CAAM 436 Notes:
Partial Differential Equations of Mathematical Physics

William W. Symes
Department of Computational and Applied Mathematics
Rice University,
Spring 2012

Contents

Preface

This course aims to make students aware of the physical origins of the main partial differential equations of classical mathematical physics, including the fundamental equations of fluid and solid mechanics, thermodynamics, and classical electrodynamics. These equations form the backbone of modern engineering and many of the sciences, and solving them numerically is a central topic in scientific computation.

The goal of these notes is to establish the essential ideas with maximum clarity and reasonable rigor, with a conventional approach to notation and nomenclature, to serve as an introduction to the enormous literature on these topics. I have not attempted a comprehensive treatment of, say, continuum mechanics - such a treatment at the upper division - beginning graduate level of this course would occupy more than the semester available for this topic in our curriculum. Many excellent comprehensive texts in continuum mechanics, as well as in classical thermodynamics and in electromagnetism, will be accessible to the student who finishes this course. My goal is rather to get across the core idea cutting across all of these subjects - that the fundamental field equations result from combining a few basic physical principles, some clearly stated assummptions and approximations, and vector calculus.

I have also not presented this subject as a branch of modern analysis, with the rigor that would entail - although such an exposition is certainly possible [Marsden and Hughes, 1994]. Moreover, the information I have developed here about the *solutions* of these equations is quite limited. The big three examples in PDE - the Laplace, wave, and heat equations - turn up as special cases, and I do establish a few basic facts about their solutions. Additional properties of solutions, and methods to approximate them numerically, are developed in other parts of the CAAM and MATH curricula:

- CAAM 437, MATH 423: analysis of PDEs;
- CAAM 452, CAAM 552: numerical methods for PDEs and their analysis.

Students should enter this course with at least a nodding acquaintance with vector calculus, as taught in the fourth semester of a calculus sequence, and some background in dif-

ferential equations, equivalent to that obtained from CAAM 336. For applied mathematics students to gain optimally from the study of this material, mastery of rigorous multivariate advanced calculus (as provided by succcessful completion of a course similar to CAAM 402) is essential.

Nothing in my exposition of this fundamental, well-aged material is novel. I have benefitted greatly from the excellent text of M. Gurtin on continuum mechanics [Gurtin, 1981], and drawn liberally from many other references, citations to which are sprinkled throughout the text. In general, I have resisted the temptation to "improve" on the notational conventions common in the subject. My principal aim is to prepare students in this course for further reading and participation in the very active fields of science and engineering that base themselves on this material. While nothing like a standard syntax for topics of these notes exists, I have chosen notation that is as close in spirit to the "norm" as I could come.

William W. Symes
Houston, Texas
August 2006

Added January 2012: this version of the notes differs from the previous in that I have made a few minor changes of notation, added some problems, and enlarged the trailing chapters in various ways. The aim and spirit of the enterprise remains precisely the same.

Chapter 1

Kinematics and Conservation

Several of the physical principles underpinning continuum mechanics take the form of conservation laws, which state that some property of the material in motion is preserved during the motion, or changes in a prescribed way. The fundamental postulate of continuum mechanics - that the properties of matter can be localized in arbitrarily small volumes - has its mathematical expression in the use of *densities*, i.e. amount per unit volume, to represent mass, momentum, energy, etc. Materials compress, expand, and change shape during a motion, so the change in a property density also has a purely kinematic component, simply due to the motion. This chapter is devoted to the expression of basic conservation principles as partial differential equations, which account for both physical and geometric causes of density changes.

1.1 Deformation and Motion

This section introduces some of the mathematics necessary to describe motion of 3D continua. These are mostly standard ideas from vector calculus, expressed in perhaps unfamiliar notation suited to the application.

A "material body" is the volume occupied by a specific piece of material. Note that this definition says both more and less than it appears to say. Namely, it implies some notion of what a "material" is, and how it may be divided into pieces - without, however, supplying a clear description of either idea. In fact, the real content of the definition is simply that pieces of material, whatever that is, are identified with the subsets of space which they occupy. A further, fundamental notion is that any given piece of material may occupy (at different times, or under different conditions) many different subsets of space. These various subsets are the possible *configurations* of the body. One of the great ideas of this subject is that one

of these possible configurations, that is, subsets of space, should be singled out as a *reference configuration.*

In this discussion, "space" is Euclidean 3-space, or $\mathbf{R}^3$ in common notation. Say $\Omega \subset \mathbf{R}^3$ - is the reference configuration of a material body. We use capital letters to denote coordinates in this reference configuration or location. In effect, these coordinate vectors $\mathbf{X} = (X_1, X_2, X_3)^T$ *label* positions within the body. Because some material is supposed to be located at each position, these vectors $\mathbf{X} \in \Omega$ are called *material points.* So, once again, a (vague) physical notion, the material located at a point in space, is confounded with a precise geometrical one, namely the point's coordinate vector.

Another brilliant idea, which came about sometime in the development of this subject, is that the various possible configurations of a material body must be related by mappings, or vector-valued functions: if ω is a configuration of the body whose reference configuration is Ω, then each point $\mathbf{x} \in \omega$ must be the position occupied by the material which is located at some other position $\mathbf{X} \in \Omega$ in the reference configuration. That is, if it's the same material, each bit ("point") of it must be in both configurations, and it must be possible to say, unambiguously, that $\mathbf{x} \in \omega$ is the *same* material ("point") as $\mathbf{X} \in \Omega$. That is, there must be a mapping $\mathbf{x} : \Omega \to \mathbf{R}^3$ so that $\mathbf{x}(\Omega) = \omega$, and any material point $\mathbf{X} \in \Omega$ represents the same material as is located at $\mathbf{x}(\mathbf{X}) \in \omega$. Such a map is called a *deformation*, as the "new" configuration ω of the material is presumably a deformed version of the reference configuration Ω.

I draw your attention to several peculiarities of the notation. I've used capital letters, as noted before, to single out vectors $\mathbf{X}$ and their coordinates X_1, X_2, X_3 in the reference configuration, and small letters $\mathbf{x} = (x_1, x_2, x_3)^T$ for vectors describing points in arbitrary (deformed) configurations. This notational device is standard in the subject - see for example [Marsden and Hughes, 1994, Chung, 2007]. Also conventional, at least implicitly, is the confusion of vector variables and vector functions: depending on context, $\mathbf{x}$ can denote either the vector variable $\mathbf{x} = (x_1, x_2, x_3)^T$ or the vector-valued function $\mathbf{x} : \Omega \to \mathbf{R}^3$. This unfortunate habit at least has the virtue of reducing the amount of notation: Marsden and Hughes, for example, write instead $\phi : \Omega \to \mathbf{R}^3$ for the deformation function, but implicitly identify ϕ with $\mathbf{x}$ whenever convenient. By simply using $\mathbf{x}$ as the name of the deformation function as well as a typical coordinate vector in the deformed configuration, I avoid introducing additional notation like Marsden-Hughes' ϕ, which soon disappears anyway, but confront you with the necessity of determining meaning from context. This will be good practice for reading the literature of continuum mechanics. Finally, I insist that vectors are columns, unless otherwise indicated - that is, $\mathbf{x} = (x_1, x_2, x_3)^T$, not $= (x_1, x_2, x_3)$. This insistence is not conventional, but has an excellent justification, to appear shortly.

Deformations must satisfy several more conditions, which express natural physical notions in geometrically precise ways. Deformations must be *invertible*: physically, it must be possible to reverse a deformation to recover the reference configuration from a deformed one.

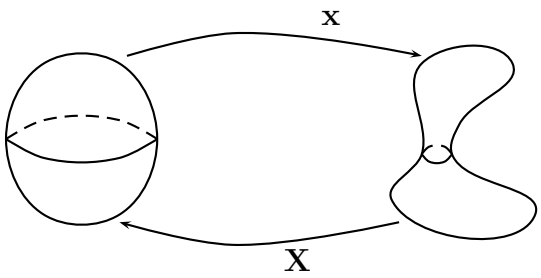

Figure 1.1: A deformation map $\mathbf{x}$ and its inverse $\mathbf{X}$.

Once again confusing variables and function names, I will use $\mathbf{X} : \omega \to \mathbf{R}^3$ as the name of the inverse of $\mathbf{x} : \Omega \to \mathbf{R}^3$, that is, $\mathbf{X} = \mathbf{x}^{-1}$. Thus if $\mathbf{Y} \in \Omega$, then $\mathbf{X}(\mathbf{x}(\mathbf{Y})) = \mathbf{Y}$. Deformations (and their inverses) must be continuous: this is a precise statement of the notion that if points $\mathbf{X}, \mathbf{X}'$ are close, then the corresponding points $\mathbf{x}(\mathbf{X}), \mathbf{x}(\mathbf{X}')$ in the deformed material should also be close - no ripping or tearing allowed! See Figure 1.1

Even more is required: until further notice, deformations are assumed (often tacitly) to be *continuously differentiable*, of class C^1, with inverses also of class C^1.

The derivative of $\mathbf{x}$, identified with its *Jacobian matrix* of partial derivatives, is conventionally denoted by $\nabla\mathbf{x}$ in this subject:

$$(\nabla\mathbf{x})_{i,j} = \frac{\partial x_i}{\partial X_j},\ i, j = 1, 2, 3$$

That is, the (i, j) element is the infinitesimal rate of change of the ith component of deformed position vector with respect to the jth component of the position vector of the same bit of material ("material point") in the undeformed configuration.

Note that many other notations ($D\mathbf{x}$, $J_{\mathbf{x}}$,...) are used in other contexts for the same thing. However, the ∇ notation is universal in continuum mechanics.

The derivative of a deformation function $\mathbf{x}$ is called the *deformation gradient* - note that this object already has a perfectly good name ("derivative"), but gets a special one for use just in continuum mechanics!

