= \mathbf{C}_t(t-s), \\ &= [\mathbf{F}_0(t)^{\mathsf{T}}]^{-1}\mathbf{C}_0(t-s)\mathbf{F}_0(t)^{-1}, \\ &= [\mathbf{F}_0(t)^{\mathsf{T}}]^{-1}\mathbf{H}(s-t)\mathbf{F}_0(t)^{-1}. \end{aligned} \qquad \text{(IV.21-5)}$$

If

$$\mathbf{E}(t) := \mathbf{Q}(t)^{\mathsf{T}}\mathbf{F}_0(t), \qquad \text{(IV.21-6)}$$

then (5) assumes the form of a difference equation:

$$\mathbf{H}(s-t) = \mathbf{E}(t)^{\mathsf{T}}\mathbf{H}(s)\mathbf{E}(t). \qquad \text{(IV.21-7)}$$

To obtain a necessary condition for a solution $\mathbf{H}(s)$, we differentiate[1] (7) with respect to t and put $t = 0$, obtaining the first-order linear differential equation

$$-\dot{\mathbf{H}}(s) = \mathbf{M}^{\mathsf{T}}\mathbf{H}(s) + \mathbf{H}(s)\mathbf{M}; \qquad \text{(IV.21-8)}$$

here $\mathbf{M} \equiv \dot{\mathbf{E}}(0)$, and the dot denotes differentiation with respect to s. The unique solution of (8) such that $\mathbf{H}(0) = \mathbf{1}$ is easily seen to be

$$\mathbf{H}(s) = e^{-s\mathbf{M}^{\mathsf{T}}}e^{-s\mathbf{M}}. \qquad \text{(IV.21-9)}$$

Since histories are defined only when $s \geqq 0$, this formula has been derived only for that domain. Nevertheless, the difference equation (7) serves to define $\mathbf{H}(s)$ for negative s as well and shows that $\mathbf{H}$ is analytic. Since the right-hand side of (9) is analytic, the principle of analytic continuation shows that (9) gives the unique solution for all s, when $\mathbf{E}(t)$ is assigned. If we substitute (9) back into (7), by putting $s = 0$ we obtain

$$[\mathbf{E}(t)e^{-t\mathbf{M}}]^{\mathsf{T}}\mathbf{E}(t)e^{-t\mathbf{M}} = \mathbf{1}. \qquad \text{(IV.21-10)}$$

Hence $\mathbf{E}(t)e^{-t\mathbf{M}}$ is an orthogonal tensor, say $\bar{\mathbf{Q}}(t)$. By (6), then,

$$\mathbf{F}_0(t) = \mathbf{Q}(t)\bar{\mathbf{Q}}(t)e^{t\mathbf{M}}. \qquad \text{(IV.21-11)}$$

We may define κ and $\mathbf{N}_0$ to within sign as follows:

$$\kappa\mathbf{N}_0 := \mathbf{M}, \qquad |\mathbf{N}_0| = 1, \qquad \text{(IV.21-12)}$$

and so (3) follows. The scalar field κ is generally called the ***shearing***. The proof reveals that the orthogonal tensor function appearing in the conclusion (3) is not generally the same as the $\mathbf{Q}$ in the hypothesis (1). Conversely, if (3) holds, an easy calculation shows that the motion is monotonous. △

Exercise IV.21.1 (NOLL). In a monotonous motion

$$\mathbf{F}_t(\tau) = \mathbf{Q}(\tau)\mathbf{Q}(t)^{\mathsf{T}}e^{(\tau-t)\kappa\mathbf{N}} = \mathbf{Q}(\tau)e^{(\tau-t)\kappa\mathbf{N}_0}\mathbf{Q}(t)^{\mathsf{T}}, \qquad \text{(IV.21-13)}$$

[1] That the assertion of the theorem remains true even if $\mathbf{H}$ is merely continuous and $\mathbf{E}$ is completely arbitrary, has been shown by W. NOLL, "The representation of monotonous processes by exponentials," *Indiana University Mathematics Journal* **25**(1976): 209–214. On pp. 338–339 of the earlier paper cited in Footnote 1 on p. 286 WANG had proved NOLL's Fundamental Theorem by use of a "minor continuity assumption" weaker than continuity.

$\mathbf{N}$ being defined as follows:

$$\mathbf{N} := \mathbf{Q}(t)\mathbf{N}_0\mathbf{Q}(t)^{\mathsf{T}}, \qquad \text{(IV.21-14)}$$

and so $|\mathbf{N}| = 1$; conversely, if $\mathbf{F}_t(\tau)$ has the form (13), any motion to which it corresponds is monotonous. In such a motion

$$\begin{aligned}
\mathbf{C}_t^t(s) &= e^{-s\kappa\mathbf{N}^{\mathsf{T}}} e^{-s\kappa\mathbf{N}}, \\
\mathbf{G} &= \kappa\mathbf{N} + \dot{\mathbf{Q}}(t)\mathbf{Q}(t)^{\mathsf{T}}, \\
\mathbf{A}_1 &= \dot{\mathbf{C}}_t^t(0) = \kappa(\mathbf{N} + \mathbf{N}^{\mathsf{T}}), \\
\mathbf{A}_2 &= \ddot{\mathbf{C}}_t^t(0) = \kappa(\mathbf{N}^{\mathsf{T}}\mathbf{A}_1 + \mathbf{A}_1\mathbf{N}) = \kappa^2(2\mathbf{N}^{\mathsf{T}}\mathbf{N} + \mathbf{N}^2 + (\mathbf{N}^{\mathsf{T}})^2), \\
\mathbf{A}_3 &= \kappa(\mathbf{N}^{\mathsf{T}}\mathbf{A}_2 + \mathbf{A}_2\mathbf{N}), \dots, \\
\mathbf{A}_k &= \kappa(\mathbf{N}^{\mathsf{T}}\mathbf{A}_{k-1} + \mathbf{A}_{k-1}\mathbf{N}),
\end{aligned} \qquad \text{(IV.21-15)}$$

the notations being those of Section II.11. A monotonous motion is isochoric if and only if

$$\operatorname{tr}\mathbf{N}_0 = 0, \qquad \text{(IV.21-16)}$$

and of course then also $\operatorname{tr}\mathbf{N} = 0$.

With the aid of these consequences the following corollary makes plain the extremely special nature of monotonous motions.

Corollary (Wang). *The relative transplacement history $\mathbf{C}_t^t$ of a monotonous motion is determined uniquely by its first three Rivlin-Ericksen tensors.*

That is, if three tensors $\mathbf{A}_1(t)$, $\mathbf{A}_2(t)$, and $\mathbf{A}_3(t)$ are given, they can be the first three Rivlin–Ericksen tensors corresponding to at most one relative deformation history $\mathbf{C}_t^t$ satisfying the defining condition (1).

The proof rests upon a simple lemma. Let $\mathbf{S}$ be a symmetric tensor and $\mathbf{W}$ a skew tensor in 3-dimensional space. Without loss of generality we can take the matrices of these tensors as having the forms

$$[\mathbf{S}] = \begin{Vmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{Vmatrix}, \qquad [\mathbf{W}] = \begin{Vmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{Vmatrix}. \qquad \text{(IV.21-17)}$$

Then

$$[\mathbf{SW}-\mathbf{WS}] = \begin{Vmatrix} 0 & (a-b)x & (a-c)y \\ (a-b)x & 0 & (b-c)z \\ (a-c)y & (b-c)z & 0 \end{Vmatrix}. \quad \text{(IV.21-18)}$$

Hence $\mathbf{S}$ and $\mathbf{W}$ commute if and only if

$$(a-b)x = 0, \qquad (a-c)y = 0, \qquad (b-c)z = 0. \quad \text{(IV.21-19)}$$

Consequently, if $\mathbf{S}$ has 3 proper numbers, it commutes with no skew tensor other than $\mathbf{0}$. If $a = b \neq c$, $\mathbf{S}$ commutes with $\mathbf{W}$ if and only if $y = z = 0$. If $a = b = c$, $\mathbf{S}$ commutes with all $\mathbf{W}$.

WANG's corollary may now be proved in stages. If two monotonous motions can correspond with $\mathbf{A}_1$ and $\mathbf{A}_2$, then because of $(15)_{4,6}$ there are tensors $\mathbf{M}$ and $\bar{\mathbf{M}}$ such that

$$\begin{aligned} \mathbf{M} + \mathbf{M}^{\mathsf{T}} &= \bar{\mathbf{M}} + \bar{\mathbf{M}}^{\mathsf{T}}, \\ \mathbf{M}^{\mathsf{T}}\mathbf{A}_1 + \mathbf{A}_1\mathbf{M} &= \bar{\mathbf{M}}^{\mathsf{T}}\mathbf{A}_1 + \mathbf{A}_1\bar{\mathbf{M}}. \end{aligned} \quad \text{(IV.21-20)}$$

The first of these equations asserts that $\mathbf{M} - \bar{\mathbf{M}}$ is skew; the second, that $\mathbf{M} - \bar{\mathbf{M}}$ commutes with $\mathbf{A}_1$. If $\mathbf{A}_1$ has 3 proper numbers, the lemma shows that $\mathbf{M} - \bar{\mathbf{M}} = \mathbf{0}$.

Suppose now that $\mathbf{A}_1$ has 2 proper numbers. Then relative to a suitable orthonormal basis

$$[\mathbf{A}_1] = \begin{Vmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{Vmatrix}, \qquad a \neq b. \quad \text{(IV.21-21)}$$

Case 1. Relative to the same basis,

$$[\mathbf{A}_2] = \begin{Vmatrix} u & 0 & 0 \\ 0 & u & 0 \\ 0 & 0 & v \end{Vmatrix}. \quad \text{(IV.21-22)}$$

The most general $\mathbf{M}$ compatible with $(15)_4$ and (21) is given by

$$\kappa[\mathbf{M}] = \begin{Vmatrix} \frac{1}{2}a & x & y \\ -x & \frac{1}{2}a & z \\ -y & -z & \frac{1}{2}b \end{Vmatrix}. \tag{IV.21-23}$$

By (21) and (22)

$$\kappa[\mathbf{M}^{\mathsf{T}}\mathbf{A}_1 + \mathbf{A}_1\mathbf{M}] = \begin{Vmatrix} a^2 & 0 & (a-b)y \\ 0 & a^2 & (a-b)z \\ (a-b)y & (a-b)z & b^2 \end{Vmatrix}. \tag{IV.21-24}$$

Since $a \neq b$, it follows from $(15)_6$ and (22) that

$$u = a^2, \qquad v = b^2, \qquad y = 0, \qquad z = 0. \tag{IV.21-25}$$

Exercise IV.21.2. Use of (23) and (25) shows that $\mathbf{M}$ commutes with $\mathbf{M}^{\mathsf{T}}$; hence by $(15)_{1,3}$

$$\mathbf{C}_t^t(s) = e^{-s\mathbf{A}_1}. \tag{IV.21-26}$$

Case 2. Still on the supposition that $\mathbf{A}_1$ is of the form (21), but regardless of whether (22) does or does not hold, we assume that two monotonous motions can correspond with $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$. Then again, there are tensors $\mathbf{M}$ and $\bar{\mathbf{M}}$ such as to satisfy (20). Since $\mathbf{M} - \bar{\mathbf{M}}$ is a skew tensor that commutes with $\mathbf{A}_1$ as given by (21), the lemma shows that

$$[\mathbf{M} - \bar{\mathbf{M}}] = \begin{Vmatrix} 0 & x & 0 \\ -x & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{IV.21-27}$$

But also by $(15)_7$

$$\mathbf{M}^{\mathsf{T}}\mathbf{A}_2 + \mathbf{A}_2\mathbf{M} = \bar{\mathbf{M}}^{\mathsf{T}}\mathbf{A}_2 + \mathbf{A}_2\bar{\mathbf{M}}, \tag{IV.21-28}$$

and so $\mathbf{M} - \bar{\mathbf{M}}$ commutes with $\mathbf{A}_2$.

Exercise IV.21.3. If (22) does not hold, $\mathbf{M} = \bar{\mathbf{M}}$ in Case 2. Finally, if $\mathbf{A}_1 = \alpha\mathbf{1}$, (26) holds. $\triangle$

Accordingly, then, three given tensors $\mathbf{A}_1(t)$, $\mathbf{A}_2(t)$, and $\mathbf{A}_3(t)$ can be the Rivlin–Ericksen tensors corresponding to *at most one* $\mathbf{C}_t^t(s)$ belonging to a monotonous motion. In general, three symmetric tensors taken arbitrarily will fail to be the first three Rivlin–Ericksen tensors of any motion at all, let alone a monotonous one, since they will fail to satisfy conditions of compatibility[1] expressing the fact that they derive from a velocity field in a region. We shall not take up those conditions because our interest lies in simplifying a constitutive relation when the motion is known to be monotonous.

While NOLL's theorem is independent of dimension, WANG's corollary rests heavily on use of the dimension 3.

NOLL's theorem (3), when applied to a space of 3 dimensions, suggests an invariant, exhaustive classification of all monotonous motions:

Type 1. $\mathbf{N}_0^2 = \mathbf{0}$. These motions are called *viscometric flows*.[2]
Type 2. $\mathbf{N}_0^3 = \mathbf{0}$ but $\mathbf{N}_0^2 \neq \mathbf{0}$.
Type 3. $\mathbf{N}_0$ is not nilpotent.

There are interesting examples of all three types, but the simplest, the viscometric flows, are used most in applications.

Exercise IV.21.4. In types 1 and 2 the motion is isochoric, and also $\operatorname{tr}\mathbf{N}_0^2 = 0$.

Exercise IV.21.5. The relative transplacement gradient $\mathbf{F}_t$ of a viscometric flow has the form

$$\mathbf{F}_t(\tau) = \mathbf{Q}(\tau)\mathbf{Q}(t)^{\mathsf{T}}[\mathbf{1} + (\tau - t)\kappa\mathbf{Q}(t)\mathbf{N}_0\mathbf{Q}(t)^{\mathsf{T}}], \tag{IV.21-29}$$

$$\mathbf{N}_0 = \text{const.} \qquad |\mathbf{N}_0| = 1, \qquad \mathbf{N}_0^2 = \mathbf{0}, \qquad \kappa = \text{a scalar field.}$$

Conversely, any relative transplacement gradient of this form corresponds with a viscometric flow. An expression for $\mathbf{F}_t(\tau)$ which is quadratic in $\kappa(\tau - t)$ characterizes motions of type 2.

Exercise IV.21.6. In any monotonous motion

$$\mathbf{A}_2 - \mathbf{A}_1^2 = \kappa^2(\mathbf{N}^{\mathsf{T}}\mathbf{N} - \mathbf{N}\mathbf{N}^{\mathsf{T}}), \tag{IV.21-30}$$

[1] Conditions necessary and sufficient that given functions $\mathbf{A}_1(t)$, $\mathbf{A}_2(t)$, $\mathbf{A}_3(t)$ be the first three Rivlin–Ericksen tensors of a monotonous motion are obtained by C.-C. WANG in Section 3 of his memoir, "A representation theorem for the constitutive equation of a simple material in motions with constant stretch history," *Archive for Rational Mechanics and Analysis* **20**(1965): 329–340.

[2] A different but equivalent definition of viscometric flow was introduced and developed by A. C. PIPKIN, "Controllable viscometric flow," *Quarterly of Applied Mathematics* **16**(1968): 87–100.

and hence

$$\operatorname{tr}\mathbf{A}_1^2 = \operatorname{tr}\mathbf{A}_2 = 2\kappa^2(1+\operatorname{tr}\mathbf{N}^2). \tag{IV.21-31}$$

Thus in a viscometric flow and in a motion of type 2

$$\kappa^2 = \tfrac{1}{2}\operatorname{tr}\mathbf{A}_1^2 = \tfrac{1}{2}\operatorname{tr}\mathbf{A}_2\,. \tag{IV.21-32}$$

The flow to which a monotonous motion gives rise is a *monotonous flow*. Many of the preceding statements refer to the motions only through their flows; more statements of that kind follow now.

Exercise IV.21.7 (NOLL, COLEMAN & NOLL). The homogeneous flow whose cartesian components are

$$\dot{x}_1 = 0, \qquad \dot{x}_2 = \mu x_1\,, \qquad \dot{x}_3 = \lambda x_1 + \nu x_2\,, \tag{IV.21-33}$$

λ, μ, and ν being constants, is monotonous of type 1 if $\mu \neq 0$, $\nu = 0$; (*cf.* II.11-11); of type 2 if $\mu \neq 0$. The flow whose cartesian components are

$$\dot{x}_k = a_k x_k\,, \qquad a_k = \text{const.}, \qquad k = 1, 2, 3, \tag{IV.21-34}$$

is monotonous of type 3 if $a_1 a_2 a_3 \neq 0$ and is isochoric if and only if $a_1 + a_2 + a_3 = 0$. The flows (33) of type 2 are not universal. The isochoric instances of (34) are universal flows for homogeneous, incompressible, isotropic bodies. $= 1$ for (33), while (34) is irrotational.

Exercise IV.21.8 (TRUESDELL). A monotonous motion with spin $\mathbf{W}$ and vorticity number $\mathfrak{W}$ may be regarded as the superposition of a rigid motion whose spin $\mathbf{W}_r = \dot{\mathbf{Q}}(t)\mathbf{Q}(t)^T$ upon a motion, quantities associated with which are distinguished by subscript 0, such that

$$|\mathbf{W}_0|^2 = \tfrac{1}{2}\kappa^2(1-\operatorname{tr}\mathbf{N}^2), \qquad \mathfrak{W}_0^2 = \frac{1-\operatorname{tr}\mathbf{N}^2}{1+\operatorname{tr}\mathbf{N}^2}. \tag{IV.21-35}$$

Hence for types 1 and 2

$$|\mathbf{W}_0| = |\mathbf{D}_0| = \frac{1}{\sqrt{2}}|\kappa|, \qquad \mathfrak{W}_0 = 1, \tag{IV.21-36}$$

and so

$$|\mathbf{W}|^2\left(1-\frac{1}{\mathfrak{W}^2}\right) = |\mathbf{W}_r|(|\mathbf{W}_r| + \sqrt{2}|\kappa|\cos\theta), \tag{IV.21-37}$$

θ being defined in Exercise II.11.6. Hence for types 1 and 2

$$\begin{aligned} \mathfrak{W} > 1 \quad &\Leftrightarrow \quad |\mathbf{W}_r| > -\sqrt{2}|\kappa| \cos\theta, \\ \mathfrak{W} = 1 \quad &\Leftrightarrow \quad \begin{cases} \mathbf{W}_r = \mathbf{0} \quad \text{or} \\ |\mathbf{W}_r| = -\sqrt{2}|\kappa| \cos\theta, \end{cases} \\ \mathfrak{W} < 1 \quad &\Leftrightarrow \quad 0 < |\mathbf{W}_r| \leqq -\sqrt{2}|\kappa| \cos\theta. \end{aligned} \tag{IV.21-38}$$

In particular, $\mathfrak{W} \geqq 1$ if $0 \leqq \theta \leqq \frac{1}{2}\pi$, while in order that $\mathfrak{W} < 1$ it is necessary that $\frac{1}{2}\pi < \theta \leqq \pi$. Not only for flows of type 1 may $\mathfrak{W}$ take any value in $[0, \infty[$ (Exercise II.11.4) but also for those of type 3.

A theorem on nilpotent tensors tells us that for a viscometric flow there is an orthonormal basis with respect to which

$$[\mathbf{N}] = \begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{IV.21-39}$$

The basis that gives $[\mathbf{N}]$ this special form generally changes in time and varies from one place to another; it need not be the natural basis of any co-ordinate system. It is called the *viscometric basis* of the flow.

If we write $\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$ for the members of this basis, we see from $(15)_3$ that the axis of $\mathbf{W}_0$ in Exercise IV.21.8 is parallel to $\mathbf{i}_3$. Thus the angle θ that appears in (37) and (38) is the angle subtended upon $\mathbf{i}_3$ by the axis of $\mathbf{W}_r$.

The theory of nilpotent tensors provides also a basis with respect to which for flows of type 2

$$[\mathbf{N}] = \frac{1}{\sqrt{2}} \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{IV.21-40}$$

The axis of $\mathbf{W}_0$ is then parallel to $\mathbf{i}_1 + \mathbf{i}_3$, and θ is the angle subtended upon $\mathbf{i}_1 + \mathbf{i}_3$ by $\mathbf{W}_r$.

We consider next some special viscometric flows of interest particularly in application to viscometric experiments.

The *steady, lineal flows* are an example:

$$\dot{x}_1 = 0, \qquad \dot{x}_2 = v(x_1), \qquad \dot{x}_3 = 0. \tag{IV.21-41}$$

We have studied already a special instance of this class, simple shearing, defined by (II.11-17); for it the shearing v is linear.

Exercise IV.21.9. For (41)

$$\mathbf{F}_t^t(s) = \mathbf{1} - \kappa s \mathbf{N} = e^{-\kappa s \mathbf{N}}; \tag{IV.21-42}$$

cf. $(29)_1$. The flow is viscometric; the co-ordinate basis is a viscometric basis; $\mathbf{N}$ has the constant matrix (39); and the shearing is given by

$$\kappa = v'(x_1). \tag{IV.21-43}$$

Also $\mathfrak{W} = 1$. The material points move in straight lines at uniform speed; the principal stretchings are 0 and $\pm \frac{1}{2}\kappa$.

The flow whose contravariant components in a cylindrical polar co-ordinate system r, θ, z are given as follows in terms of arbitrary functions ω and u is called a *helical flow*:

$$\dot{r} = 0, \qquad \dot{\theta} = \omega(r), \qquad \dot{z} = u(r). \tag{IV.21-44}$$

Each material point remains upon a fixed cylinder r = const., on which it describes a helix, whose pitch is the same for all material points on any one cylinder. We have already encountered a special flow of this kind, the simple vortex, which is defined by (II.11.12). We set

$$f(r) := \omega'(r), \qquad h(r) := u'(r). \tag{IV.21-45}$$

Exercise IV.21.10 (Rivlin, Coleman & Noll). A helical flow is a viscometric flow, and

$$\kappa^2 = r^2 f(r)^2 + h(r)^2. \tag{IV.21-46}$$

Let $\{\mathbf{e}_k(\mathbf{x})\}$ be a natural basis for the co-ordinate system at $\mathbf{x}$, and let

$$\mathbf{i}_1 := \mathbf{e}_1, \qquad \mathbf{i}_2 := \alpha \mathbf{e}_2 + \beta \mathbf{e}_3, \qquad \mathbf{i}_3 := -\beta \mathbf{e}_2 + \alpha \mathbf{e}_3, \tag{IV.21-47}$$

the functions α and β being defined as follows:

$$\alpha := \frac{r}{\kappa} f(r), \qquad \beta := \frac{1}{\kappa} h(r), \qquad \alpha^2 + \beta^2 = 1. \tag{IV.21-48}$$

$\mathbf{N}_0$ has the form (39) with respect to the orthonormal basis $\{\mathbf{i}_k(\mathbf{x})\}$, which is not the natural basis of any co-ordinate system unless $\alpha = 0$ or $\beta = 0$. (*Cf.* the end of Section App.IIC.7.) Also

$$\mathfrak{W}^2 = 1 + \frac{4\omega(r\omega)'}{\kappa^2}. \qquad \text{(IV.21-49)}$$

(*Cf.* Exercise II.11.4, in which it is shown that even very special instances of (44) give $\mathfrak{W}$ any value in $[0, \infty[$.)

Exercise IV.21.11 (PIPKIN). The flows whose contravariant components in cylindrical co-ordinates are

$$\dot{r} = 0, \qquad \dot{\theta} = 0, \qquad \dot{z} = A\theta, \qquad A = \text{const.}, \qquad \text{(IV.21-50)}$$

and

$$\dot{r} = 0, \qquad \dot{\theta} = \kappa \log \frac{r}{R}, \qquad \dot{z} = 0, \qquad \kappa = \text{const.}, \qquad \text{(IV.21-51)}$$

are viscometric and are universal for homogeneous, incompressible, isotropic bodies. The first represents an accelerationless shearing of fanned planes; the second, a flow at uniform shearing between rotating cylinders.

Exercise IV.21.12. The flow (50) is the only monotonous flow included in the family (IV.18-2), while the family (IV.18-1) includes no monotonous flow.

In a major memoir on the kinematics and dynamics of viscometric flows YIN & PIPKIN[1] proved that those may be regarded as the effect of sliding inextensible material surfaces upon one another. While in the commonest examples these *slip surfaces* are rigid, typically they are flexible. Indeed, if they are rigid, they must be cylinders, surfaces of revolution, or helicoids, not necessarily co-axial. Some motions with rigid slip surfaces are intrinsically unsteady. YIN & PIPKIN give the following unsteady viscometric flow as an example: in cylindrical polar co-ordinates

$$r = \frac{R}{\sqrt{1 + \kappa^2 t^2}}, \qquad \theta = (1 + \kappa^2 t^2)\Theta + \kappa t \log R - \arctan(\kappa t), \qquad z = Z, \qquad \text{(IV.21-52)}$$

[1] W.-L. YIN & A. C. PIPKIN, "Kinematics of viscometric flow," *Archive for Rational Mechanics and Analysis* 37(1970): 111–133.

the constant κ being the shearing. As t increases, a material cylinder with the z-axis as central line shrinks inward toward that axis, and its length increases. The motion is kinematically admissible for only some finite interval of time.

YIN & PIPKIN prove also that the only viscometric flows such as to be universal for homogeneous, isotropic, incompressible fluids are the steady, simple shearing (II.11-17) and PIPKIN's flows (50) and (51), to within arbitrary rigid translations and certain rotations. In fact they are universal also for incompressible, isotropic solids. Whether any monotonous, isochoric flows beyond the homogeneous ones are universal for isotropic bodies seems not to be known.

While the term "viscometric flow" was intended to suggest a motion appropriate to an instrument for measuring a fluid's viscosity or more general properties of a similar kind, and the old viscometers did indeed presume one or another motion of the class here called "viscometric", recently some motions not in this class have been shown to lend themselves to such studies. One of these is described by the following steady, isochoric flow, introduced by BERKER in his researches on the Navier–Stokes theory:

$$\dot{x}_1 = -\Omega(x_2 - g(x_3)), \quad \dot{x}_2 = \Omega(x_1 - f(x_3)), \quad \dot{x}_3 = 0, \quad \Omega = \text{const.} \neq 0, \tag{IV.21-53}$$

in which f and g are differentiable functions, not both constant. The plane $x_3 =$ const. rotates with angular speed Ω about the point $x_1 = f(x_3)$, $x_2 = g(x_3)$. The locus of these points is a curve which crosses each plane $x_3 =$ const. just once.

Exercise IV.21.13 (RAJAGOPAL). For (53)

$$[\mathbf{G}] = \Omega \begin{Vmatrix} 0 & -1 & g' \\ 1 & 0 & -f' \\ 0 & 0 & 0 \end{Vmatrix},$$

$$\left.\begin{aligned} \mathbf{G}^{2n+1} &= (-1)^n \Omega^{2n} \mathbf{G}, \\ \mathbf{G}^{2n+2} &= (-1)^n \Omega^{2n} \mathbf{G}^2, \end{aligned}\right\} \quad n = 1, 2, \ldots, \tag{IV.21-54}$$

and

$$[\mathbf{F}_t(\tau)] = \begin{Vmatrix} C & S & -Sg' + (1-C)f' \\ -S & C & (1-C)g' + Sf' \\ 0 & 0 & 1 \end{Vmatrix}, \tag{IV.21-55}$$

$$C := \cos(\Omega(t-\tau)), \qquad S := \sin(\Omega(t-\tau)).$$

Exercise IV.21.14 (RAJAGOPAL). For (53)

$$\mathbf{F}_0(\tau) = e^{\tau\mathbf{G}}, \qquad \mathbf{G}^3 \neq 0, \tag{IV.21-56}$$

and so (53) is monotonous of NOLL's type 3. Also

$$\mathfrak{W}^2 = 1 + 4/(f'^2 + g'^2) > 1; \tag{IV.21-57}$$

and

$$\begin{aligned} \mathbf{A}_{2n+1} &= (-1)^n \Omega^{2n} \mathbf{A}_1 , \\ \mathbf{A}_{2n+2} &= (-1)^n \Omega^{2n} \mathbf{A}_2 . \end{aligned} \qquad n = 1, 2, \ldots, \tag{IV.21-58}$$

Thus Ω, $\mathbf{A}_1$, and $\mathbf{A}_2$ determine all the Rivlin–Ericksen tensors of (53).

22. Reduction of the Constitutive Relation for a Simple Material in a Monotonous Motion

In view of WANG's corollary, any information that can be determined from $\mathbf{C}_t^t$ in a monotonous motion can be determined also from $\mathbf{A}_1(t)$, $\mathbf{A}_2(t)$, $\mathbf{A}_3(t)$. Therefore, the values of a functional of $\mathbf{C}_t^t$ equal, in these motions, the values of a function of $\mathbf{A}_1(t)$, $\mathbf{A}_2(t)$, $\mathbf{A}_3(t)$. Consequently the general constitutive relation (IV.5-15) may be replaced, as far as monotonous motions are concerned, by

$$\mathbf{R}^\mathsf{T}\mathbf{T}\mathbf{R} = \mathfrak{f}(\mathbf{R}^\mathsf{T}\mathbf{A}_1(t)\mathbf{R}, \mathbf{R}^\mathsf{T}\mathbf{A}_2(t)\mathbf{R}, \mathbf{R}^\mathsf{T}\mathbf{A}_3(t)\mathbf{R}, \mathbf{C}(t)), \tag{IV.22-1}$$

$\mathfrak{f}$ being a function. A material whose constitutive relation is (1) is called *a material of differential type of complexity* 3. By (1), then, we have the following

Theorem. *In undergoing a monotonous motion, a simple body is subject to the same stress as is a body of some material of complexity* 3 *undergoing the same motion.*

Consequently, *no measurement of stress in a monotonous motion can distinguish a general simple material from a material of differential type of complexity* 3. As we shall see in the next chapter, the special flows most commonly used to describe the properties of natural fluids are of the kind considered here and hence are of limited service in exploring the physical properties of those fluids.

An isotropic material of differential type is called a *Rivlin–Ericksen material*. For it, (1) becomes

$$\mathbf{T}(t) = \mathfrak{f}(\mathbf{A}_1(t), \mathbf{A}_2(t), \mathbf{A}_3(t), \mathbf{B}(t)), \qquad \text{(IV.22-2)}$$

and when the isotropic material is fluid,

$$\mathbf{T}(t) = -p(\rho)\mathbf{1} + \mathfrak{f}(\mathbf{A}_1(t), \mathbf{A}_2(t), \mathbf{A}_3(t), \rho); \qquad \text{(IV.22-3)}$$

the functions $\mathfrak{f}$, in the two cases, are isotropic in the sense that for all symmetric $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, $\mathbf{B}$, and for all orthogonal $\mathbf{Q}$

$$\mathfrak{f}(\mathbf{QA}_1\mathbf{Q}^\mathsf{T}, \mathbf{QA}_2\mathbf{Q}^\mathsf{T}, \mathbf{QA}_3\mathbf{Q}^\mathsf{T}, \mathbf{QBQ}^\mathsf{T} \text{ or } \rho) = \mathbf{Q}\mathfrak{f}(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{B} \text{ or } \rho)\mathbf{Q}^\mathsf{T}, \qquad \text{(IV.22-4)}$$

this being the functional equation to which (IV.14-3) reduces in the present instance. Moreover, for a fluid $\mathfrak{f}(\mathbf{0}, \mathbf{0}, \mathbf{0}, \rho) = \mathbf{0}$.

The statements in Section IV.7 enable the student to write down at once the constitutive relations for incompressible Rivlin–Ericksen materials.

The reductions just given may be interpreted in two ways. On the one hand, they enable us to solve easily various special problems concerned with monotonous motions. However complicated may be in general the response of a material, in these particular motions we need consider only a simple, special, constitutive equation. On the other hand, they show that observation of this class of flows is insufficient to tell us much about a material, since most of the complexities of material response are prevented from manifesting themselves.

In Section VI.1 we shall discuss materials of the differential type in somewhat more detail, but in the next chapter we shall exploit the present theorems so as to obtain specific solutions for viscometric flows of simple fluids.

If in the constitutive relation (1) the numbers 1, 2, 3 are replaced by 1, $2, \ldots, n$, the material so defined is called *a material of differential type of complexity n*.

In a viscometric flow, by definition, $\mathbf{N}_0^2 = \mathbf{0}$, and hence by (IV.21-14) and (IV.21-15)

$$\mathbf{A}_3 = \mathbf{A}_4 = \cdots = \mathbf{0}. \qquad \text{(IV.22-5)}$$

Therefore, *in a viscometric flow a simple fluid cannot be distinguished from some Rivlin–Ericksen fluid of complexity* 2. The condition (5) is merely sufficient, not necessary for this deduction. If $\mathbf{A}_3$ is determined by $\mathbf{A}_1$ and $\mathbf{A}_2$ for the class of flows considered, the same conclusion follows.

Exercise IV.22.1 (RAJAGOPAL). In the class of motions (IV.21-53) a simple fluid cannot be distinguished from a Rivlin–Ericksen fluid of complexity 2.

For a monotonous motion we see from $(\mathrm{IV}.21\text{-}15)_1$ that

$$\mathbf{R}^{\mathsf{T}}\mathbf{C}_t^t(s)\mathbf{R} = \exp[-s\kappa(\mathbf{R}^{\mathsf{T}}\mathbf{N}\mathbf{R})^{\mathsf{T}}]\,\exp[-s\kappa\mathbf{R}^{\mathsf{T}}\mathbf{N}\mathbf{R}]. \qquad (\mathrm{IV}.22\text{-}6)$$

Hence any quantity determined by $\mathbf{R}^{\mathsf{T}}\mathbf{C}_t^t\mathbf{R}$ in general is determined here by $\kappa\mathbf{R}^{\mathsf{T}}\mathbf{N}\mathbf{R}$. Referring to the frame-indifferent constitutive relation (IV.5-15) of a simple material, we may set

$$\mathfrak{f}(\kappa, \mathbf{R}^{\mathsf{T}}\mathbf{N}\mathbf{R}, \mathbf{C}) := \mathfrak{R}(\mathbf{R}^{\mathsf{T}}\mathbf{C}_t^t\mathbf{R}, \mathbf{C}) \qquad (\mathrm{IV}.22\text{-}7)$$

and so obtain

$$\mathbf{R}^{\mathsf{T}}\mathbf{T}\mathbf{R} = \mathfrak{f}(\kappa, \mathbf{R}^{\mathsf{T}}\mathbf{N}\mathbf{R}, \mathbf{C}) \qquad (\mathrm{IV}.22\text{-}8)$$

as an expression for it when restricted to monotonous motions. The student will see at once the simpler forms to which (8) reduces for isotropic solids and fluids. For an incompressible fluid the reduction of the determinate stress (IV.17-8) is

$$\mathbf{S} = \mathfrak{f}(\kappa, \mathbf{N}), \qquad (\mathrm{IV}.22\text{-}9)$$

the function $\mathfrak{f}$ being subject to the requirement that

$$\mathfrak{f}(\kappa, \mathbf{Q}\mathbf{N}\mathbf{Q}^{\mathsf{T}}) = \mathbf{Q}\mathfrak{f}(\kappa, \mathbf{N})\mathbf{Q}^{\mathsf{T}} \qquad (\mathrm{IV}.22\text{-}10)$$

for every orthogonal tensor $\mathbf{Q}$ and for all $\mathbf{N}$ such that $|\mathbf{N}| = 1$ and $\mathbf{N}^2 = \mathbf{0}$. From (6) we see that $\mathbf{R}^{\mathsf{T}}\mathbf{C}_t^t\mathbf{R}$ is unchanged when κ and $\mathbf{N}$ are replaced by $-\kappa$ and $-\mathbf{N}$. Hence

$$\mathfrak{f}(-\kappa, -\mathbf{N}) = \mathfrak{f}(\kappa, \mathbf{N}). \qquad (\mathrm{IV}.22\text{-}11)$$

The relations (9), (10), and (11) provide the starting point for the analysis in the following chapter.

General References

Sections 26–30 of NFTM.

W. NOLL, "A mathematical theory of the mechanical behavior of continuous media," *Archive for Rational Mechanics and Analysis* **2**(1958): 197–226. Reprinted in *Continuum Mechanics II, The Rational Mechanics of Materials*, ed. C. TRUESDELL, New York, Gordon and Breach, 1965, and in *Continuum Theory of Inhomogeneities in Simple Bodies*, New York, Springer-Verlag, 1968, and in *The Foundations of Mechanics and Thermodynamics*, New York, Heidelberg, and Berlin, Springer-Verlag, 1974.

B. D. COLEMAN & W. NOLL, "Material symmetry and thermostatic inequalities in finite elastic deformations," *Archive for Rational Mechanics and Analysis* **15**(1964): 87–111, reprinted in NOLL's *The Foundations of Mechanics and Thermodynamics*.

W. NOLL, "A new mathematical theory of simple materials," *Archive for Rational Mechanics and Analysis* **48**(1972): 1–50, reprinted with the preceding.

W. NOLL, "Lectures on the foundations of continuum mechanics and thermodynamics," *Archive for Rational Mechanics and Analysis* **52**(1973): 62–92, reprinted with the preceding.

This Page Intentionally Left Blank

Appendix I

General Scheme of Notation

Departures from the general scheme occur here and there, usually only within single sections.

Script majuscules: $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}, \ldots, \mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$, denote bodies, sets, regions of space, curves, and surfaces. $\mathscr{A}^e$ is the exterior of $\mathscr{A}$. $\mathscr{R}$ is the real line.

If $\mathscr{S}$ is a set in a topological space, then $\operatorname{int}\mathscr{S}$ and $\operatorname{clo}\mathscr{S}$ denote the interior and closure, respectively, of $\mathscr{S}$.

Lightfaced italics, both majuscule and minuscule, stand for scalars and scalar-valued functions: $A, B, C, \ldots, X, Y, Z, a, b, c, \ldots, x, y, z$. Included are the components of vectors and tensors with respect to particular bases.

Exception: X usually stands for a substantial point, $A(\,)$ means "area of" and $V(\,)$ means "volume of".

Special letters: t always denotes the time, and n usually denotes the dimension of a vector space.

Note also the uses of o and O explained below in Section C1 of Appendix II.

Boldfaced roman minuscules stand for vectors and vector-valued functions: $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{u}, \mathbf{v}$, except that $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are places and $\mathbf{n}$ always denotes an oriented unit normal to a surface.

Boldfaced greek minuscules $\boldsymbol{\vartheta}$, $\boldsymbol{\lambda}$, $\boldsymbol{\tau}$, *etc.*, denote mappings other than functions of vectors.

Special greek letters:

χ always denotes the motion of a substantial point or of a body (Section I.7).

κ always denotes a reference placement of a body (Section II.3).

χ_κ always denotes the transplacement of a substantial point or body from the reference placement κ to the actual placement (Section II.3).

Boldfaced majuscules $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{U}, \mathbf{V}, \mathbf{W}$ denote linear transformations (second-order tensors) over finite-dimensional (usually three-dimensional) vector spaces.

Exception: $\mathbf{X}$ is always the place of the substantial point X in a reference placement.

Special letters:

$\mathbf{Q}$ and $\mathbf{R}$ are always orthogonal.

$\mathbf{W}$ is always skew.

$\mathbf{F}$ is always an invertible tensor which can be interpreted as a transplacement gradient.

If $\mathbf{A}$ is a tensor, $\operatorname{sym}\mathbf{A}$ and $\operatorname{skw}\mathbf{A}$ denote its symmetric and skew parts, while $\operatorname{adj}\mathbf{A}$, $\mathbf{A}^{\mathsf{T}}$, $\operatorname{tr}\mathbf{A}$ and $\det\mathbf{A}$ denote its adjugate, transpose, trace, and determinant. The matrices of a tensor and the determinant of a tensor and of a matrix are defined in Sections II A.3 and 4.

Lightfaced greek minuscules are used for three different kinds of quantities:

1. For angles, rates of change of angles, and other pure rates.
2. For scalar potentials.
3. For scalar moduli or scalar-valued material functions of a real variable.

Exceptions:

ρ is always the mass-density (Section II.5), v is always the specific volume $1/\rho$.

δ and ϵ are usually scalars which can be chosen arbitrarily small; they are used also for some particular functions, *e.g.* alternators.

Fraktur letters, both majuscule and minuscule, denote constitutive mappings (responses). Lightfaced $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \ldots, \mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$ are used if the values of the mappings are scalars; boldfaced $\boldsymbol{\mathfrak{a}}, \boldsymbol{\mathfrak{b}}, \boldsymbol{\mathfrak{c}}, \ldots, \boldsymbol{\mathfrak{A}}, \boldsymbol{\mathfrak{B}}, \boldsymbol{\mathfrak{C}}, \ldots$, if the values are vectors or tensors, respectively.

Black letter ("old English") majuscules denote scaling parameters, in this volume only $\mathfrak{W}$.

Script minuscules denote groups of tensors.

Special letters:

$\mathscr{o}$ is the full orthogonal group.

$\mathscr{u}$ is the full unimodular group.

$\mathscr{g}$ is a subgroup of $\mathscr{u}$.

Lightfaced greek majuscules Θ, Φ, . . . , are used for certain angles in a reference placement.

Boldfaced sans-serif majuscules A, B, C, L denote third-order or fourth-order tensors or affine mappings of tensors over a three-dimensional vector space.

Astronomical symbols ♆ , ♃ , ♂ , *etc.*, stand for quantities of arbitrary tensorial order; scalars, vectors, *etc.*

Special symbols: $\oint$ denotes a frame (Section I.6), and $\varnothing$ denotes the null set.

Indices:

The uses of subscripts and superscripts are standard. A few examples will suffice, but the list is far from exhaustive.

If $\mathbf{a}$, $\mathbf{A}$, and A are vectors and tensors denoted as above, then their components with respect to a basis are denoted by adjoined indices, for example a_k, A_{mp}, A_{qrsu}. The particular basis is always specified. In the case of curvilinear co-ordinates, the usual notations of contravariant and covariant components such as a^k and a_k are employed once in a while. Physical components, which are components with respect to an orthonormal basis other than a cartesian basis, are denoted by indices following the letter at middle height: $T^{\langle rr \rangle}$, $T^{\langle \theta\theta \rangle}$, *etc.*

Superscript T (sans-serif) always indicates transposition.

Greek minuscule indices refer to co-ordinate systems in the reference shape $\boldsymbol{\kappa}(\mathscr{B})$. For example, F^k_α is the component of $\mathbf{F}$ that corresponds with x^k and X^α. Both systems of co-ordinates may be curvilinear if so desired.

A boldfaced subscript kappa, as on $\chi_{\boldsymbol{\kappa}}$ and $\mathbf{A}_{\boldsymbol{\kappa}}$, reminds the reader that the reference placement $\boldsymbol{\kappa}$ is being used.

Roman letters associated with mathematical symbols are labels. Examples: min and max in the obvious senses, superscript c for "convected", e superscript for "exterior", subscript B for "body", C for "contact", *etc.*

General relations:

$A \Rightarrow B$	Proposition A implies Proposition B.
$A \Leftrightarrow B$	A holds if and only if B holds.
$\mathscr{A} \vee \mathscr{B}, \mathscr{A} \wedge \mathscr{B}$	join and meet, respectively, of the bodies $\mathscr{A}$ and $\mathscr{B}$
$\mathscr{A} \prec \mathscr{B}, \mathscr{B} \succ \mathscr{A}$	$\mathscr{A}$ is a part of $\mathscr{B}$.
$\mathscr{A} \cup \mathscr{B}, \mathscr{A} \cap \mathscr{B}$	union and intersection, respectively, of the sets $\mathscr{A}$ and $\mathscr{B}$
$\mathscr{A} \backslash \mathscr{B}$	$\mathscr{A} \cap \mathscr{B}^{e}$
$\mathscr{A} \subset \mathscr{B}, \mathscr{B} \supset \mathscr{A}$	$\mathscr{A}$ is a subset of $\mathscr{B}$.
$x \in \mathscr{A}$	x is an element of the set $\mathscr{A}$.
$f: \mathscr{A} \rightarrow \mathscr{B}$	f maps the set $\mathscr{A}$ into or onto the set $\mathscr{B}$.
$f: x \mapsto y$	f maps the element x onto the element y; that is, $f(x) = y$.
$f \circ g$	composition of the mappings g and f; that is, $(f \circ g)(x) = f(g(x))$.
$\forall x \in \mathscr{A}$	"for every x that is an element of the set $\mathscr{A}$"
$\{x : x \in \mathscr{A}\}$	the set of all x that are elements of $\mathscr{A}$
$\{x_1, x_2, \ldots, x_n\}$	the set consisting in the elements $x_1, x_2, \ldots, x_n$

After ∂, a subscript indicates the variable on which ∂ operates; for example, ∂_θ is the partial derivative with respect to θ.

A superimposed dot always indicates a time derivative in some sense. For example, $\dot{\chi}$ is the velocity field over a body or a placement of a body and $\dot{\mathbf{x}}$ is the corresponding velocity field over the present shape of that body. *Cf.* Section II.4.

Functions:

$f = g$	The numbers f and g are the same; the functions f and g are the same. *E.g.*, for functions $f = 0$ means that $f(x) = 0$ for all x in the domain of f; in other words, f is the zero function.
$f := g$	The function or number f is by definition the same as the function or number g.
$f =: g$	The function or number g is by definition the same as the function or number f.
$f(x) = g(X)$	when x is an assigned, invertible function of X. Then the value of f at x equals the value of g at X (*cf. e.g.* Sections II.2, II.4, and II.6).
$f(x) = 0$	The value of f at x is 0.

Operations on vectors and tensors: see Appendix II.

Appendix II

Some Definitions and Theorems of Algebra, Geometry, and Calculus[1]

A. Algebra

1. Vector Spaces, Bases

With a few specified exceptions, the *vector spaces* $\mathscr{V}$, $\mathscr{V}'$ *etc.*, that are recognized explicitly in this book are of *finite dimension* n, usually three, and their field of scalars is the real field. Their elements are denoted by bold-faced minuscule letters $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{u}, \mathbf{v}, \mathbf{w}$, and their scalars by light-faced italics $a, b, \ldots, A, B, \ldots$ The null vector is $\mathbf{0}$.

The set of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ is a *linearly dependent* set if scalars $a^1, a^2, \ldots, a^m$, not all null, can be found such that

$$a^1\mathbf{u}_1 + a^2\mathbf{u}_2 + \cdots + a^m\mathbf{u}_m = \mathbf{0}.$$

Otherwise the set is *linearly independent*. The expression on the left-hand side is called a *linear combination* of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$. The set of values of all linear combinations of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ is a *subspace*, which the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ are said to *span*; the subspace itself is called the *span* of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$.

[1] The material listed here is drawn largely from unpublished notes leading to W. NOLL's *Finite-Dimensional Spaces*, Volume 1, *Algebra, Geometry, and Analysis*, Dordrecht *etc.*, Martinus Nijhoff, 1987. Excellent treatments are included also in two other books: R. M. BOWEN & C.-C. WANG, *Introduction to Vectors and Tensors*, 2 vols., New York & London, Plenum, 1976, and M. E. GURTIN, *An Introduction to Continuum Mechanics*, New York *etc.*, Academic Press, 1981.

The dimension n of a space is the number of vectors in the smallest set that spans it. The dimension of a subspace of an n-dimensional vector space is at most n. (Of course vector spaces of infinite dimension whose elements are functions occur implicitly in this book, but they are only rarely considered as such in the presentation.)

Any indexed set $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ of n linearly independent vectors spans an n-dimensional vector space and hence is called a *basis* in it. If $\mathbf{u}$ is any vector, then

$$\mathbf{u} = u^k \mathbf{e}_k ,$$

in which $u^1, u^2, \ldots, u^n$ are uniquely determined scalars. In this expression, and subsequently, diagonally repeated indices are to be summed from 1 to n. The n scalars u^k are the *components* of $\mathbf{u}$ relative to the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Here and henceforth in this appendix a free index such as k is understood to run through the numbers $1, 2, \ldots, n$ unless a different range is specified.

If $\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n$ is another basis, of course the vector $\mathbf{u}$ has components relative to it:

$$\mathbf{u} = \bar{u}^k \bar{\mathbf{e}}_k .$$

Also $\bar{\mathbf{e}}_p$ has components, say A^q_p, relative to $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$:

$$\bar{\mathbf{e}}_p = A^q_p \mathbf{e}_q , \qquad p = 1, 2, \ldots, n.$$

Likewise

$$\mathbf{e}_m = \bar{A}^q_m \bar{\mathbf{e}}_q , \qquad m = 1, 2, \ldots, n,$$

and hence, since the vectors of both bases are linearly independent,

$$A^q_k \bar{A}^m_q := \begin{cases} 1 & \text{if } m = k, \\ 0 & \text{if } m \neq k. \end{cases}$$

Therefore the components of any vector $\mathbf{u}$ relative to the two bases are determined from each other as follows:

$$\bar{u}^q = \bar{A}^q_m u^m , \qquad u^p = A^p_q \bar{u}^q .$$

Persons who prefer numerical to geometrical treatments may use this *transformation law* for components to define vectors. They may choose to specify a vector by prescribing its components relative to some one basis and then use the transformation law to calculate its components relative to every other basis. Alternatively, they may start with lists of numbers $u^1, u^2, \ldots, u^n$ and $\bar{u}^1, \bar{u}^2, \ldots, \bar{u}^n, \ldots$, associated with various bases and then say that these lists do or do not constitute components of one and the same vector relative to the respective bases according as they are or are not related by the transformation law.

Every vector space of dimension n is isomorphic to the "cartesian space" $\mathscr{R}_n$ (defined below in Section A5 of this appendix), but usually a conceptual argument is clearer if it does not employ co-ordinates.

2. *Linear Mappings*

The concept of "mapping" and the terms associated with it are presumed familiar. In this book the words "into" and "onto" retain their senses in brief, idiomatic English. For example, a vector may be mapped onto a vector but cannot be mapped into a vector, for a vector has no inside. The statement "**L** maps vectors onto vectors" means "**L** maps a set of vectors (specified by the context) into a vector space" (also specified by the context); equivalently, "the domain and codomain of **L** are subsets of vector spaces."

A mapping **L** of a vector space $\mathscr{V}$ into a vector space $\mathscr{V}'$ is *linear* if

$$\mathbf{L}(a\mathbf{u} + b\mathbf{v}) = a\mathbf{L}(\mathbf{u}) + b\mathbf{L}(\mathbf{v})$$

for all **u** and **v** in $\mathscr{V}$ and all scalars a and b. The *scalar multiple* $a\mathbf{L}$ of **L** by a and the *sum* $\mathbf{L} + \mathbf{M}$ of such mappings **L** and **M** are defined as follows:

$$(a\mathbf{L})(\mathbf{u}) := a(\mathbf{L}(\mathbf{u})),$$

$$(\mathbf{L} + \mathbf{M})(\mathbf{u} + \mathbf{v}) := \mathbf{L}(\mathbf{u} + \mathbf{v}) + \mathbf{M}(\mathbf{u} + \mathbf{v}).$$

It is easy to show that $a\mathbf{L}$ and $\mathbf{L} + \mathbf{M}$ are themselves linear mappings of $\mathscr{V}$ into $\mathscr{V}'$.

The *nullspace* of a linear mapping **L** is the set of vectors that **L** maps onto **0**. The *range* of **L** is the set of all values $\mathbf{L}(\mathbf{u})$. These sets are subspaces of $\mathscr{V}$ and $\mathscr{V}'$, respectively, and

$$\dim \text{Nullspace } \mathbf{L} + \dim \text{Range } \mathbf{L} = \dim \mathscr{V}.$$

If $\dim \mathscr{V} = \dim \mathscr{V}'$, the linear mapping **L** may have an *inverse* $\mathbf{L}^{-1}$; such an **L** is *invertible*. If $\dim \mathscr{V} = \dim \mathscr{V}'$, any of the following statements is a necessary and sufficient condition that $\mathbf{L}$ have an inverse: $\mathbf{L}$ is one-to-one, $\mathbf{L}$ maps $\mathscr{V}$ onto $\mathscr{V}'$, the nullspace of $\mathbf{L}$ consists in $\mathbf{0}$ alone. The inverse, if it exists, is itself a linear invertible mapping. Of course $(\mathbf{L}^{-1})^{-1} = \mathbf{L}$.

A mapping $\mathbf{A}$ is *affine* if it is the sum of a linear mapping and a constant mapping:

$$\mathbf{A}(\mathbf{v}) := \mathbf{L}(\mathbf{v}) + \mathbf{a},$$

$\mathbf{L}$ being a linear mapping of $\mathscr{V}$ into $\mathscr{V}'$ and $\mathbf{a}$ being a particular element of $\mathscr{V}'$.

3. *Tensors*

A linear mapping $\mathbf{L}$ of a vector space into itself is called a *tensor* (of second order). The value $\mathbf{L}(\mathbf{u})$ of the tensor $\mathbf{L}$ at $\mathbf{u}$ is written like a multiplication: $\mathbf{L}\mathbf{u} := \mathbf{L}(\mathbf{u})$.

The mapping whose value for every vector is $\mathbf{0}$ is a tensor; it is called the *zero tensor* and is denoted by $\mathbf{0}$. The identity mapping is a tensor; it is called the *unit tensor* or *identity tensor* and is denoted by $\mathbf{1}$. Thus for all vectors $\mathbf{u}$

$$\mathbf{0}\mathbf{u} = \mathbf{0}, \qquad \mathbf{1}\mathbf{u} = \mathbf{u}.$$

The tensor that transforms every vector into its opposite is called the *central inversion* and is denoted by $-\mathbf{1}$:

$$(-\mathbf{1})\mathbf{v} = -\mathbf{v}.$$

If $\mathbf{L}$ and $\mathbf{M}$ are tensors, so is their composition, which we denote by $\mathbf{ML}$ and call the *product* of $\mathbf{L}$ by $\mathbf{M}$. The set of all tensors forms an *algebra* under the operations denoted by $a\mathbf{L}$, $\mathbf{L} + \mathbf{M}$, and $\mathbf{LM}$. Clearly $\mathbf{1L} = \mathbf{L1} = \mathbf{L}$ for every $\mathbf{L}$. As usual for algebras, $-\mathbf{L}$ is written for $(-1)\mathbf{L}$; of course $(-\mathbf{1})\mathbf{L} = \mathbf{L}(-\mathbf{1}) = (-1)\mathbf{L} = -\mathbf{L}$ for every $\mathbf{L}$, and likewise $\mathbf{L0} = \mathbf{0L} = 0\mathbf{L} = \mathbf{0}$, but if $\mathbf{ML} = \mathbf{0}$, neither $\mathbf{M}$ nor $\mathbf{N}$ need be $\mathbf{0}$.

Generally $\mathbf{LM} \neq \mathbf{ML}$. If $\mathbf{LM} = \mathbf{ML}$, the tensors $\mathbf{L}$ and $\mathbf{M}$ *commute*. For clarity we sometimes state that one tensor commutes with another. We have seen that $\mathbf{1}$, $-\mathbf{1}$, and $\mathbf{0}$ commute with every $\mathbf{L}$.

The *powers* of a tensor $\mathbf{L}$ are defined as follows:

$$\mathbf{L}^0 := \mathbf{1}, \qquad \mathbf{L}^1 := \mathbf{L}, \qquad \mathbf{L}^2 := \mathbf{LL}, \qquad etc.;$$

these obey the usual *rules of exponentiation*:

$$\mathbf{L}^m\mathbf{L}^q = \mathbf{L}^{m+q} = \mathbf{L}^q\mathbf{L}^m, \qquad (a\mathbf{L})^m = a^m\mathbf{L}^m,$$

$$(\mathbf{L}^m)^q = \mathbf{L}^{mq},$$

if $m \geqq 0$ and $q \geqq 0$.

If $\mathbf{L}^m = \mathbf{0}$ for some positive integer m but $\mathbf{L}^p \neq \mathbf{0}$ if $0 < p < m$, the tensor $\mathbf{L}$ is *nilpotent* of *order m*. Nilpotent tensors of orders $1, 2, 3, \ldots, n$ exist, but not of any greater order. That is, if $\mathbf{L}^n \neq \mathbf{0}$, then $\mathbf{L}$ is not nilpotent.

If $\mathbf{L}$ is invertible,

$$\mathbf{L}\mathbf{L}^{-1} = \mathbf{L}^{-1}\mathbf{L} = \mathbf{1}.$$

Also if there is a tensor $\mathbf{M}$ such that $\mathbf{LM} = \mathbf{ML} = \mathbf{1}$, then $\mathbf{L}$ is invertible, and $\mathbf{M} = \mathbf{L}^{-1}$. Clearly, $\mathbf{1}^{-1} = \mathbf{1}$, $(-\mathbf{1})^{-1} = -\mathbf{1}$. If $\mathbf{L}$ and $\mathbf{M}$ are invertible, and so is $\mathbf{LM}$, and

$$(\mathbf{LM})^{-1} = \mathbf{M}^{-1}\mathbf{L}^{-1}.$$

Thus the invertible tensors form a group under multiplication and a subalgebra of the algebra of tensors. The tensor $\mathbf{0}$ is not invertible, nor is any nilpotent tensor. A tensor that is not invertible is sometimes called *singular*.

The product of two invertible tensors is invertible, and so is the multiple of an invertible tensor by any scalar other than 0, and also $(a\mathbf{L})^{-1} = a^{-1}\mathbf{L}^{-1}$. If $\mathbf{L}$ is invertible, then

$$(\mathbf{L}^{-1})^n = (\mathbf{L}^n)^{-1},$$

and the rules of exponentiation extend to negative powers.

If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is a basis of the vector space, the conditions

$$\mathbf{L}\mathbf{e}_k = L^q{}_k\mathbf{e}_q$$

define unique scalars $L^q{}_k$, which are called the *components* of $\mathbf{L}$ relative to the basis. The matrix $\|L^q{}_k\|$ of the components $L^q{}_k$ is called the *matrix* of $\mathbf{L}$ relative to $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ and is denoted by $[\mathbf{L}]$. That is,

$$[\mathbf{L}] := \|L^q{}_k\| := \begin{Vmatrix} L^1{}_1 & L^1{}_2 & \cdots & L^1{}_n \\ L^2{}_1 & & & \\ \vdots & & & \\ L^n{}_1 & L^n{}_2 & \cdots & L^n{}_n \end{Vmatrix}.$$

Of course $[a\mathbf{L}] = a[\mathbf{L}]$. Also $[\mathbf{LM}] = [\mathbf{L}][\mathbf{M}]$; the components of $\mathbf{LM}$ are $L^p{}_k M^k{}_q$. No matter what the choice of basis,

$$[\mathbf{0}] = \begin{Vmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{Vmatrix}, \qquad [\mathbf{1}] = \begin{Vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{Vmatrix},$$

$$[-\mathbf{1}] = \begin{Vmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & -1 \end{Vmatrix}.$$

If the bases $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ and $\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n$ are related as in Section IIA.1, the components $L^m{}_k$ and $\bar{L}^p{}_q$ of $\mathbf{L}$ relative to them are determined uniquely in terms of each other by the following transformation law:

$$\bar{L}^p{}_q = \bar{A}^p_r A^s_q L^r{}_s, \qquad L^r{}_s = A^r_k \bar{A}^m_s \bar{L}^k{}_m,$$

in which the scalars A^q_p and $\bar{A}^k_m$ are the coefficients defining the change of basis, introduced in Section App. IIA.1.

The set of all tensors over a vector space of dimension n is itself a vector space of dimension n^2 under the operations of addition and scalar multiplication already introduced. The vector $\mathbf{0}$ in the space of n^2 dimensions is simply the tensor $\mathbf{0}$.

Over the vector space of dimension n^2 so obtained, we may consider linear transformations in just the same way as before. If $\mathbf{M}$ is such a tensor, its *components* $M^h{}_f{}^k{}_g$ relative to the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ can be determined from definitions already given. Under change of basis those components transform as follows:

$$\bar{M}^p{}_q{}^r{}_s = \bar{A}^p_h A^f_q \bar{A}^r_k A^g_s M^h{}_f{}^k{}_g\,.$$

Rules of this kind may be used, alternatively, to define tensors in terms of their components.

Tensors of order higher than two are used in this book only a few special contexts. The considerations concerning them may be understood either abstractly or in terms of components, as the reader prefers.

For the result of operating with a fourth-order tensor **K** upon a second-order tensor **L** we use the special notation **K**[**L**]. The components of **K**[**L**] are $K^k{}_m{}^p{}_q L_p{}^q$. Sometimes the same notation is used to indicate the linear part of an affine mapping of tensors onto tensors.

4. Determinant and Adjugate of a Tensor

The determinant of the matrix whose components are $L^p{}_q$ may be defined as follows:

$$\begin{aligned}\det \|L^p{}_q\| &:= \epsilon^{k_1 k_2 \ldots k_n} L^1{}_{k_1} L^2{}_{k_2} \cdots L^n{}_{k_n}\\ &= \epsilon_{k_1 k_2 \ldots k_n} L^{k_1}{}_1 L^{k_2}{}_2 \cdots L^{k_n}{}_n,\end{aligned}$$

The symbol $\epsilon^{k_1 k_2 \ldots k_q}$ denotes 1 if $k_1, k_2, \ldots, k_q$ are obtained from $1, 2, \ldots, q$ by an even permutation; -1, if by an odd permutation; and 0 otherwise. If **L** is a tensor, the determinants of its matrices of components $\|L^p{}_q\|$ all have a common value, irrespective of the choice of basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ used to define those components. This common value $\det \mathbf{L}$ is called the *determinant* of the tensor:

$$\det \mathbf{L} := \det \|L^p{}_q\|.$$

It follows that

$$\det(\mathbf{LM}) = (\det \mathbf{L})(\det \mathbf{M}) = \det(\mathbf{ML}), \qquad \det(a\mathbf{L}) = a^n \det \mathbf{L},$$

and of course $\det \mathbf{1} = 1$, $\det(-\mathbf{1}) = (-1)^n$.

The adjugate $\operatorname{adj} \mathbf{L}$ of a tensor **L** is that tensor whose components are the cofactors of the elements of [**L**]:

$$(\operatorname{adj} \mathbf{L})^p_q = \epsilon_{k_1 k_2 \ldots k_n} L^{k_1}{}_1 \ldots L^{k_{p-1}}{}_{p-1} \delta^{k_p}{}_q L^{k_{p+1}}{}_{p+1} \ldots L^{k_n}{}_n,$$

Thus

$$\mathbf{L} \operatorname{adj} \mathbf{L} = (\det \mathbf{L})\mathbf{1} = (\operatorname{adj} \mathbf{L})\mathbf{L}.$$

Therefore **L** is invertible if and only if $\det \mathbf{L} \neq 0$. If so,

$$\mathbf{L}^{-1} = (\det \mathbf{L})^{-1} \operatorname{adj} \mathbf{L}.$$

The invertible tensors constitute a group under multiplication.

A tensor $\mathbf{L}$ such that $\det \mathbf{L} = \pm 1$ is *unimodular*. The unimodular tensors constitute a subgroup of the group of invertible tensors. That group is called the *unimodular group* $\mathscr{U}$.

5. *Inner-Product Spaces*

The vector spaces encountered in this book are endowed with an *inner product*, denoted by a dot. The *magnitude* or *length* $|\mathbf{u}|$ of a vector is defined in terms of the inner product:

$$|\mathbf{u}| := \sqrt{\mathbf{u}\cdot\mathbf{u}}.$$

The elements $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* or *perpendicular* if

$$\mathbf{u}\cdot\mathbf{v} = 0.$$

The only vector orthogonal to all vectors is $\mathbf{0}$. In fact, if $\mathbf{u}$ is such that $\mathbf{u}\cdot\mathbf{v}$ is bounded above for all $\mathbf{v}$, then $\mathbf{u} = \mathbf{0}$. Also

$$|\mathbf{u}\cdot\mathbf{v}| \leqq |\mathbf{u}|\,|\mathbf{v}|,$$
$$|\mathbf{u}+\mathbf{v}| \leqq |\mathbf{u}| + |\mathbf{v}|.$$

In the former inequality, which is called *Cauchy's inequality*, the sign $=$ is valid if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

The set of all vectors perpendicular to a given set of vectors forms a subspace. It is called the *orthogonal complement* of the subspace spanned by the given set. The vector space itself is the *direct sum* of any of its subspaces and the orthogonal complement of that subspace. This statement means that if $\mathbf{u}$ is any vector, it can be expressed as the sum of a uniquely determined vector from any desired subspace and another vector, also uniquely determined, from the orthogonal complement of that subspace.

If g is a linear function of vectors whose domain is the whole vector space and whose values are scalars, there is one and only one vector $\mathbf{f}$ such that

$$g(\mathbf{u}) = \mathbf{f}\cdot\mathbf{u}.$$

This statement is the *representation theorem* for linear, scalar-valued functions.

If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is a basis, another basis $\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n$ is determined uniquely by the conditions

$$\mathbf{e}_q\cdot\mathbf{e}^k = \delta_q^k.$$

The basis $\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n$ is *reciprocal* to the original one. The conditions

$$\mathbf{u} = u^p \mathbf{e}_p \qquad \text{and} \qquad \mathbf{u} = u_r \mathbf{e}^r$$

are equivalent, respectively, to

$$u^k = \mathbf{e}^k \cdot \mathbf{u}, \qquad u_k = \mathbf{e}_k \cdot \mathbf{u}.$$

The components u^k are called *contravariant*, while the components u_k are called *covariant*, both relative to the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, and

$$\mathbf{u} \cdot \mathbf{v} = u^q v_q = u_p v^p.$$

A basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is *orthonormal* if

$$\mathbf{e}_q \cdot \mathbf{e}_k = \delta_{qk} := \begin{cases} 1 & \text{if } q = k \\ 0 & \text{if } q \neq k. \end{cases}$$

A basis is orthonormal if and only if it is its own reciprocal. Corresponding contravariant and covariant components relative to an orthonormal basis equal one another, and so when an orthonormal basis is used, one speaks simply of "components".