I'll use the same notation for arbitrary vector-valued functions of vector arguments, with arbitrary numbers of components in domain and range: that is, ∇ denotes the derivative, expressed as a matrix relative to some coordinate system, and the ith row contains the partial derivatives of the ith component, whereas the jth column contains the jth partial derivatives of all of the components.

With this notation, the chain rule takes a simple form: if $\mathbf{u} : \mathbf{R}^m \to \mathbf{R}^p$ (p-vector function of an m-vector), and $\mathbf{v} : \mathbf{R}^n \to \mathbf{R}^m$ (m-vector function of an n-vector) are differentiable,

then the *composition* $\mathbf{u}(\mathbf{v}(\mathbf{x}))$ is also differentiable, and

$$\nabla(\mathbf{u}(\mathbf{v}(\mathbf{x}))) = (\nabla\mathbf{u})(\mathbf{v}(\mathbf{x}))\nabla\mathbf{v}(\mathbf{x})$$

i.e. the derivative of the composition of two functions is just the matrix product of the derivatives of the two components in the order of composition.

Note that $\mathbf{u}$ and $\mathbf{v}$ don't need to be defined in all of the respective Euclidean spaces for the composition be well defined and for the chain rule to hold. It's enough that the *range* of $\mathbf{v}$ be contained in the *domain* of $\mathbf{u}$, so that the composition makes sense.

[In particular, if $\mathbf{X} : (a, b) \to \Omega$ is *differentiable curve* defined on the interval $(a, b) \subset \mathbf{R}$, then

$$\frac{d}{dt}(\mathbf{x}(\mathbf{X}(t)) = \nabla\mathbf{x}(\mathbf{X}(t))\frac{d\mathbf{X}}{dt}(t).$$

This works because coordinate vectors like $\mathbf{X}$ are *columns* - that's why it's the right convention!]

Notational Remark: A more satisfactory way to write this relationship suppresses the role of the dummy argument ($\mathbf{x}$ above). The composition of $\mathbf{u}$ and $\mathbf{v}$ is denoted $\mathbf{u} \circ \mathbf{v}$: that is, for any $\mathbf{x}$ in the domain of $\mathbf{v}$, $\mathbf{u} \circ \mathbf{v}(\mathbf{x}) \equiv \mathbf{u}(\mathbf{v}(\mathbf{x}))$. In this more natural notation, which gives a name to the composition, the chain rule reads

$$\nabla(\mathbf{u} \circ \mathbf{v}) = (\nabla\mathbf{u} \circ \mathbf{v})\nabla\mathbf{v}$$

Unfortunately this more rational approach to notation for continuum mechanics is confined to a few authors - for example [Marsden and Hughes, 1994]. The vast bulk of the literature on continuum mechanics uses the older standard syntax, with dummy variables and indices galore. Since the purpose of these notes is to prepare you, dear reader, for the literature as it actually is, rather than as the author wishes it were, I will stick with the notational conventions introduced in this chapter - less than ideal, but representative of what you will find in papers, books, and presentations on this subject.

There is one annoying exception to this nice organization of partial derivatives into arrays. For a scalar function $\Phi(\mathbf{x})$, the symbol $\nabla\Phi(\mathbf{x})$ denotes the *gradient*, which is the *column* vector of partial derivatives. That is, the gradient is an object of the same type as a position vector. This is not the natural specialization of the vector ∇ notation: a scalar function would naturally be identified with a 1-vector-valued function, and its derivative represented as a *row* vector. A row vector is the matrix representation of a linear map, from $\mathbf{R}^3$ to $\mathbf{R}$. The derivative of a real-valued function is exactly such a linear map, so indeed the natural representation of the derivative should in principle be a row vector. However the transpose of this row vector is the *Riesz representer* of the derivative: its inner product with a direction (vector) $\mathbf{w} \in \mathbf{R}^3$ gives the directional derivative in that direction,

$$(\nabla\Phi(\mathbf{x}))^T\mathbf{w} = \lim_{\epsilon\to 0}\frac{\Phi(\mathbf{x} + \epsilon\mathbf{w}) - \Phi(\mathbf{x})}{\epsilon}$$

For various reasons, the Riesz representer (the fastest increase direction of Φ, in the Euclidean metric) is more useful than the natural matrix representation of the derivative, and is conventionally signified by ∇ (or grad).

So $\nabla\Phi$ is the *transpose* of what you would expect, and the chain rule reads

$$\nabla(\Phi(\mathbf{v}(\mathbf{x})) = \nabla\mathbf{v}(\mathbf{x})^T(\nabla\Phi)(\mathbf{v}(\mathbf{x})).$$

You should satisfy yourself that this matrix product actually does express the componentwise chain rule

$$\frac{\partial}{\partial x_i}\Phi(\mathbf{v}(\mathbf{x})) = \sum_j \frac{\partial\Phi}{\partial v_j}(\mathbf{v}(\mathbf{x}))\frac{\partial v_j}{\partial x_i}(\mathbf{x})$$

I'll also need the change of variables formula for multiple integrals: if ω is the image of Ω under the deformation $\mathbf{x}$, then

$$\int_\omega dx\,\Phi(\mathbf{x}) = \int_\Omega dX\,\Phi(\mathbf{x}(\mathbf{X}))\,\det\nabla\mathbf{x}(\mathbf{X})$$

[Note that on the LHS of this equation, $\mathbf{x}$ is a dummy variable representing an arbitrary point in ω, whereas on the RHS it refers to the deformation (vector-valued function) mapping Ω to ω, and $\mathbf{X}$ is the dummy variable. This type of contextual reading will occur over and over again.]

Finally, I will often use the following *Averaging Theorem*: if f and g are (scalar or vector valued) continuous functions defined on the same ("reasonable") domain $\Omega \subset \mathbf{R}^3$, and for any *cube* $C \subset \Omega$

$$\int_C f = \int_C g$$

then $f(\mathbf{X}) = g(\mathbf{X})$ for all $\mathbf{X} \in \Omega$.

Note that if you divide both sides by the volume of ω, the hypothesis becomes: the averages of f and g over any cube are the same - thus the name of the theorem.

Other classes of sets, other than cubes, for example all balls, or even more general classes of sets, will work as well.

The meaning of "reasonable" here, i.e. the sets over which integration has a well-defined meaning, is a complicated issue which I am not going to address in this course. Cubes, balls, images of these under continuously differentiable and invertible maps, and unions and intersections of finite numbers of these, are certainly "reasonable". Courses on measure and integration, such as MATH 425 at Rice, make these notions precise. Good texts are [Rudin, 1986, Royden, 1988, LANG, 1993, **?**].

One final condition on deformations: they are required to be *orientation preserving*, which means that

$$\det \nabla \mathbf{x} > 0$$

This condition prevents deformation from turning the material "inside out". An example of a non-orientation preserving map figures in one of the exercises - according to this condition, it cannot define a deformation.

A lot of this subject will concern *motions* of material continua. A motion of a material body is simply a time-parameterized family of deformations, $\mathbf{x}(\cdot, t)$, defined for an interval (t_0, t_1) of "time" ($\mathbf{R}$!) - that is, for each $t \in (t_0, t_1)$. $\mathbf{x}(\cdot, t) : \Omega \to \mathbf{R}^3$ is a deformation. The range, or image, of the reference configuration under the motion is a time-parameterized family $\omega_t = \mathbf{x}(\Omega, t)$ of subsets of $\mathbf{R}^3$, that is, a curve in the set of configurations.

Motions will be assumed differentiable (continuously, if necessary several times) without further mention, as functions of all arguments.

1.2 Conservation of Mass, Material Description

For present purposes, mass is simply an attribute of material bodies (that is, subsets of $\mathbf{R}^3$) with certain properties, to be enunciated. Most of the properties flow from assuming the existence of a *mass density* function $\rho_{\text{ref}} : \Omega \to \mathbf{R}$ on the reference configuration. This function is assumed to be continuously differentiable and positive. It gives the mass of any "reasonable" part $p \subset \Omega$ through

$$m_0(p) = \int_p dX \, \rho_{\text{ref}}(\mathbf{X}).$$

It follows from what we have assumed so far that there also exists a mass density for the deformed states, $\rho(\mathbf{x}, t)$. The integral of which over a part of the deformed body gives the mass of that part: for $p \subset \omega_t$,

$$\int_p dx \, \rho(\mathbf{x}, t)$$

is the mass in p at time t.

This is particularly true if $p = p_t$, where $p_t = \mathbf{x}(p, t) \equiv \{\mathbf{x}(\mathbf{X}, t) : \mathbf{X} \in p\}$ is a *moving part* of the body, i.e. the volume occupied at time t by the material which occupied some part p of the undeformed body. See Figure 1.2.

VERY IMPORTANT: Throughout this section p_t will denote such a moving part - i.e. it is a time-parameterized family of subsets of $\mathbf{R}^3$ corresponding to a part of the *material*, moving with the material. The same material is in p_t for all t. In particular p_t is not a fixed

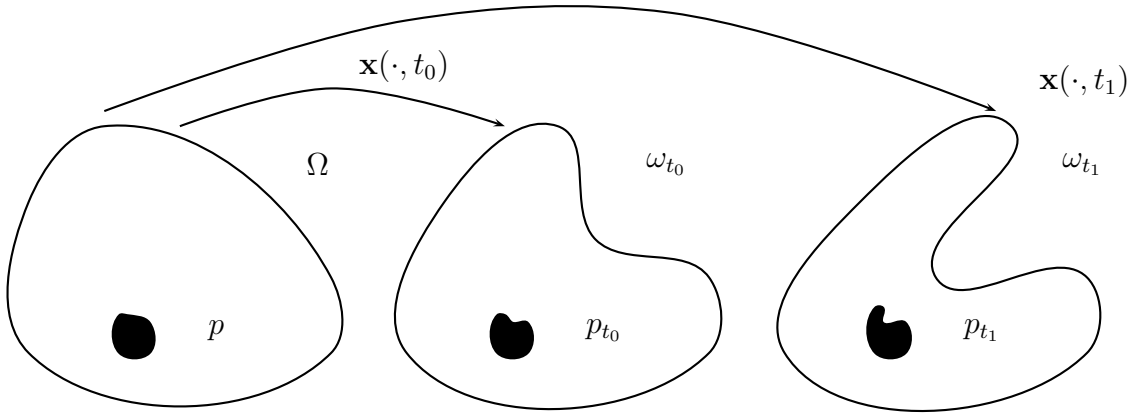

Figure 1.2: A body in the undeformed state (Ω) and a part of it (p), and at two times $t_0 < t_1$ during a motion. The moving part p_t, or $t = t_0, t_1$ and any other t, is presumed to contain *exactly the same material* as the corresponding part p in the undeformed configuration of the body.

subset of $\mathbf{R}^3$ except in case the motion preserves it, and its variation with t is determined by the motion and is not arbitrary.

So we can state a (material) description of the conservation of mass as follows:

$$\int_p dX\, \rho_{\text{ref}}(\mathbf{X}) = \int_{p_t} dx\, \rho(\mathbf{x}, t)$$

The change-of-variable rule for multiple integrals gives an alternate form in terms on integration over p only:

$$\int_p dX\, \rho_{\text{ref}}(\mathbf{X}) = \int_p dX\, \det\nabla\mathbf{x}(\mathbf{X}, t)\rho(\mathbf{x}(\mathbf{X}, t), t)$$

This relation must hold for "any" subset $p \subset \Omega$. Certainly it holds when p is a sufficiently small cube. The Averaging Theorem then implies that the (continuous) integrands must actually be identical:

$$\rho_{\text{ref}}(\mathbf{X}) = \det\nabla\mathbf{x}(\mathbf{X}, t)\rho(\mathbf{x}(\mathbf{X}, t), t).$$

Since the $\mathbf{x}(\cdot, t)$ is invertible and its Jacobian is everywhere positive, this implies in turn that

$$\rho(\mathbf{x}, t) = \rho_{\text{ref}}(\mathbf{X}(\mathbf{x}, t))\det\nabla\mathbf{X}(\mathbf{x}, t).$$

[Recall that $\mathbf{X}$ is the *inverse deformation*, i.e. the inverse function to $\mathbf{x}$.]

The preceding equation is the *material description* of the conservation-of-mass principle: it defines a quantity (mass density) in the deformed configuration explicitly in terms of the reference configuration and the relation between the two. It is also possible to produce a

spatial description of this conservation law, that is, one that refers *only* to quantities defined in the deformed configuration (in this case, $\rho(\mathbf{x}, t)$, not mentioning the reference configuration at all. The principle is general, and is explained in the next section.

1.3 Reynolds' Transport Theorem

Reynolds' Transport Theorem provides a general rule for passing between material and spatial descriptions of conservation principles.

Suppose that $\mathbf{x} : \Omega \times (t_0, t_1) \to \mathbf{R}^3$ is a motion, $p_t = \mathbf{x}(p, t) \subset \omega_t = \mathbf{x}(\Omega, t)$ is a moving part, and $\Phi(\mathbf{x}, t)$ is a continuously differentiable function on an open set in $\mathbf{R}^4$ containing $\{\mathbf{x}, t : \mathbf{x} \in p_t, t_0 < t < t_1\}$. Then the function

$$Q(t) \equiv \int_{p_t} dx\, \Phi(\mathbf{x}, t)$$

is differentiable. The main result of this section is an expression for its rate of change in terms of partial derivatives of Φ and of the motion.

The easiest way to proceed is to refer the integration back to the undeformed configuration, so that at least the domain does not depend on t:

$$Q(t) = \int_p dX\, \det\nabla\mathbf{x}(\mathbf{X}, t)\Phi(\mathbf{x}(\mathbf{X}, t), t)$$

The t-derivative will then simply pass under the integral sign:

$$\frac{dQ}{dt}(t) = \int_p dX\, \left(\frac{\partial}{\partial t}\det\nabla\mathbf{x}(\mathbf{X}, t)\right)\Phi(\mathbf{x}(\mathbf{X}, t), t) + \det\nabla\mathbf{x}(\mathbf{X}, t)\left(\frac{\partial}{\partial t}\Phi(\mathbf{x}(\mathbf{X}, t), t)\right)$$

Evidently we will need to understand how to differentiate determinants of matrix-valued functions of t. In fact

$$\frac{\partial}{\partial t}\det\nabla\mathbf{x}(\mathbf{X}, t) = \det\nabla\mathbf{x}(\mathbf{X}, t)\mathrm{tr}\left(\frac{\partial}{\partial t}\nabla\mathbf{x}(\mathbf{X}, t)(\nabla\mathbf{x}(\mathbf{X}, t))^{-1}\right),$$

where tr denotes the *trace*, i.e. the sum of the diagonal entries. I defer the proof of this formula to the next section.

Since $\mathbf{x}$ and $\mathbf{X}$ are inverse functions, their Jacobians are inverse matrices as noted above, so

$$= \det\nabla\mathbf{x}(\mathbf{X}, t)\mathrm{tr}\left(\nabla_{\mathbf{X}}\frac{\partial\mathbf{x}}{\partial t}(\mathbf{X}, t)(\nabla\mathbf{X})(\mathbf{x}(\mathbf{X}, t), t)\right)$$

The chain rule says that for any vector-valued function $\mathbf{w}$,

$$\nabla \mathbf{w}(\mathbf{X}(\mathbf{x},t)) = (\nabla \mathbf{w})(\mathbf{X}(\mathbf{x},t))\nabla \mathbf{X}(\mathbf{x},t)$$

so substituting $\mathbf{x} = \mathbf{x}(\mathbf{X},t)$ on both sides,

$$(\nabla \mathbf{w}(\mathbf{X}(\mathbf{x},t)))_{\mathbf{x}=\mathbf{x}(\mathbf{X},t)} = \nabla \mathbf{w}(\mathbf{X})(\nabla \mathbf{X})(\mathbf{x}(\mathbf{X},t),t)$$

The right hand side of the last expression for the time derivative of the Jacobian determinant has the form of the right hand side above, with $w = \partial \mathbf{x}/\partial t$. Thus

$$\frac{\partial}{\partial t}\det \nabla \mathbf{x}(\mathbf{X},t) = \det \nabla \mathbf{x}(\mathbf{X},t)\mathrm{tr}\nabla \left(\frac{\partial \mathbf{x}}{\partial t}(\mathbf{X}(\mathbf{x},t),t) \right)_{\mathbf{x}=\mathbf{x}(\mathbf{X},t)}$$

An interesting piece of this expression has a significance of its own: define the *spatial velocity field*:

$$\mathbf{v}(\mathbf{x},t) \equiv \frac{\partial \mathbf{x}}{\partial t}(\mathbf{X}(\mathbf{x},t),t).$$

This is actually the velocity of the "material point" which happens to be located at $\mathbf{x}$ at time t. Indeed, the velocity of the point corresponding to location $\mathbf{X}$ in the reference configuration is

$$\mathbf{V}(\mathbf{X},t) = \frac{\partial \mathbf{x}}{\partial t}(\mathbf{X},t);$$

this field is known as the *material velocity*. If a "material point" is at $\mathbf{x}$ at time t, then it was at $\mathbf{X}(\mathbf{x},t)$ in the reference configuration, so the velocity is precisely $\mathbf{v}(\mathbf{x},t)$ given above. That is,

$$\mathbf{v}(\mathbf{x},t) = \mathbf{V}(\mathbf{X}(\mathbf{x},t),t).$$

Recall from vector calculus the *divergence operator* $\nabla\cdot$: for any 3-vector field $\mathbf{w}$,

$$\nabla \cdot \mathbf{w} \equiv \sum_{i=1}^{3} \frac{\partial w_i}{\partial x_i} = \mathrm{tr}\nabla \mathbf{w}.$$

With these notations, the expression for the time derivative of the Jacobian determinant turns into the remarkable formula

$$\frac{\partial}{\partial t}\det \nabla \mathbf{x}(\mathbf{X},t) = \det \nabla \mathbf{x}(\mathbf{X},t)(\nabla \cdot \mathbf{v})(\mathbf{x}(\mathbf{X},t),t)$$

This handles the first half of the expression for the time derivative of Q. The second half involves

$$\begin{aligned}\frac{\partial}{\partial t}\Phi(\mathbf{x}(\mathbf{X},t),t) &= \frac{\partial \mathbf{x}}{\partial t}(\mathbf{X},t)^T(\nabla \Phi)(\mathbf{x}(\mathbf{X},t),t) + \frac{\partial \Phi}{\partial t}(\mathbf{x}(\mathbf{X},t),t) \\ &= \left(\frac{\partial \Phi}{\partial t} + \mathbf{v}\cdot\nabla\Phi \right)(\mathbf{x}(\mathbf{X},t),t)\end{aligned}$$

This combination of partial derivatives occurs so often that it gets a special name and notation: it is called the *material derivative*, represents the derivative along the trajectory $\mathbf{x}(\mathbf{X}, t)$ of a material particle, and is denoted

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$$

Putting all of this together,

$$\frac{dQ}{dt}(t) = \int_p dX \, \det \nabla \mathbf{x}(\mathbf{X}, t) \left(\frac{D\Phi}{Dt} + (\nabla \cdot \mathbf{v})\Phi \right) (\mathbf{x}(\mathbf{X}, t), t)$$

This is a setup for changing the integration back to p_t. This leads to the final expression for the rate of change of Q, which is *Reynolds' Transport Theorem*:

$$\frac{d}{dt} \int_{p_t} dx \, \Phi(\mathbf{x}, t) = \int_{p_t} dx \, \left(\frac{D\Phi}{Dt} + (\nabla \cdot \mathbf{v})\Phi \right) (\mathbf{x}, t)$$

$$= \int_{p_t} dx \, \left(\frac{\partial \Phi}{\partial t} + \nabla \cdot (\mathbf{v}\Phi) \right) (\mathbf{x}, t)$$

The last equality follows from one of the vector calculus identities which you will establish in Problem Set 1.