A familiar example of an n-dimensional inner-product space is the *cartesian space* $\mathscr{R}_n$, the vectors of which are lists of n real numbers v_k:

$$\mathbf{v} := (v_1, v_2, \ldots, v_n),$$

provided addition and scalar multiplication are defined by the corresponding operations on the entries in the list. The *standard basis* is defined as follows:

$$\mathbf{e}_k := (0, 0, \ldots, 0, 1, 0, \ldots, 0),$$

the 1 being the k^{th} entry.

With no loss in generality any limit process on an n-dimensional vector space may be expressed in terms of cartesian co-ordinates in $\mathscr{R}_n$. That usage is familiar from old treatises on mechanics and many engineering textbooks today. $\mathscr{R}_1$ is the real line.

6. *Tensor Products. Tensors of Orders Greater than 2*

If $\mathbf{a}$ and $\mathbf{b}$ are vectors, their *tensor product* is the tensor $\mathbf{a} \otimes \mathbf{b}$ such that

$$(\mathbf{a} \otimes \mathbf{b})\mathbf{u} = (\mathbf{u} \cdot \mathbf{b})\mathbf{a} \qquad \forall \mathbf{u} \in \mathscr{V}.$$

In components,

$$(\mathbf{a} \otimes \mathbf{b})^k{}_m = a^k b_m .$$

If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ and $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ are bases of the vector space, the set of tensor products

$$\mathbf{e}_q \otimes \mathbf{f}_k$$

form a basis for the space of tensors:

$$\mathbf{L} = L^{qk} \mathbf{e}_q \otimes \mathbf{f}_k .$$

The scalars L^{qk} are the *contravariant components* of $\mathbf{L}$ with respect to the basis. Commonly $\mathbf{f}_k$ is chosen for $\mathbf{e}_k$. Then

$$\mathbf{L} = L^{qk} \mathbf{e}_q \otimes \mathbf{e}_k = \mathbf{L}^q{}_k \mathbf{e}_q \otimes \mathbf{e}^k = L_{qk} \mathbf{e}^q \otimes \mathbf{e}^k = L_q{}^k \mathbf{e}^q \otimes \mathbf{e}_k .$$

The scalars $L^r{}_s$ here are the same as those denoted previously by the same symbol and called simply "components" of $\mathbf{L}$ relative to $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. They are called also *mixed components* relative to that basis and its reciprocal. In the same terms, the $L_u{}^v$ are the mixed components relative to the basis $\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n$ and its reciprocal. The scalars L^{rw} and L_{hm} are the *contravariant* and *covariant* components, respectively, of $\mathbf{L}$ relative to the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. If the basis is orthonormal, then $L^{qk} = L^q{}_k = L_q{}^k = L_{qk}$.

We note that

$$\mathbf{1} = \mathbf{e}_k \otimes \mathbf{e}^k = \mathbf{e}^k \otimes \mathbf{e}_k .$$

The introduction of tensor products affords another method of defining tensors. For example, if we use the symbol $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$ to denote the linear mapping of the given vector space into (second-order) tensors such that

$$(\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c})\mathbf{d} = (\mathbf{c} \cdot \mathbf{d})(\mathbf{a} \otimes \mathbf{b}),$$

we can prove that the products $\mathbf{e}_k \otimes \mathbf{e}_p \otimes \mathbf{e}_q$ of elements of a basis for the given vector space form a basis for the set of such transformations. That is, if N is any linear mapping of the given vector space into the space of (second-order) tensors, it may be expressed in the form

$$\mathsf{N} = N^{kmp}\mathbf{e}_k \otimes \mathbf{e}_m \otimes \mathbf{e}_p ,$$

and its *contravariant components* N^{qrs} obey the transformation law

$$\overline{N}^{qrs} = \overline{A}^q_u \overline{A}^r_m \overline{A}^s_p N^{ump} .$$

The tensors so defined are of *third* order; the method illustrated serves to define tensors of any order.

7. *Transposition. Symmetric and Skew Tensors*

If $B(\mathbf{u}, \mathbf{v})$ is a scalar-valued bilinear function defined for all vectors $\mathbf{u}$ and $\mathbf{v}$, there is a unique tensor $\mathbf{L}$ such that

$$B(\mathbf{u}, \mathbf{v}) = \mathbf{u}\cdot\mathbf{L}\mathbf{v}.$$

This statement is the *representation theorem for bilinear functions*. If $\mathbf{L}$ is determined in this way by B, we can determine another tensor $\mathbf{L}^{\mathsf{T}}$, called the *transpose* of $\mathbf{L}$, by the requirement that

$$B(\mathbf{v}, \mathbf{u}) = \mathbf{u}\cdot\mathbf{L}^{\mathsf{T}}\mathbf{v}.$$

Then

$$(\mathbf{L}+\mathbf{M})^{\mathsf{T}} = \mathbf{L}^{\mathsf{T}} + \mathbf{M}^{\mathsf{T}},$$

$$(\mathbf{L}\mathbf{M})^{\mathsf{T}} = \mathbf{M}^{\mathsf{T}}\mathbf{L}^{\mathsf{T}},$$

$$(\mathbf{L}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{L},$$

$$(\mathbf{a}\otimes\mathbf{b})^{\mathsf{T}} = \mathbf{b}\otimes\mathbf{a}.$$

If $\mathbf{L}$ is invertible, then so is $\mathbf{L}^{\mathsf{T}}$, and

$$(\mathbf{L}^{\mathsf{T}})^{-1} = (\mathbf{L}^{-1})^{\mathsf{T}} =: \mathbf{L}^{-\mathsf{T}} .$$

In mixed components,

$$(L^{\mathsf{T}})^q{}_k = L_k{}^q.$$

In terms of matrices, $[\mathbf{L}^{\mathsf{T}}] = [\mathbf{L}]^{\mathsf{T}}$ if the components are taken relative to an orthonormal basis, but otherwise in general $[\mathbf{L}^{\mathsf{T}}] \neq [\mathbf{L}]^{\mathsf{T}}$. Of course

$$(L^{\mathsf{T}})^{kq} = L^{qk}, \qquad (L^{\mathsf{T}})_{qk} = L_{kq};$$

that is, the matrices of contravariant and covariant components of $\mathbf{L}^{\mathsf{T}}$ are the transposes of the respective matrices of $\mathbf{L}$.

Tensors $\mathbf{S}$ and $\mathbf{W}$ such that

$$\mathbf{S} = \mathbf{S}^{\mathsf{T}}, \qquad \mathbf{W} = -\mathbf{W}^{\mathsf{T}},$$

are called *symmetric* and *skew*, respectively. The conditions are expressed as follows in terms of components:

$$\begin{aligned} S_{qk} &= S_{kq}, & S^{qk} &= S^{kq}, & S_q{}^k &= S^k{}_q, \\ W_{qk} &= -W_{kq}, & W^{qk} &= -W^{kq}, & W_q{}^k &= -W^k{}_q. \end{aligned}$$

The set of all symmetric tensors is a $\frac{1}{2}n(n+1)$-dimensional subspace of the space of tensors; the set of skew tensors, a $\frac{1}{2}n(n-1)$-dimensional subspace. Bases for these two subspaces are formed by the following sets of products of the vectors of a basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$:

$$\mathbf{e}_k \otimes \mathbf{e}_m + \mathbf{e}_m \otimes \mathbf{e}_k, \qquad k \leqq m,$$

and

$$\mathbf{e}_k \wedge \mathbf{e}_m, \qquad k < m,$$

the *wedge product* or *exterior product* being the skew tensor defined as follows:

$$\mathbf{a} \wedge \mathbf{b} := \mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a}.$$

Any tensor $\mathbf{L}$ has a unique representation as the sum of a symmetric and a skew tensor, both unique:

$$\mathbf{L} = \operatorname{sym}\mathbf{L} + \operatorname{skw}\mathbf{L},$$

$$\operatorname{sym}\mathbf{L} := \tfrac{1}{2}(\mathbf{L} + \mathbf{L}^{\mathsf{T}}), \qquad \operatorname{skw}\mathbf{L} := \tfrac{1}{2}(\mathbf{L} - \mathbf{L}^{\mathsf{T}}).$$

If $\mathbf{S}$ is either symmetric or skew, $\mathbf{S}^2$ is symmetric, and $\mathbf{S}^2$ and $\mathbf{S}$ have the same nullspace. If $\mathbf{T}$ and $\mathbf{U}$ are either both symmetric or both skew, $(\mathbf{TU})^{\mathsf{T}} = \mathbf{UT}$. Hence $\mathbf{T}$ and $\mathbf{U}$, if both symmetric or both skew, commute if and only if $\mathbf{TU}$ is symmetric. If $\mathbf{W}$ is skew and $\dim \mathscr{V}$ is odd, $\det \mathbf{W} = 0$.

8. Orthogonal Tensors

A mapping $\mathbf{Q}$ of an inner-product space onto itself is *orthogonal* if it preserves the inner product:

$$(\mathbf{Qu})\cdot(\mathbf{Qv}) = \mathbf{u}\cdot\mathbf{v}.$$

This condition is satisfied if and only if $\mathbf{Q}$ is a tensor such that

$$\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}.$$

Hence

$$\det \mathbf{Q} = \pm 1.$$

If $\det \mathbf{Q} = 1$, the orthogonal tensor $\mathbf{Q}$ is called *proper*, or equivalently, a *rotation*. The *central inversion* $-\mathbf{1}$ is orthogonal; it is a rotation if and only if n is even. If n is odd, either $\mathbf{Q}$ or $-\mathbf{Q}$ is a rotation, while the other is the product of a rotation by the central inversion.

The orthogonal tensors constitute a proper subgroup of $\mathscr{u}$ called the (full) *orthogonal group* $\mathscr{o}$; if n is odd, the rotations form a proper subgroup of the orthogonal group.

Some special properties of orthogonal tensors over a 3-dimensional space are listed below in Section App. II.A.14.

9. Trace, Inner Product of Tensors

The *trace* $\operatorname{tr}\mathbf{A}$ of the tensor $\mathbf{A}$ is defined uniquely by the following two requirements: It is a linear function whose domain is the set of all tensors and whose values are scalars, and

$$\operatorname{tr}(\mathbf{u}\otimes\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}.$$

Hence

$$\operatorname{tr}\mathbf{A} = A^k{}_k = A_k{}^k = \operatorname{tr}\mathbf{A}^{\mathsf{T}};$$

that is, the trace of a tensor is the trace of the matrix of either of its arrays of mixed components. Of course $\operatorname{tr}\mathbf{1} = n$, $\operatorname{tr}\mathbf{0} = 0$. If $\mathbf{W}$ is skew, $\operatorname{tr}\mathbf{W} = 0$.

The *inner product* $\mathbf{A}\cdot\mathbf{B}$ of the tensors $\mathbf{A}$ and $\mathbf{B}$ is defined as follows:

$$\mathbf{A}\cdot\mathbf{B} := \operatorname{tr}(\mathbf{A}\mathbf{B}^{\mathsf{T}}) = \mathbf{B}\cdot\mathbf{A}.$$

With this definition the set of all tensors $\mathbf{A}$, regarded as a vector space of dimension n^2, becomes an inner-product space. The *magnitude* $|\mathbf{A}|$ of the tensor $\mathbf{A}$ is defined from the inner product in the usual way:

$$|\mathbf{A}| := \sqrt{\mathbf{A}\cdot\mathbf{A}} = \sqrt{\operatorname{tr}\mathbf{A}\mathbf{A}^{\mathsf{T}}}.$$

In components,

$$\mathbf{A}\cdot\mathbf{B} = A^k{}_q B_k{}^q,$$

$$|\mathbf{A}| = \sqrt{A^k{}_q A_k{}^q}.$$

If $\mathbf{S}$ is symmetric and $\mathbf{W}$ is skew, then $\mathbf{S}\cdot\mathbf{W} = 0$; also for any tensor $\mathbf{L}$

$$\mathbf{S}\cdot\mathbf{L} = \mathbf{S}\cdot\operatorname{sym}\mathbf{L}, \qquad \mathbf{W}\cdot\mathbf{L} = \mathbf{W}\cdot\operatorname{skw}\mathbf{L},$$

and so

$$\mathbf{W}\cdot\mathbf{a}\wedge\mathbf{b} = -2\mathbf{b}\cdot\mathbf{W}\mathbf{a} = 2\mathbf{a}\cdot\mathbf{W}\mathbf{b},$$

$$\tfrac{1}{2}(\mathbf{a}\wedge\mathbf{b})\cdot(\mathbf{c}\wedge\mathbf{d}) = \mathbf{a}\cdot\mathbf{d}\,\mathbf{b}\cdot\mathbf{c} - \mathbf{a}\cdot\mathbf{c}\,\mathbf{b}\cdot\mathbf{d}.$$

10. Invariant Subspaces, Projections, Proper Vectors, Proper Numbers

If the tensor $\mathbf{A}$ maps a certain subspace into itself, that subspace is *invariant* under $\mathbf{A}$. Every tensor $\mathbf{A}$ has invariant subspaces, among which are the whole vector space, $\{\mathbf{0}\}$, Range $\mathbf{A}$, and Nullspace $\mathbf{A}$.

A tensor $\mathbf{E}$ is called a *projection* if it is idempotent: $\mathbf{E}^2 = \mathbf{E}$. A projection is called a *perpendicular projection* if also it is symmetric: $\mathbf{E}^{\mathsf{T}} = \mathbf{E}$. If $\mathbf{E}$ is a projection, there is a basis relative to which

$$[\mathbf{E}] = \begin{Vmatrix} 1 & 0 & 0 & & \cdots & & & 0 \\ 0 & 1 & & & & & & \vdots \\ & & \ddots & & & & & \\ & & & 1 & 0 & 0 & & 0 \\ \vdots & & & 0 & 0 & & & \\ & & & \vdots & & 0 & & \vdots \\ & & & & & & \ddots & \\ 0 & & \cdots & 0 & & & & 0 \end{Vmatrix},$$

the number of 1s being equal to the dimension of the invariant subspace of $\mathbf{E}$. If $\mathbf{E}$ is a perpendicular projection, the basis may be chosen orthonormal.

If $\mathbf{e}$ is a unit vector, any vector $\mathbf{v}$ has a unique decomposition as the sum of a vector $\mathbf{P_e v}$ parallel to $\mathbf{e}$ and another $\mathbf{P^e v}$ perpendicular to it:

$$\mathbf{v} = \mathbf{P_e v} + \mathbf{P^e v}, \qquad \mathbf{e} \cdot \mathbf{P^e v} = 0.$$

Both $\mathbf{P_e}$ and $\mathbf{P^e}$ are perpendicular projections, and

$$\mathbf{P_e} = \mathbf{e} \otimes \mathbf{e}, \qquad \mathbf{P^e} = \mathbf{1} - \mathbf{e} \otimes \mathbf{e}.$$

$\mathbf{P_e}$ is the *projection onto the span* of $\mathbf{e}$; $\mathbf{P^e}$ is the *projection onto the plane normal to* $\mathbf{e}$. The *reflection* $\mathbf{R_e}$ *across the plane normal to* $\mathbf{e}$ is the orthogonal tensor defined as follows:

$$\mathbf{R_e} := -\mathbf{P_e} + \mathbf{P^e} = \mathbf{1} - 2\mathbf{e} \otimes \mathbf{e}.$$

Therefore $\mathbf{v}$ is parallel to $\mathbf{e}$ if and only if $\mathbf{R_e v} = -\mathbf{v}$; perpendicular to $\mathbf{e}$ if and only if $\mathbf{R_e v} = \mathbf{v}$.

If x is any scalar, the nullspace of $\mathbf{A} - x\mathbf{1}$ is an invariant subspace of $\mathbf{A}$. It is called the *proper space* of $\mathbf{A}$ corresponding with x, and its dimension is the *multiplicity* of x for $\mathbf{A}$. The scalar x is a *proper number* of $\mathbf{A}$ if any one, and hence all, of the following equivalent conditions holds:

1. There is a vector **u** other than **0** such that

$$\mathbf{A}\mathbf{u} = x\mathbf{u}.$$

2. The proper space of **A** corresponding with x contains one vector besides **0**.
3. The multiplicity of x for **A** is not 0.

The elements of the proper space are the *proper vectors* corresponding to that proper number. A proper number is *simple* if its multiplicity is 1; that is, if its proper space is one-dimensional. The set of all proper numbers of **A** is called the *spectrum* of **A**. By definition, the scalars constituting the spectrum are distinct.

The *characteristic polynomial* $P_{\mathbf{A}}(x)$ of **A** is defined as follows:

$$P_{\mathbf{A}}(x) := x^n - I_1 x^{n-1} + I_2 x^{n-2} + \cdots (-1)^n I_n ,$$

the signs being alternately $-$ and $+$; the *principal invariants* I_k of **A** are defined as follows:

$$I_k := \frac{1}{k!}\delta^{s_1 s_2 \ldots s_k}_{m_1 m_2 \ldots m_k} A^{m_1}{}_{s_1} A^{m_2}{}_{s_2} \cdots A^{m_k}{}_{s_k} , \qquad k = 1, 2, \ldots, n.$$

The symbol $\delta^{s_1 s_2 \ldots s_k}_{m_1 m_2 \ldots m_k}$ denotes 0 if any superscript or subscript is repeated, or if the subscripts fail to be the same numbers as the superscripts; otherwise it denotes ± 1 according as an even or odd permutation is needed to bring the subscripts into the same order as the superscripts. While this definition of the I_k seems to depend upon a basis, the value of I_k so obtained is the same for all bases. For example,

$$I_1 = \operatorname{tr} \mathbf{A},$$

$$I_n = \det \mathbf{A}.$$

Because the field of scalars is the real field, the principal invariants of tensors are real numbers.

The characteristic polynomial $P_{\mathbf{A}}(x)$ has real co-efficients and is a function of a real variable x. We obtain from it a complex polynomial if we replace x by a complex variable z. There are exactly n unique numbers $a_1, a_2, \ldots, a_n$, possibly complex, such that

$$P_{\mathbf{A}}(z) = \prod_{k=1}^{n}(z - a_k).$$

The equation $P_{\mathbf{A}}(z) = 0$ is the *characteristic equation* of $\mathbf{A}$, and the numbers $a_1, a_2, \ldots, a_n$ are the *latent roots* of $\mathbf{A}$. If q of the latent roots are equal, their common value is said to be a latent root of *algebraic multiplicity* q. Since the principal invariants of $\mathbf{A}$ are real, such latent roots of $\mathbf{A}$ as are not real occur in complex-conjugate pairs. I_k is the sum of the products of the latent roots taken k at a time, $k = 1, 2, \ldots, n$. Thus, for example, if $n = 3$,

$$\begin{aligned} I_{\mathbf{A}} &:= I_1 = a_1 + a_2 + a_3\,, \\ II_{\mathbf{A}} &:= I_2 = a_2 a_3 + a_3 a_1 + a_1 a_2\,, \\ III_{\mathbf{A}} &:= I_3 = a_1 a_2 a_3\,. \end{aligned}$$

Every proper number of $\mathbf{A}$ is a latent root, and every real latent root is a proper number. If $\mathbf{A}\mathbf{A}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}}\mathbf{A}$, the multiplicity of a proper number of $\mathbf{A}$ is the same as its algebraic multiplicity as a latent root. Such is the case, therefore, for tensors that are symmetric, skew, or orthogonal. If n is odd, $\mathbf{A}$ has at least one proper number, but if n is even, $\mathbf{A}$ need have none.

The *Hamilton–Cayley Theorem* states that the tensor $\mathbf{A}$ satisfies an equation having the same form as its characteristic equation:

$$\mathbf{A}^n - I_1\mathbf{A}^{n-1} + \cdots (-1)^n I_n \mathbf{1} = \mathbf{0}.$$

It is possible, of course, that $\mathbf{A}$ may satisfy also a polynomial equation of degree less than n. If $\mathbf{A}$ is invertible,

$$\mathbf{A}^{-1} = \frac{(-1)^{n-1}}{I_n}[\mathbf{A}^{n-1} - I_1\mathbf{A}^{n-2} + \cdots + (-1)^{n-1} I_{n-1}\mathbf{1}].$$

Thus

$$\operatorname{adj}\mathbf{A} = (-1)^{n-1}[\mathbf{A}^{n-1} - I_1\mathbf{A}^{n-2} + \cdots + (-1)^{n-1} I_{n-1}\mathbf{1}].$$

For a given $\mathbf{A}$ and $\mathbf{B}$ the equation

$$\mathbf{AX} + \mathbf{XB} = \mathbf{0} \qquad (*)$$

has the solution $\mathbf{X} = \mathbf{0}$. A theorem of SYLVESTER[1] states that there is another solution if and only if $\mathbf{A}$ and $-\mathbf{B}$ have a common proper number. If the equation

$$\mathbf{AX} + \mathbf{XB} = \mathbf{C}$$

for given $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ other than $\mathbf{0}$ has one solution $\mathbf{X}$, then its other solutions, if any, are obtained by adding to $\mathbf{X}$ all solutions of $(*)$ beyond $\mathbf{0}$.

11. Spectral Decomposition of Symmetric Tensors

Every symmetric tensor $\mathbf{S}$ has at least one proper number. In fact, the least and greatest proper numbers of $\mathbf{S}$ are the least and greatest values, respectively, of $\mathbf{u}\cdot\mathbf{Su}$ as $\mathbf{u}$ ranges over all unit vectors. Every latent root of a symmetric tensor is real and hence is a proper number. The proper spaces of a symmetric tensor are mutually orthogonal. Any vector may be expressed as a linear combination of vectors, each of which belongs to one (and of course, if it is not the vector $\mathbf{0}$, to only one) of the proper spaces of $\mathbf{S}$.

If $s_1, s_2, \ldots, s_p$ are the proper numbers of $\mathbf{S}$, then there is a unique set of perpendicular projections $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_p$ such that

$$\mathbf{E}_k\mathbf{E}_q = \mathbf{0} \qquad \text{if } k \neq q,$$

$$\sum_{k=1}^{p} \mathbf{E}_k = \mathbf{1},$$

and

$$\mathbf{S} = \sum_{k=1}^{p} s_k \mathbf{E}_k .$$

Hence there is at least one orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, each member of which is a proper vector of $\mathbf{S}$:

$$\mathbf{Se}_q = s_q \mathbf{e}_q , \qquad q = 1, 2, \ldots, n,$$

[1] This theorem follows from a more general one, likewise due to SYLVESTER, which is stated and proved in outline at the beginning of Chapter VIII of C.-C. MACDUFFEE's *The Theory of Matrices*, Volume 2 of *Ergebnisse der Mathematik und ihre Grenzgebiete*, Berlin, Springer, 1933, reprinted in 1946 by Chelsea Publishing Co., N.Y.

s_q being a proper number of **S**, repeated a number of times equal to its multiplicity. Such a basis is called *principal*. The matrix of components of **S** relative to this basis is diagonal. Indeed,

$$[\mathbf{S}] = \begin{Vmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & & \\ \vdots & & \ddots & \\ 0 & \cdots & & s_n \end{Vmatrix},$$

where again each proper number occurs a number of times equal to its multiplicity. With the same convention of multiplicity,

$$\mathbf{S} = \sum_{k=1}^{n} s_k \mathbf{P}_{\mathbf{e}_k} = \sum_{k=1}^{n} s_k (\mathbf{e}_k \otimes \mathbf{e}_k).$$

This statement presents the *spectral decomposition* of **S**.

In order that two symmetric tensors **S** and **T** have the same proper numbers, each with the same multiplicity, it is necessary and sufficient that there be an orthogonal tensor **Q** such that

$$\mathbf{T} = \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathsf{T}}.$$

If this condition holds, the proper spaces of **T** are the images under **Q** of the proper spaces of **S**. An orthogonal tensor **Q** commutes with the symmetric tensor **S** if and only if the proper spaces of **S** are invariant subspaces of **Q**.

12. Positive Tensors

A tensor **S** is *positive* if

$$\mathbf{u}\cdot\mathbf{S}\mathbf{u} > 0 \qquad \text{unless} \qquad \mathbf{u} = \mathbf{0},$$

not negative if

$$\mathbf{u}\cdot\mathbf{S}\mathbf{u} \geqq 0 \qquad \text{for all vectors } \mathbf{u}.$$

If $-\mathbf{S}$ is positive, **S** is *negative*; if $-\mathbf{S}$ is not negative, **S** is not positive. We abbreviate these terms by the notations $\mathbf{S} > 0$, $\mathbf{S} \geqq 0$, $\mathbf{S} < 0$, $\mathbf{S} \leqq 0$, respectively. If $\mathbf{L}$ is any tensor, $\mathbf{LL}^T \geqq 0$ and $\mathbf{L}^T\mathbf{L} \geqq 0$. If $\mathbf{L}$ is invertible, $\mathbf{LL}^T > 0$ and $\mathbf{L}^T\mathbf{L} > 0$. If $\mathbf{L}$ is symmetric, $\mathbf{L} > 0$ if and only if all of its proper numbers are positive, and $\mathbf{L} \geqq 0$ if and only if none of its proper numbers is negative.

If $\mathbf{S}$ is symmetric and positive, there is one and only one symmetric and positive $\mathbf{T}$ such that $\mathbf{T}^2 = \mathbf{S}$. We denote this tensor by $\sqrt{\mathbf{S}}$ and call it the *square root* of $\mathbf{S}$. The proper numbers of $\sqrt{\mathbf{S}}$ are the positive square roots of those of $\mathbf{S}$.

If $\mathbf{A}$ is a given symmetric tensor that is either positive or negative, the equation

$$\mathbf{AX} + \mathbf{XA} = \mathbf{C}$$

has a unique solution $\mathbf{X}$ for given $\mathbf{C}$. If $\mathbf{C} = 0$, then $\mathbf{X} = 0$. If $\mathbf{C}$ is symmetric, so is $\mathbf{X}$; if $\mathbf{C}$ is skew, so is $\mathbf{X}$. GUO[1] has shown that if $\mathbf{X}$ and $\mathbf{C}$ are skew and $\mathbf{A}$ is positive, and if $n = 3$, then

$$\mathbf{X} = \frac{1}{I_{\mathbf{A}}II_{\mathbf{A}} - III_{\mathbf{A}}}[(I_{\mathbf{A}}^2 - II_{\mathbf{A}})\mathbf{C} - (\mathbf{A}^2\mathbf{C} + \mathbf{CA}^2)].$$

13. Polar Decomposition

If $\mathbf{L}$ is an invertible tensor, then there are unique, positive, symmetric tensors $\mathbf{S}$ and $\mathbf{T}$ and a unique, orthogonal tensor $\mathbf{Q}$ such that

$$\mathbf{L} = \mathbf{QS} = \mathbf{TQ}.$$

$\mathbf{T}$ and $\mathbf{S}$ determine each other as follows if $\mathbf{Q}$ is known:

$$\mathbf{T} = \mathbf{QSQ}^T.$$

If $\mathbf{L}$ is not invertible, the polar decomposition still holds with symmetric $\mathbf{S}$ or $\mathbf{T}$ not negative and with $\mathbf{Q}$ not unique.

14. Structure of Orthogonal Tensors over a 3-Dimensional Vector Space

In this book we need to analyse orthogonal tensors only when $\dim \mathscr{V} = 3$. Then the central inversion $-\mathbf{1}$ is orthogonal but not a rotation. Every orthogonal tensor $\mathbf{Q}$ is either a rotation $\mathbf{R}$ or the product $-\mathbf{R}$ of a rotation $\mathbf{R}$ by $-\mathbf{1}$. Thus the structure of orthogonal tensors is determined by the structure of rotations.

[1] *Cf.* GUO ZHONG-HENG, "Rates of stretch tensors," *Journal of Elasticity* **14**(1984): 263–267.

The latent roots of $\mathbf{R}$ are $1, e^{i\theta}, e^{-i\theta}$ for some real number θ, called an *angle of rotation*. If θ is an angle of rotation, then so is $\pm\theta \pm 2n\pi$ for any integer n and any combination of signs. There is exactly one angle of rotation in the interval $0 \leqq \theta < \pi$.

If $\mathbf{R} \neq \mathbf{1}$, the proper space that corresponds with the proper number 1 is 1-dimensional. It is called the *axis* of the rotation $\mathbf{R}$. There are two unit vectors in the axis of $\mathbf{R}$. Corresponding with any angle of rotation θ there is a unique unit vector $\mathbf{e}$ in the axis of $\mathbf{R}$ such that

$$\begin{aligned} \mathbf{Re}_1 &= \cos\theta\,\mathbf{e}_1 + \sin\theta\,\mathbf{e}_2, \\ \mathbf{Re}_2 &= -\sin\theta\,\mathbf{e}_1 + \cos\theta\,\mathbf{e}_2, \\ \mathbf{Re} &= \mathbf{e}. \end{aligned} \tag{$*$}$$

for any $\mathbf{e}_1$ and $\mathbf{e}_2$ such that $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}$ is a right-handed orthonormal basis. Then for any vector $\mathbf{v}$ orthogonal to $\mathbf{e}$, θ is the angle between $\mathbf{v}$ and $\mathbf{Rv}$, measured counter-clockwise from $\mathbf{v}$. It follows from $(*)$ that θ is an angle of rotation of $\mathbf{R}$ if and only if it is a root of the equation

$$\cos\theta = \tfrac{1}{2}(\operatorname{tr}\mathbf{R} - 1).$$

Keeping $\mathbf{e}$ fixed, we see that $(*)$ holds if θ is replaced by $\theta \pm 2n\pi$, but $(*)$ does not hold if θ is replaced by $-\theta \pm 2n\pi$. In other words, if $\mathbf{e}$ is the unit vector in the axis that corresponds with the angle of rotation θ, then the other unite vector in the axis, namely $-\mathbf{e}$, corresponds with the angles of rotation $-\theta \pm 2n\pi$. If $\mathbf{R} = \mathbf{1}$, then $\theta = 0$, and $(*)$ holds for every basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}$. The matrix of $\mathbf{R}$ with respect to the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}$ is given by

$$[\mathbf{R}] = \begin{Vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{Vmatrix}.$$

In general co-ordinates

$$R^k{}_m = \cos\theta\,\delta^k_m + (1-\cos\theta)e^k e_m - \sin\theta\,\epsilon^k{}_{mp}e^p.$$

If $\mathbf{Q} = -\mathbf{R}$, the axis of $\mathbf{R}$ is called the *axis* of $\mathbf{Q}$, and an angle of rotation of $\mathbf{R}$ is called an *angle of rotation* of $\mathbf{Q}$.

$\mathbf{R} = \mathbf{R}^{\mathsf{T}}$ if and only if $\theta = 0$ or π.

15. Structure of Skew Tensors over a Three-Dimensional Vector Space. Three-Dimensional Vector Algebra[1]

In this book we need analyse skew tensors only if $\dim \mathscr{V} = 3$. Then the nullspace of a not null, skew tensor $\mathbf{W}$, namely, the subspace of vectors $\mathbf{n}$ such that

$$\mathbf{W}\mathbf{n} = \mathbf{0},$$

is 1-dimensional; it is called the *axis* of $\mathbf{W}$. If $\mathbf{W} \neq \mathbf{0}$, its only proper number is 0, and its axis is its only invariant subspace. Two skew tensors, neither of which is $\mathbf{0}$, have the same axis if and only if they are proportional to one another.

Let $\mathbf{n}$ be one of the two unit vectors lying on the axis of a skew tensor $\mathbf{W}$ other than $\mathbf{0}$, and let $\mathbf{e}$ be normal to $\mathbf{n}$. Then of the two unit vectors normal to the span of $\mathbf{e}$ and $\mathbf{n}$ we may choose one, say $\mathbf{f}$, such that

$$\mathbf{W} = \frac{1}{\sqrt{2}}|\mathbf{W}|\mathbf{e} \wedge \mathbf{f}.$$

Equivalently, there is an orthonormal basis such that

$$[\mathbf{W}] = \begin{Vmatrix} 0 & \frac{1}{\sqrt{2}}|\mathbf{W}| & 0 \\ -\frac{1}{\sqrt{2}}|\mathbf{W}| & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Two orthonormal bases $\{\mathbf{e}_k\}$ and $\{\bar{\mathbf{e}}_k\}$ are said to have *the same orientation* if they are obtainable from one another by a rotation: $\bar{\mathbf{e}}_k = \mathbf{R}\mathbf{e}_k$, $k = 1, 2, 3$. Since every orthogonal transformation is either a rotation or the negative of one, there are exactly two distinct classes of bases having the same orientation. In three-dimensional vector algebra one of these is set down and fixed. Of the two possible isomorphisms between skew tensors and vectors, one is specified by use of a particular orthonormal basis. "The Gibbsian cross" $\mathbf{T}_\times$ of a tensor $\mathbf{T}$ is the vector defined as follows in terms of components with respect to any such basis:

$$T_{\times 3} := T_{12} - T_{21}, \qquad T_{\times 1} := T_{23} - T_{32}, \qquad T_{\times 2} := T_{31} - T_{13}.$$

[1] *Cf.* Sections 7.16, 8.15, and 8.16 of W. H. Greub, *Linear Algebra*, 3rd ed., Berlin and New York, Springer-Verlag, 1967.