1.4 Differentiating Determinants

Suppose that $A(t)$ is a continuously differentiable $n \times n$ matrix valued function of t, defined on an open interval containing $t = 0$. In this section I'll derive a surprisingly simple formula for the derivative of $\det A(t)$ at $t = 0$. This formula played an important role in the derivation of Reynolds' Transport Theorem.

Before establishing determinant derivative formula in general, I should mention a special case in which you can see the result without any trouble at all. Suppose that $A(t)$ is continuously differentiable and *diagonalizable*, with continuously differentiable diagonalizing matrix $T(t)$ so that

$$T^{-1}(t)A(t)T(t) = \text{diag}(\alpha_1(t), ..., \alpha_n(t))$$

with continuously differentiable and positive eigenvalues $\alpha_1, ..., \alpha_n$. Then

$$\det A(t) = \alpha_1(t)\alpha_2(t)...\alpha_n(t)$$

so

$$\frac{d}{dt} \det A(t) = \frac{d\alpha_1}{dt}(t)\alpha_2(t)...\alpha_n(t)$$

$$+... + \alpha_1(t)\alpha_2(t)...\frac{d\alpha_n}{dt}(t)$$
$$= \det A(t)\left(\frac{1}{\alpha_1(t)}\frac{d\alpha_1}{dt}(t) + ... + \frac{1}{\alpha_n(t)}\frac{d\alpha_n}{dt}(t)\right)$$
$$= \det A(t)\mathrm{tr}\left(\frac{dA}{dt}(t)A^{-1}(t)\right).$$

In fact the formula just derived is true, regardless of whether A is diagonalizable or not (though without that assumption, it requires a different proof!).

It is necessary to assume that $\det A(0) > 0$, from which it follows that $\det A(t)$ is positive for small t. Note that for our application, in which $A(t) = \nabla\mathbf{x}(\mathbf{X}, t)$, this is one of the defining properties of a motion.

The first order Taylor expansion of A is

$$A(t) = A(0) + tA'(0) + R(t)$$

where $R(t) = o(t)$. Thus

$$\det A(t) = \det A(0)\det(I + t(A^{-1}(0)A'(0) + t^{-1}A^{-1}(0)R(t)))$$

Set $\lambda = -1/t$. The scaling property of determinants ($\det(cA) = c^n \det A$) gives

$$\det A(t) = (-t)^n \det A(0)\det(\lambda I - A^{-1}(0)(A'(0) + t^{-1}R(t)))$$

Fact from Linear Algebra: the coefficient of λ^{n-1} in the characteristic polynomial $\det(\lambda I - M)$ of an $n \times n$ matrix M is the negative of the trace $\mathrm{tr}M = \sum M_{ii}$. So

$$\det A(t) = (-t)^n \det A(0)(\lambda^n - \lambda^{n-1}\mathrm{tr}A(0)^{-1}(A'(0) + t^{-1}R(t)) + P(\lambda)$$

where $P(\lambda)$ represents a polynomial in λ of degree $\leq n-2$. Multiplying through bt $(-t)^n = \lambda^{-n}$,

$$= \det A(0)(1 + t\mathrm{tr}A(0)^{-1}(A'(0) + t^{-1}R(t)) + O(t^2))$$
$$= \det A(0)(1 + t\mathrm{tr}A(0)^{-1}A'(0) + o(t))$$

Comparing with the Taylor series

$$= \det A(0) + t\frac{d\det A}{dt}(0) + o(t)$$

and noting that there is nothing special about $t = 0$ gives the desired identity:

$$\frac{d\det A}{dt} = \det A\mathrm{tr}(A^{-1}\frac{dA}{dt})$$

$$= \det A \mathrm{tr}(\frac{dA}{dt}A^{-1})$$

(since $\mathrm{tr}AB = \mathrm{tr}BA$ for any $n \times n$ matrices A, B). Otherwise put,

$$\frac{d}{dt}\log\det A = \mathrm{tr}(A^{-1}\frac{dA}{dt})$$

1.5 Conservation of Mass, Spatial Description

This follows directly from the Transport Theorem. Simply take $\Phi(\mathbf{x}, t) = \rho(\mathbf{x}, t)$ - from the material description of conservation of mass,

$$Q(t) = \int_{p_t} \rho = \int_p \rho_{\mathrm{ref}}$$

is independent of time, being the mass "transported" in the deforming part p_t of the material. So

$$0 = \frac{d}{dt}\int_{p_t} \rho = \int_{p_t} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{v}\rho)\right).$$

This must hold for "all" (regular enough?) parts p_t - it certainly holds when p_t is a small cube. That is enough (Averaging Theorem!) to show that

$$0 = \frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{v}\rho) = \frac{D\rho}{Dt} + (\nabla \cdot \mathbf{v})\rho$$

which gives two forms of mass conservation in the spatial description.

1.6 Volume change and the significance of the velocity divergence

You get the volume of a region in $\mathbf{R}^3$ by integrating the constant function 1 over it. So for our moving part p_t,

$$V(t) = \int_{p_t} dx\, 1$$

You can apply Reynolds' Transport Theorem to this integral too, as $\Phi \equiv 1$ is a perfectly legitimate choice of function. Of course there is not reason for the volume to remain constant, and indeed moving parts may expand or contract during a motion. The rate of change of volume is

$$\frac{dV}{dt}(t) = \int_{p_t} dx\, \nabla \cdot \mathbf{v}(\mathbf{x}, t).$$

That is, *the velocity divergence is the density for volume rate of change!*

With this insight, you can see the meaning of the terms in the RHS of the Transport Theorem: the first ($D\Phi/Dt\,dx$) is the intrinsic rate of change of the density ($\Phi\,dx$) itself along material particle trajectories, whereas the second ($(\nabla\cdot\mathbf{v})\Phi\,dx$) reflects the rate of density change due solely to changing geometry, i.e. contraction or expansion of the volume occupied by the material.

A very important class of motions also reveals its characterisation courtesy of this identity. A motion is *volume preserving* if and only if for *any* part p_t, the volume is constant in time: $dV/dt \equiv 0$. The identity above and the Averaging Theorem imply that this is the case exactly when

$$\nabla\cdot\mathbf{v}(\mathbf{x},t) = 0$$

A material is called *incompressible* if it admits only volume-preserving motions. Many common fluids, eg. water, are approximately incompressible under ordinary "room" conditions.

1.7 Conservation of Linear Momentum

We don't need to assume the existence of a momentum density - this is assured by the existence of a mass density, and even uniquely determined. The mass per unit volume at $\mathbf{x}$ at time t is $\rho(\mathbf{x},t)$, and the velocity of material at that time and place is $\mathbf{v}(\mathbf{x},t)$. Therefore the momentum density, i.e. momentum per unit volume, of the material at $\mathbf{x}$ at time t is $\rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)$, and so the momentum of a part p_t of the body at time t is

$$\mathbf{p}_{p_t}(t) = \int_{p_t} dx\,\rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)$$

Note that this is already a spatial description, not referring explicitly to the location of material particles in the undeformed body. As an excercise, find a material description of momentum.

Momentum is not an invariant of motion - instead, its evolution is governed by Newton's second law of motion, the famous "F=ma", interpreted as applying to a material body or part of it. "Mass times acceleration" is the same as "rate of change of momentum". Thus Newton's law for the part p_t of the material reads

$$\frac{d\mathbf{p}_{p_t}}{dt}(t) = \mathbf{F}[p_t]$$

where the right hand side is simply shorthand for the sum of all forces acting on the part p_t at time t. These forces will get careful attention shortly. Meanwhile express this relation in

terms of ρ and $\mathbf{v}$, using the vector version of Reynolds' Transport Theorem (Problem 1-9):

$$\mathbf{F}[p_t] = \int_{p_t} dx \, \frac{D}{Dt}(\rho \mathbf{v}) + (\nabla \cdot \mathbf{v})\rho \mathbf{v}$$

This statement uses the material derivative applied to a vector field (the momentum density $\rho\mathbf{v}$: for any vector field (vector-valued function) $\mathbf{w}$,

$$\frac{D\mathbf{w}}{Dt} \equiv \frac{\partial \mathbf{w}}{\partial t} + (\nabla \mathbf{w})\mathbf{v}$$

An alternate statement of the vector Transport Theorem is also possible, as in the scalar case.

If the force acting on p_t were given by integration of a density, $\mathbf{f}$ say, over p_t, then you could use the Averaging Theorem: since p_t is arbitrary, you would conclude that

$$\mathbf{f} = \frac{D}{Dt}(\rho \mathbf{v}) + (\nabla \cdot \mathbf{v})\rho \mathbf{v}$$

In Problem Set 1 you will show that a consequence of mass conservation is the identity

$$\frac{D}{Dt}(\rho \Phi) + (\nabla \cdot \mathbf{v})\rho \Phi = \rho \frac{D\Phi}{Dt}$$

which holds for *any* scalar- or vector-valued function Φ. Thus momentum balance takes the quite pleasing form

$$\mathbf{f} = \rho \frac{D\mathbf{v}}{Dt}$$

which is about as close to "F=ma" as one ought to expect to get.

This is the form that most of the famous PDEs ("laws of motion") take, such as the Euler and Navier-Stokes equations and the various versions of elastodynamics. It will take some work to understand the internal structure of the force density $\mathbf{f}$.

In fact it's actually possible to view momentum conservation in a way that looks even more like classical mechanics. The *center of mass* of the body is the average position within the body, weighted (as it were!) by the mass density:

$$\mathbf{q}(t) = \frac{1}{m} \int_{\omega_t} dx \, \mathbf{x} \rho(\mathbf{x}, t)$$

where

$$m = \int_{\omega_t} dx \, \rho(\mathbf{x}, t)$$

is the mass of the object. Note that we have integrated over the *entire body*, occupying volume ω_t at time t, so no mass is lost or gained - and the m in the last equation is really as independent of t as the notation indicates!.

The rate of change of the center of mass is

$$\frac{d\mathbf{q}}{dt}(t) = \frac{1}{m}\int_{\omega_t} dx\, \rho(\mathbf{x},t)\frac{D\mathbf{x}}{Dt}$$

According to the definition of the material derivative for vector-valued functions (like $\mathbf{x}$ - here a dummy variable denoting position, not the material particle trajectory $\mathbf{x}(\mathbf{X},t)$!),

$$\frac{D\mathbf{x}}{Dt} = \frac{\partial \mathbf{x}}{\partial t} + (\nabla \mathbf{x})\mathbf{v} = \mathbf{v}$$

since the Jacobian of the position coordinate vector $\mathbf{x}$ is I. Thus

$$\frac{d\mathbf{q}}{dt}(t) = \frac{1}{m}\int_{\omega_t} dx\, \rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)$$

Recalling the definition of momentum $\mathbf{p}(t)$ (in this case $p_t = \omega_t$) and clearing the denominator, you get

$$m\frac{d\mathbf{q}}{dt} = \mathbf{p}$$

which reads "mass times velocity equals momentum". Thus from this point of view, the body behaves as if it were a classical mechanical particle with mass m concentrated at the center of mass, whose momentum is the total momentum of the body. In this notation, conservation of momentum amounts to

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = m\frac{d^2\mathbf{q}}{dt^2}$$

which really *is* "F=ma"!

1.8 Problems

Problem 1-1: Suppose that Ω (the volume occupied by the undeformed body) is the unit cubic meter: $\{\mathbf{X} : 0\text{m} \leq x_{0,i} \leq 1\text{m}, i = 1, 2, 3\}$, which deforms by stretching uniformly under the motion

$$\mathbf{x}(\mathbf{X}, t) = (1 + t)\mathbf{X}.$$

(a uniform expansion or dilatation). The initial density is homogeneous, $\rho_{\text{ref}} \equiv 1000\text{kg/m}^3$. (a) **Verify** that this formula defines a motion for $t \geq 0$. **Compute** (b) $\mathbf{v}(\mathbf{x}, t)$ (units of m/s), (c) $\rho(\mathbf{x}, t)$.

Problem 1-2: Define $f : \mathbf{R}^3 \to \mathbf{R}^3$ by

$$f(\mathbf{x}) = \begin{bmatrix} x_1^2 - 2x_3 \\ x_3^3 - x_1^3 \\ 2x_2^2 \end{bmatrix}$$

and set

$$x_0 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

Find the best linear approximation (i.e. the first order Taylor series) of f at $\mathbf{X}$.

Problem 1-3: Define the map $\mathbf{x} : \mathbf{B}^3 \times \mathbf{R} \to \mathbf{R}^3$ by

$$\mathbf{x}(\mathbf{X}, t) = \begin{bmatrix} x_{0,1} + tx_{0,1}^2 \\ x_{0.2} + t(x_{0.2} + 2x_{0,3}) \\ x_{0,3} + tx_{0,1} \end{bmatrix}$$

where $\mathbf{B}^3 = \{\mathbf{X} \in \mathbf{R}^3 : \|\mathbf{X}\| \leq 1\}$ is the unit ball in $\mathbf{R}^3$.

(i) **Show** that $\mathbf{x}$ is a motion for small t. Is $\mathbf{x}$ a motion for all t? For all positive t? [**Hint:** The issue is whether $\mathbf{x}$ is invertible *on the reference configuration*, i.e. with its given domain. If $\mathbf{X} \in \mathbf{B}^3$, then $|x_{0,1}| \leq 1$. If t is small enough, the "other" root of the quadratic giving $x_{0,1}$ in terms of x_1 is too big to be the first component of a vector in the unit ball.]

(ii) Define $\phi : \mathbf{R}^3 \times \mathbf{R} \to \mathbf{R}$ by

$$\phi(\mathbf{x}, t) = 2x_1^2 + x_2 + tx_3$$

Compute (a) $(\partial\phi/\partial t)(\mathbf{x}(\mathbf{X}, t), t)$, (b) $\partial/\partial t(\phi(\mathbf{x}(\mathbf{X}, t), t))$ (i.e. formulas, good for all $\mathbf{X}, t$ - note carefully the order of the parentheses! In one case you compute the partial derivative,

then substitute $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$, whereas in the other you substitute first then compute the partial derivative).

(iii) What would be involved in computing $\mathbf{v}(\mathbf{x}, t)$ for this motion? Can you give a formula for $\mathbf{v}$? ["No" is an acceptable answer, if you can justify it - otherwise produce $\mathbf{v}(\mathbf{x}, t)$!]

Problem 1-4: Take Ω to be all of $\mathbf{R}^3$ except the origin, and define

$$\mathbf{x}(\mathbf{X}) = \frac{\mathbf{X}}{\|\mathbf{X}\|^2}.$$

Note that this map sends points inside the unit sphere to points outside, and points near the origin to points very far away - it turns Ω inside out! Show that

$$\det \nabla \mathbf{x}(\mathbf{X}) = -\|\mathbf{X}\|^{-6}$$

and conclude that this map is *not* a deformation, even though its Jacobian determinant does not vanish anywhere in its domain - it is *not* orientation preserving!

Problem 1-5: Suppose that the motion $\mathbf{x}(\mathbf{X}, t)$ is volume-preserving, and that the density at time $t = 0$ is independent of position $\mathbf{x}$ (*homogeneous*), i.e. $\rho(\mathbf{x}, 0) =$ constant (independent of $\mathbf{x}$). **Show** that the density remains homogenous at all times, i.e. $\rho(\mathbf{x}, t) =$ constant (independent of $\mathbf{x}$ and t). **Hint:** Use the spatial description of the conservation of mass to conclude that $D\rho/Dt = 0$. Then **interpret** the material derivative as rate of change along the trajectory of a material particle.

Problem 1-6: Suppose that $\Phi(\mathbf{x}, t)$ is a differentiable scalar- or vector-valued function, defined on a subset of $\mathbf{R}^4$ including all positions of the body. **Show** that for any part p_t,

$$\frac{d}{dt}\int_{p_t} dx\, \rho(\mathbf{x}, t)\Phi(\mathbf{x}, t) = \int_{p_t} dx\, \rho(\mathbf{x}, t)\frac{D\Phi}{Dt}(\mathbf{x}, t)$$

Deduce that

$$\frac{D}{Dt}(\rho\Phi) + (\nabla \cdot \mathbf{v})\rho\Phi = \rho\frac{D\Phi}{Dt}.$$

Problem 1-7: Suppose that Ω (the volume occupied by the undeformed body) is the cylinder: $\{\mathbf{X} : 0\text{m} \le x_{0,3} \le 1\text{m}, x_{0,1}^2 + x_{0.2}^2 \le 1\}$, and that the body undergoes the rotation about the $x_{0,3}$ axis with angular frequency ω:

$$x_1(\mathbf{X}, t) = x_{0,1}\cos\omega t + x_{0.2}\sin\omega t$$

$$x_2(\mathbf{X}, t) = -x_{0,1} \sin \omega t + x_{0,2} \cos \omega t$$
$$x_3(\mathbf{X}, t) = x_{0,3}$$

The initial density is homogeneous, $\rho_{\text{ref}} \equiv 1000\text{kg/m}^3$. (a) **Verify** that this formula defines a motion for all t. **Compute** (b) $\mathbf{v}(\mathbf{x}, t)$ (units of m/s), (c) $\nabla \cdot \mathbf{v}(\mathbf{x}, t)$ (what units?), (d) $\rho(\mathbf{x}, t)$.

Problem 1-8: **Derive** the following vector calculus identities, from the definition of gradient and divergence. All, when reduced to components, are straightforward consequences of the Leibnitz rule for partial derivatives of scalar-valued functions. In each statement, f and g denote scalar-valued functions and $\mathbf{u}$ and v (3-)vector-valued functions. Recall that the gradient is defined (inconsistently!) for a scalar-valued function f as a *column* vector with ith row entry $(\nabla f)_i = \frac{\partial f}{\partial x_i}$, and for a vector-valued function $\mathbf{u}$ as the 3×3 matrix

$$(\nabla \mathbf{u})_{i,j} = \frac{\partial u_i}{\partial x_j}.$$

The divergence of a vector-valued function $\mathbf{u}$ is

$$\nabla \cdot \mathbf{u} = \sum_{i=1}^{3} \frac{\partial u_i}{\partial x_i}$$

I'll use both $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{u}^T \mathbf{v}$ to denote the standard Euclidean inner product $\sum u_i v_i$ of vectors $\mathbf{u}, \mathbf{v}$.