If $\mathbf{S}$ is skew, then

$$S_{\times 3} = 2S_{12}, \qquad S_{\times 1} = 2S_{23}, \qquad S_{\times 2} = 2S_{31}.$$

The *cross product* $\mathbf{u} \times \mathbf{w}$ of two vectors $\mathbf{u}$ and $\mathbf{w}$ is defined thus:

$$\mathbf{u} \times \mathbf{w} := \tfrac{1}{2}(\mathbf{u} \wedge \mathbf{w})_{\times},$$

and so

$$(\mathbf{u} \times \mathbf{w})_3 = u_1 w_2 - u_2 w_1, \qquad \text{etc.}$$

and

$$|\mathbf{u} \wedge \mathbf{w}| = \sqrt{2}|\mathbf{u} \times \mathbf{w}|.$$

Also

$$2\,\mathbf{S}\mathbf{u} = -\mathbf{S}_{\times} \times \mathbf{u}, \qquad \mathbf{S}\cdot(\mathbf{u} \wedge \mathbf{v}) = -\mathbf{S}_{\times}\cdot(\mathbf{u} \times \mathbf{v}),$$

and if $\mathbf{S}$ and $\mathbf{T}$, neither of them $\mathbf{0}$, are both skew,

$$\mathbf{S}\cdot\mathbf{T} = \tfrac{1}{2}\mathbf{S}_{\times}\cdot\mathbf{T}_{\times}, \qquad (\mathbf{S}\mathbf{T})_{\times} = \tfrac{1}{2}(\mathbf{S}\mathbf{T} - \mathbf{T}\mathbf{S})_{\times} = -\tfrac{1}{4}\mathbf{S}_{\times} \times \mathbf{T}_{\times}.$$

The first of these relations shows that $\mathbf{S}\cdot\mathbf{T} = 0$ if and only if the nullspaces of $\mathbf{S}$ and $\mathbf{T}$ are perpendicular; the second shows that $\mathbf{S}$ and $\mathbf{T}$ commute if and only if the nullspace of one of them contains the nullspace of the other. If $\mathbf{L}$ is invertible,

$$\mathbf{L}\mathbf{u} \times \mathbf{L}\mathbf{v} = (\det \mathbf{L})(\mathbf{L}^{-1})^{\mathsf{T}}(\mathbf{u} \times \mathbf{v}).$$

B. Geometry

1. Euclidean Point Spaces

While in this book there are allusions to rather general manifolds, the only specific geometry employed is that of Euclidean space.

A set $\mathscr{E}$ of elements $\mathbf{x}$, $\mathbf{y}$ is an n-dimensional *Euclidean point space* or *Euclidean manifold* if it is endowed with a structure defined in reference to an inner-product space $\mathscr{V}$ of n dimensions by the following axioms:

1. Each vector $\mathbf{u} \in \mathscr{V}$ maps $\mathscr{E}$ onto itself

$$\mathbf{u}(\mathbf{x}) \in \mathscr{E} \qquad \forall \mathbf{x} \in \mathscr{E}.$$

2. The composition of the mappings $\mathbf{u}$ and $\mathbf{v}$ is their vector sum:

$$(\mathbf{u} + \mathbf{v})(\mathbf{x}) = \mathbf{u}(\mathbf{v}(\mathbf{x})) \qquad \forall \mathbf{x} \in \mathscr{E}.$$

3. For given $\mathbf{x}$ and $\mathbf{y}$ there is exactly one vector $\mathbf{u}$ such that

$$\mathbf{u}(\mathbf{x}) = \mathbf{y}.$$

The elements of $\mathscr{E}$ are called *points*. The vector space $\mathscr{V}$ is the *translation space* of $\mathscr{E}$, and its elements are called *translations* of $\mathscr{E}$. The translations may be visualized as arrows; if the butt of the arrow $\mathbf{u}$ is put at $\mathbf{x}$, its sharp end distinguishes $\mathbf{u}(\mathbf{x})$. Thus we say that $\mathbf{u}$ *translates* $\mathbf{x}$ into $\mathbf{y}$; of course $-\mathbf{u}$ translates $\mathbf{y}$ into $\mathbf{x}$.

The second axiom, in view of the first, asserts that the result of applying first the translation $\mathbf{v}$ and then the translation $\mathbf{u}$ is the same as the result of applying $\mathbf{u} + \mathbf{v}$ to start with; thus it expresses the *axiom of resultant displacements*, familiar from elementary geometry and mechanics, and it suggests the notation

$$\mathbf{x} + \mathbf{u} := \mathbf{u}(\mathbf{x}).$$

Thus we use the plus sign to denote not only addition of vectors to each other but also addition of vectors to points. Axiom 3 enables us to extend the interpretation by writing $\mathbf{y} - \mathbf{x}$ for the unique vector $\mathbf{u}$ that maps $\mathbf{x}$ into $\mathbf{y}$:

$$\mathbf{y} - \mathbf{x} := \mathbf{u}.$$

Thus the difference of points is defined, and

$$\mathbf{x} + (\mathbf{y} - \mathbf{x}) = \mathbf{y}.$$

Let $\mathscr{U}$ be a subspace of $\mathscr{V}$ having positive dimension, and let some point $\mathbf{y}_0$ of $\mathscr{E}$ be selected. If $\mathscr{F} := \{\mathbf{y}_0 + \mathbf{u},\ \mathbf{u} \in \mathscr{U}\}$, then $\mathscr{F}$ is a *flat* of $\mathscr{E}$ *parallel* to $\mathscr{U}$. Two flats that are both parallel to the same subspace of $\mathscr{V}$ are *parallel* to each other. The *dimension* of $\mathscr{F}$ is the dimension of the $\mathscr{U}$ that defines it. A 1-dimensional flat is a *straight line*, while a 2-dimensional flat is a *plane*, and an $(n-1)$-dimensional flat is a *hyperplane*. If $n = 3$, planes and hyperplanes are the same thing. An equation for a straight line is

$$\mathbf{y} = \mathbf{y}_0 + s\mathbf{e},$$

in which $\mathbf{e}$ is some vector and s runs from $-\infty$ to ∞, while if $\mathbf{e}$ and $\mathbf{f}$ are linearly independent, an equation for a plane is $\mathbf{y} = \mathbf{y}_0 + s\mathbf{e} + r\mathbf{f}$, r and s running from $-\infty$ to ∞.

2. Distance, Isometry

The *distance* between the points $\mathbf{x}$ and $\mathbf{y}$ is the magnitude of the vector $\mathbf{u}$ that translates $\mathbf{x}$ into $\mathbf{y}$, that is,

$$|\mathbf{y} - \mathbf{x}|.$$

It is easy to see that this function of pairs of points satisfies the axioms of a metric, and in particular that it obeys the *triangle axiom*:

$$|\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}| \geqq |\mathbf{x} - \mathbf{z}|.$$

A mapping $\boldsymbol{\alpha}$ of $\mathscr{E}$ onto itself is called an *isometry* if it preserves distances. The *representation theorem* for isometries asserts that to each isometry $\boldsymbol{\alpha}$ of $\mathscr{E}$ corresponds a unique orthogonal tensor $\mathbf{Q}$ over $\mathscr{V}$ such that

$$\boldsymbol{\alpha}(\mathbf{x}) = \boldsymbol{\alpha}(\mathbf{x}_0) + \mathbf{Q}(\mathbf{x} - \mathbf{x}_0)$$

for each pair of points $\mathbf{x}_0$ and $\mathbf{x}$. Thus each isometry may be regarded as the succession, in either order, of a translation of an arbitrarily selected point and a uniquely determined orthogonal transformation of the vectors that translate that point into the other points of space.

3. Topology, Figures

The topology of a Euclidean space is defined in the standard way by the metric $|\mathbf{x} - \mathbf{y}|$. The definitions of spheres, cubes, parallelepipeds, open and closed sets, neighborhoods, interiors, closures, and boundaries as those concepts are used in this book may be found in any text on elementary analysis.

C. Calculus

1. Limits, Orders

Euclidean point space has a metric; likewise, the magnitude of a difference of vectors or tensors, $|\mathbf{u} - \mathbf{v}|$ or $|\mathbf{T} - \mathbf{S}|$, serves as a metric. In terms of the topologies defined by these metrics, standard procedure defines continuity, convergence, limits, boundedness, compactness, *etc.*, in the respective spaces. Standard theorems of calculus, such as those on subsequences, Cauchy's criterion, covering theorems, the theorem of the maximum and minimum on a compact set, are easy to extend to $\mathscr{E}$ and to vector spaces.

The *order symbols* O and o are defined as follows for scalar-valued functions of a scalar variable.

If there is a constant K such that

$$|f(x)| < K|g(x)|$$

when x is sufficiently near to a, we write

$$f = O(g) \qquad \text{as } x \to a.$$

If

$$\frac{f(x)}{g(x)} \to 0 \qquad \text{as } x \to a,$$

we write

$$f = o(g) \qquad \text{as } x \to a.$$

For example, $O(1)$ stands for a function that is bounded near a, and $o(1)$ stands for a function that tends to 0 as $x \to a$.

The statement that f is continuous at $x = a$ may be put as follows:

$$f(x) = f(a) + o(1) \qquad \text{as } x \to a.$$

These definitions are easily extended to functions of points whose values are scalars, vectors, or tensors. In estimating vectors or tensors we write $\mathbf{o}$ and $\mathbf{O}$ instead of o and O. For example, if f maps a domain of a normed vector space into a normed vector space, we write $f = o(v)$ as $v \to \mathbf{0}$ if

$$\lim \frac{|f(v)|}{|v|} = 0 \qquad \text{as } |v| \to 0.$$

2. *Differentiation*

If $\mathbf{f}$ is a function of a real variable t whose values are points or vectors, its *derivative* $\dot{\mathbf{f}}(t)$ at t is defined as follows: If there is a vector $\mathbf{g}$ such that

$$\mathbf{f}(t+s) = \mathbf{f}(t) + s\mathbf{g}(t) + \mathbf{o}(s) \qquad \text{as } s \to 0,$$

then $\mathbf{f}$ is *differentiable* at t, and $\mathbf{g}$ is the *derivative* of $\mathbf{f}$ at t. The standard notation for the derivative is

$$\dot{\mathbf{f}}(t) := \mathbf{g}(t).$$

Thus the derivative $\dot{\mathbf{f}}(t)$ defines a linear function that approximates the function $\mathbf{f}(t + \cdot) - \mathbf{f}(t)$ near $s = 0$. For a function whose values are tensors a similar definition and notation may be used.

There are simple rules for interchanging the order of differentiation and other operations. A few of these, in a notation which may confuse functions with their values, are listed below.

$$\begin{aligned}
(\mathbf{u} \otimes \mathbf{v})^{\cdot} &= \dot{\mathbf{u}} \otimes \mathbf{v} + \mathbf{u} \otimes \dot{\mathbf{v}}. \\
(\mathbf{LM})^{\cdot} &= \dot{\mathbf{L}}\mathbf{M} + \mathbf{L}\dot{\mathbf{M}}. \\
(\mathbf{L}^{\mathsf{T}})^{\cdot} &= (\dot{\mathbf{L}})^{\mathsf{T}}. \\
(\mathbf{L}^{m})^{\cdot} &= \sum_{k=1}^{m} \mathbf{L}^{k-1}\dot{\mathbf{L}}\mathbf{L}^{m-k}. \\
(\mathbf{L}^{-1})^{\cdot} &= -\mathbf{L}^{-1}\dot{\mathbf{L}}\mathbf{L}^{-1}. \\
(\det \mathbf{L})^{\cdot} &= (\det \mathbf{L})\dot{\mathbf{L}} \cdot \mathbf{L}^{-1}.
\end{aligned}$$

For the last two rules to hold, it is necessary that $\mathbf{L}$ be invertible. To prove the last one, do Exercise II.5.1.

If $\mathbf{Q}$ is a function whose values are orthogonal tensors, then the values of $\dot{\mathbf{Q}}\mathbf{Q}^{\mathsf{T}}$ are skew.

3. *Gradients*

A function $\mathbf{f}$ that maps points in a Euclidean space $\mathscr{E}$ into a vector space $\mathscr{W}$ is called a *vector field*. A vector field $\mathbf{f}$ is said to be *differentiable* at $\mathbf{x}$ if there is a linear mapping $\nabla\mathbf{f}(\mathbf{x})$ of n-dimensional vectors onto m-dimensional vectors such that

$$\mathbf{f}(\mathbf{x}+\mathbf{u}) = \mathbf{f}(\mathbf{x}) + \nabla\mathbf{f}(\mathbf{x})(\mathbf{u}) + \mathbf{o}(\mathbf{u}) \qquad \text{as } \mathbf{u} \to \mathbf{0}.$$

The function of $\mathbf{x}$ whose value at $\mathbf{x}$ is $\nabla\mathbf{f}(\mathbf{x})$ is called the *derivative* (or *gradient*) of $\mathbf{f}$ at $\mathbf{x}$. Equivalently, the gradient $\nabla\mathbf{f}(\mathbf{x})$, if it exists, is a linear mapping such that

$$\frac{d}{dt}\mathbf{f}(\mathbf{x}+t\mathbf{u})|_{t=0} = \nabla\mathbf{f}(\mathbf{x})(\mathbf{u}).$$

A function on an open set is *differentiable* thereon if it is differentiable at each point of that set. A function $\mathbf{f}$ that has a continuous derivative $\nabla\mathbf{f}$ on an open set is sometimes called *smooth* on that set.

Two special cases deserve notice. First, if $\mathscr{W} = \mathscr{V}$, the translation space of $\mathscr{E}$, then $\nabla\mathbf{f}(\mathbf{x})$ is a tensor, and $\nabla\mathbf{f}(\mathbf{x})\mathbf{u}$ is written for $\nabla\mathbf{f}(\mathbf{x})(\mathbf{u})$. Second, if $\mathscr{W}$ is the set of real numbers, the field f is called a *scalar field*. By the representation theorem for functions of vectors whose values are scalars we know that $\nabla f(\mathbf{x})(\mathbf{u})$ equals the inner product of some vector and $\mathbf{u}$. In this sense we say that the gradient of a scalar field at a point is a vector. Writing $\nabla f(\mathbf{x})$ for that vector, we have

$$f(\mathbf{x}+\mathbf{u}) = f(\mathbf{x}) + \nabla f(\mathbf{x})\cdot\mathbf{u} + \mathbf{o}(\mathbf{u}).$$

Similar definitions can be framed for functions of points whose values are points, vectors, or tensors.

Among the rules for taking the gradients of products of various kinds are

$$\nabla(fg) = f\,\nabla g + g\,\nabla f,$$
$$\nabla(\mathbf{f}\cdot\mathbf{g}) = (\nabla\mathbf{f})^{\mathsf{T}}\mathbf{g} + (\nabla\mathbf{g})^{\mathsf{T}}\mathbf{f},$$
$$\nabla(f\mathbf{g}) = \mathbf{g}\otimes\nabla f + f\,\nabla\mathbf{g}.$$

There is also the *chain rule* for taking the gradient of a composite function. If $\mathbf{f}\circ\mathbf{g}$ denotes the composition of $\mathbf{g}$ with $\mathbf{f}$, the rule can be written symbolically as

$$\nabla(\mathbf{f}\circ\mathbf{g}) = ((\nabla\mathbf{f})\circ\mathbf{g})\,\nabla\mathbf{g}.$$

If χ maps points onto points, and if f maps points onto scalars, then

$$\nabla(f \circ \chi) = (\nabla\chi)^{\mathsf{T}}((\nabla f) \circ \chi).$$

4. *Other Differential Operators*

The repeated or *second* gradient is the mapping that results from taking the gradient twice. It is denoted by ∇^2. If the values of f are scalars, the value of $\nabla^2 f$ is a symmetric tensor.

The operators *divergence* div and *laplacian* Δ upon vector fields and scalar fields, respectively, are defined as follows:

$$\operatorname{div}\mathbf{f} := \operatorname{tr}\nabla\mathbf{f},$$
$$\Delta f := \operatorname{div}\nabla f = \operatorname{tr}\nabla^2 f.$$

If $\mathbf{L}$ is a tensor field and $\mathbf{a}$ is a fixed vector, then $\mathbf{L}^{\mathsf{T}}\mathbf{a}$ is a vector field, and it is easy to see that the values of $\operatorname{div}(\mathbf{L}^{\mathsf{T}}\mathbf{a})$ are linear functions of the vector $\mathbf{a}$ whose values are scalars. Thus the divergence $\operatorname{div}\mathbf{L}$ of a tensor field $\mathbf{L}$ can be defined by the requirement that

$$\mathbf{a}\cdot\operatorname{div}\mathbf{L} = \operatorname{div}(\mathbf{L}^{\mathsf{T}}\mathbf{a}).$$

Among the rules for calculating divergences and gradients of products are the following:

$$\operatorname{div}(f\mathbf{g}) = \mathbf{g}\cdot\nabla f + f\operatorname{div}\mathbf{g},$$

$$\operatorname{div}(\mathbf{L}\mathbf{g}) = (\operatorname{div}\mathbf{L}^{\mathsf{T}})\cdot\mathbf{g} + \operatorname{tr}(\mathbf{L}\,\nabla\mathbf{g}),$$

$$\operatorname{div}(\nabla\mathbf{g})^{\mathsf{T}} = \nabla\operatorname{div}\mathbf{g},$$

$$\operatorname{div}[\nabla\mathbf{g} \pm (\nabla\mathbf{g})^{\mathsf{T}}] = \Delta\mathbf{g} \pm \nabla\operatorname{div}\mathbf{g}.$$

5. *Special Kinds of Vector Fields*

A vector field whose divergence vanishes is called *solenoidal*; whose laplacian vanishes, *harmonic*. The label "laplacian" is merely traditional, not an attribution.

A vector field $\mathbf{f}$ is *lamellar* if there is a scalar field P such that for every sufficiently short curve $\mathscr{C}$ that connects two sufficiently near points $\mathbf{x}_1$ and $\mathbf{x}_2$

$$\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x} = P(\mathbf{x}_1) - P(\mathbf{x}_2),$$

the sense of $\mathscr{C}$ being from $\mathbf{x}_1$ to $\mathbf{x}_2$. The function P is a *potential* of $\mathbf{f}$; the surfaces $P = \text{const.}$, which are called *equipotential surfaces*, are normal to $\mathbf{f}$. Conversely, if there is a scalar field P such that

$$\mathbf{f} = -\nabla P,$$

the field $\mathbf{f}$ is lamellar. P is determined by $\mathbf{f}$ only to within an additive constant. If $\mathbf{f}$ is differentiable, it is lamellar if and only if $\nabla\mathbf{f}$ is symmetric:

$$\operatorname{skw} \nabla\mathbf{f} = \mathbf{0}.$$

If the domain of $\mathbf{f}$ is a simply connected, open set, the restriction to sufficiently near points and sufficiently short curves, imposed as part of the definition of lamellar, is unnecessary.

If the domain of the lamellar field $\mathbf{f}$ is multiply connected, a potential exists locally, but the line integral $\int_{\mathscr{C}} \mathbf{f}\cdot d\mathbf{x}$ is not generally independent of the path $\mathscr{C}$ connecting two given points. If the two curves $\mathscr{C}_1$ and $\mathscr{C}_2$ connect $\mathbf{x}_1$ to $\mathbf{x}_2$, then

$$\int_{\mathscr{C}_1} \mathbf{f}\cdot d\mathbf{x} - \int_{\mathscr{C}_2} \mathbf{f}\cdot d\mathbf{x} = \sum_{k=1}^{q} n_k K_k\,;$$

the "cyclic constants" K_k are determined by $\mathbf{f}$ and its domain alone, and the numbers n_k are integers. The concept of potential may be extended to lamellar fields on multiply connected domains by introducing "cyclic functions," which map each point onto a set[1] of the form $\{P_0 + \sum_{k=1}^{q} n_k K_k\}$.

A vector field $\mathbf{f}$ is *complex-lamellar* if and only if it is non-trivially proportional to a lamellar field: There are scalar vields K and P, neither of them

[1]The classical treatment of cyclic potentials, due to KELVIN, is most easily available in Sections 49–51 of H. LAMB's *Hydrodynamics*, Cambridge, Cambridge University Press, 2nd–6th editions, 1895/1932. It is not easy to find a simple treatment that satisfies modern standards of rigor.

An elegant, rigorous treatment of lamellar fields that need not be differentiable, and also of solenoidal fields, may be found in a paper by H. WEYL, "The method of orthogonal projection in potential theory," *Duke Mathematical Journal* **7** (1940): 411–444.

constant, such that $\mathbf{f} = K\nabla P$. A theorem of EULER and KELVIN[1] asserts that a continuously differentiable field $\mathbf{f}$ is complex-lamellar if and only if it is not lamellar and $\mathbf{f} \cdot \operatorname{curl} \mathbf{f} = 0$.

6. *Curves. Vector Lines*[2]

A *curve* is a mapping, twice continuously differentiable, of an interval of $\mathscr{R}^1$ into $\mathscr{E}^3$, say $\mathbf{x} = \mathbf{g}(s)$; the parameter s, which increases monotonically from one end of the interval to the other, may be taken as arc-length. The *unit tangent* $\mathbf{t}$ at s is $\mathbf{g}'(s)$; we may write $\mathbf{t} := \mathbf{g}'$. The *curvature* κ is the scalar arc-rate at which the tangent turns; that is, $\kappa\mathbf{n} := \mathbf{t}'$, in which the *unit principal normal* $\mathbf{n}$ at $\mathbf{g}(s)$ is taken as one of the two unit vectors normal to $\mathbf{t}$ that lie in the *osculating plane* at s, namely the plane determined by three distinct values of $\mathbf{g}$ confluent at $\mathbf{g}(s)$. One of the two unit vectors normal to the osculating plane at $\mathbf{g}(s)$ is taken as the *unit binormal* $\mathbf{b}$. Thus $\mathbf{t}$, $\mathbf{n}$, $\mathbf{b}$ form an orthogonal triad at each point on the curve. Differentiation of $\mathbf{t} \cdot \mathbf{b} = 0$ yields $\mathbf{t} \cdot \mathbf{b}' = 0$, and so, since $\mathbf{b}'$ is perpendicular to $\mathbf{b}$, it must lie in the direction of $\mathbf{n}$. Writing $-\tau$ for the magnitude of $\mathbf{b}'$, we obtain $\mathbf{b}' = -\tau\mathbf{n}$. The quantity τ, which is called the *torsion*, is the arc-rate at which the osculating plane rotates around the tangent. Finally, $\mathbf{n}' = (\mathbf{b} \times \mathbf{t})' = -\tau\mathbf{n} \times \mathbf{t} + \kappa\mathbf{b} \times \mathbf{n} = \tau\mathbf{b} - \kappa\mathbf{t}$. The formulae for $\mathbf{t}'$ and $\mathbf{b}'$ are due to EULER and CAUCHY, respectively, while that for $\mathbf{n}'$ is an easy consequence of them. The set of three is called "the Serret–Frenet formulae".

The definition of "curve" can be broadened, typically by allowing piecewise smoothness, but points where differentiability fails always require special treatment.

At a given time, the *vector lines* of a vector field $\mathbf{c}$ are the curves everywhere tangent to $\mathbf{c}$. At each point the tangent of the vector line has the same direction as the value of $\mathbf{c}$ at that point. A vector field continuous in a closed region possesses at least one vector line through each interior point of the region; moreover, if the field satisfies a Lipschitz condition, it has exactly one vector line through each point $\mathbf{x}$ at which $\mathbf{c}(\mathbf{x}) \neq \mathbf{0}$.

The unit tangents of the vector lines in a region form a field $\mathbf{t}$, a function of $\mathbf{x}$ and t. The same is true of $\mathbf{n}$ and $\mathbf{b}$. The field $\mathbf{t}$ has two important scalar invariants:

$$\Theta := \operatorname{div} \mathbf{t}, \qquad \Omega := \mathbf{t} \cdot \operatorname{curl} \mathbf{t}.$$

[1] *Cf.* Section 105 of L. BRAND's *Vector and Tensor Analysis*, New York, Wiley, 1947, and Section 52 of *Introduction to Vectors and Tensors* by R. M. BOWEN & C. C. WANG, New York and London, Plenum Press, 1976.

[2] *Cf.* CFT, Section App. Va.

The field Ω, called the *abnormality* of the vector lines, which was introduced by ZHUKOVSKY and named by LEVI-CIVITA, gives measure to the motion's departure from being lamellar or complex lamellar. Since

$$\Omega = \frac{1}{c^2}\mathbf{c} \cdot \operatorname{curl}\mathbf{c}, \qquad c := |\mathbf{c}|,$$

$\mathbf{c}$ is lamellar or complex-lamellar if and only if $\Omega = 0$.

MASOTTI derived an intrinsic representation for $\operatorname{curl}\mathbf{c}$:

$$\operatorname{curl}\mathbf{c} = c\Omega\mathbf{t} + (\mathbf{b} \cdot \nabla c)\mathbf{n} + (c\kappa - \mathbf{n} \cdot \nabla c)\mathbf{b},$$

in which κ is the curvature of the vector lines. Putting $c = 1$ yields

$$\operatorname{curl}\mathbf{t} = \Omega\mathbf{t} + \kappa\mathbf{b}.$$

BJØRGUM obtained intrinsic expressions for the gradients and curls of the fields $\mathbf{t}$, $\mathbf{n}$, $\mathbf{b}$; corresponding conditions of compatibility were obtained by YIN & PIPKIN in the paper cited above in the footnote on p. 290. A convenient display of all of these are found in the paper by MARRIS & WANG cited above in Footnote 2 on p. 143.

7. *Co-ordinates*

A *co-ordinate system* on an open set of an n-dimensional Euclidean space is a one-to-one mapping of that set into $\mathscr{R}_n$, a mapping which has an invertible gradient and a continuous second gradient. If $\bar{\mathbf{x}}$ is such a mapping,

$$\bar{\mathbf{x}}(\mathbf{x}) = (\bar{x}^1(\mathbf{x}), \bar{x}^2(\mathbf{x}), \ldots, \bar{x}^n(\mathbf{x})),$$

in which $\bar{x}^k$ is a scalar field having the same degree of smoothness as that assumed for $\bar{\mathbf{x}}$. The number $\bar{x}^k(\mathbf{x})$ is the k^{th} *co-ordinate* of the point $\mathbf{x}$ in the co-ordinate system $\bar{\mathbf{x}}$.

If $\hat{\mathbf{x}}$ denotes the inverse of $\bar{\mathbf{x}}$, then

$$\bar{x}^k(\hat{\mathbf{x}}(x^1, x^2, \ldots, x^n)) = x^k, \qquad k = 1, 2, \ldots, n,$$

for all lists $(x^1, x^2, \ldots, x^n)$ that lie in the range of $\bar{\mathbf{x}}$. We set

$$\mathbf{e}^k(\mathbf{x}) := \nabla\bar{x}^k(\mathbf{x}), \qquad \mathbf{e}_m(\mathbf{x}) := \partial_{x^m}\hat{\mathbf{x}}(x^1, x^2, \ldots, x^n)|_{x^j = \bar{x}^j(\mathbf{x})},$$

where ∂_{x^m} denotes the partial derivative with respect to x^m of a point-valued function of the n real variables $x^1, x^2, \ldots, x^n$. The vector $\mathbf{e}^k(\mathbf{x})$ is normal at $\mathbf{x}$ to the *co-ordinate surface* $x^k(\mathbf{y}) = \text{const.}$ that passes through $\mathbf{x}$. The vector $\mathbf{e}_m(\mathbf{x})$ is tangent to the m^{th} *co-ordinate curve* at $\mathbf{x}$, that curve being the set of points near $\mathbf{x}$ for which every co-ordinate but x^m has the same value as it does at $\mathbf{x}$.

The sets of vectors $\mathbf{e}^1(\mathbf{x}), \mathbf{e}^2(\mathbf{x}), \ldots, \mathbf{e}^n(\mathbf{x})$ and $\mathbf{e}_1(\mathbf{x}), \mathbf{e}_2(\mathbf{x}), \ldots, \mathbf{e}_n(\mathbf{x})$ are reciprocal bases of the translation space of $\mathscr{E}$. The basis $\mathbf{e}_1(\mathbf{x}), \mathbf{e}_2(\mathbf{x}), \ldots, \mathbf{e}_n(\mathbf{x})$ is called the *natural basis* of the co-ordinate system $\bar{\mathbf{x}}$ at $\mathbf{x}$, and $\mathbf{e}^1(\mathbf{x}), \mathbf{e}^2(\mathbf{x}), \ldots, \mathbf{e}^n(\mathbf{x})$ is the *reciprocal natural basis* there. As the point $\mathbf{x}$ varies over the domain of $\bar{\mathbf{x}}$, fields of natural bases and their reciprocals are obtained. In general, these bases are not orthonormal. If the co-ordinate surfaces are mutually orthogonal, the co-ordinate curves are normal to the co-ordinate surfaces, and so $\mathbf{e}^k$ is parallel to $\mathbf{e}_k$, but generally the two are not the same. Indeed, the natural basis field is orthonormal only if it is a constant field, in which case the co-ordinates are called *cartesian*. The values of the cartesian co-ordinate fields may be interpreted as distances from a particular set of n mutually orthogonal $(n-1)$-dimensional flats, or as distances measured parallel to a particular set of n mutually orthogonal lines, as we please. (We refer to rectangular rectilinear co-ordinates as "cartesian", but as the baroque savant DESCARTES never used them, in this book we adjust fact to tradition by writing the initial letter minuscule.)

Two other systems are commonly used in three-dimensional space. The *cylindrical co-ordinates* (r, θ, z) of $\mathbf{x}$, are, respectively, the distance of $\mathbf{x}$ from a chosen line called the *axis*, the angle subtended upon a particular plane through that line by a chosen plane through the axis at $\mathbf{x}$, and the distance of $\mathbf{x}$ from a particular plane perpendicular to the axis. Hence

$$\mathbf{e}_r = \frac{\partial \mathbf{x}}{\partial r} = \mathbf{e}^r, \qquad \mathbf{e}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} = r^2 \mathbf{e}^\theta, \qquad \mathbf{e}_z = \frac{\partial \mathbf{x}}{\partial z} = \mathbf{e}^z.$$

The *spherical co-ordinates* (r, θ, φ) are, respectively, the distance of $\mathbf{x}$ from a certain point, an angle between planes through an axis through that particular point, and an angle subtended upon the axis by a line from the particular point to $\mathbf{x}$. Hence

$$\mathbf{e}_r = \frac{\partial \mathbf{x}}{\partial r} = \mathbf{e}^r, \qquad \mathbf{e}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} = r^2 \mathbf{e}^\theta, \qquad \mathbf{e}_\varphi = \frac{\partial \mathbf{x}}{\partial \varphi} = r^2 \sin^2 \theta \, \mathbf{e}^\varphi.$$

Thus far in this section we have used co-ordinate surfaces to define components. The student will recall that components may be defined relative to any basis, and that if $n > 2$ a vector field is not generally normal to any family of surfaces. In particular, an arbitrary field of bases will not generally be the field of natural bases of any co-ordinate system. Components that do not derive from a co-ordinate basis are called *anholonomic*. For some purposes anholonomic components are more convenient than components with respect to a co-ordinate system. For an example, see Exercise IV.21.10.

In works on differential geometry may be found necessary and sufficient conditions that an orthogonal basis field be locally the natural basis of some co-ordinate system.

8. Contravariant, Covariant, and Mixed Components Relative to a Co-ordinate System

The value $\mathbf{v}(\mathbf{x})$ of a vector field at $\mathbf{x}$ is a vector and hence has unique components relative to any basis (above, Section App. IIA.1), and in particular relative to the natural and reciprocal bases of a co-ordinate system $\bar{\mathbf{x}}$. Thus

$$\mathbf{v} = v^k \mathbf{e}_k = v_k \mathbf{e}^k .$$

The scalar fields $v^1, v^2, \ldots, v^n$ are the *contravariant component fields of* $\mathbf{v}$ *relative to the co-ordinate system* $\bar{\mathbf{x}}$; likewise, the fields $v_1, v_2, \ldots, v_k$ are the *covariant component fields* relative to that system. When a particular co-ordinate system is set down for use, we usually speak simply of *contravariant and covariant components*, respectively.