Here are the identities you need to justify:

1. $\nabla(fg) = f\nabla g + g\nabla f$
2. $\nabla \cdot (f\mathbf{u}) = \nabla f \cdot \mathbf{u} + f \nabla \cdot \mathbf{u}$
3. $\nabla(\mathbf{u} \cdot \mathbf{v}) = (\nabla \mathbf{u})^T \mathbf{v} + (\nabla \mathbf{v})^T \mathbf{u}$
4. $\nabla(f\mathbf{u}) = \mathbf{u}(\nabla f)^T + f \nabla \mathbf{u}$

Problem 1-9: Derive the *vector version of Reynold's Transport Theorem*: for a vector-valued function $\mathbf{w}(\mathbf{x}, t)$, and a moving part p_t within its domain of definition,

$$\frac{d}{dt} \int_{p_t} dx\, \mathbf{w} = \int_{p_t} dx\, \frac{D\mathbf{w}}{Dt} + (\nabla \cdot \mathbf{v})\mathbf{w}$$

in which the material derivative $D\mathbf{w}/Dt$ is defined by

$$\frac{D\mathbf{w}}{Dt} \equiv \frac{\partial \mathbf{w}}{\partial t} + (\nabla \mathbf{w})\mathbf{v}.$$

Chapter 2

Strain

Deformations which don't change the internal distances in a body are in some sense trivial: they should not give rise to changes in the internal forces which each part of the body exerts on its complement. Thus there is need for a construction which expresses not deformation *per se*, but the extent to which a deformation or motion gives rise to changes in the internal geometry of a body.

2.1 Rigid Deformation and the Definition of Strain

The official adjective for a deformation which does not change internal distances is *rigid*. To say that $\mathbf{x} : \Omega \to \mathbf{R}^3$ does not change internal distances means

$$\|\mathbf{x}(\mathbf{X}_1) - \mathbf{x}(\mathbf{X}_2)\| = \|\mathbf{X}_1 - \mathbf{X}_2\|$$

where the vertical double bars denote Euclidean distance. A rigid motion is of course a motion consisting of rigid deformations.

Suppose that the two points under consideration are close, say $\mathbf{X}$ and $\mathbf{X} + \Delta\mathbf{X}$. The Taylor series of $\mathbf{x}$ gives

$$\mathbf{x}(\mathbf{X} + \Delta\mathbf{X}) - \mathbf{x}(\mathbf{X}) = \nabla\mathbf{x}(\mathbf{X})\Delta\mathbf{X} + R$$

with $R = o(\|\Delta\mathbf{X}\|)$. So the change in length (squared) is

$$\|\mathbf{x}(\mathbf{X} + \Delta\mathbf{X}) - \mathbf{x}(\mathbf{X})\|^2 - \|\Delta\mathbf{X}\|^2$$

$$= \Delta\mathbf{X}^T(\nabla\mathbf{x}(\mathbf{X})^T\nabla\mathbf{x}(\mathbf{X}) - I)\Delta\mathbf{X} + o(\|\Delta\mathbf{X}\|^2)$$

$$= 2\Delta\mathbf{X}^T L(\mathbf{X})\Delta\mathbf{X} + o(\|\Delta\mathbf{X}\|^2)$$

where $L(\mathbf{X})$ is the (Lagrangian or) material *strain tensor*

$$L(\mathbf{X}) = \frac{1}{2}(\nabla \mathbf{x}(\mathbf{X})^T \nabla \mathbf{x}(\mathbf{X}) - I)$$

As with everything else in this subject you can ask where the stuff currently at $\mathbf{x}$ came from, instead of where the stuff originally at $\mathbf{X}$ went to. That is, express the change of length as

$$\|\Delta \mathbf{x}\|^2 - \|\mathbf{X}(\mathbf{x} + \Delta \mathbf{x}) - \mathbf{X}(\mathbf{x})\|^2 = 2\Delta \mathbf{x}^T e(\mathbf{x}) \Delta \mathbf{x} + o(\|\Delta \mathbf{x}\|^2)$$

where the (Eulerian or) spatial strain tensor is

$$E(\mathbf{x}) = \frac{1}{2}(I - \nabla \mathbf{X}(\mathbf{x})^T \nabla \mathbf{X}(\mathbf{x}))$$

The following result expresses the relation between the strain tensors and rigidity:

Theorem Suppose that the interior Ω° is non-empty. Then the following conditions are equivalent:

1. $\mathbf{x} : \Omega \to \mathbf{R}^3$ is a rigid deformation
2. $$\mathbf{x}(\mathbf{X}) = \mathbf{x}(\mathbf{Y}) + M(\mathbf{X} - \mathbf{Y})$$

   where $\mathbf{Y} \in \Omega^\circ$ and M is a rotation matrix, i.e. an orthogonal matrix with positive determinant, $M^T M = M M^T = I$, $\det M = 1$ - that is, $\mathbf{x}$ is a *Euclidean affine transformation*;
3. The Lagrangian strain tensor L vanishes;
4. The Eulerian strain tensor E vanishes.

Proof: $1 \Rightarrow 2$: Pick an arbitrary $\mathbf{Y} \in \Omega^\circ$. Then the hypothesis means that for any $\mathbf{X} \in \Omega^\circ$,

$$(\mathbf{x}(\mathbf{Y}) - \mathbf{x}(\mathbf{X}))^T (\mathbf{x}(\mathbf{Y}) - \mathbf{x}(\mathbf{X})) - (\mathbf{Y} - \mathbf{X})^T (\mathbf{Y} - \mathbf{X}).$$

Both sides are differentiable in a neighborhood of $\mathbf{Y}$; a calculation gives the derivatives (with respect to $\mathbf{Y}$):

$$\nabla \mathbf{x}(\mathbf{Y})^T (\mathbf{x}(\mathbf{Y}) - \mathbf{x}(\mathbf{X})) = \mathbf{Y} - \mathbf{X}$$

Since $\mathbf{x}$ is a deformation, $\nabla \mathbf{x}(\mathbf{Y})^T$ is invertible. Write $M = (\mathbf{x}(\mathbf{Y})^T)^{-1}$. Then solving the above equation for $\mathbf{x}(\mathbf{X})$ gives

$$\mathbf{x}(\mathbf{X}) = \mathbf{x}(\mathbf{Y}) + M(\mathbf{X} - \mathbf{Y}).$$

Invoking the rigidity condition again, for any $\mathbf{X} \in \Omega$,

$$(M(\mathbf{X}-\mathbf{Y}))^T(M(\mathbf{X}-\mathbf{Y})) = (\mathbf{X}-\mathbf{Y})^T(M^TM)(\mathbf{X}-\mathbf{Y})$$
$$= (\mathbf{X}-\mathbf{Y})^T(\mathbf{X}-\mathbf{Y})$$

Since $\mathbf{Y}$ is interior to Ω, this equation holds for $\mathbf{X}$ in a sufficiently small ball centered at $\mathbf{Y}$. Since any vector is a multiple of $\mathbf{X}-\mathbf{Y}$ for some $\mathbf{X}$ in such a ball, it follows that M is an orthogonal matrix: $M^TM = I$. Since f is a deformation, its gradient must have positive determinant, therefore $\det M > 0$.

$2 \Rightarrow 3$: the deformation gradient $\nabla\mathbf{x}(\mathbf{X})$ is constant and $= M$, so,

$$L = M^TM - I = 0$$

because M is orthogonal. You can also easily see that $\nabla\mathbf{X}(\mathbf{x}) \equiv M^T$, so $E = 0$. q.e.d.

Thus the strain tensors really do measure the amount by which a deformation deviates from rigid motion - this deviation is *strain*. Other quantities expressing changes in the internal state of the body, such as internal forces, should depend on the strain, not on the deformation.

2.2 Strain and Displacement

It's also convenient to express strain in terms of the *displacement* of points during the deformation, defined as

$$\mathbf{w}(\mathbf{X}) = \mathbf{x}(\mathbf{X}) - \mathbf{X}$$

(Lagrangian displacement) or

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} - \mathbf{X}(\mathbf{x})$$

(Eulerian displacement). The expressions are

$$L = \frac{1}{2}(\nabla\mathbf{w} + \nabla\mathbf{w}^T + \nabla\mathbf{w}^T\nabla\mathbf{w})$$

$$E = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T - \nabla\mathbf{u}^T\nabla\mathbf{u})$$

Homogeneous deformations are described by

$$\mathbf{x}(\mathbf{X}) = \mathbf{x}_a + A(\mathbf{X} - \mathbf{x}_b)$$

for some choice of vectors $\mathbf{x}_a, \mathbf{x}_b$ and a matrix $A \in \mathbf{R}^3$. Since $\nabla\mathbf{x}(\mathbf{X}) = A$, this formula defines a deformation if $\det A > 0$. Then

$$\mathbf{X}(\mathbf{x}) = \mathbf{X}_a + A^{-1}(\mathbf{x} - \mathbf{X}_b)$$

(What are $\mathbf{X}_a, \mathbf{X}_b$ in terms of $\mathbf{x}_a, \mathbf{x}_b$?) It's easy to compute the strains:

$$L = \frac{1}{2}(A^T A - I),\ E = \frac{1}{2}(I - (A^{-1})^T A^{-1})$$

where I've left off $\mathbf{X}$ and $\mathbf{x}$ because both of these matrix-valued functions are constant!

The *Polar Decomposition Theorem* of linear algebra states that for any square matrix A, there are symmetric matrices S and T and an orthogonal matrix R of the same size as A so that

$$A = RS = TR.$$

If $\det A > 0$ then R is a rotation matrix, and $\det S,\ .\det T > 0$ also.

Substitution in the formulae for the strains and use of the identities satisfied by rotation matrices results in the formulae

$$L = \frac{1}{2}(S^2 - I),\ E = \frac{1}{2}(I - T^{-2})$$

Similarly, for an arbitrary deformation you can write

$$\nabla \mathbf{x}(\mathbf{X}) = R(\mathbf{X})S(\mathbf{X}) = T(\mathbf{X})R(\mathbf{X})$$

in which the symmetric and rotation parts of the polar decomposition are now matrix-valued functions. The same manipulations lead to

$$L(\mathbf{X}) = \frac{1}{2}(S(\mathbf{X})^2 - I),\ E(\mathbf{x}) = \frac{1}{2}(I - T(\mathbf{q}(\mathbf{x}))^{-2})$$

That is: *the strain tensors depend only on the symmetric parts of the polar decomposition of the deformation gradient* $\nabla \mathbf{x}(\mathbf{X})$.

Since the strains are symmetric matrices, the *Spectral Theorem* of linear algebra asserts the existence of an orthonormal basis of eigenvectors, called the principal axes of Eulerian or Lagrangian strain. The continuum mechanics name for the eigenvalues is *principal strains*. Some of the problems in the next set should give you a sense of the meanings of principal strains and axes.

2.3 Problems

Problem 2-1. A *uniform extension* in the direction $\mathbf{e} \in \mathbf{R}^3$ is a homogeneous deformation of the form

$$\mathbf{x} = \mathbf{x}_a + A(\mathbf{X} - \mathbf{x}_b), \ A = I + (\lambda - 1)\mathbf{e}\mathbf{e}^T$$

where λ is a positive scalar. Suppose that $\mathbf{e} = (1, 0, 0)^T$. **Compute** the strain tensors L and E explicitly (i.e. give the matrix entries in terms of the scalar λ), and **compute** the principal strains and axes (eigenvalues, eigenvectors of L, E). **Draw** the unit cube $\{\mathbf{X} : 0 \leq x_{0,i} \leq 1\}$ and its image under a uniform extension (do your best a perspective drawing) - can you see why this deformation is called an extension?

Problem 2-2. A homogeneous deformation with

$$A = \begin{bmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where γ is a scalar, is called *pure shear* in the (1,2) (or (x,y)) plane. **Compute** L, E, and the principal strains and axes. **Draw** the image of the unit cube under this deformation. Is "shear" the right word?

Problem 2-2 - Extra Credit. **Conjecture** a relation between Eulerian and Lagrangian principal strains and axes, based on the result of the last two problems. **Prove** it.

Problem 2-3. Suppose that $f : \mathbf{R} \to \mathbf{R}$, $f(0) = 1$, and

$$\mathbf{x}(\mathbf{X}, t) = \begin{bmatrix} 1 + t & 0 & 0 \\ 0 & f(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{X}.$$

Specify (i.e. give a formula for) the function $f(t)$ so that this motion is volume-preserving. **Describe** the motion thus defined in words, as best you can.

Problem 2-4. Suppose that A is any 3×3 matrix, i.e. $A \in \mathbf{R}^{3\times 3}$. **Explain why** the formula

$$\mathbf{x}(\mathbf{X}, t) = \mathbf{x}_a + (I + tA)(\mathbf{X} - \mathbf{x}_b)$$

defines a motion for small t. **Compute** $v(\mathbf{x}, t)$.

Problem 2-4 - Extra Credit. Under what conditions on the matrix A is the motion defined above volume-preserving?

Problem 2-5. Suppose that $\phi : \Omega \to \mathbf{R}^3$ is any continuously differentiable function on the undeformed material volume, and define

$$\Phi(\mathbf{x}, t) = \phi(\mathbf{q}(\mathbf{x}, t))\det\nabla\mathbf{q}(\mathbf{x}, t)$$

Show that

$$\frac{D\Phi}{Dt}(\mathbf{x}, t) + (\nabla \cdot v)(\mathbf{x}, t)\Phi(\mathbf{x}, t) = 0$$

(**N.B.:** this is a review problem, so it really is as easy as it looks!)

Chapter 3

Stress

3.1 Zoology of forces

The forces acting on continuum materials divide logically into those exerted in the material by itself, and those which are the result of some external agency. For a part occupying subvolume ω of the material in Ω, we can write

$$\mathbf{F}_{\text{total}}[\omega] = \mathbf{F}_{\text{int}}[\omega] + \mathbf{F}_{\text{ext}}[\omega]$$

As was explained in Ch. 1, momentum balance takes the form of a differential equation if the force can be expressed as an integral of a force density over ω, so it is interesting to see when that is the case.

I will have little to say about external forces: I will simply *assume* that they are given in terms of a force density $\mathbf{f}$ having units of force per unit volume:

$$\mathbf{F}_{\text{ext}}[\omega] = \int_{\omega} \mathbf{f}$$

Newtonian gravity is an important example of such a force; near the surface of the Earth the gravitational force density is $-\rho g \mathbf{e}_r$, in which $\mathbf{e}_r$ is the "radial" direction, i.e. the unit vector (in the units used to define ρ and g) on the line through the point on the surface and the Earth's center, pointing outward.

The standard analysis of internal forces, as presented in texts like [Gurtin, 1981], depends on two fundamental hypotheses, both asserting the locality of internal forces, i.e. the absence (or unimportance!) of action at a distance:

Internal Forces are Contact Forces: There is a *contact force density* function $\mathbf{t}[\omega]$ on the boundary $\partial\omega$, having units of force per unit area, giving the force exerted by the

remainder of the body (material in $\Omega - \omega$) on the material in ω:

$$\mathbf{F}_{\text{int}}[\omega] = \int_{\partial\omega} dA(\mathbf{x})\, \mathbf{t}[\omega](\mathbf{x})$$

Here $dA(\mathbf{x})$ is the area element on $\partial\omega$.

This equation expresses the conviction that *all influence of one part of the material on the remainder takes place at the its boundary* - what is going on in the remainder away from the boundary has absolutely no influence in the force experienced by the part.

Cauchy's Hypothesis on the Form of Contact Forces: The contact force density depends only on the tangent plane, or equivalently the (outward) normal vector $\mathbf{n}(\mathbf{x})$ at each point of the boundary $\partial\omega$:

$$\mathbf{t}[\omega](\mathbf{x}) = \mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x}))$$

for a function $\mathbf{T} : \Omega \times \mathbf{R}^3 \to \mathbf{R}^3$, called the *traction*.

This apparently innocent hypothesis has profound consequences. Note that whereas the contact force density of the first hypothesis depends on the part ω in some apparently very general way, Cauchy's hypothesis ties it to a much more restricted form: the contact force at a point on the boundary $\mathbf{x} \in \partial\omega$ does not depend on the shape or curvature of the boundary at $\mathbf{x}$ but only on its orientation, ie. on the direction of the normal at $\mathbf{x}$. This implies that all parts whose boundaries are tangent at $\mathbf{x}$ experience the same contact force density at $\mathbf{x}$. This is clearly a very great restriction in the form of contact forces - just how great will become apparent shortly.

The intuitive meaning of Cauchy's hypothesis is a little harder to divine, but it also expresses locality of influence. For purposes of computing contact force, it is as if at each point of the boundary, one replaces the boundary by a small flat piece of the tangent plane. If the curvature or higher order shape descriptors were involved, then the contact force would depend *less locally* on the disposition of material. Since the orientation of the boundary is clearly a factor in the contact force experienced by real materials (think of sticking your hand out the window of a speeding car - the force exerted by the airstream depends dramatically on its orientation!), Cauchy's hypothesis asserts the *maximum locality* of contact forces consistent with experience.

Remark: In all of this discussion, it has gone without saying that the boundary $\partial\omega$ is sufficiently regular to have a well-defined normal. I will make this assumption throughout.

3.2 Covariance of forces

Up to now, the discussion has treated the coordinate system as given. Of course coordinates are more or less arbitrary, and "physics" should not depend on the choice of coordinates. These notes will make use of only a part of the invariance of classical mechanics: all of the laws stated so far, and those to come, must be unchanged in their meaning under Euclidean changes of coordinates, i.e. new coordinates $\mathbf{x}'$ related to old coordinates $\mathbf{x}$ by

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}) = \mathbf{x}'_0 + R(\mathbf{x} - \mathbf{x}_0)$$

where R is an orthogonal matrix.

This is indeed a coordinate transformation, i.e. it is invertible with smooth inverse: the inverse Euclidean transformation is

$$\mathbf{x} = \mathbf{x}(\mathbf{x}') = \mathbf{x}_0 + R^{\mathbf{T}}(\mathbf{x}' - \mathbf{x}'_0)$$

Note the similarity in form, but not in meaning, to the definition of rigid deformation: here no deformation occurs, i.e. "the same stuff is in the same place" - just the coordinate description of physical location has changed.

This difference in meaning has an important consequence: it is not necessary to require that R is a rotation matrix, i.e. has a positive determinant. It is perfectly acceptable to use R with $\det R = -1$, as occurs for instance changing from a right-handed to a left-handed coordinate system. In particular $R = -I$ is fine (note that $-I$ is not a rotation in 3D!).

The velocity vectors of a material point in the two coordinate descriptions are related:

$$\frac{\partial \mathbf{x}'}{\partial t} = R\frac{\partial \mathbf{x}}{\partial t}$$

so

$$\mathbf{v}'(\mathbf{x}', t) = R\mathbf{v}(\mathbf{x}, t)$$

Note what's going on here. The velocity vector $\mathbf{v}$ does not simply compose with the change of coordinate - i.e. the velocity in the new coordinate system is not $\mathbf{v}(\mathbf{x}', t')$. Instead the *value* of this vector quantity also rotates, as described by the above equation, and in the same way that the position vector $\mathbf{x} - \mathbf{x}_0$ rotates into $\mathbf{x}' - \mathbf{x}'_0$. A vector quantity which changes in this systematic way is *covariant*.

A covariant vector field (vector-valued function) is not just a vector field - it is one vector field for each coordinate system, i.e. a vector-valued function on the *Cartesian product* of the Euclidean transformations and $\mathbf{R}^3$. Moreover, all of these vector-valued functions are related, as just described for the velocity. This is a subtle idea, and the notation does not support it well.