The covariant and contravariant *metric components*, g_{km} and g^{km}, are the scalar fields defined as follows:

$$g_{km} := \mathbf{e}_k \cdot \mathbf{e}_m , \qquad g^{km} := \mathbf{e}^k \cdot \mathbf{e}^m ,$$

and so

$$\mathbf{e}_k = g_{km}\mathbf{e}^m , \qquad \mathbf{e}^k = g^{km}\mathbf{e}_m , \qquad g_{hk}g^{kq} = \delta_h^q .$$

For cartesian co-ordinates

$$g_{km} = \delta_{km} = g^{km},$$

for cylindrical co-ordinates

$$\|g_{km}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \qquad \|g^{km}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

and for spherical ones

$$\|g_{km}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{Vmatrix}, \qquad \|g^{km}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & r^{-2}/\sin^2\theta \end{Vmatrix}.$$

In terms of the metric components, it is easy to relate covariant and contravariant components of one and the same vector field $\mathbf{v}$:

$$v^k = g^{km} v_m, \qquad v_q = g_{qs} v^s.$$

Similar definitions and rules hold for the components of tensor fields, *e.g.*

$$L^{km} = g^{kp} L_p{}^m = g^{pm} L^k{}_p = g^{kp} g^{mq} L_{pq}.$$

Let $\bar{\mathbf{x}}$ and $\tilde{\mathbf{x}}$ be co-ordinate systems. Then the co-ordinates of $\mathbf{x}$ with respect to these two systems are functionally related:

$$\tilde{x}^k(\mathbf{x}) = f^k(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^n),$$
$$\bar{x}^q(\mathbf{x}) = g^q(\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^n).$$

Let $\tilde{\mathbf{e}}_1(\mathbf{x}), \tilde{\mathbf{e}}_2(\mathbf{x}), \ldots, \tilde{\mathbf{e}}_n(\mathbf{x})$ be the natural basis of the co-ordinate system $\tilde{\mathbf{x}}$ at $\mathbf{x}$. From the definition of natural basis it follows that

$$\begin{aligned} \tilde{\mathbf{e}}_m &= \partial_{\tilde{x}^m} \hat{\mathbf{x}}(g^1(\tilde{x}^1, \ldots, \tilde{x}^n), \ldots, g^n(\tilde{x}^1, \ldots, \tilde{x}^n)), \\ &= [\partial_{\bar{x}^k} \hat{\mathbf{x}}(\bar{x}^1, \ldots, \bar{x}^n)] \partial_{\tilde{x}^m} g^k(\tilde{x}^1, \ldots, \tilde{x}^n), \\ &= (\partial_{\tilde{x}^m} g^k) \bar{\mathbf{e}}_k. \end{aligned}$$

Thus, if we set

$$A^k_m = \partial_{\tilde{x}^m} g^k, \qquad \bar{A}^p_q = \partial_{\bar{x}^q} f^p$$

(often denoted by $\partial \bar{x}^k / \partial \tilde{x}^m$ and $\partial \tilde{x}^p / \partial \bar{x}^q$), from the transformation rules in Section App.IIA.1 we may read off the relations between components of various kinds relative to different co-ordinate systems. *E.g.*, if the components of a vector $\mathbf{u}$ with respect to the two systems are distinguished by superimposed bars and tildes, then

$$\tilde{u}_k = \frac{\partial \bar{x}^m}{\partial \tilde{x}^k}\bar{u}_m , \qquad \tilde{u}^k = \frac{\partial \tilde{x}^k}{\partial \bar{x}^m}\bar{u}^m ,$$

and so on for tensors.

These transformation laws for components were used to define vector fields in some of the older literature. *E.g.*, the scalar functions $\bar{A}^p{}_r{}^q{}_s$ and $\tilde{A}^k{}_u{}^h{}_v$ are said to be components, in the co-ordinate systems $\bar{\mathbf{x}}$ and $\tilde{\mathbf{x}}$, of a tensor of order four (contravariant order two and covariant order two) if they are related as follows:

$$\tilde{A}^k{}_u{}^h{}_v = \frac{\partial \tilde{x}^k}{\partial \bar{x}^p}\frac{\partial \bar{x}^r}{\partial \tilde{x}^u}\frac{\partial \tilde{x}^h}{\partial \bar{x}^q}\frac{\partial \bar{x}^s}{\partial \tilde{x}^v}\bar{A}^p{}_r{}^q{}_s$$

the functions on the left-hand side being evaluated at the argument $\tilde{\mathbf{x}}(\mathbf{x})$, and those on the right-hand side at $\bar{\mathbf{x}}(\mathbf{x})$. The other approaches to tensors of order greater than 2 which were mentioned above in Section App.IIA.4 and App.IIA.6 may be extended to fields in a straightforward way.

Whatever be the definitions chosen, there is no doubt that specific calculations are performed most easily by means of the transformation rules. For example, it is obvious that for a cartesian co-ordinate system $g_{km} = \delta_{km} = g^{km}$. The covariant metric components $\tilde{g}_{km}$ in the co-ordinate system $\tilde{\mathbf{x}}$, therefore, are obtained as follows:

$$\tilde{g}_{km} = \frac{\partial x^p}{\partial \tilde{x}^k}\frac{\partial x^q}{\partial \tilde{x}^m}\delta_{pq} = \sum_{p=1}^{n}\frac{\partial x^p}{\partial \tilde{x}^k}\frac{\partial x^p}{\partial \tilde{x}^m},$$

where the cartesian co-ordinates x^p are presumed given as functions of the general co-ordinates $\tilde{x}^k$:

$$x^p = f^p(\tilde{x}^1, \tilde{x}^2, \cdots, \tilde{x}^n), \quad p = 1, 2, \cdots, n.$$

For example, in cylindrical co-ordinates, if we write r, θ, and z, respectively, for $\tilde{x}^1, \tilde{x}^2, \tilde{x}^3$, then

$$x^1 = x = r\cos\theta,$$
$$x^2 = y = r\sin\theta,$$
$$x^3 = z.$$

It is a trivial matter to obtain in this way the matrices $\|\tilde{g}_{km}\|$ and $\|\tilde{g}^{pq}\|$ for cylindrical co-ordinates. Likewise, the components of vectors and tensors relative to any co-ordinate system may be calculated routinely from their components relative to a cartesian system.

Tensors of order greater than 2 occur rarely in this book. A student who does not possess a technique of handling them should be able to follow all developments by simply referring them to components relative to cartesian co-ordinates. Of course, this procedure, while often inelegant, is perfectly rigorous.

9. Physical Components Relative to an Orthogonal Co-ordinate System

The vectors and tensors that occur in physical problems usually are assigned physical dimensions. For example, a velocity has the dimensions of length divided by time. The components of a velocity field with respect to a co-ordinate system do not necessarily have these same dimensions, since the dimensions of the different members of natural basis are not usually all the same. For example, in a cylindrical system $\mathbf{e}^r$ is dimensionless, but $\mathbf{e}_\theta$ has the dimension of length, and $\mathbf{e}^\theta$ has the dimension of reciprocal length. In physical problems it is often desirable to be able to interpret each component of a vector in the same terms as the vector itself, and for this reason *physical components* are used. For an orthogonal co-ordinate system these components are defined unambiguously as being the components with respect to the following orthonormal basis field:

$$\mathbf{i}_k := \frac{\mathbf{e}_k}{|\mathbf{e}_k|} = \frac{\mathbf{e}^k}{|\mathbf{e}^k|}, \qquad k = 1, 2, \dots, n;$$

here $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ is the natural basis field of the co-ordinate system, and $\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^n$ is its reciprocal natural basis field. The orthonormal basis field $\mathbf{i}_1, \mathbf{i}_2, \dots, \mathbf{i}_n$ is everywhere tangent to the co-ordinate curves and normal to the co-ordinate surfaces. Physical components are denoted by indices at middle height, neither subscript nor superscript, thus:

$$v\,r, \qquad v\,\theta, \qquad v\,z.$$

A similar notation is used for tensors, *e.g.*

$$T\,rr, \qquad T\,r\theta, \qquad \text{etc.}$$

In Section App.IIA.15 we have listed some algebraic formulae peculiar to skew tensors over a three-dimensional space. For reference we converted some of them to common special notations including the two vector products, either of which may be denoted by $\times$. Here we adjoin some differential statements in the same context. The symbols $\mathbf{u}$ and $\mathbf{v}$ denote arbitrary vector fields.

$$\operatorname{curl}\mathbf{u} := -(\nabla\mathbf{u})_\times ;$$

in cartesian components,

$$(\operatorname{curl}\mathbf{u})_3 = u_{2,1} - u_{1,2}, \qquad \text{etc.}$$

Also

$$\operatorname{div}(\mathbf{u}\wedge\mathbf{v}) = \operatorname{curl}(\mathbf{u}\times\mathbf{v}).$$

If $\mathbf{S} := \operatorname{skw}\nabla\mathbf{u}$, then

$$2\operatorname{div}\mathbf{S} = \Delta\mathbf{u} - \nabla\operatorname{div}\mathbf{u} = -\operatorname{curl}\operatorname{curl}\mathbf{u}.$$

10. Christoffel Components

The gradients $\mathbf{\Gamma}_{(k)}$ of the natural basis of a co-ordinate system exist and are continuous tensor fields:

$$\mathbf{\Gamma}_{(k)} := \nabla\mathbf{e}_k .$$

The mixed components $\Gamma_p{}^k{}_q$ of these fields, namely

$$\Gamma_p{}^k{}_q := \Gamma_{(p)}{}^k{}_q = \mathbf{e}^k\cdot\mathbf{\Gamma}_{(p)}\mathbf{e}_q ,$$

are the *Christoffel symbols* of the given co-ordinate system. It is possible to prove that

$$\Gamma_p{}^k{}_q = \Gamma_q{}^k{}_p$$

and that

$$\partial_{x^s}\mathbf{e}_r = \Gamma_r{}^k{}_s\mathbf{e}_k .$$

Furthermore, the Christoffel symbols can be calculated as follows from the metric components g^{kp} and g_{qr} of the co-ordinate system:

$$\Gamma_u{}^k{}_v = \tfrac{1}{2} g^{kp}(\partial_{x^v} g_{pu} + \partial_{x^u} g_{pv} - \partial_{x^p} g_{uv}).$$

The notation might suggest that the Christoffel symbols are components of third-order tensors, but, as their name indicates, they are not.

It can be shown that the Christoffel symbols of a co-ordinate system vanish identically if and only if its natural basis field is constant. Such is the case for a cartesian co-ordinate system.

11. Covariant Derivatives, Differential Operators

If $\mathbf{f}$ is a vector field, its gradient $\nabla \mathbf{f}$ is a tensor field. The four usual kinds of components of $\nabla \mathbf{f}$ are called *covariant derivatives* of $\mathbf{f}$. These are defined as components always are:

$$f^k{}_{,m} := \mathbf{e}^k \cdot (\nabla \mathbf{f}) \mathbf{e}_m \,,$$

$$f_{k,m} := \mathbf{e}_k \cdot (\nabla \mathbf{f}) \mathbf{e}_m \,,$$

$$f_k{}^{,m} := \mathbf{e}_k \cdot (\nabla \mathbf{f}) \mathbf{e}^m ,$$

$$f^{k,m} := \mathbf{e}^k \cdot (\nabla \mathbf{f}) \mathbf{e}^m .$$

Each covariant derivative is thus a scalar field.

To calculate the covariant derivatives of $\mathbf{f}$ in terms of the components of $\mathbf{f}$, we note first that

$$\nabla \mathbf{f} = \nabla(f^p \mathbf{e}_p) = \mathbf{e}_p \otimes \nabla f^p + f^p \, \nabla \mathbf{e}_p \,.$$

Hence

$$\begin{aligned} f^k{}_{,m} &= \mathbf{e}^k \cdot (\mathbf{e}_p \otimes \nabla f^p + f^p \boldsymbol{\Gamma}_{(p)}) \mathbf{e}_m \\ &= \partial_{x^m} f^k + f^p \Gamma_p{}^k{}_m \,. \end{aligned}$$

Likewise

$$f_{k,m} = \partial_{x^m} f_k - f_p \Gamma_k{}^p{}_m \,.$$

Each covariant derivative equals the corresponding partial derivative if the co-ordinate system is cartesian; for such equality to hold for all f, the co-ordinate system must be cartesian.

Similar formulae hold for tensors of all orders. In particular, covariant derivatives of all tensors reduce to the corresponding partial derivatives if the co-ordinate system is cartesian.

The values of all differential operators can be calculated in terms of covariant derivatives or Christoffel symbols. For example,

$$
\begin{aligned}
(\operatorname{div}\mathbf{L})^k &= L^{km}{}_{,m}\,, \\
&= \frac{1}{\sqrt{g}}\partial_{x^m}(\sqrt{g}L^{km}) + L^{pm}\Gamma_p{}^k{}_m\,,
\end{aligned}
$$

where $g := \det g_{pq}$.

The easiest way to get expressions in terms of physical components is to derive them first in terms of contravariant or covariant components, which is a simple routine matter, and then convert the results. We record here the physical components of the divergence of a symmetric tensor $\mathbf{L}$ in cylindrical co-ordinates:

$$
\begin{aligned}
(\operatorname{div}\mathbf{L})^{\langle r\rangle} &= \partial_r L^{\langle rr\rangle} + \frac{1}{r}\partial_\theta L^{\langle r\theta\rangle} + \partial_z L^{\langle rz\rangle} + \frac{L^{\langle rr\rangle} - L^{\langle \theta\theta\rangle}}{r}, \\
(\operatorname{div}\mathbf{L})^{\langle \theta\rangle} &= \partial_r L^{\langle r\theta\rangle} + \frac{1}{r}\partial_\theta L^{\langle \theta\theta\rangle} + \partial_z L^{\langle \theta z\rangle} + \frac{2}{r}L^{\langle r\theta\rangle}, \\
(\operatorname{div}\mathbf{L})^{\langle z\rangle} &= \partial_r L^{\langle rz\rangle} + \frac{1}{r}\partial_\theta L^{\langle \theta z\rangle} + \partial_z L^{\langle zz\rangle} + \frac{1}{r}L^{\langle rz\rangle}.
\end{aligned}
$$

In spherical co-ordinates they are

$$
\begin{aligned}
(\operatorname{div}\mathbf{L})^{\langle r\rangle} &= \partial_r L^{\langle rr\rangle} + \frac{1}{r}\partial_\theta L^{\langle r\theta\rangle} + \frac{1}{r\sin\theta}\partial_\varphi L^{\langle r\varphi\rangle} + \frac{1}{r}(2L^{\langle rr\rangle} - L^{\langle \theta\theta\rangle} - L^{\langle \varphi\varphi\rangle} + L^{\langle r\theta\rangle}\cot\theta), \\
(\operatorname{div}\mathbf{L})^{\langle \theta\rangle} &= \partial_r L^{\langle \theta r\rangle} + \frac{1}{r}\partial_\theta L^{\langle \theta\theta\rangle} + \frac{1}{r\sin\theta}\partial_\varphi L^{\langle \theta\varphi\rangle} + \frac{1}{r}[(L^{\langle \theta\theta\rangle} - L^{\langle \varphi\varphi\rangle})\cot\theta + 3L^{\langle r\theta\rangle}], \\
(\operatorname{div}\mathbf{L})^{\langle \varphi\rangle} &= \partial_r L^{\langle \varphi r\rangle} + \frac{1}{r}\partial_\theta L^{\langle \varphi\theta\rangle} + \frac{1}{r\sin\theta}\partial_\varphi L^{\langle \varphi\varphi\rangle} + \frac{1}{r}(3L^{\langle r\varphi\rangle} + 2L^{\langle \theta\varphi\rangle}\cot\theta).
\end{aligned}
$$

For a skew tensor $\mathbf{S}$, in cylindrical co-ordinates,

$$(\operatorname{div}\mathbf{S})^r = \frac{1}{r}\partial_\theta S^{r\theta} + \partial_z S^{rz},$$

$$(\operatorname{div}\mathbf{S})^\theta = \partial_z S^{\theta z} - \partial_r S^{r\theta} - \frac{1}{r}S^{r\theta},$$

$$(\operatorname{div}\mathbf{S})^z = -\left(\partial_r S^{rz} + \frac{1}{r}S^{rz} + \frac{1}{r}S^{\theta z}\right),$$

and in spherical co-ordinates

$$(\operatorname{div}\mathbf{S})^r = \frac{1}{r}S^{r\theta}\cot\theta + \frac{1}{r}\partial_\theta S^{r\theta} + \frac{1}{r\sin\theta}\partial_\varphi S^{r\varphi},$$

$$(\operatorname{div}\mathbf{S})^\theta = -\frac{2}{r}S^{r\theta} - \partial_r S^{r\theta} + \frac{1}{r\sin\theta}\partial_\varphi S^{\theta\varphi},$$

$$(\operatorname{div}\mathbf{S})^\varphi = -\frac{2}{r}S^{r\varphi} - \frac{1}{r}S^{\theta\varphi}\cot\theta - \partial_r S^{r\varphi} - \frac{1}{r}\partial_\theta S^{\theta\varphi}.$$

This Page Intentionally Left Blank

Appendix III

Solutions of the Exercises

Note: Solutions immediate by merely following the directions given in their statements are omitted from this list.

I.2.1 Use $(\mathrm{I}.2\text{-}5)_3$, Axiom B3, and the definition of meet to prove the first implication. Use $(\mathrm{I}.2\text{-}5)_4$ and the definition of join to prove the second.

I.2.2 Adopting the first of the two possible hypotheses, we set

$$\mathscr{P}_1 := \mathscr{B} \wedge \mathscr{C}, \qquad \mathscr{P}_2 := \mathscr{C} \wedge \mathscr{D}, \qquad \mathscr{P}_3 := (\mathscr{B} \wedge \mathscr{C}) \wedge \mathscr{D}.$$

By the definition of meet,

$$\mathscr{P}_3 \prec \mathscr{B}, \qquad \mathscr{P}_3 \prec \mathscr{C}, \qquad \mathscr{P}_3 \prec \mathscr{D}.$$

Thus $\mathscr{P}_3$ is a part of $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{D}$. Now suppose that

$$\mathscr{X} \prec \mathscr{B}, \qquad \mathscr{X} \prec \mathscr{C}, \qquad \mathscr{X} \prec \mathscr{D}.$$

Then, again by the definition of meet,

$$\mathscr{X} \prec \mathscr{P}_1, \qquad \mathscr{X} \prec \mathscr{P}_2, \qquad \mathscr{X} \prec \mathscr{P}_3.$$

Thus any part $\mathscr{X}$ of $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{D}$ is a part of $\mathscr{P}_3$. By the definition of the meet of three bodies, then,

$$\mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D} = \mathscr{P}_3 .$$

From this last, we see that

$$\mathscr{P}_3 \prec \mathscr{B}, \qquad \mathscr{P}_3 \prec \mathscr{C} \wedge \mathscr{D}.$$

Suppose now that

$$\mathscr{Y} \prec \mathscr{B}, \qquad \mathscr{Y} \prec \mathscr{C} \wedge \mathscr{D}.$$

Since $\mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D}$ exists,

$$\mathscr{Y} \prec \mathscr{B} \wedge \mathscr{C} \wedge \mathscr{D} = \mathscr{P}_3 .$$

The definition of meet shows that

$$\mathscr{P}_3 = \mathscr{B} \wedge (\mathscr{C} \wedge \mathscr{D}).$$

A similar proof holds if we assume that $\mathscr{B} \wedge (\mathscr{C} \wedge \mathscr{D})$ exists.

I.2.3 Use $(\text{I.2-8})_1$ and $(\text{I.2-12})_2$.

I.2.4 Use $(\text{I.2-8})_1$, $(\text{I.2-16})_1$, and $(\text{I.2-12})_2$.

I.3.1 Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be bodies of $\mathbf{\Omega}_{\text{O}}$. Then $\mathscr{B}_i = \operatorname{int}\operatorname{clo}\mathscr{B}_i$, $i = 1, 2$. We know also that $\operatorname{int}\operatorname{clo}(\mathscr{B}_1 \cup \mathscr{B}_2)$ belongs to $\mathbf{\Omega}_{\text{O}}$. Consider any $\mathscr{D} \in \mathbf{\Omega}_{\text{O}}$ such that $\mathscr{B}_1 \subset \mathscr{D}$ and $\mathscr{B}_2 \subset \mathscr{D}$; then $\mathscr{D} = \operatorname{int}\operatorname{clo}\mathscr{D}$ and $\mathscr{B}_1 \cup \mathscr{B}_2 \subset \mathscr{D}$, and so $\operatorname{int}\operatorname{clo}(\mathscr{B}_1 \cup \mathscr{B}_2) \subset \mathscr{D}$. Since $\mathscr{D}$ is arbitrary, the conclusion follows.

I.4.1 Use (I.2-38) to write $\mathscr{B} \vee \mathscr{C}$ as the join of separate bodies. Then use Axiom M3, (I.4-3), and Axiom M1 (if necessary).

I.5.1 Substitute (I.5-16) into (I.5-26), then use (I.5-20).

I.5.3 Expand $\mathbf{f}(\mathscr{B} \vee \mathscr{C}, \mathscr{B} \vee \mathscr{C})$ with the aid of Axioms FE2 and FE3, then similarly expand the results.

I.8.1 $(\text{I.8-6})_1$ follows simply by differentiating (I.8-4). To derive $(\text{I.8-6})_2$, use the definition of $\mathbf{M}_{\mathbf{x}_1}$ so as to get $\dot{\mathbf{M}}_{\mathbf{x}_1}$, and adjust the terms.

I.8.2 Write $\mathbf{F}(\mathscr{B}, \mathscr{B}^e) + \mathbf{F}(\mathscr{B}, \mathscr{B})$ explicitly, and use (I.5-26).

I.8.3 In (I.8-6) take $\mathbf{x}_c$ for $\mathbf{x}_1$, and use (I.8-29). The second term in (I.8-30) is the rate of change of rotational momentum with respect to the fixed place $\mathbf{x}_0$ of a mass-point located at the center of mass of $\boldsymbol{\chi}(\mathscr{B}, t)$ and endowed with the linear momentum of $\mathscr{B}$.

I.9.1 Let $\mathscr{U}$ and $\mathscr{V}$ be two n-dimensional inner-product spaces over $\mathscr{R}$. Let $\mathbf{h}$: $\mathscr{U} \to \mathscr{V}$ be an isometry. Let $\mathbf{Q}$: $\mathscr{U} \to \mathscr{V}$ be defined by $\mathbf{Q}(\mathbf{u}) =$

$\mathbf{h}(\mathbf{u}) - \mathbf{h}(\mathbf{0})$ for each $\mathbf{u}$ in $\mathscr{U}$. Then $\mathbf{Q}(\mathbf{0}) = \mathbf{0}$, and we conclude that (i) $|\mathbf{Q}(\mathbf{u})|_{\mathscr{V}} = |\mathbf{u}|_{\mathscr{U}}$ for each $\mathbf{u}$ in $\mathscr{U}$. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathscr{U}$. By multiplying out both sides of $|\mathbf{Q}(\mathbf{u}) - \mathbf{Q}(\mathbf{v})|^2_{\mathscr{V}} = |\mathbf{u} - \mathbf{v}|^2_{\mathscr{U}}$ and using (i) to simplify the resulting expression, we obtain (ii) $\langle \mathbf{Q}(\mathbf{u}), \mathbf{Q}(\mathbf{v})\rangle_{\mathscr{V}} = \langle \mathbf{u}, \mathbf{v}\rangle_{\mathscr{U}}$ for every $\mathbf{u}$ and $\mathbf{v}$ in $\mathscr{U}$. By (i) and (ii), it is easy to show that (iii) $|\mathbf{Q}(\mathbf{u}+\mathbf{v}) - (\mathbf{Q}(\mathbf{u}) + \mathbf{Q}(\mathbf{v}))|^2_{\mathscr{V}} = 0$ and (iv) $|\mathbf{Q}(\lambda\mathbf{u}) - \lambda\mathbf{Q}(\mathbf{u})|^2_{\mathscr{V}} = 0$, for each scalar λ in $\mathscr{R}$ and all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathscr{U}$. It follows from (iii) and (iv) that $\mathbf{Q}$ is linear; (ii) states that $\mathbf{Q}$ is orthogonal. Moreover, by definition $\mathbf{Q}$ satisfies the equation $\mathbf{Q}\mathbf{u} = \mathbf{h}(\mathbf{u}) - \mathbf{h}(\mathbf{0})$ for each $\mathbf{u}$ in $\mathscr{U}$. The uniqueness of such a $\mathbf{Q}$ is easy to prove.

I.9.2

$$
\begin{aligned}
\mathbf{x}^* &= \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\mathbf{x} - \mathbf{x}_0),\\
&= \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\mathbf{x} - \bar{\mathbf{x}}_0(t) + \bar{\mathbf{x}}_0(t) - \mathbf{x}_0),\\
&= \mathbf{x}_0^*(t) + \mathbf{Q}(t)(\bar{\mathbf{x}}_0(t) - \mathbf{x}_0) + \mathbf{Q}(t)(\mathbf{x} - \bar{\mathbf{x}}_0(t)),\\
&= \bar{\mathbf{x}}_0^*(t) + \mathbf{Q}(t)(\mathbf{x} - \bar{\mathbf{x}}_0(t)),
\end{aligned}
$$

say. Since a transformation of this kind, along with $t^* = t + a$, preserves the metrics in $\mathscr{E}$ and $\mathscr{R}$, it defines a change of frame. To prove the group property, use the fact that the orthogonal tensors form a group.

I.9.3 Differentiate $\mathbf{Z} = \mathbf{Y}\mathbf{Y}^{\mathsf{T}}$ and use (I.9-16) and (I.9-15). A solution of (I.9-17) is furnished by $\mathbf{Z}(t) = \mathbf{1}$; by uniqueness it is the only solution satisfying $\mathbf{Z}(t_0) = \mathbf{1}$.

I.9.4 $\mathbf{A}^*$ is formed from the tensor $\mathbf{Q}^*$ that enters the inverse of (I.9-4). Since $\mathbf{Q}^* = \mathbf{Q}^{\mathsf{T}}$, (I.9-18) follows. To derive (I.9-19), write out the equations for the three changes of frame. $\mathbf{A}_3 = \dot{\mathbf{Q}}_3\mathbf{Q}_3^{\mathsf{T}}$; simplification by $(\text{I.9-19})_1$ yields $(\text{I.9-19})_2$. When $\oint_3$ and $\oint_2$ coincide at some instant, $\mathbf{Q}_2 = \mathbf{Q}_2^{\mathsf{T}} = \mathbf{1}$ at that instant.

I.9.5 If $\mathbf{e}$ is a unit vector on the axis of rotation, $\mathbf{Q}\mathbf{e} = \pm\mathbf{e}$. Hence $\dot{\mathbf{Q}}\mathbf{e} + \mathbf{Q}\dot{\mathbf{e}} = \pm\dot{\mathbf{e}}$; that is,

$$\mathbf{A}\mathbf{e} + (\mathbf{Q} \mp \mathbf{1})\dot{\mathbf{e}} = \mathbf{0}.$$

Thus if $\dot{\mathbf{e}} = \mathbf{0}$, it follows that $\mathbf{A}\mathbf{e} = \mathbf{0}$. The relation between ω and $\dot{\theta}$ can be proved by use of an explicit representation of the orthonormal components of rotations over a three-dimensional space:

$$[\mathbf{Q}] = \begin{Vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

where θ is the angle of rotation and the basis vector $\mathbf{e}_3$ lies in the axis of rotation.

I.9.6

$$\begin{aligned}\operatorname{tr}[\dot{\mathbf{R}}\mathbf{R}^{\mathsf{T}}(\mathbf{R}-\mathbf{R}^{\mathsf{T}})] &= \operatorname{tr}[\dot{\mathbf{R}}+(\dot{\mathbf{R}}\mathbf{R}^{\mathsf{T}})^{\mathsf{T}}\mathbf{R}^{\mathsf{T}}],\\ &= 2\operatorname{tr}\dot{\mathbf{R}}.\end{aligned}$$

In a space of 3 dimensions

$$\operatorname{tr}\dot{\mathbf{R}} = -2(\sin\theta)\dot{\theta}.$$

Also, by appeal to the explicit representation given in Section App. IIA.14,

$$\begin{aligned}\operatorname{tr}[\dot{\mathbf{R}}\mathbf{R}^{\mathsf{T}}(\mathbf{R}-\mathbf{R}^{\mathsf{T}})] &= -2(\sin\theta)A^{KL}\epsilon_{LKP}e^{P},\\ &= -4(\sin\theta)\boldsymbol{\omega}\cdot\mathbf{e}.\end{aligned}$$

If $\theta = 0$, no relation between $\boldsymbol{\omega}\cdot\mathbf{e}$ and $\dot{\theta}$ can hold, because $\mathbf{e}$ can be any unit vector. If $\theta = \pi$, the foregoing argument delivers nothing, but $\boldsymbol{\omega}\cdot\mathbf{e}$ is a continuous function of $\mathbf{e}$, and so we may infer (I.9-20) by a passage to the limit, using the conclusion established for values of θ near π.

I.10.1 The remark just preceding the exercise solves half of it, for in any frame that gives rise to a spin $\bar{\mathbf{W}}$ having the same axis as $\mathbf{W}$ at each t the points on that axis will maintain their mutual distances. Conversely, suppose that

$$\bar{\mathbf{c}} + \bar{\mathbf{W}}(\mathbf{x}-\bar{\mathbf{x}}_0) = \mathbf{c} + \mathbf{W}(\mathbf{x}-\mathbf{x}_0) \qquad \forall\mathbf{x}\in\chi(\mathscr{B}, t).$$

Then

$$(\bar{\mathbf{W}}-\mathbf{W})(\mathbf{x}_0-\mathbf{x}) = \bar{\mathbf{W}}(\mathbf{x}_0-\bar{\mathbf{x}}_0) + \mathbf{c} - \bar{\mathbf{c}}.$$

Thus $\bar{\mathbf{W}}-\mathbf{W}$ is a skew tensor such that

$$(\bar{\mathbf{W}}-\mathbf{W})(\mathbf{x}-\mathbf{y}) = \mathbf{0} \qquad \text{if} \qquad \mathbf{x}\in\chi(\mathscr{B}, t) \qquad \text{and} \qquad \mathbf{y}\in\chi(\mathscr{B}, t).$$

For any given $\mathbf{x}$ and $\mathbf{y}$ there are infinitely many skew tensors $\mathbf{S}$ such that $\mathbf{S}(\mathbf{x}-\mathbf{y}) = \mathbf{0}$. Such $\mathbf{S}$ belong in common to all $\mathbf{x}$ and $\mathbf{y}$ that lie upon the same straight line. If $\chi(\mathscr{B}, t)$ contains three not collinear places, $\bar{\mathbf{W}} - \mathbf{W}$ is a skew tensor whose nullspace contains two distinct straight lines; since the nullspace of a skew tensor other than $\mathbf{0}$ is 1-dimensional, $\bar{\mathbf{W}} - \mathbf{W} = \mathbf{0}$.

I.10.2 (I.10-1)$_2$ shows that $\dot{\mathbf{x}}_0 = \mathbf{c}$. Therefore $\dot{\mathbf{p}}_i = \mathbf{W}\mathbf{p}_i$, $i = 1, 2$. Compute $\frac{d}{dt}(\mathbf{p}_1 \cdot \mathbf{p}_2)$, and use $\mathbf{W} = -\mathbf{W}^{\mathsf{T}}$.

I.10.4 Use (I.8-2)$_1$, (I.10-1)$_2$, (I.10-5), and the fact that $\mathbf{c}$ is a function of t only. To obtain (I.10-8)$_2$, note that $\mathbf{Q}(\mathbf{a} \wedge \mathbf{b})\mathbf{Q}^{\mathsf{T}} = \mathbf{Qa} \wedge \mathbf{Qb}$, and use (I.10-3) and (I.10-7). To obtain (I.10-9), use Guo's formula in Section App.IIA.12.

I.10.5 Use $\mathbf{W} = -\mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q}$ and (I.9-15).

I.10.7 If $\dot{\mathbf{e}} = \mathbf{0}$, then $\mathbf{W}\mathbf{e} = \mathbf{0}$ if and only if $(\dot{\mathbf{W}} + \mathbf{W}^2)\mathbf{e} = \mathbf{0}$.

I.10.8 Use (I.8-3), (I.10-1), and (I.10-4) to get (I.10-15). If $\mathbf{x}_0(t)$ is chosen as directed, then $\dot{\mathbf{x}}_0(t) = \mathbf{c}$, and $\bar{\mathbf{p}} = \mathbf{0}$. (I.10-16) follows with the aid of a little manipulation in the third term on the right-hand side of (I.10-15).

I.11.1 Let $\mathbf{T}_T$: $\mathscr{V}_T \to \mathscr{V}_T$ be the linear transformation in question. Then $\mathbf{T} = (D\oint_T)\mathbf{T}_T(D\oint_T)^{-1}$, and $\mathbf{T}^* = (D\oint_T^*)\mathbf{T}_T(D\oint_T^*)^{-1}$. Thence

$$
\begin{aligned}
\mathbf{T}^* &= (D\oint_T^*)(D\oint_T)^{-1}\mathbf{T}(D\oint_T)(D\oint_T^*)^{-1}, \\
&= D(\oint_T^* \circ \oint_T^{-1})\mathbf{T}(D(\oint_T \circ \oint_T^{*-1})) = \mathbf{QTQ}^{\mathsf{T}}.
\end{aligned}
$$

I.11.2 Since under galilean transformations $\ddot{\mathbf{x}}^* = \mathbf{Q}\ddot{\mathbf{x}}$ and $M^* = M$, (I.8-5)$_1$ shows that $\dot{\mathbf{m}}^* = \mathbf{Qm}$.

I.11.3 The first statement follows at once from the chain rule. The fourth statement is proved as follows from (I.11-2):

$$
\begin{aligned}
\det \mathbf{T}^* &= \det(\mathbf{QTQ}^{\mathsf{T}}) = (\det \mathbf{Q})(\det \mathbf{T})(\det \mathbf{Q}^{\mathsf{T}}), \\
&= \det \mathbf{T}.
\end{aligned}
$$

This being so, $\det(\mathbf{T} - r\mathbf{1})$ is frame-indifferent, no matter what be the number r. Therefore, $\mathbf{T}^*$ and $\mathbf{T}$ have the same latent roots. Consequently they have the same trace and the same proper numbers. If $\mathbf{e}$ is a proper vector corresponding to the proper number r, then

$$
\mathbf{Te} = r\mathbf{e},
$$

so that

$$
(\mathbf{Q}^{\mathsf{T}}\mathbf{T}^*\mathbf{Q})\mathbf{e} = r\mathbf{e},
$$

and hence

$$\mathbf{T}^*(\mathbf{Qe}) = r(\mathbf{Qe}).$$

Thus $\mathbf{Qe}$ is the corresponding proper vector of $\mathbf{T}^*$.

I.11.4 A given smooth surface can be embedded in a smooth family of surfaces f = const., where f is a frame-indifferent scalar. Then, by conclusions of Exercise I.11.3, both ∇f and $|\nabla f|$ are frame-indifferent, and $\mathbf{n} = \nabla f/|\nabla f|$. A better proof can be constructed by writing an equation for a single surface as

$$\mathbf{x} - \mathbf{x}_0 = \mathbf{g}(a, b),$$

where a and b are parameters. For each a and b, the left-hand side is a frame-indifferent vector, and so the vector-valued function $\mathbf{f}$ is frame-indifferent. Accordingly, $\partial_a\mathbf{g}$ and $\partial_b\mathbf{g}$ are frame-indifferent. They span the tangent plane at (a, b). The line normal to the tangent plane contains exactly two unit vectors. Their construction as above shows that both are frame-indifferent.

I.12.1 Use Axiom A2 and conclusions from Exercise I.11.3.

I.12.2 In (I.9-13) suppose that $\dot{\mathbf{x}}^* = \mathbf{0}$, $\dot{\mathbf{Q}} = \mathbf{0}$. Then $\dot{\chi}^* = \mathbf{Q}\chi$, and so $W^* = W$ implies that

$$\int_{\mathscr{B}} \mathbf{Q}\dot{\chi} \cdot d\mathbf{f}^*_{\mathscr{B}}\mathrm{e} = \int_{\mathscr{B}} \dot{\chi} \cdot d\mathbf{f}_{\mathscr{B}}\mathrm{e},$$

whence

$$\mathbf{Q}^{\mathsf{T}}\, d\mathbf{f}^*_{\mathscr{B}}\mathrm{e} = d\mathbf{f}_{\mathscr{B}}\mathrm{e}.$$

I.13.1 We note that (I.5-22) holds as long as the system of forces is balanced. Since the axioms of inertia as applied to analytical dynamics respect the requirement that the forces be balanced, they do not alter the requirement (I.5-22).

I.13.2 Note that $(\mathbf{f} \wedge \mathbf{g})^{\cdot} = \mathbf{0}$. The spectral decomposition of $\mathbf{E}^*_{\mathbf{x}_0^*}$ is

$$\mathbf{E}^*_{\mathbf{x}_0^*} = E_1\mathbf{e} \otimes \mathbf{e} + E_2\mathbf{f} \otimes \mathbf{f} + E_3\mathbf{g} \otimes \mathbf{g}.$$

It is easy to show that

$$\mathbf{E}^*_{\mathbf{x}_0^*}\mathbf{A}^2 - \mathbf{A}^2\mathbf{E}^*_{\mathbf{x}_0^*} = \mathbf{0},$$

and

$$\mathbf{E}^*_{\mathbf{x}_0*}\dot{\mathbf{A}} + \dot{\mathbf{A}}\mathbf{E}^*_{\mathbf{x}_0*} = \dot{\omega}(E_2 + E_3)\mathbf{f} \wedge \mathbf{g}.$$

Thus (I.13-22) reduces to (I.13-24).

I.14.1 Invariance under translation is equivalent to a representation

$$U_{qk}(\mathbf{x}_q, \mathbf{x}_k) = W_{qk}(\mathbf{x}_k - \mathbf{x}_q),$$

W_{qk} being a frame indifferent function of vectors. A theorem of CAUCHY [NFTM, p. 29] tells us that

$$W_{qk}(\mathbf{v}^*) = W_{qk}(\mathbf{v})$$

for all $\mathbf{v}$ and all changes of frame if and only if $W_{qk}(\mathbf{v}) = V_{qk}(\mathbf{v}\cdot\mathbf{v})$. Calculation of $\mathbf{f}_{qk}$ and V_{qk} yields

$$\mathbf{f}_{qk} = \left.\frac{dV_{qk}(a)}{da}\right|_{a=|\mathbf{x}_q-\mathbf{x}_k|^2}.$$

The conclusion about $\mathbf{f}_k^0$ follows by a similar argument.

I.15.1 The heating Q obeys the identity (I.5-2).

II.1.1 If $\partial\mathscr{C}$ is of class C^1, by use of the divergence theorem it follows from the definition of perimeter that

$$\operatorname{per}(\mathscr{C}) \leqq A(\partial\mathscr{C}).$$

If $\partial\mathscr{C}$ is of class C^2, one can easily arrive at the inequality

$$\operatorname{per}(\mathscr{C}) \geqq A(\partial\mathscr{C})$$

by using the divergence theorem to compute $\int \operatorname{div}\mathbf{g}\,dV$ for $\mathbf{g} \in C_0^1(\mathscr{E}, \mathscr{V})$ such that $|\mathbf{g}| \leqq 1$ and $\mathbf{g}|_{\partial\mathscr{C}} = \mathbf{n}_{\mathscr{C}}$. Comparison of the two inequalities delivers the desired conclusion if $\partial\mathscr{C}$ is of class C^2. If $\partial\mathscr{C}$ is merely of class C^1, the same conclusion follows from a rather more delicate argument which is sketched on p. 157 of the book by VOL'PERT & HUDJAEV, cited in Footnote 1 on p. 88.

II.5.1 From the definition of a determinant, or by use of the characteristic polynomial, it follows that for a tensor $\mathbf{A}$

$$\det(\mathbf{1} + \mathbf{A}) = 1 + \operatorname{tr}\mathbf{A} + o(\mathbf{A}) \qquad \text{as } \mathbf{A} \to \mathbf{0}.$$

Thus if $\mathbf{L}$ is a fixed invertible tensor

$$\begin{aligned}\det(\mathbf{L}+\mathbf{U}) &= \det\mathbf{L}\det(\mathbf{1}+\mathbf{L}^{-1}\mathbf{U})\\ &= \det\mathbf{L}[1+\operatorname{tr}(\mathbf{U}\mathbf{L}^{-1})]+o(\mathbf{U}) \qquad \text{as } \mathbf{U}\to\mathbf{0}.\end{aligned}$$

If an invertible tensor $\mathbf{F}$ is a differentiable function of a parameter, we may put $\mathbf{F}$ for $\mathbf{L}$ and $\epsilon\dot{\mathbf{F}}$ for $\mathbf{U}$ and so obtain

$$\begin{aligned}\det(\mathbf{F}+\epsilon\dot{\mathbf{F}})-\det\mathbf{F} &= \det\mathbf{F}(1+\epsilon\operatorname{tr}\dot{\mathbf{F}}\mathbf{F}^{-1})-\det\mathbf{F}+o(\epsilon)\\ &= \epsilon(\det\mathbf{F})\operatorname{tr}\dot{\mathbf{F}}\mathbf{F}^{-1}+o(\epsilon).\end{aligned}$$

Divide by ϵ and then let $\epsilon\to 0$ to conclude that

$$(\det\mathbf{F})^{\cdot} = (\det\mathbf{F})\operatorname{tr}(\dot{\mathbf{F}}\mathbf{F}^{-1}).$$

Interpret $\mathbf{F}$ as being the transplacement gradient and use the chain rule to show that $\dot{\mathbf{F}}\mathbf{F}^{-1} = \nabla\dot{\mathbf{x}}$, or simply use (II.11-5) and (II.11-7).

II.5.2 (An easier problem of this kind is given below as Exercise II.6.3.) Consider the linear partial differential equation

$$\sum_{i=0}^{n} P_i\,\partial_{x_i} Z = R, \tag{L}$$

where $P_0, P_1, \ldots, P_n$, R are given functions of $x_0, x_1, \ldots, x_n$. The *characteristics* of (L) are the integral curves of the system

$$\frac{dx_0}{P_0} = \frac{dx_1}{P_1} = \cdots = \frac{dZ}{R}. \tag{C}$$

A *characteristic integral* is a function $f_i(x_0, x_1, \ldots, x_n, Z)$ such that $f_i = \text{const.}$ on every characteristic curve. The formal statement of LAGRANGE's theorem is that if $f_1, \ldots, f_n$ are any n functionally independent, characteristic integrals of (C), then the general solution of (L) is

$$F(f_1, f_2, \ldots, f_n, Z) = 0.$$

To treat (II.5.7) in n dimensions, let $x_0 = t$, write $\mathbf{x}$ for $(x_1, x_2, \ldots, x_n)$, $Z := \log\rho$. Then $P_0 = 1$, $P_i = \dot{x}_i$, and, by (II.5-6),

$R = -\dot{J}/J$. Hence n members of (C) can be written in the form $d\mathbf{x} = \dot{\mathbf{x}}\,dt$, and so n families of characteristic curves are provided by the path-lines of the substantial points. Thus χ_κ^{-1} denotes n characteristic integrals. An $(n+1)^{\text{st}}$ integral can be obtained by integrating $dt = dZ/R = -d\log\rho/d\log J$, the resulting integral being ρJ. Thus the general solution of (II.5-7) is

$$F(\chi_\kappa^{-1}, \rho J) = 0,$$

and this is (II.5-4).

Note: The method of characteristics for linear partial differential equations of first order was invented by LAGRANGE on the basis of this example and the one in Exercise II.6.3, both of these having arisen in hydrodynamics. The trivial generalization of the particular case (II.5-6) from 3 dimensions to n was obtained by LIOUVILLE; it is the only one of the several statements physicists call "Liouville's theorem in statistical mechanics" that has any connection with LIOUVILLE.

A rigorous treatment of LAGRANGE's theory in the large is intricate. Most modern books on partial-differential equations omit it. A clear and precise treatment of the local theory may be found in Chapter 2 of P. R. GARABEDIAN's *Partial Differential Equations*, New York, John Wiley & Sons, 1964, reprinted New York, Chelsea Publications, 1986. A simple treatment of characteristics is given by C.-C. WANG in the appendix to his *Mathematical Principles of Mechanics and Electromagnetism, Part A*, N.Y. & London, Plenum, 1979.

II.5.3 By the theorem of integral calculus used to derive (II.2-6), the volume of $\chi_1(\mathscr{P}, t)$ is given by

$$\int_{\chi_1(\mathscr{P},t)} dV = \int_{\chi_2(\mathscr{P},t)} J\,dV.$$

The condition of isochoric motion is therefore locally equivalent to $J = 1$. To complete the exercise, use (II.5-6) and (II.5-7).

II.5.4 For a plane motion (II.5-8) becomes

$$\partial_x \dot{x} + \partial_y \dot{y} = 0,$$

in which x, y are cartesian co-ordinates and $\dot{x}$, $\dot{y}$ are the corresponding components of the velocity field. This is a necessary and sufficient condition that in each simply connected region there be a function q such that

$$\dot{x} = -\partial_y q, \qquad \dot{y} = \partial_x q,$$

and this is (II.5-11). For the extension to multiply connected domains see CFT, Section 161. Clearly $\dot{\mathbf{x}}\cdot\nabla q = 0$, and so the stream lines are normal to the normals of the curves $q(\cdot, t) = \text{const.}$

II.5.5

$$n_k\, dA(\mathbf{x}) = \epsilon_{kpq}\, dx^p\, dx^q = \epsilon_{kpq} F^p{}_\alpha F^q{}_\beta\, dX^\alpha\, dX^\beta .$$

(An interpretation for this transformation law is given in Section II.13.) The conclusion follows by comparing both sides of (II.5-12).

II.6.1 As (II.6-11) suggests, take $\mathcal{U}/\rho$ for Ψ in (II.6-9), then use (II.5-7), (II.6-3), and the divergence theorem. The value of the left-hand side of (II.6-10) is the time derivative of $\int \mathcal{U}\, dV$ for a given part $\mathscr{P}$ of $\mathscr{B}$; the operation denoted by a prime is the time derivative of $\int \mathcal{U}\, dV$ obtained if, neglecting the motion χ, we confuse $\mathscr{P}$ with its present shape $\chi(\mathscr{P}, t)$. The difference between these is explained and evaluated by the third term, which gives the rate of increase of $\int \mathcal{U}\, dV$ for $\mathscr{P}$ effected by the motion of substantial points out of or into the present shape of $\mathscr{P}$. To complete the exercise, refer to the definitions (I.8-1) and (I.8-2), and take for $\mathcal{U}$ first $\rho\dot{\mathbf{x}}$ and then $(\mathbf{x} - \mathbf{x}_0) \wedge \rho\dot{\mathbf{x}}$.

II.6.2 For a moving surface $\mathscr{S}$, choose a particular parametric representation:

$$\mathbf{x} = \mathbf{g}(\mathbf{A}, t),$$

and think of $\mathbf{A}$ as being attached permanently to a point on $\mathscr{S}$ as it progresses. Then the velocity $\mathbf{u}$ at $\mathbf{A}$ is given by

$$\mathbf{u} := \partial_t \mathbf{g}.$$

Of course the field $\mathbf{u}$ so defined on $\mathscr{S}$ depends upon the particular parametrization used to describe $\mathscr{S}$. Now suppose the parameter $\mathbf{A}$ to have been eliminated, so that an equation for $\mathscr{S}$ is

$$f(\mathbf{x}, t) = 0.$$

All the infinitely many different parametrizations of $\mathscr{S}$ will lead to one and the same set of points satisfying a relation of this kind, and this relation characterizes $\mathscr{S}$ over an interval of time:

$$f(\mathbf{g}(\mathbf{A}, t), t) = 0 \qquad (1)$$

for each fixed $\mathbf{A}$ in any parametrization. If $h(\mathbf{x}, t) = 0$ is another equation for $\mathscr{S}$, then h is an invertible function of f. Differentiation of (1) yields

$$f' + (\operatorname{grad} f)\cdot\mathbf{u} = 0. \tag{2}$$

Now the unit normal to $\mathscr{S}$ in the direction of increasing f is given by

$$\mathbf{n} = \frac{\operatorname{grad} f}{|\operatorname{grad} f|}.$$

Therefore (2) asserts that

$$\mathbf{u}\cdot\mathbf{n} = -\frac{f'}{|\operatorname{grad} f|} = -\frac{h'}{|\operatorname{grad} h|},$$

h being any differentiable function of f. Because the right-hand side is independent of the parametrization, so is the left-hand side. Thus what we have defined as the speed of displacement S_n is in fact the common normal speed of advance of all possible assignments of velocity to points on $\mathscr{S}$.

II.6.3 For the method of characteristics, see Exercise II.5.2. In the present instance $R = 0$, and f is the unknown function.

II.6.4 Note that $g(\mathbf{x}, t) = g(\chi_\kappa(\mathbf{X}, t), t) := G(\mathbf{X}, t) = 0$, and

$$\mathbf{n}_\kappa = \frac{\operatorname{Grad} G(\mathbf{X}, t)}{|\operatorname{Grad} G(\mathbf{X}, t)|} = \frac{\mathbf{F}^{\mathsf{T}} \operatorname{grad} g}{|\operatorname{Grad} G(\mathbf{X}, t)|}.$$

A little simplification gives (II.6-21). If $\mathscr{S}_\kappa$ is not a substantial surface, then at different times different substantial points will lie upon it. Of course (II.6-22) is merely an application of (II.6-16). To get (II.6-23), use $(\text{II.6-3})_1$ and (II.6-16).

II.9.1 The common proof starts from the assumption $\mathbf{F}\mathbf{F}^{\mathsf{T}} = \mathbf{1}$ and by differentiating it and using the fact that $F^k_{\alpha,\beta} = F^k_{\beta,\alpha}$ concludes that $\mathbf{F} = \text{const.}$ Gurtin & Williams have found an elegant proof that does not require $\mathbf{F}$ to be differentiable. Let $\mathbf{f}$ be a differentiable function of place $\mathbf{z}$ in some open, connected set $\mathscr{T}$ on which $(\nabla\mathbf{f})(\nabla\mathbf{f})^{\mathsf{T}} = \mathbf{1}$. Then $\det \nabla\mathbf{f} = \pm 1$. If $\mathbf{x}_0 \in \mathscr{T}$, there is an open ball $\mathscr{S}$ such that $\mathbf{x}_0 \in \mathscr{S} \subset \mathscr{T}$ and that $\mathbf{f}$ is invertible in $\mathscr{S}$. If $\mathbf{x} \in \mathscr{S}$ and $\mathbf{y} \in \mathscr{S}$, let $\mathscr{C}$ be the line segment from $\mathbf{y}$ to $\mathbf{x}$. Then

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) = \int_{\mathscr{C}} \nabla\mathbf{f}(\mathbf{z})\, d\mathbf{z}.$$

Because $|\mathbf{Qu}| = |\mathbf{u}|$ for any orthogonal tensor $\mathbf{Q}$ and any vector $\mathbf{u}$,

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leqq \int_{\mathscr{C}} |\nabla\mathbf{f}(\mathbf{z})\,d\mathbf{z}| = \int_{\mathscr{C}} ds = |\mathbf{x} - \mathbf{y}|.$$

Just the same argument applies to $\mathbf{f}^{-1}$:

$$|\mathbf{x} - \mathbf{y}| = |\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) - \mathbf{f}^{-1}(\mathbf{f}(\mathbf{y}))| \leqq |\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})|.$$

Comparison of these two inequalities yields

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|.$$

Therefore $\mathbf{f}$ preserves distances in $\mathscr{S}$. Since $\mathscr{T}$ is connected, the assertion follows. If $\bar{\mathbf{U}} = \mathbf{U}$, then $\operatorname{grad}(\bar{\boldsymbol{\chi}}_\kappa \circ \boldsymbol{\chi}_\kappa^{-1}) = \bar{\mathbf{F}}\mathbf{F}^{-1} = \bar{\mathbf{R}}\mathbf{R}^{\mathsf{T}}$, which must be constant in virtue of the preceding.

II.9.2 $\mathbf{C}(\tau) = \mathbf{F}^{\mathsf{T}}(\tau)\mathbf{F}(\tau)$. Use (II.8-7), and simplify.

II.9.3 The principal stretches v are the roots of $\det(\mathbf{B} - v^2\mathbf{1}) = 0$. $\mathbf{B} = \mathbf{RCR}^{\mathsf{T}}$, and

$$[\mathbf{R}] = \begin{Vmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{Vmatrix}$$

since the principal axes are rotated about the x_3-axis.

II.9.4

$$\|g_{\alpha\beta}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & R^2 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \qquad \|g^{\alpha\beta}\| = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & R^{-2} & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \qquad \text{etc.}$$

II.9.5 Calculate the physical components of $\mathbf{B}$ by employing

$$B\langle km\rangle = B^{km}\sqrt{g_{kk}g_{mm}} \qquad \text{(no summation)},$$

and compare with $(\text{II}.9\text{-}13)_1$.

II.9.6 Use (II.6-5) and $(\text{II}.9\text{-}5)_4$ to get (II.9-19).

II.11.3 EULER proved the statement by first differentiating the component equations $\dot{x}_{k,m} + \dot{x}_{m,k} = 0$. The elegant proof of GURTIN & WILLIAMS does not require that $\mathbf{G}$ be differentiable. Let the notations be as in

Exercise II.9.1. Then

$$[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})]\cdot[\mathbf{x} - \mathbf{y}] = \int_{\mathscr{C}} [\mathbf{x} - \mathbf{y}]\cdot\nabla\mathbf{f}(\mathbf{z})\, d\mathbf{z}.$$

Since $\mathscr{C}$ is a straight line, $d\mathbf{z}$ is parallel to $\mathbf{x} - \mathbf{y}$, and so the integrand is 0 if $\nabla\mathbf{f}$ is skew. Therefore

$$[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})]\cdot[\mathbf{x} - \mathbf{y}] = 0. \tag{A}$$

This condition is equivalent to (I.10.1) in the present notation:

$$\mathbf{f} = \mathbf{c} + \mathbf{W}(\mathbf{x} - \mathbf{x}_0), \qquad \mathbf{W}^{\mathsf{T}} = -\mathbf{W} = \text{const.} \tag{B}$$

Indeed, that (B) $\Rightarrow$ (A) is immediate. Conversely, by differentiating (A) with respect to $\mathbf{x}$ we obtain

$$\nabla\mathbf{f}(\mathbf{x})^{\mathsf{T}}(\mathbf{x} - \mathbf{y}) + \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) = \mathbf{0}.$$

Differentiation with respect to $\mathbf{y}$ yields

$$-\nabla\mathbf{f}(\mathbf{x})^{\mathsf{T}} - \nabla\mathbf{f}(\mathbf{y}) = \mathbf{0}.$$

Thus $\nabla\mathbf{f}$ is both constant and skew.

II.11.5 Show first that $\operatorname{div}\ddot{\mathbf{x}} = \dot{E} + \operatorname{tr}\mathbf{G}^2$.

II.11.6 The preceding exercise shows that in a rigid motion $\operatorname{div}\ddot{\mathbf{x}} = -|\mathbf{W}|^2$. Use (II.11-11) and (II.11-22) to obtain (II.11-23), and then use Section App.IIA.15.

II.11.7 $\mathbf{G} = \partial_u\mathbf{F}(u)\mathbf{F}^{-1}(t)|_{u=t}$. Use the polar decompositions of $\mathbf{F}(u)$ and $\mathbf{F}(t)$, and then carry out the indicated differentiation to obtain

$$\mathbf{G} = \dot{\mathbf{R}}\mathbf{R}^{-1} + \mathbf{R}\dot{\mathbf{U}}\mathbf{U}^{-1}\mathbf{R}^{-1}.$$

This equation can be written as

$$\mathbf{D} + \mathbf{W} = \dot{\mathbf{R}}\mathbf{R}^{-1} + \tfrac{1}{2}\mathbf{R}(\dot{\mathbf{U}}\mathbf{U}^{-1} - \mathbf{U}^{-1}\dot{\mathbf{U}})\mathbf{R}^{\mathsf{T}}$$
$$+ \tfrac{1}{2}\mathbf{R}(\dot{\mathbf{U}}\mathbf{U}^{-1} + \mathbf{U}^{-1}\dot{\mathbf{U}})\mathbf{R}^{\mathsf{T}}.$$

Use uniqueness of the additive decomposition of a tensor into symmetric and skew parts to get $(\text{II.11-26})_{2,3}$. To get $(\text{II.11-26})_1$, start from (II.11-26)$_3$, and use $\mathbf{C} = \mathbf{U}^2$ and the polar decomposition theorem. (In fact, (II.11-26)$_1$ is easy to derive directly, but from it as a starting point there seems to be no obvious way to reach (II.11-26)$_2$.) The last relation follows from $\mathbf{B} = \mathbf{F}\mathbf{F}^{\mathsf{T}}$ and $\dot{\mathbf{F}}|_{\mathbf{F}=\mathbf{1}} = \mathbf{G}$.

II.11.8 $\dot{\mathbf{G}} = (\dot{\mathbf{F}}\mathbf{F}^{-1})^{\cdot} = \ddot{\mathbf{F}}\mathbf{F}^{-1} + \dot{\mathbf{F}}(\mathbf{F}^{-1})^{\cdot}$, $(\mathbf{F}\mathbf{F}^{-1})^{\cdot} = \mathbf{0} = \dot{\mathbf{F}}(\mathbf{F}^{-1}) + \mathbf{F}(\mathbf{F}^{-1})^{\cdot}$; hence $(\mathbf{F}^{-1})^{\cdot} = \mathbf{F}^{-1}\dot{\mathbf{F}}\mathbf{F}^{-1}$.

II.11.9 To get (II.11-32), use Leibniz's rule to differentiate $\mathbf{F}_t(\tau)^{\mathsf{T}}\mathbf{F}_t(\tau)$. To get (II.11-33), first prove that

$$\overset{(n)}{\mathbf{C}} = \mathbf{F}^{\mathsf{T}}\mathbf{A}_n\mathbf{F}.$$

(A prescription for proving this formula is given in the text of Section II.14, where it is listed as (II.14-16).) Hence

$$\begin{aligned}\overset{(n+1)}{\mathbf{C}} &= \mathbf{F}^{\mathsf{T}}\mathbf{A}_{n+1}\mathbf{F},\\ &= \mathbf{F}^{\mathsf{T}}\dot{\mathbf{A}}_n\mathbf{F} + \dot{\mathbf{F}}^{\mathsf{T}}\mathbf{A}_n\mathbf{F} + \mathbf{F}^{\mathsf{T}}\mathbf{A}_n\dot{\mathbf{F}}.\end{aligned}$$

Now use (II.11-5).

II.11.10 A formula for the derivative of the determinant of an invertible tensor is given in Exercise II.5.1. Differentiating it yields

$$\begin{aligned}(\det\mathbf{L})^{\cdot\cdot} &= (\det\mathbf{L})\operatorname{tr}[\ddot{\mathbf{L}}\mathbf{L}^{-1} - (\dot{\mathbf{L}}\mathbf{L}^{-1})^2] + (\det\mathbf{L})^{\cdot}\operatorname{tr}(\dot{\mathbf{L}}\mathbf{L}^{-1}),\\ (\det\mathbf{L})^{\cdot\cdot\cdot} &= (\det\mathbf{L})\operatorname{tr}[\dddot{\mathbf{L}}\mathbf{L}^{-1} - 3\ddot{\mathbf{L}}\mathbf{L}^{-1}\dot{\mathbf{L}}\mathbf{L}^{-1} + 2(\dot{\mathbf{L}}\mathbf{L}^{-1})^3]\\ &\quad + (\det\mathbf{L})^{\cdot}(\cdots) + (\det\mathbf{L})^{\cdot\cdot}(\cdots),\end{aligned}$$

etc. If $\det\mathbf{L} = 1$ always, these relations reduce to

$$\begin{aligned}\operatorname{tr}[\ddot{\mathbf{L}}\mathbf{L}^{-1} - (\dot{\mathbf{L}}\mathbf{L}^{-1})^2] &= 0,\\ \operatorname{tr}[\dddot{\mathbf{L}}\mathbf{L}^{-1} - 3\ddot{\mathbf{L}}\mathbf{L}^{-1}\dot{\mathbf{L}}\mathbf{L}^{-1} + 2(\dot{\mathbf{L}}\mathbf{L}^{-1})^3] &= 0,\end{aligned}$$

etc. In an isochoric motion, we may substitute $\mathbf{C}_t(u)$ for $\mathbf{L}$. Putting u for t, followed by use of the definition (II.11-31), yields (II.11-34)$_{2,3}$. The term involving the time derivative of highest order in the formula for $(\det\mathbf{L})^{(n)}$ is $(\det\mathbf{L})\operatorname{tr}\left(\overset{(n)}{\mathbf{L}}\mathbf{L}^{-1}\right)$, and so the general assertion of the exercise follows.

II.11.11 Let $\mathbf{x}$ be a point of $\mathscr{S}$, and let $\mathbf{k}$ be a vector in the tangent plane of $\mathscr{S}$ at $\mathbf{x}$. Then there are points $\mathbf{y}(h)$ on $\mathscr{S}$ such that

$$\mathbf{y}(h) = \mathbf{x} + h\mathbf{k} + \mathbf{o}(h) \qquad \text{as } h \to 0.$$

Therefore,

$$\mathbf{Gk} := \lim_{h\to 0}\frac{\dot{\mathbf{x}}(\mathbf{x}+h\mathbf{k}) - \dot{\mathbf{x}}(\mathbf{x})}{h},$$

$$= \lim_{h\to 0}\frac{\dot{\mathbf{x}}(\mathbf{y}(h)) - \dot{\mathbf{x}}(\mathbf{x})}{h}.$$

If $\dot{\mathbf{x}}$ vanishes on $\mathscr{S}$, the difference quotient on the right-hand side vanishes, and so (II.11-38) follows. Now in (II.11-13) replace $\mathbf{e}$ by $\mathbf{n}$, a unit normal to the tangent plane at $\mathbf{x}$; then denote by $\mathbf{e}$ a unit vector in the axis of $\mathbf{W}$, so that $\mathbf{W} = \frac{1}{2}w\mathbf{n}\wedge\mathbf{f}$, $\mathbf{We} = \mathbf{0}$. Because of (II.11-38), $\mathbf{Ge} = \mathbf{Gf} = \mathbf{0}$, and so

$$\mathbf{De} = \mathbf{0}, \qquad \mathbf{Df} = -\mathbf{Wf} = -\tfrac{1}{2}w\mathbf{n},$$

the last equation being a consequence of $(\text{II.11-14})_2$. Now since

$$\begin{aligned}\mathbf{Dn} &= (\mathbf{n}\cdot\mathbf{Dn})\mathbf{n} + (\mathbf{e}\cdot\mathbf{Dn})\mathbf{e} + (\mathbf{f}\cdot\mathbf{Dn})\mathbf{f} = (\mathbf{n}\cdot\mathbf{Dn})\mathbf{n} + (\mathbf{n}\cdot\mathbf{Df})\mathbf{f},\\ E &= \mathbf{n}\cdot\mathbf{Dn} + \mathbf{e}\cdot\mathbf{De} + \mathbf{f}\cdot\mathbf{Df},\end{aligned}$$

it follows that

$$\begin{aligned}\mathbf{Dn} &= (\mathbf{n}\cdot\mathbf{Dn})\mathbf{n} - \tfrac{1}{2}w\mathbf{f},\\ E &= \mathbf{n}\cdot\mathbf{Dn}.\end{aligned}$$

Hence (II.11-39) follows. Because $\mathbf{De} = \mathbf{0}$, $\mathbf{e}$ is a principal axis of stretching, the corresponding principal stretching is 0, and $\det\mathbf{D} = 0$. The second principal invariant of $\mathbf{D}$ is $-\frac{1}{4}w^2$. Thus the characteristic equation of $\mathbf{D}$ is

$$D(D^2 - ED - \tfrac{1}{4}w^2) = 0,$$

the solutions of which are (II.11-40).

II.11.12 From (II.11-40) we see that

$$\mathfrak{W}^2 = 1/[1 + (E/w)^2].$$

II.11.13 $(\mathbf{F}^{\mathsf{T}}\dot{\mathbf{F}})^{\cdot} = \mathbf{F}^{\mathsf{T}}\ddot{\mathbf{F}} +$ a symmetric tensor. By choosing n successively as 1 and 2 in (II.11-28) show that $(\mathbf{F}^{\mathsf{T}}(\nabla\dot{\mathbf{x}})\mathbf{F})^{\cdot} = \mathbf{F}^{\mathsf{T}}(\nabla\ddot{\mathbf{x}})\mathbf{F} +$ a symmetric tensor. Take the skew part of this relation to get (II.11-42).

II.11.14 By (II.11-44)

$$\mathbf{W} = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{W}_{\kappa} = \mathbf{0}.$$

II.11.15 Use (II.11-28) to show that

$$\mathbf{G}_2 = \ddot{\mathbf{F}}\mathbf{F}^{-1} = (\mathbf{G}\mathbf{F})^{\cdot}\mathbf{F}^{-1} = \dot{\mathbf{G}} + \mathbf{G}^2.$$

The skew part is (II.11-48).

II.11.16 By (II.11-48)

$$(\tfrac{1}{2}|\mathbf{W}|^2)^{\cdot} = \mathbf{W}\cdot(\mathbf{W}_{\mathbf{a}} - \mathbf{D}\mathbf{W} - \mathbf{W}\mathbf{D}).$$

If $\dim \mathscr{V} = 3$, then for any skew tensor $\mathbf{W}$ and any symmetric tensor $\mathbf{D}$

$$\mathbf{W}\cdot(\mathbf{D}\mathbf{W} + \mathbf{W}\mathbf{D}) = |\mathbf{W}|^2(\operatorname{tr}\mathbf{D} - \mathbf{n}\cdot\mathbf{D}\mathbf{n}),$$

in which $\mathbf{n}$ is either unit vector in the nullspace of $\mathbf{W}$. Use of (II.11-15) yields (II.11-49), from which the conclusion of the exercise is obvious.

II.11.18 In the proof of the theorem of Kelvin and Helmholtz replace the assumption of steady density by the general equation $(\text{II.6-6})_2$ and so obtain

$$2K = -\int_{\mathscr{R}} \rho' P_{\mathbf{v}}\, dV - \int_{\partial\mathscr{R}} \rho P_{\mathbf{v}} \dot{\mathbf{x}}\cdot\mathbf{n}\, dA,$$

generalizing (II.11-58). Under the conditions stated in the exercise the surface integral vanishes. In unbounded domains the condition (II.11-59) suffices to make the surface integral vanish. (The isochoric instance is more fruitful because the student has at his disposition the developed discipline called "potential theory", while conditions at ∞ for a mass density that depends upon $\mathbf{x}$ and t are difficult to ascertain in practice.)

II.12.1 The ellipsoid in $\kappa(\mathscr{B})$ is swept out by the termini of vectors $\mathbf{m}_{\kappa}$ such that

$$\text{const.} = |\mathbf{m}|^2 = |\mathbf{F}\mathbf{m}_{\kappa}\cdot\mathbf{F}\mathbf{m}_{\kappa}| = \mathbf{m}_{\kappa}\cdot\mathbf{C}\mathbf{m}_{\kappa}.$$

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be an orthonormal set of unit proper vectors of $\mathbf{C}$, so that $\mathbf{C}\mathbf{e}_j = v_j^2\mathbf{e}_j$, where v_j is the principal stretch corresponding to $\mathbf{e}_j$. Let the co-ordinates of $\mathbf{m}_\kappa$ with respect to this basis be m^k. Then the above equation for the ellipsoid assumes in cartesian co-ordinates the form

$$\sum_{k=1}^{3}(m^k)^2 v_k^2 = \text{const.}$$

Therefore, the principal axes of the ellipsoid are the principal axes of strain at $\mathbf{X}$, and the lengths of the semi-axes are inversely proportional to the corresponding squared principal stretches. The extremal properties of the principal stretches correspond inversely to the extremal properties of the lengths of vectors to points on the ellipsoid.

That the principal axes are not sheared, is the same as the statement $\cos\theta_{(\mathbf{e}_i,\mathbf{e}_j)} = \delta_{ij}$, which is an immediate consequence of (II.12-6).

Since (II.12-1) can be written in the form

$$\begin{aligned}\chi_\kappa(\mathbf{X}, t) &= \mathbf{X}_0 + (\mathbf{x}_0(t) - \mathbf{X}_0) + \mathbf{R}(t)\mathbf{U}(t)(\mathbf{X} - \mathbf{X}_0),\\ &= \mathbf{X}_0 + (\mathbf{x}_0(t) - \mathbf{X}_0) + \mathbf{V}(t)\mathbf{R}(t)(\mathbf{X} - \mathbf{X}_0),\end{aligned}$$

the last statement follows immediately by aid of (II.9-4).

II.12.3 Differentiate (II.12-6) after writing it as

$$v_{\mathbf{n}_\kappa} v_{\mathbf{m}_\kappa} \cos\theta_{(\mathbf{n}_\kappa,\mathbf{m}_\kappa)} = \mathbf{n}_\kappa \cdot \mathbf{U}^2\mathbf{m}_\kappa .$$

II.13.1 $\int_{\chi(\mathscr{C},t)} \mathbf{f}\cdot d\mathbf{x} = \int_{\kappa(\mathscr{C})} \mathbf{f}(\mathbf{X}, t)\cdot\mathbf{F}(\mathbf{X}, t)\,d\mathbf{X}$. Now on the right-hand side differentiation can be performed under the integral sign.

II.13.2 The volume V of a tetrahedron whose vertices are $\mathbf{x}_0(t)$ and the termini of $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$, is given in terms of the components p_a^k as follows:

$$V = \epsilon_{mnq} p_1^m p_2^n p_3^q .$$

Hence

$$\begin{aligned}\dot{V} &= \epsilon_{mnq}[G_r^m p_1^r p_2^n p_3^q + G_r^n p_1^m p_2^r p_3^q + G_r^q p_1^m p_2^n p_3^r],\\ &= (\operatorname{tr}\mathbf{G})V.\end{aligned}$$

II.13.3 Put $d\mathbf{x} = \mathbf{t}\,ds$, $\mathbf{f} = f\mathbf{t}$ in (II.13-1).

II.13.4 By (II.13-7) and (II.11-8)

$$\mathbf{e}^c = \dot{\mathbf{e}} - (\mathbf{D} + \mathbf{W})\mathbf{e}. \tag{A}$$

Since $\mathbf{e}\cdot\dot{\mathbf{e}} = 0$ and $\mathbf{W}$ is skew,

$$\mathbf{e}\cdot\mathbf{e}^c = -\mathbf{e}\cdot\mathbf{De}.$$

Substituting this formula into (II.13-14) and then putting the outcome into (A) yields (II.13-15). If $\mathbf{De} = d\mathbf{e}$, (II.13-15) reduces to

$$\dot{\mathbf{e}} = \mathbf{We}.$$

Cf. the discussion of rigid motion in Section I.10. For any vector $\mathbf{m}$ we obtain from (II.13-15)

$$\mathbf{m}\cdot\dot{\mathbf{e}} = \mathbf{m}\cdot\mathbf{De} + \mathbf{m}\cdot\mathbf{We} - (\mathbf{e}\cdot\mathbf{De})(\mathbf{m}\cdot\mathbf{e}).$$

Hence (II.13-16) follows. All these conclusions apply to the position vectors $\mathbf{p}$ in a homogeneous motion because $\mathbf{p}^c = \mathbf{0}$. Since $\mathbf{m}\cdot\mathbf{n} = \cos\theta_{(\mathbf{m},\mathbf{n})}$, (II.13-16)$_1$ reduces to (II.12-10) if $\mathbf{m}\cdot\mathbf{n} = 0$ at the instant in question. Likewise (II.13-16)$_2$ reduces to (II.12-15).

II.13.5 Let the vector field $\mathbf{f}$ be tangent to a vector line of $\mathbf{S}$ at the time t_0. Then $\mathbf{f}$ is tangent to a vector line of $\mathbf{S}$ for all $t \geqq t_0$ if and only if

$$(\mathbf{Sf})^{\cdot} = \mathbf{0}.$$

Hence by use of (II.13-4) and (II.13-7)

$$\mathbf{S}^c\mathbf{f} + \mathbf{S}\mathbf{f}^c = \mathbf{0}.$$

By (II.13-12), there is a scalar field α such that $\mathbf{f}^c = \alpha\mathbf{f}$. Thus, $\mathbf{f}$ too, has to be tangent to a vector line of $\mathbf{S}^c$: $\mathbf{S}^c\mathbf{f} = \mathbf{0}$. Now recall that two non-null skew tensors have one and the same vector lines if and only if they commute.

II.13.7 The argument is phrased in terms of *vortex tubes*. These are surfaces swept out by the vortex lines through the points of some circuit nowhere tangent to the axes of spin. The flux of spin has the same value, at a given instant, for all like-oriented surfaces bounded by circuits embracing the tube just once (HELMHOLTZ's *First Vorticity Theorem*). KELVIN's argument, a classic example of conceptual mathematics, may be found in LAMB's treatise, in Section 128 of CFT, and elsewhere.

II.13.8 POINCARÉ's Theorem makes the hypothesis equivalent to $\mathbf{W}\mathbf{W}_a = \mathbf{W}_a\mathbf{W}$. Exercise II.11.12 makes the differential relation equivalent to $\mathbf{W}\cdot\mathbf{W}_a = 0$. If neither $\mathbf{W}$ nor $\mathbf{W}_a$ vanishes, the two requirements are incompatible, for one requires the axes of the two tensors to coincide and the other requires that they be perpendicular to each other. If $\mathbf{W} = \mathbf{0}$, then $\mathbf{W}_a = \mathbf{0}$, as is shown in Exercise II.11.12. Thus $\mathbf{W}_a = \mathbf{0}$ is the only possibility, and clearly it is sufficient that the two conditions be compatible.

II.13.9 Take $\dot{\mathbf{x}}$ for $\mathbf{c}$ in MASOTTI's formula (Section App.IIC.6) to obtain for $\mathbf{w} - \Omega\dot{\mathbf{x}}$ a vector which for a screw motion must vanish because $\mathbf{w}$ and $\dot{\mathbf{x}}$ are collinear.

II.13.10 Put $\dot{\mathbf{x}}$ for $\mathbf{c}$ in the second formula for Ω in Section App.IIC.6, then use (II.13-23).

II.13.11 Take the curl of (II.13-23); then the inner product of the result and $\dot{\mathbf{x}}$, and use (II.13-23).

II.13.12 Inspect (II.11-9).

II.13.13 As was remarked just after (II.6-6), in a motion with steady density $\operatorname{div}(\rho\dot{\mathbf{x}}) = 0$. Using (II.13-23) delivers $\operatorname{div}[(\rho/\Omega)\mathbf{w}] = 0$, and so $\mathbf{w}\cdot\operatorname{grad}(\rho/\Omega) = 0$. Thus $\dot{\mathbf{x}}$ is the tangent to one of the surfaces $\rho/\Omega = \text{const.}$

II.13.14 If a screw motion preserves circulation, then taking the screw part of the gradient of (II.11-9) yields $\mathbf{w}' = \mathbf{0}$. The conclusion follows from (II.13-23) and (II.13-24).

II.13.15 Begin as in Exercise II.13.11. For the first statement the condition $\operatorname{grad}\Omega \times \dot{\mathbf{x}} = \mathbf{0}$ is necessary and sufficient. The others follow by taking curls of the preceding.

II.14.1 Referring to (I.9-14), for $\mathbf{A}(\boldsymbol{\chi}^* - \mathbf{x}_0^*)$ write $\mathbf{w} \times \mathbf{p}^*$ and note that

$$\begin{aligned}\mathscr{C}_{\oint^*}(C_{\oint^*}) - \mathscr{C}_{\oint}(C_{\oint}) &= \int_{\mathscr{C}_{\oint}^*} \mathbf{w} \times \mathbf{p}^*\cdot d\mathbf{s}^*,\\ &= \mathbf{w}\cdot\int_{\mathscr{C}_{\oint}^*} \mathbf{p}^* \times d\mathbf{s}^*.\end{aligned}$$

For a plane circuit $\mathscr{C}$ the vector delivered by the latter integral is normal to the plane, and its magnitude is the area of the region bounded by $\mathscr{C}_{\oint^*}$.

II.14.2 By (II.8-7) and (II.14-6),

$$\begin{aligned}\mathbf{F}_t^*(\tau) &= \mathbf{F}^*(\tau)(\mathbf{F}^*(t))^{-1},\\ &= \mathbf{Q}(\tau)\mathbf{F}(\tau)\mathbf{F}^{-1}(t)\mathbf{Q}(t)^{\mathsf{T}},\\ &= \mathbf{Q}(\tau)\mathbf{F}_t(\tau)\mathbf{Q}(t)^{\mathsf{T}}.\end{aligned}$$

Hence

$$\mathbf{R}_t^*(\tau)\mathbf{U}_t^*(\tau) = \mathbf{Q}(\tau)\mathbf{R}_t(\tau)\mathbf{Q}(t)^{\mathsf{T}}[\mathbf{Q}(t)\mathbf{U}_t(\tau)\mathbf{Q}(t)^{\mathsf{T}}].$$

The conclusion (II.14-20) follows by the uniqueness of a polar decomposition.

III.1.1

$$\begin{aligned}\mathbf{f}_{\mathrm{B}}^{\mathrm{a}} &= \int_{\chi(\mathscr{B},t)} \mathbf{b}\,dM = M\mathbf{b},\\ \mathbf{F}_{\mathrm{B}}^{\mathrm{a}} &= \int_{\chi(\mathscr{B},t)} (\mathbf{x}-\mathbf{x}_0)\wedge\mathbf{b}\,dM,\\ &= \int_{\chi(\mathscr{B},t)} \mathbf{p}\,dM\wedge\mathbf{b} = \bar{\mathbf{p}}\wedge M\mathbf{b},\end{aligned}$$

by (I.8-28).

III.1.2 By (III.1-50)$_2$ and (III.1-46),

$$\int_{\mathscr{S}} \mathbf{t}_{\mathscr{S}}\,dA = -\int_{-\mathscr{S}} \mathbf{t}_{-\mathscr{S}}\,dA.$$

The Lebesgue differentiation theorem gives $\mathbf{t}_{\mathscr{S}} = -\mathbf{t}_{-\mathscr{S}}$ a.e.

III.2.1 Expand $\rho(\dot{\mathbf{x}}-\mathbf{v})\otimes(\dot{\mathbf{x}}-\mathbf{v})\mathbf{n}$ and $\mathbf{p}\wedge[\rho(\dot{\mathbf{x}}-\mathbf{v})\otimes(\dot{\mathbf{x}}-\mathbf{v})\mathbf{n}]$; integrate over the shape of $\mathscr{B}$; use the divergence theorem; note that

$$\operatorname{div}(\rho\mathbf{p}\otimes\dot{\mathbf{x}}) = \rho\dot{\mathbf{x}} + \mathbf{p}\operatorname{div}(\rho\dot{\mathbf{x}});$$

use (II.6-6) for a motion with steady density, and note that $\dot{\mathbf{x}}\cdot\mathbf{n} = 0$ on the obstacle.

III.3.1 Choose Δr such that $A(\Delta\mathscr{A}) \leqq A(\Delta\mathscr{A}') + A(\Delta\mathscr{A}^*)$. Then (III.3-11) and (III.3-9)$_2$ yield (III.3-12). Hence

$$\begin{aligned}A(\partial\,\Delta\mathscr{D}) &= A(\Delta\mathscr{A}) + A(\Delta\mathscr{A}^*) + A(\Delta\mathscr{A}'),\\ V(\Delta\mathscr{D}) &= o(\Delta r^3) \qquad \text{as } \Delta r \to 0.\end{aligned}$$

III.3.2 The student would be well advised to draw a figure. Let $\mathscr{T}'$ be the tangent plane to $\mathscr{S}$ and $\mathscr{T}$ at $\mathbf{x}$. With respect to cartesian coordinates (x, y, z) with x-axis and y-axis in $\mathscr{T}'$, let $z = f(x, y)$ and $z = g(x, y)$ be the representations of $\mathscr{S}$ and $\mathscr{T}$ near $\mathbf{x}$. Choose Δr such that when $x^2 + y^2 \leqq \Delta r^2$, $\mathscr{S}$ and $\mathscr{T}$ lie entirely between two paraboloids $z = \pm K(x^2 + y^2)$, where

$$K := \max(|\partial_x^2 f|, |\partial_x\partial_y f|, |\partial_y^2 f|, |\partial_x^2 g|, |\partial_x\partial_y g|, |\partial_y^2 g|).$$

Follow the same procedure as before.

III.4.1 Let C be a cube, and let two of its faces be normal to $\mathbf{k}$. Then

$$\int_{\partial \mathrm{C}} \mathbf{t}\, dA = 2(V(\mathrm{C}))^{2/3}\mathbf{k},$$

and

$$\lim_{V(\mathrm{C})\to 0} \int_{\partial \mathrm{C}} \mathbf{t}\, dA = \mathbf{0},$$

but

$$\lim_{V(\mathrm{C})\to 0} \frac{\left|\int_{\partial \mathrm{C}} \mathbf{t}\, dA\right|}{V(\mathrm{C})} = +\infty,$$

and so (III.1-58) is violated, while (III.1-59) is not.

III.4.2 Immediate from (III.4-1) and the definition of the transpose.

III.6.1 Prove the identity

$$\operatorname{div}(\mathbf{v} \otimes \mathbf{S}) = \mathbf{v} \otimes \operatorname{div} \mathbf{S} + (\nabla \mathbf{v})\mathbf{S}^{\mathsf{T}},$$

take the skew part, set $\mathbf{v} = \mathbf{x} - \mathbf{x}_0$, and apply the divergence theorem.

III.6.2 Hold $\mathbf{x}$ fixed, and drop it from the notation; do not assume that $\mathbf{T}^{\mathsf{T}} = \mathbf{T}$. Trivially (III.6-10) $\Leftrightarrow$ (III.6-11), and

$$\text{(III.6-11)} \qquad \Rightarrow \qquad \text{(III.6-8) \& (III.6-9).}$$

Write (III.6-9) in the form $\mathbf{n}\cdot\mathbf{T}\mathbf{n} = -p$ for all unit vectors $\mathbf{n}$; let $\mathbf{n} = \cos\theta\, \mathbf{n}_1 + \sin\theta\, \mathbf{n}_2$, $\mathbf{n}_1$ and $\mathbf{n}_2$ being unit vectors, and show that $\mathbf{n}_1\cdot\mathbf{T}\mathbf{n}_2 = -\mathbf{n}_2\cdot\mathbf{T}\mathbf{n}_1$. Hence conclude that

$$\text{(III.6-9)} \qquad \Leftrightarrow \qquad \mathbf{T} = -p\mathbf{1} + \mathbf{S},$$

$\mathbf{S}$ being a skew tensor. If $\mathbf{T}$ is symmetric, (III.6-11) follows. Otherwise it does not.

(III.6-8) requires that

$$|\mathbf{t}(\mathbf{n})|^2 = \mathbf{n}\cdot\mathbf{T}^{\mathsf{T}}\mathbf{T}\mathbf{n} = p^2.$$

The conclusion just reached in regard to (III.6-9) shows that $\mathbf{T}^{\mathsf{T}}\mathbf{T} = p^2\mathbf{1}$ + a skew tensor, but this latter is $\mathbf{0}$ because $\mathbf{T}^{\mathsf{T}}\mathbf{T}$ is symmetric. If $p = 0$, (III.6-11) holds trivially; if $p \neq 0$, we have shown that $p^{-1}\mathbf{T}$ is an orthogonal tensor, say $-\mathbf{Q}$. Then

$$\mathbf{n}\cdot\mathbf{t}(\mathbf{n}) = -p\mathbf{n}\cdot\mathbf{Q}\mathbf{n}.$$

If $\mathbf{R}$ is the rotation such that $\mathbf{Q} = \pm\mathbf{R}$, show that

$$\mathbf{n}\cdot\mathbf{R}\mathbf{n} = 1 - 2n_{\perp}^2 \sin^2 \tfrac{1}{2}\theta,$$

$n_{\perp}$ being the magnitude of the component of $\mathbf{n}$ normal to the axis of $\mathbf{R}$, and θ being the angle of $\mathbf{R}$. In order that $\mathbf{n}\cdot\mathbf{R}\mathbf{n} > 0\ \forall\mathbf{n}$, it is necessary and sufficient that $1 - 2\sin^2 \tfrac{1}{2}\theta > 0$. Thus (III.6-8) and

$$\mathbf{T} = -p\mathbf{R}, \qquad 0 \leqq \theta < \tfrac{1}{2}\pi \quad \text{or} \quad \tfrac{3}{2}\pi < \theta \leqq 2\pi,$$

are equivalent. If $\mathbf{T}$ is symmetric, $\mathbf{R} = \mathbf{R}^{\mathsf{T}}$, and hence $\theta = 0$ or π. Since the latter alternative is excluded by the conclusion just drawn, $\mathbf{R} = \mathbf{1}$. Thus (III.6-11) follows from (III.6-8) if $\mathbf{T}$ is symmetric. Otherwise it does not.

III.6.3 The statement is really an instance of NOLL's theorem in Section I.12 but is more than a century older. For an independent proof, hold t fixed and consider the rigid transplacement defined as follows by a constant vector $\mathbf{v}_0$, the position vector $\mathbf{p}$, and a constant skew tensor $\mathbf{W}_0$: $\mathbf{v} = \mathbf{v}_0 + \mathbf{W}_0\mathbf{p}$. Since $\mathbf{a}\cdot\mathbf{S}\mathbf{b} = \mathbf{S}\cdot(\mathbf{a}\otimes\mathbf{b})$ for any skew tensor $\mathbf{S}$,

$$P = \mathbf{v}_0\cdot\left[\int_{\partial\chi(\mathscr{P},t)} \mathbf{t}\,dA + \int_{\chi(\mathscr{P},t)} \rho\mathbf{b}\,dV\right] + \mathbf{W}_0\cdot\left[\int_{\partial\chi(\mathscr{P},t)} \mathbf{t}\otimes\mathbf{p}\,dA + \int_{\chi(\mathscr{P},t)} \rho\mathbf{b}\otimes\mathbf{p}\,dV\right].$$

In order that $P = 0$ for all choices of $\mathbf{v}_0$ and $\mathbf{W}_0$, it is necessary and sufficient that the first bracket vanish and the second bracket be symmetric.

III.6.4 Use (I.14-1), (III.6-12), the divergence theorem, (III.6-1), and (II.11-8). In a rigid motion $\mathbf{D} = \mathbf{0}$. In an isochoric motion $\operatorname{tr}\mathbf{D} = 0$.

III.6.5 Substitute from (III.1-7) and (III.6-15) and (III.6-12) and use (II.6-9) to obtain (III.6-17). Then (III.6-18) and (III.6-19) are easy to obtain.

III.6.6 The existence of h, the *influx of heating*, first demonstrated by STOKES, follows by use of arguments parallel to those that deliver CAUCHY's Fundamental Theorem (Section III.4). The differential equation (III.6-20) follows by appropriate substitutions in (III.5-1), which delivers (III.5-4).

III.7.1

$$\left(\int_{\mathscr{A}} \mathbf{p} \otimes \mathbf{T}\mathbf{e}\, dA\right) \mathbf{e} = \int_{\mathscr{A}} (\mathbf{e}\cdot\mathbf{T}\mathbf{e})\mathbf{p}\, dA,$$
$$= \frac{F}{A(\mathscr{A})}\int_{\mathscr{A}} \mathbf{p}\, dA,$$
$$= F\mathbf{p}_0(\mathscr{A}).$$

Using subscript 1 and 2 to refer to quantities associated to the two plane, parallel faces, if $\mathbf{n}_1 = \mathbf{e}$ we must take $\mathbf{n}_2$ as $-\mathbf{e}$, and so (III.7-14) yields

$$\mathbf{e}\cdot\mathbf{T}\mathbf{e} = \frac{1}{V(\mathscr{S})}[F_1\mathbf{p}_0(\mathscr{A}_1) - F_2\mathbf{p}_0(\mathscr{A}_2)]\cdot\mathbf{e}.$$

Equilibrium of forces requires that $F_1 = F_2$; equilibrium of moments, that $\mathbf{p}_0(\mathscr{A}_1) - \mathbf{p}_0(\mathscr{A}_2)$ be parallel to $\mathbf{e}$.

III.7.2 Taking Ψ as $\mathbf{p}\otimes\mathbf{p}$ in (III.7-3) yields at once

$$\overline{T_{rk}p_q} + \overline{T_{rq}p_k} = M_{kqr}\,.$$

Forming the combination indicated by (III.7-10)$_1$ and then using Cauchy's Second Law yields

$$\overline{p_kT_{qr}} = L_{kqr}\,.$$

III.8.1 Let the constant g denote the gravitational acceleration, let ρ denote the density of the heavy liquid, and let z denote the distance downward from the surface of the liquid. Then $p = \rho g z$ on the surface of the submerged part of the body, say $\mathscr{S}$. The resultant surface force and surface torque upon $\chi(\mathscr{B}, t)$ are, respectively,

$$\mathbf{f}_{\mathrm{C}}^{\mathrm{a}} = -\rho g \int_{\mathscr{S}} z\mathbf{n}\, dA,$$

$$\mathbf{F}_{\mathrm{C}}^{\mathrm{a}} = -\rho g \int_{\mathscr{S}} (\mathbf{x} - \mathbf{x}_0) \wedge z\mathbf{n}\, dA.$$

The formulae used to denote the two integrands serve to extend them smoothly to the interior of the part of $\chi(\mathscr{B}, t)$ below the plane $z := 0$. (This fact expresses STEVIN's *Principle of Solidification*: The load exerted by one part of a heavy fluid body upon another is unchanged if either is replaced by a rigid solid.) Thus we can apply Green's transformation to express the two surface integrals as volume integrals over the submerged part $\mathscr{V}$. Since $\operatorname{grad} z = \mathbf{k}$, a unit vector pointing downwards,

$$\int_{\mathscr{S}} z\mathbf{n}\, dA = \left(\int_{\mathscr{V}} dV\right)\mathbf{k} = V(\mathscr{V})\mathbf{k}.$$

Likewise

$$\begin{aligned}\int_{\mathscr{S}} (\mathbf{x} - \mathbf{x}_0) \wedge z\mathbf{n}\, dA &= \int_{\mathscr{V}} (\mathbf{x} - \mathbf{x}_0)\, dV \wedge \mathbf{k},\\ &= \mathbf{p}_{\mathrm{m}} \wedge V(\mathscr{V})\mathbf{k},\end{aligned}$$

$\mathbf{p}_{\mathrm{m}}$ being the position vector of the center of buoyancy.[1]

To consider a heavy body, invoke the result of Exercise III.1.1 to conclude that the load on that body is equipollent to two parallel forces: the weight of the body, acting downward at its center of mass, and the weight of the displaced fluid, acting upward at the center of buoyancy. Consideration of a simple vector diagram suffices to conclude the exercise.

III.9.1 Since $\partial\chi(\mathscr{B}, t)$ has a differentiable unit normal field $\mathbf{n}$, that field can be extended smoothly into a small region containing $\partial\chi(\mathscr{B}, t)$ in its interior. A standard theorem of differential geometry asserts then that the mean curvature

$$k = \tfrac{1}{2} \operatorname{div} \mathbf{n}. \tag{A}$$

[1] To define the centroid of a region, in (I.8-28) replace M by V and $\mathscr{B}$ by the region considered.

If $\mathbf{n}$ is any differentiable field of unit vectors, and if $\mathbf{c}$ is a constant vector field, then in a 3-dimensional space

$$\mathbf{n}\cdot\operatorname{curl}(\mathbf{n}\times\mathbf{c}) = -\mathbf{c}\cdot(\operatorname{div}\mathbf{n})\mathbf{n}, \tag{B}$$

$$\mathbf{n}\cdot\operatorname{curl}\{\mathbf{n}\times[\mathbf{c}\times(\mathbf{x}-\mathbf{x}_0)]\} = -\mathbf{c}\cdot[(\mathbf{x}-\mathbf{x}_0)\times(\operatorname{div}\mathbf{n})\mathbf{n}].$$

By Kelvin's transformation the integral of $\mathbf{n}\cdot\operatorname{div}(\mathbf{p}\wedge\mathbf{q})$ over a surface $\mathscr{S}$ is equal to the value of a line integral around $\partial\mathscr{S}$. If $\mathscr{S}$ is a surface without boundary, that value is 0. Thus the integrals of the right-hand sides of (B) over $\partial\chi(\mathscr{B}, t)$ both equal 0. Use of (A) completes the proof.

IV.4.1 By hypothesis, $\mathbf{QK} = \mathbf{KQ}\ \forall\mathbf{Q}$. For $\mathbf{Q}$ take the reflection $\mathbf{R}_e$ in the plane normal to $\mathbf{e}$. Then $\mathbf{R}_e\mathbf{v} = -\mathbf{v}$ if and only if $\mathbf{v}$ is proportional to $\mathbf{e}$. But $\mathbf{R}_e\mathbf{K}\mathbf{e} = \mathbf{K}\mathbf{R}_e\mathbf{e} = -\mathbf{K}\mathbf{e}$, so $\mathbf{K}\mathbf{e}$ is proportional to $\mathbf{e}$, no matter what be $\mathbf{e}$. Suppose now that $\mathbf{K}\mathbf{e} = \alpha\mathbf{e}$, $\mathbf{K}\mathbf{f} = \beta\mathbf{f}$, $\mathbf{K}(\mathbf{e}+\mathbf{f}) = \gamma(\mathbf{e}+\mathbf{f})$. Then $\alpha\mathbf{e} + \beta\mathbf{f} = \gamma(\mathbf{e}+\mathbf{f})$. Choosing $\mathbf{e}$ and $\mathbf{f}$ as linearly independent shows that $\alpha = \beta = \gamma$.

Note. We may ask if the condition $\mathbf{RK} = \mathbf{KR}$ for all rotations $\mathbf{R}$ implies (IV.4-23). The answer is no if the dimension of the vector space is 2, for then all rotations commute. If the dimension of the vector space is odd, the answer is obviously yes, since the tensors $\pm\mathbf{R}$ exhaust the orthogonal tensors. The answer is yes also for vector spaces of even dimension greater than 2 but is not so obvious. It is easy to prove that a symmetric tensor which commutes with every rotation is proportional to $\mathbf{1}$.

IV.4.2 Follow the procedure given in the proof of (IV.4-2).

IV.4.3 If $\mathbf{g}$ in (IV.4-1) is an affine function of $\mathbf{F}$,

$$\mathbf{T} = \mathbf{A} + \mathbf{B}[\mathbf{F}],$$

$\mathbf{A}$ being a constant tensor and $\mathbf{B}$ being a tensor-valued linear function of tensors. In order for this constitutive equation to satisfy the Principle of Material Frame-Indifference, it is necessary and sufficient that

$$\mathbf{Q}(\mathbf{A} + \mathbf{B}[\mathbf{F}])\mathbf{Q}^{\mathsf{T}} = \mathbf{A} + \mathbf{B}[\mathbf{QF}] \tag{$*$}$$

for all invertible $\mathbf{F}$ and all orthogonal $\mathbf{Q}$. Put $\mathbf{F} = C\mathbf{1}$, $C \neq 0$, to obtain

$$\mathbf{QAQ}^{\mathsf{T}} - \mathbf{A} + C\mathbf{f}(\mathbf{B}, \mathbf{Q}) = \mathbf{0}$$

Since the real number C is arbitrary, both terms in this affine function of C vanish. Because $\mathbf{A}$ commutes with all orthogonal tensors, $\mathbf{A} = A\mathbf{1}$. The functional equation (*) reduces to

$$\mathbf{Q}(\mathsf{B}[\mathbf{F}])\mathbf{Q}^{\mathsf{T}} = \mathsf{B}[\mathbf{QF}].$$

Taking $\mathbf{Q} = -\mathbf{1}$ shows that $\mathsf{B}[\mathbf{F}] = \mathbf{0}$. (*Note.* $\mathbf{T} = \alpha\mathbf{1} + \beta\mathbf{V}$ does satisfy the principle, but the right-hand side is not an affine function of $\mathbf{F}$.)

IV.6.1. (IV.6-2) and (IV.6-3) are straightforward. A frame-indifferent, elastic constraint equivalent to (IV.6-1) is of the form $\mu(\mathbf{C}) = 0$, where μ vanishes if and only if γ vanishes. This statement is logically equivalent to the last sentence of the exercise. (Note that μ and γ are not claimed to be functionally dependent, though of course they may be.)

IV.7.1 Only (IV.7-14) requires care, because CAUCHY's Theorem in Section IV.4 refers to a function whose domain is the space of all symmetric tensors. If the domain of $\mathfrak{g}$ is the subspace of traceless, symmetric tensors, we define as follows a function $\mathfrak{f}$ on all symmetric tensors:

$$\mathfrak{f}(\mathbf{D}) := \mathfrak{g}(\mathbf{D} - \tfrac{1}{3}(\operatorname{tr}\mathbf{D})\mathbf{1}).$$

If $\mathfrak{g}$ is affine and isotropic, so is $\mathfrak{f}$. Thus CAUCHY's Theorem applies to $\mathfrak{f}$. Specializing the conclusion to traceless tensors $\mathbf{D}$ yields (IV.7-14).

IV.7.2 Since $\mathbf{t} = -p\mathbf{n}$, the last statement follows at once from the formula proved in Exercise III.6.5. More generally, (III.6-18) becomes

$$\dot{K} + \dot{V} = P_{\mathrm{C}} = -\int_{\partial\chi(\mathscr{P},\,t)} p\dot{\mathbf{x}}\cdot\mathbf{n}\,dV.$$

If $p = \text{const.}$ on $\partial\chi(\mathscr{P}, t)$, the right-hand side reduces to

$$-p\int_{\chi(\mathscr{P},\,t)} \operatorname{div}\dot{\mathbf{x}}\,dV,$$

which vanishes since the flow is isochoric.

IV.9.1 A glance at (II.11-5) and $(\mathrm{IV.9\text{-}7})_1$ gives the assertion.

IV.9.2

$$\det\mathbf{F}(t) = \det[\mathbf{F}_0(\mathbf{1} + t\mathbf{F}_1)] = \det\mathbf{F}_0\det(\mathbf{1} + t\mathbf{F}_1),$$

$$\det(\mathbf{1} + t\mathbf{F}_1) = 1 + t(\operatorname{tr}\mathbf{F}_1) + \tfrac{1}{2}t^2[(\operatorname{tr}\mathbf{F}_1)^2 - \operatorname{tr}\mathbf{F}_1^2] + t^3\det\mathbf{F}_1,$$

because the second principal invariant of a tensor $\mathbf{S}$ equals $\frac{1}{2}[(\operatorname{tr}\mathbf{S})^2 - \operatorname{tr}\mathbf{S}^2]$.

IV.9.3 Generalizing the spectral decomposition (IV.9-13), we allow not only the latent roots of $\mathbf{U}$ but also the orthonormal basis vectors to depend differentiably upon t. Then

$$\dot{\mathbf{U}} = \sum_{k=1}^{3}(\dot{u}_k \mathbf{e}_k \otimes \mathbf{e}_k + u_k \dot{\mathbf{e}}_k \otimes \mathbf{e}_k + u_k \mathbf{e}_k \otimes \dot{\mathbf{e}}_k),$$

and hence

$$\dot{\mathbf{U}}\mathbf{U} - \mathbf{U}\dot{\mathbf{U}} = \sum_{k=1}^{3} u_k^2(\dot{\mathbf{e}}_k \otimes \mathbf{e}_k - \mathbf{e}_k \otimes \dot{\mathbf{e}}_k)$$
$$+ \sum_{k,q=1}^{3} u_k u_q(\dot{\mathbf{e}}_q \cdot \mathbf{e}_k)(\mathbf{e}_q \otimes \mathbf{e}_k - \mathbf{e}_k \otimes \mathbf{e}_q).$$

Because $\dot{\mathbf{e}}_k \cdot \mathbf{e}_q + \mathbf{e}_k \cdot \dot{\mathbf{e}}_q = 0$, $k, q = 1, 2, 3$, a calculation yields

$$\dot{\mathbf{U}}\mathbf{U} - \mathbf{U}\dot{\mathbf{U}} = -\frac{1}{2} \sum_{k,q=1}^{3} \dot{\mathbf{e}}_k \cdot \mathbf{e}_q (u_k - u_q)^2 (\mathbf{e}_k \otimes \mathbf{e}_q - \mathbf{e}_q \otimes \mathbf{e}_k).$$

Consequently a necessary and sufficient condition that $\dot{\mathbf{U}}\mathbf{U} = \mathbf{U}\dot{\mathbf{U}}$ is

$$\dot{\mathbf{e}}_k \cdot \mathbf{e}_q (u_k - u_q)^2 = 0, \qquad k, q = 1, 2, 3.$$

If the orthonormal basis is constant in time, the statement of the exercise follows. Another sufficient condition, obviously, is that $\mathbf{U}$ shall have only one proper number. The student will distinguish and assemble other solutions of the above differential equation.

For the first instance, write

$$\mathbf{x} = \mathbf{x}_0(t) + \sum_{k=1}^{3} u_k(t) \mathbf{e}_k \otimes \mathbf{e}_k (\mathbf{X} - \mathbf{X}_0),$$

and conclude the first statement following (IV.9-13). Taking the $\mathbf{e}_i$ as the axes, let the block be the region included by the planes $X_k = \pm a_k$. Show that it is deformed into a similar block. Since it is already proved that $u_k = a_k + b_k t$ if $a_k > 0$ and $b_k > 0$, the motion will be isochoric if and only if

$$\prod_{k=1}^{3}(a_k + b_k t) = 1 \qquad \forall t.$$

This condition holds if and only if $a_1 a_2 a_3 = 1$, $b_1 = b_2 = b_3 = 0$.

IV.10.2 For a pure stretch $\dot{\mathbf{R}} = \mathbf{0}$, $\mathbf{R} = \mathbf{1}$. Use (II.11-26)$_2$ and (II.11-42).

IV.10.3 For an unconstrained simple body $\mathbf{T} = \mathfrak{G}(\mathbf{F}^t)$, and for the corresponding incompressible simple body $\mathbf{T} = \rho(\varpi - h)\mathbf{1} + \mathfrak{G}(\mathbf{F}^t)$. $\mathbf{F}(t)$ is given by (IV.9-10). In the unconstrained body, every component of $\mathbf{T}$ is determined uniquely. In the incompressible body, the function h is arbitrary. For example, if $\varpi = 0$, then by choice of h we may let any one of the tractions T_{xx}, T_{yy}, and T_{zz} be any function we please, *e.g.* 0.

IV.12.1 If $\mathbf{H}_1, \mathbf{H}_2 \in \mathscr{g}_\kappa$, then

$$\mathfrak{G}_\kappa(\mathbf{F}^t\mathbf{H}_1\mathbf{H}_2) = \mathfrak{G}_\kappa(\mathbf{F}^t\mathbf{H}_1) = \mathfrak{G}_\kappa(\mathbf{F}^t),$$

where the first step follows because $\mathbf{H}_2 \in \mathscr{g}_\kappa$, and the second because $\bar{\bar{\mathbf{H}}}_1 \in \mathscr{g}_\kappa$. Thus $\mathbf{H}_1\mathbf{H}_2 \in \mathscr{g}_\kappa$. Similar arguments verify the other axioms of a group.

IV.12.2 Since $\mathfrak{G}_\kappa$ satisfies Axiom N3,

$$\mathfrak{G}_\kappa(\mathbf{Q}\mathbf{F}^t) = \mathbf{Q}\mathfrak{G}_\kappa(\mathbf{F}^t)\mathbf{Q}^\mathsf{T}.$$

This statement when combined with (IV.12-6) implies that $\mathbf{Q}^\mathsf{T} \in \mathscr{g}_\kappa$.

IV.15.5 The change of reference placement is described as follows by use of the cartesian co-ordinates $\bar{X}, \bar{Y}, \bar{Z}$:

$$R^2 = 2\bar{A}\bar{X} + \bar{B}, \qquad \Theta = \bar{C}\bar{Y} + \bar{D}\bar{Z} + \bar{E},$$

$$Z = \bar{F}\bar{Y} + \bar{G}\bar{Z} + \bar{H};$$

$\bar{A}, \bar{B}, \dots \bar{H}$ are constants, and $\bar{A}(\bar{C}\bar{G} - \bar{D}\bar{F}) = 1$.

IV.16.1 By NOLL's rule $\mathbf{Q}^* = \mathbf{P}\mathbf{Q}\mathbf{P}^{-1}$, which can be written as

$$\mathbf{Q}^*\mathbf{R}_0\mathbf{U}_0 = \mathbf{R}_0\mathbf{Q}\mathbf{Q}^\mathsf{T}\mathbf{U}_0\mathbf{Q}.$$

Use the uniqueness of the polar decomposition.

IV.16.2 If

$$X_1^* = \lambda X_1, \qquad X_2^* = \lambda X_2, \qquad X_3^* = \mu X_3, \qquad \lambda \neq \mu,$$

then $[\mathbf{P}] = \operatorname{diag}(\lambda, \lambda, \mu)$, and so if $\mathbf{H} \in \mathscr{g}_\kappa$, then NOLL's rule gives

$$[\mathbf{H}^*] = \begin{Vmatrix} H_{11} & H_{12} & \frac{\lambda}{\mu} H_{13} \\ H_{21} & H_{22} & \frac{\lambda}{\mu} H_{23} \\ \frac{\mu}{\lambda} H_{31} & \frac{\mu}{\lambda} H_{32} & H_{33} \end{Vmatrix}.$$

If

$$\mathbf{H} = \begin{Vmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

then $\mathbf{H}^* = \mathbf{H}$. If

$$[\mathbf{H}] = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{Vmatrix},$$

then

$$[\mathbf{H}^*] = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \frac{\lambda}{\mu}\sin\varphi \\ 0 & -\frac{\mu}{\lambda}\sin\varphi & \cos\varphi \end{Vmatrix}.$$

IV.16.3 Since in general $\mathscr{g}_\kappa \neq \mathscr{o}$, $\mathscr{g}_\kappa$ will not be an invariant subgroup of $\mathscr{o}$. Thus $\mathscr{g}_\kappa^* (= \mathbf{R}\mathscr{g}_\kappa\mathbf{R}^{-1})$ will not be equal to $\mathscr{g}_\kappa$.

IV.16.4 Put $\mathbf{U}_0 = K\mathbf{1}$ in Theorem 2.

IV.17.2 Note that the right-hand side of (IV.17-3) depends upon κ only through ρ_κ / J.

IV.17.3 Use (II.11-53).

IV.18.1 For (IV.18-1) $[\mathbf{D}] = \operatorname{diag}(-(a + b/r^2), b/r^2, a)$. For (IV.18-2)

$$[\mathbf{D}] = \begin{Vmatrix} -h/r^2 & 0 & 0 \\ \cdot & h/r^2 & \frac{1}{2}l/r \\ \cdot & \cdot & 0 \end{Vmatrix}.$$

Clearly $-h/r^2$ is a principal stretching, and it is constant if and only if $h = 0$. In that case the other principal stretchings are $\pm\frac{1}{2}l/r$.

IV.18.3 In (IV.18-1) the constants c and g may be removed by superposing a rigid motion. Then

$$A = e^{-2as}, \qquad B = \frac{b}{a}(1 - e^{-2as}), \qquad C = e^{as}, \qquad D = 0,$$

$$K = \frac{c}{2a}(e^{as} - 1), \qquad E = 0, \qquad F = e^{as}, \qquad L = \frac{g}{a}(e^{as} - 1),$$

while for (6)

$$A = 1, \qquad B = 2hs, \qquad C = 1, \qquad D = 0, \qquad K = ks,$$

$$E = ls, \qquad F = 1, \qquad L = ms + \tfrac{1}{2}lks^2.$$

IV.21.1 Substitute (IV.21-3) into (II.8-8) to get (IV.21.13). The other relations follow easily from the definitions (II.11-2) and (II.11-31). The condition $\operatorname{tr}\mathbf{G} = 0$ is necessary and sufficient that the motion be isochoric, so (IV.21-16) follows.

IV.21.2 If $\mathbf{A}$ and $\mathbf{B}$ commute, then $e^{\mathbf{A}}e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}$.

IV.21.3 The most general form of $\mathbf{A}_2$ is

$$[\mathbf{A}_2] = \begin{Vmatrix} u & a & b \\ a & v & c \\ b & c & w \end{Vmatrix}.$$

Remembering that (IV.21-22) does not hold, show that this $\mathbf{A}_2$ commutes with $\mathbf{M} - \overline{\mathbf{M}}$ as given by (IV.21-27) if and only if $x = 0$. When $\mathbf{A}_1 = \alpha\mathbf{1}$, by the lemma $\mathbf{A}_1$ commutes with every skew tensor. Therefore $(\mathbf{M} - \mathbf{M}^{\mathsf{T}})\mathbf{A}_1 = \mathbf{A}_1(\mathbf{M} - \mathbf{M}^{\mathsf{T}})$. Using $(\text{IV.21-15})_3$, conclude that $\mathbf{M}\mathbf{M}^{\mathsf{T}} = \mathbf{M}^{\mathsf{T}}\mathbf{M}$, and then arrive at (IV.21-26).

IV.21.4 If $\mathbf{N}^3 = \mathbf{0}$, the Hamilton-Cayley equation reduces to

$$(\operatorname{tr}\mathbf{N})\mathbf{N}^2 = \tfrac{1}{2}[(\operatorname{tr}\mathbf{N})^2 - \operatorname{tr}\mathbf{N}^2]\mathbf{N}.$$

Thus

$$\mathbf{N} \neq \mathbf{0} \ \&\ \mathbf{N}^2 = \mathbf{0} \quad \Rightarrow \quad \operatorname{tr}\mathbf{N}^2 = 0 \ \&\ \operatorname{tr}\mathbf{N} = \mathbf{0},$$

while

$$\mathbf{N}^2 \neq \mathbf{0} \quad \Rightarrow \quad (\operatorname{tr}\mathbf{N})\mathbf{N}^3 = \mathbf{0} = \tfrac{1}{2}[(\operatorname{tr}\mathbf{N})^2 - \operatorname{tr}\mathbf{N}^2]\mathbf{N}^2$$
$$\Rightarrow [(\operatorname{tr}\mathbf{N})^2 = \operatorname{tr}\mathbf{N}^2] \quad \Rightarrow \quad \operatorname{tr}\mathbf{N} = 0 \quad \Rightarrow \quad \operatorname{tr}\mathbf{N}^2 = 0.$$

IV.21.6 Use $(\text{IV.21-15})_{4,7}$ to get (IV.21-30). Then (IV.21-31) follows by use of $(\text{IV.21-14})_2$ and $(\text{IV.21-15})_7$.

IV.21.7 Use (II.8-3) and (II.8-4) to obtain the relative description of the motion whose spatial velocity field is (IV.21-33):

$$\xi_1 = x_1, \qquad \xi_2 = (\tau - t)\mu x_1 + x_2,$$

$$\xi_3 = (\tau - t)(\lambda x_1 + \nu x_2) + \tfrac{1}{2}(t - \tau)^2 \mu\nu x_1 + x_3.$$

Hence $\mathbf{F}_t(\tau)$ assumes the form (IV.21-13) with the special values $\mathbf{Q} = \mathbf{1}$ and

$$[\kappa\mathbf{N}_0] = \begin{Vmatrix} 0 & 0 & 0 \\ \mu & 0 & 0 \\ \lambda & \nu & 0 \end{Vmatrix} = [\mathbf{G}],$$

components being taken with respect to the cartesian co-ordinate basis. Since $(\kappa\mathbf{N}_0)^3 = \mathbf{0}$ and $(\kappa\mathbf{N}_0)^2 = \mathbf{0} \Leftrightarrow \mu\nu = 0$, the first two assertions of the exercise follow. The relative description of the motion whose spatial velocity field is (IV.21-34) is

$$\xi_k = x_k e^{a_k(\tau - t)}, \qquad k = 1, 2, 3.$$

Thus

$$[\mathbf{F}] = \begin{Vmatrix} e^{a_1 t} & 0 & 0 \\ 0 & e^{a_2 t} & 0 \\ 0 & 0 & e^{a_3 t} \end{Vmatrix} = [\mathbf{U}];$$

(IV.9-15) is satisfied; and the last sentence of the exercise follows by the conclusion of Exercise IV.10.2.

IV.21.8 Use $(\text{IV.21-15})_2$, (II.11-22), and the statement in Exercise IV.21.4. A non-vanishing isochoric dilatation (IV.21.34) superposed on a rigid motion of spin $\mathbf{W}_\mathrm{r}$ provides a monotonous motion of NOLL's third class. For it $\mathfrak{W} = |\mathfrak{W}_\mathrm{r}|/\sqrt{2(a_1^2 + a_2^2 + a_1 a_2)}$, which for a fixed $\mathbf{W}_\mathrm{r}$ by choice of a_1 and a_2 gives $\mathfrak{W}$ an arbitrary value in $]0, \infty[$.

IV.21.9 (II.8-3), (II.8-5), and (IV.21-39) show that

$$\mathbf{F}_t(\tau) = \mathbf{1} + (\tau - t)v'(x_1)\mathbf{N}.$$

Cf. $(\text{IV.21-29})_1$.

IV.21.10 Using (II.8-3) and (II.8-4), integrate (IV.21-44). So as to calculate physical components of $\mathbf{F}_t(\tau)$ with respect to $\{\mathbf{e}_i(\boldsymbol{\xi})\}$ and $\{\mathbf{e}_i(\mathbf{x})\}$, evaluate the quantities $\mathbf{e}_i(\boldsymbol{\xi})\cdot\mathbf{F}_t(\boldsymbol{\xi})\mathbf{e}_j(\mathbf{x})$. Since the bases $\{\mathbf{e}_i(\boldsymbol{\xi})\}$ and $\{\mathbf{e}_j(x)\}$ are orthonormal, there is an orthogonal tensor function $\mathbf{Q}$ such that $\mathbf{e}_i(\boldsymbol{\xi}(\tau)) = \mathbf{Q}(\tau)\mathbf{e}_i(\mathbf{x})$. Thus

$$\mathbf{e}_i(\boldsymbol{\xi})\cdot\mathbf{F}_t(\tau)\mathbf{e}_j(\mathbf{x}) = \mathbf{Q}(\tau)\mathbf{e}_i(\mathbf{x})\cdot\mathbf{F}_t(\tau)\mathbf{e}_j(\mathbf{x}) = \mathbf{e}_i(\mathbf{x})\cdot\mathbf{Q}(\tau)^\mathsf{T}\mathbf{F}_t(\tau)\mathbf{e}_j(\mathbf{x}).$$

Now show that $\mathbf{F}_0(\tau) = \mathbf{Q}(\tau)(1 + \tau\kappa\mathbf{N}_0)$, where

$$[\mathbf{N}_0] = \begin{Vmatrix} 0 & 0 & 0 \\ \alpha & 0 & 0 \\ \beta & 0 & 0 \end{Vmatrix} \qquad \text{with respect to the basis } \{\mathbf{e}_i(\mathbf{x})\}.$$

Writing (IV.21-47) as $\mathbf{i}_k = \mathbf{R}\mathbf{e}_k$, calculate the matrix of $\mathbf{N}_0$ with respect to the basis $\{\mathbf{i}_k\}$.

IV.21.11 Inspect (IV.15-32) and (IV.15-61) to show that (IV.21-50) and (IV.21-51) are universal. Proceed as in Exercise IV.21.7 to show by use of (II.8-7) that the flows are monotonous and that $\mathbf{N}^2 = \mathbf{0}$.

IV.21.12 Inspect the solution of Exercise IV.18.3.

IV.21.13 Use (IV.21-54)$_1$ to perceive (IV.21-54)$_{2,3}$. Follow the method of Exercise IV.21.7 and use its notations to obtain

$$\begin{aligned}
\xi_1 &= (x_1 - f(x_3))C + (x_2 - g(x_3))S + f(x_3),\\
\xi_2 &= -(x_1 - f(x_3))S + (x_2 - g(x_3))C + g(x_3),\\
\xi_3 &= x_3\,,
\end{aligned}$$

which delivers (IV.21-55).

IV.21.14 From (IV.21-54)$_{2,3}$ conclude that

$$e^{\tau\mathbf{G}} = \mathbf{1} + \frac{1}{\Omega}(\sin(\Omega\tau))\mathbf{G} + \frac{1}{\Omega^2}(\cos(\Omega\tau) - 1)\mathbf{G}^2$$

and then use (IV.21-54)$_1$ and (IV.21-55). Since $\mathbf{G}^3 = -\Omega^2\mathbf{G}$, comparison of (IV.21-54) with (IV.21-3) shows that $\mathbf{N}_0$ is not nilpotent unless $\Omega = 0$. Use (IV.21-54) to calculate the components of $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, and $\mathbf{A}_4$ and so establish (IV.21-58) when $n = 1$. Next, suppose that if $n \geqq 4$

$$\mathbf{A}_{n-1} = -\Omega^2\mathbf{A}_{n-3}\,. \tag{$*$}$$

If so, (II.11-33) shows that

$$\mathbf{A}_n = -\Omega^2(\dot{\mathbf{A}}_{n-3} + \mathbf{G}^{\mathsf{T}}\mathbf{A}_{n-3} + \mathbf{A}_{n-3}\mathbf{G}) = -\Omega^2\mathbf{A}_{n-2},$$

whence by induction $(*)$ is proved to hold if $n \geqq 4$.

IV.22.1 The conclusions of Exercise IV.21.14 show that $\mathbf{A}_3 = -\Omega^2\mathbf{A}_1$, and so the third argument of (IV.22-1) may be replaced by Ω. Calculation of $\mathbf{A}_1$ and $\mathbf{A}_2$ shows that they together determine Ω unless $f' = g' = 0$. In that case the motion is rigid, and so the assertion of the exercise becomes trivial.

This Page Intentionally Left Blank

Gazetteer

of Axioms, Laws, Theorems in the order of their first appearances

This Page Intentionally Left Blank

Index

Names of Persons are cited only when they serve to identify theorems or concepts, not for attributions. The contents of the appendices are not indexed.

This Page Intentionally Left Blank

PURE AND APPLIED MATHEMATICS

Vol. 1	Arnold Sommerfeld, *Partial Differential Equations in Physics*
Vol. 2	Reinhold Baer, *Linear Algebra and Projective Geometry*
Vol. 3	Herbert Busemann and Paul Kelly, *Projective Geometry and Projective Metrics*
Vol. 4	Stefan Bergman and M. Schiffer, *Kernel Functions and Elliptic Differential Equations in Mathematical Physics*
Vol. 5	Ralph Philip Boas, Jr., *Entire Functions*
Vol. 6	Herbert Busemann, *The Geometry of Geodesics*
Vol. 7	Claude Chevalley, *Fundamental Concepts of Algebra*
Vol. 8	Sze-Tsen Hu, *Homotopy Theory*
Vol. 9	A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces, Third Edition of Solution of Equations and Systems of Equations*
Vol. 10	J. Dieudonné, *Treatise on Analysis*: *Volume I, Foundations of Modern Analysis*; *Volume II*; *Volume III*; *Volume IV*; *Volume V*; *Volume VI*; *Volume VII*
Vol. 11*	S. I. Goldberg, *Curvature and Homology*
Vol. 12†	Sigurdur Helgason, *Differential Geometry and Symmetric Spaces*
Vol. 13	T.H. Hildebrandt, *Introduction to the Theory of Integration*
Vol. 14	Shreeram Abhyankar, *Local Analytic Geometry*
Vol. 15*	Richard L. Bishop and Richard J. Crittenden, *Geometry of Manifolds*
Vol. 16*	Steven A. Gaal, *Point Set Topology*
Vol. 17	Barry Mitchell, *Theory of Categories*
Vol. 18	Anthony P. Morse, *A Theory of Sets*
Vol. 19	Gustave Choquet, *Topology*
Vol. 20	Z. I. Borevich and I. R. Shafarevish, *Number Theory*
Vol. 21	José Luis Massera and Juan Jorge Schaffer, *Linear Differential Equations and Function Spaces*
Vol. 22	Richard D. Schafer, *An Introduction to Nonassociative Algebras*
Vol. 23*	Martin Eichler, *Introduction to the Theory of Algebraic Numbers and Functions*
Vol. 24	Shreeram Abhyankar, *Resolution of Singularities of Embedded Algebraic Surfaces*

* Presently out of print.

† Out of print; please see Vol. 80.

Vol. 25 François Treves, *Topological Vector Spaces, Distributions, and Kernels*

Vol. 26 Peter D. Lax and Ralph S. Phillips, *Scattering Theory*

Vol. 27 Oystein Ore, *The Four Color Problem*

Vol. 28* Maurice Heins, *Complex Function Theory*

Vol. 29 R. M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory*

Vol. 30 L. J. Mordell, *Diophantine Equations*

Vol. 31 J. Barkley Rosser, *Simplified Independence Proofs: Boolean Valued Models of Set Theory*

Vol. 32 William F. Donoghue, Jr., *Distributions and Fourier Transforms*

Vol. 33 Marston Morse and Stewart S. Cairns, *Critical Point Theory in Global Analysis and Differential Topology*

Vol. 34* Edwin Weiss, *Cohomology of Groups*

Vol. 35 Hans Freudenthal and H. DeVries, *Linear Lie Groups*

Vol. 36 Laszlo Fuchs, *Infinite Abelian Groups*

Vol. 37 Keio Nagami, *Dimensional Theory*

Vol. 38 Peter L. Duren, *Theory of H^p Spaces*

Vol. 39 Bodo Pareigis, *Categories and Functors*

Vol. 40* Paul L. Butzer and Rolf J.Nessel, *Fourier Analysis and Approximation: Volum I, One-Dimensional Theory*

Vol. 41* Eduard Prugovečki, *Quantum Mechanics in Hilbert Space*

Vol. 42 D. V. Widder, *An Introduction to Transform Theory*

Vol. 43 Max D. Larsen and Paul J. McCarthy, *Multiplicative Theory of Ideals*

Vol. 44 Ernst-August Behrens, *Ring Theory*

Vol. 45 Morris Newman, *Integral Matrices*

Vol. 46 Glen E. Bredon, *Introduction to Compact Transformation Groups*

Vol. 47 Werner Greub, Stephen Halperin, and Ray Vanstone, *Connections, Curvature, and Cohomology: Volume I, De Rham Cohomology of Manifolds and Vector Bundles; Volume II, Lie Groups, Principal Bundles, and Characteristic Classes; Volume III, Cohomology of Principal Bundles and Homogeneous Spaces*

Vol. 48 Xia Dao-xing, *Measure and Integration Theory of Infinite-Dimensional Spaces: Abstract Harmonic Analysis*

Vol. 49 Ronald G. Douglas, *Banach Algebra Techniques in Operator Theory*

Vol. 50 Willard Miller, Jr., *Symmetry Groups and Theory Applications*

Vol. 51 Arthur A. Sagle and Ralph E. Walde, *Introduction to Lie Groups and Lie Algebras*

Vol. 52 T. Benny Rushing, *Topological Embeddings*
Vol. 53* James W. Vick, *Homology Theory: An Introduction to Algebraic Topology*
Vol. 54 E. R. Kolchin, *Differential Algebra and Algebraic Groups*
Vol. 55 Gerald J. Janusz, *Algebraic Number Fields*
Vol. 56 A. S. B. Holland, *Introduction to the Theory of Entire Functions*
Vol. 57 Wayne Roberts and Dale Varberg, *Convex Functions*
Vol. 58 H. M. Edwards, *Riemann's Zeta Function*
Vol. 59 Samuel Eilenberg, *Automata, Languages, and Machines: Volume A, Volume B*
Vol. 60 Morris Hirsch and Stephen Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*
Vol. 61 Wilhelm Magnus, *Noneuclidean Tesselations and Their Groups*
Vol. 62 François Treves, *Basic Linear Partial Differential Equations*
Vol. 63* William M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*
Vol. 64 Brayton Gray, *Homotopy Theory: An Introduction to Algebraic Topology*
Vol. 65 Robert A. Adams, *Sobolev Spaces*
Vol. 66 John J. Benedetto, *Spectral Synthesis*
Vol. 67 D. V. Widder, *The Heat Equation*
Vol. 68 Irving Ezra Segal, *Mathematical Cosmology and Extragalactic Astronomy*
Vol. 69 I. Martin Isaacs, *Character Theory of Finite Groups*
Vol. 70 James R. Brown, *Ergodic Theory and Topological Dynamics*
Vol. 71 C. Truesdell, *A first Course in Rational Continuum Mechanics: Volume 1, General Concepts; Second Edition, Corrected, Revised, and Augmented*
Vol. 72 K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the Theory of Infinitesimals*
Vol. 73 B. M. Puttaswamaiah and John D. Dixon, *Modular Representations of Finite Groups*
Vol. 74 Melvyn Berger, *Nonlinearity and Functional Analysis: Lectures on Nonlinearity Problems in Mathematical Analysis*
Vol. 75 George Gratzer, *Lattice Theory*
Vol. 76 Charalambos D. Aliprantis and Owen Burkinshaw, *Locally Solid Riesz Spaces*
Vol. 77 Jan Mikysinski, *The Bochner Integral*
Vol. 78 Michiel Hazewinkel, *Formal Groups and Applications*
Vol. 79 Thomas Jech, *Set Theory*
Vol. 80 Sigurdur Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*

Vol. 81 Carl L. DeVito, *Functional Analysis*
Vol. 82 Robert B. Burckel, *An Introduction to Classical Complex Analysis*
Vol. 83 C. Truesdell and R. G. Muncaster, *Fundamentals of Maxwell's Kinetic Theory of a Simple Monatomic Gas: Treated as a Branch of Rational Mechanics*
Vol. 84 Louis Halle Rowen, *Polynomial Identities in Ring Theory*
Vol. 85 Joseph J. Rotman, *An Introduction to Homological Algebra*
Vol. 86 Barry Simon, *Functional Integration and Quantum Physics*
Vol. 87 Dragos M. Cvetkovic, Michael Doob, and Horst Sachs, *Spectra of Graphs*
Vol. 88 David Kinderlehrer and Guido Stampacchia, *An Introduction to Variational Inequalities and Their Applications*
Vol. 89 Herbert Seifert, W. Threlfall, *A Textbook of Topology*
Vol. 90 Grzegorz Rozenberg and Arto Salomaa, *The Mathematical Theory of L Systems*
Vol. 91 Donald W. Kahn, *Introduction to Global Analysis*
Vol. 92 Eduard Prugovečki, *Quantum Mechanics in Hilbert Space, Second Edition*
Vol. 93 Robert M. Young, *An Introduction to Nonharmonic Fourier Series*
Vol. 94 M. C. Irwin, *Smooth Dynamical Systems*
Vol. 96 John B. Garnett, *Bounded Analytic Functions*
Vol. 97 Jean Dieudonné, *A Panorama of Pure Mathematics: As Seen by N. Bourbaki*
Vol. 98 Joseph G. Rosenstein, *Linear Orderings*
Vol. 99 M. Scott Osborne and Garth Warner, *The Theory of Eisenstein Systems*
Vol. 100 Richard V. Kadison and John R. Ringrose, *Fundamentals of the Theory of Operator Algebras: Volume 1, Elementary Theory; Volume 2, Advanced Theory*
Vol. 101 Howard Osborn, *Vector Bundles: Volume 1, Foundations and Stiefel-Whitney Classes*
Vol. 102 Avraham Feintuch and Richard Saeks, *System Theory: A Hilbert Space Approach*
Vol. 103 Barrett O'Neill, *Semi-Riemannian Geometry: With Applications to Relativity*
Vol. 104 K. A. Zhevlakov, A. M. Slin'ko, I. P. Shestakov, and A. I. Shirshov, *Rings That Are Nearly Associative*
Vol. 105 Ulf Grenander, *Mathematical Experiments on the Computer*
Vol. 106 Edward B. Manoukian, *Renormalization*
Vol. 107 E. J. McShane, *Unified Integration*

Vol. 108 A. P. Morse, *A Theory of Sets, Revised and Enlarged Edition*
Vol. 109 K. P. S. Bhaskara-Rao and M. Bhaskara-Rao, *Theory of Charges: A Study of Finitely Additive Measures*
Vol. 110 Larry C. Grove, *Algebra*
Vol. 111 Steven Roman, *The Umbral Calculus*
Vol. 112 John W. Morgan and Hyman Bass, editors, *The Smith Conjecture*
Vol. 113 Sigurdur Helgason, *Groups and Geometric Analysis: Integral Geometry, Invariant Differential Operators, and Spherical Functions*
Vol. 114 E. R. Kolchin, *Differential Algebraic Groups*
Vol. 115 Isaac Chavel, *Eigenvalues in Riemannian Geometry*
Vol. 116 W. D. Curtis and F. R. Miller, *Differential Manifolds and Theoretical Physics*
Vol. 117 Jean Berstel and Dominque Perrin, *Theory of Codes*
Vol. 118 A. E. Hurd and P. A. Loeb, *An Introduction to Nonstandard Real Analysis*
Vol. 119 Charalambos D. Aliprantis and Owen Burkinshaw, *Positive Operators*
Vol. 120 Willaim M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry, Second Edition*
Vol. 121 Douglas C. Ravenel, *Complex Cobordism and Stable Homotopy Groups of Spheres*
Vol. 122 Sergio Albeverio, Jens Erik Fenstad, Raphael Høegh-Krohn, and Tom Lindstrøm, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*
Vol. 123 Alberto Torchinsky, *Real-Variable Methods in Harmonic Analysis*
Vol. 124 Robert J. Daverman, *Decomposition of Manifolds*
Vol. 125 J. M. G. Fell and R. S. Doran, *Representations of *-Algebras, Locally Compact Groups, and Banach *-Algebraic Bundles: Volume 1, Basic Representation Theory of Groups and Algebras*
Vol. 126 J. M. G. Fell and R. S. Doran, *Representations of *-Algebras, Locally Compact Groups, and Banach *-Algebraic Bundles: Volume 2, Induced Representations, the Imprimitivity Theorem, and the Generalized Mackey Analysis*
Vol. 127 Louis H. Rowen, *Ring Theory, Volume I*
Vol. 128 Louis H. Rowen, *Ring Theory, Volume II*
Vol. 129 Colin Bennett and Robert Sharpley, *Interpolation of Operators*
Vol. 130 Jürgen Pöschel and Eugene Trubowitz, *Inverse Spectral Theory*
Vol. 131 Jens Carsten Jantzen, *Representations of Algebraic Groups*
Vol. 132 Nolan R. Wallach, *Real Reductive Groups I*

Vol. 133 Michael Sharpe, *General Theory of Markov Processes*

Vol. 134 Igor Frenkel, James Lepowsky, and Arne Meurman, *Vertex Operator Algebras and the Monster*

Vol. 135 Donald Passman, *Infinite Crossed Products*

Vol. 136 Heinz-Otto Kreiss and Jens Lorenz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*

Vol. 137 Jean-Dominique Deuschel and Daniel W. Stroock, *Large Deviations*