The Jacobian matrix of $\mathbf{x}'$ with respect to $\mathbf{x}$ is simply R, so the chain rule for a vector-valued function $\mathbf{\Phi}$ reads

$$\nabla_{\mathbf{x}}\mathbf{\Phi}(\mathbf{x}'(\mathbf{x})) = (\nabla_{\mathbf{x}'}\mathbf{\Phi})(\mathbf{x}'(\mathbf{x}))R$$

If $\mathbf{\Phi}$ is a covariant vector field, then

$$\mathbf{\Phi}'(\mathbf{x}'(\mathbf{x})) = R\mathbf{\Phi}(\mathbf{x}),\ \mathbf{\Phi}(\mathbf{x}) = R^T\mathbf{\Phi}'(\mathbf{x}'(\mathbf{x}))$$

so

$$\nabla_{\mathbf{x}}\mathbf{\Phi}(\mathbf{x}) = \nabla_{\mathbf{x}}(R^T\mathbf{\Phi}'(\mathbf{x}'(\mathbf{x}))) = (R^T\nabla_{\mathbf{x}'}\mathbf{\Phi}'R)(\mathbf{x}'(\mathbf{x}))$$

whence

$$\begin{aligned}\frac{D\mathbf{\Phi}}{Dt} &= \frac{\partial\mathbf{\Phi}}{\partial t} + \nabla_{\mathbf{x}}\mathbf{\Phi}v \\ &= \frac{\partial(R^T\mathbf{\Phi}')}{\partial t} + \nabla_{\mathbf{x}}(R^T\mathbf{\Phi}')R^Tv' \\ &= R^T\frac{\partial\mathbf{\Phi}'}{\partial t} + R^T\nabla_{\mathbf{x}'}\mathbf{\Phi}'RR^Tv' = R^T\frac{D\mathbf{\Phi}'}{Dt}\end{aligned}$$

That is: the material derivative of a covariant vector field is also a covariant vector field.

This applies in particular to the velocity field. So the left-hand side of the momentum balance law is also covariant (since ρ is scalar, and changes simply by composition with the change of coordinates):

$$\rho\frac{D\mathbf{v}}{Dt} = R^T\left(\rho'\frac{D\mathbf{v}'}{Dt}\right)$$

Therefore the momentum balance law can only hold regardless of coordinate system if *force densities are covariant vector fields.*

To see how this rule applies to traction, you need to understand how normal vectors change under changes of coordinate. Suppose that $P = \{\delta\mathbf{x} : \mathbf{n}\cdot\delta\mathbf{x} = 0\}$ describes the tangent plane to a surface at some position: i.e. $\mathbf{n}$ is the normal there. That is, each $\delta\mathbf{x} \in P$ is the tangent vector of some curve lying inside the surface. Then the tangent plane to the same surface in the new (primed) coordinates is $P' = \{\delta\mathbf{x}' : \delta\mathbf{x}' = R\delta\mathbf{x}\,\text{forsome}\,\delta\mathbf{x} \in P\}$, as you can see by looking at the coordinates of the curves. To be perpindicular to all of these vectors, $\mathbf{n}'$ must satisfy $\mathbf{n}' = R\mathbf{n}$, i.e. $\mathbf{n} = R\mathbf{n}'$, so normal vectors are also covariant.

Therefore traction, being a force (density), must satisfy

$$\mathbf{T}(\mathbf{x},\theta) = R^T\mathbf{T}'(\mathbf{x}',\theta') = R^T\mathbf{T}'(\mathbf{x}',R\theta)$$

i.e.

$$\mathbf{T}'(\mathbf{x}',\theta) = R\mathbf{T}(\mathbf{x},R^T\theta)$$

for any unit vector θ.

Curious example: In particular, taking $R = -I$, $\mathbf{x}_0 = \mathbf{x}'_0$ we get

$$\mathbf{T}(\mathbf{x}_0, \theta) = -\mathbf{T}'(\mathbf{x}_0, -\theta)$$

This *looks* like Newton's Law of Action and Reaction - i.e. the force exerted by the material outside a surface on the material inside is the same in magnitude as, and the opposite in sign to, the force exerted by the material inside the surface on the material outside - but isn't quite, since the right hand side is the traction field in a *different coordinate system*. Newton's Law in the case of tractions will follow from momentum balance, as a consequence of Cauchy's Theorem in the next section.

The careful reader will have noticed that the traction field appears only with unit vectors (namely, outward unit normals) in its second argument, in the statement of momentum balance. Therefore we are free to define its values for non-unit vectors in any way we see fit. The following convention is convenient:

Scaling Convention for Traction: For any real $\alpha > 0$, and $\theta \in \mathbf{R}^3$:

$$\mathbf{T}(\mathbf{x}, \alpha\theta) = \alpha\mathbf{T}(\mathbf{x}, \theta)$$

i.e. tractions are homogeneous of degree 1 in their second arguments.

3.3 Cauchy's Theorem

Cauchy's Theorem on the Linearity of Traction: There exists a 3×3 matrix valued function $S(\mathbf{x})$, called the *Stress Tensor* (or just *stress*) so that

$$\mathbf{T}(\mathbf{x}_0, \theta) = S(\mathbf{x}_0)\theta, \ \mathbf{x}_0 \in \Omega, \ \theta \in \mathbf{R}^3$$

Proof: This is Cauchy's famous "tetrahedron argument", see [Gurtin, 1981] pp. 97-106 for example. It uses a number of facts about multidimensional integrals which are established in advanced calculus.

The Scaling Convention, explained in the last subsection, implies that it is sufficient to consider only unit vectors θ. Since the traction (like everything else, whenever convenience dictates) is continuous at least, it also suffices to consider only θ with nonzero components, the other cases following by continuity. After a Euclidean change of coordinates with diagonal matrix leaving $\mathbf{x}_0$ fixed:

$$\mathbf{x} = \mathbf{x}_0 + R(\mathbf{x}' - \mathbf{x}_0), \ R_{ii} = \frac{\theta_i}{|\theta_i|}, \ R_{ij} = 0, \ 1 \leq i, j \leq 3$$

$\theta' = R\theta$ in the new coordinates has positive components. Since T transforms linearly, as estabished in the last section, the general case will follow from this one.

Construct a tetrahedron Γ with one face in the plane $\{\mathbf{x}_0 + \delta\mathbf{x} : \delta\mathbf{x} \cdot \theta = 0\}$ (i.e. with normal parallel to θ) and the other three faces parallel to the three coordinate directions. As it is no trouble to do so, I'll assume that the deformation in question is part of a motion, though I will suppress t from the notation. This covers static deformations, which are snapshots of trivial motions (for which rates of change, including velocities, are zero).

The balance of linear momentum states that

$$\int_\Gamma dx\, \rho \frac{Dv}{Dt} = \mathbf{F}_{\text{total}}[\Gamma]$$

$$= \int_{\partial\Gamma} dA(\mathbf{x})\mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x})) + \int_\Gamma dx\, \mathbf{f}(\mathbf{x})$$

where the first term on the right hand side represents the (internal) force exerted by the rest of the body on the stuff in the tetrahedron, and the second term represents all external forces.

For each $R > 0$ there is exactly one such tetrahedron for which the smallest circumscribed sphere centered at $\mathbf{x}_0$ has radius R; denote this tetrahedron by Γ_R. Essentially R is the diameter of Γ_R.

Assume as always that all functions appearing in the integrals above are at least once continuously differentiable. Then the volume integrals are bounded by a constant times the volume of the tetrahedron, which is proportional to R^3. The standard notation to express this fact concisely is that the volume integrals are $O(R^3)$.

By Taylor's Theorem,

$$\mathbf{T}(\mathbf{x}, \mathbf{n}) = \mathbf{T}(\mathbf{x}_0, \mathbf{n}) + O(|\mathbf{x} - \mathbf{x}_0|) = \mathbf{T}(\mathbf{x}_0, \mathbf{n}) + O(R)$$

Since an area integral of a continuous function over the bounded surface $\partial\Gamma_R$ is bounded by the area of $\partial\Gamma_R$ times the max of the integrand,

$$\int_{\partial\Gamma_R} dA(\mathbf{x})\mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x})) = \int_{\partial\Gamma_R} dA(\mathbf{x})\mathbf{T}(\mathbf{x}_0, \mathbf{n}(\mathbf{x})) + O(R^3)$$

The boundary $\partial\Gamma_R$ is the union of four faces: F_R^0, perpindicular to θ, and F_R^i, $i = 1, 2, 3$, perpindicular to $\mathbf{e}_1 = (1, 0, 0)^T$, $\mathbf{e}_2 = (0, 1, 0)^T$, and $\mathbf{e}_3 = (0, 0, 1)^T$ respectively. In particular the (outward!) normal is constant on each face: on F_R^0, $\mathbf{n}(\mathbf{x}) = \theta$, and on F_R^i, $i = 1, 2, 3$, $\mathbf{n}(\mathbf{x}) = -\mathbf{e}_i$. So

$$\int_{\partial\Gamma_R} dA(\mathbf{x})\mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x})) = \int_{F_R^0} dA(\mathbf{x})\mathbf{T}(\mathbf{x}_0, \theta) +$$

$$\int_{F_R^1} dA(\mathbf{x})\mathbf{T}(\mathbf{x}_0, -\mathbf{e}_1) + \int_{F_R^2} dA(\mathbf{x})\mathbf{T}(\mathbf{x}_0, -\mathbf{e}_2) + \int_{F_R^3} dA(\mathbf{x})\mathbf{T}(\mathbf{x}_0, -\mathbf{e}_3) + O(R^3)$$

$$= A_R^0\mathbf{T}(\mathbf{x}_0,\theta) + A_R^1\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_1) + A_R^2\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_2) + A_R^3\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_3) + O(R^3)$$

where A_R^i is the area of face F_R^i. Since $A_R^i = R^2 A_1^i$, this is the same as

$$= R^2(A_1^0\mathbf{T}(\mathbf{x}_0,\theta) + A_1^1\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_1) + A_1^2\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_2) + A_1^3\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_3)) + O(R^3)$$

Plug this expression for the integral of traction over the boundary of Γ_R into the momentum balance law, use the already-noted fact that the volume integrals are all $O(R^3)$, and divide by R^2 to get

$$A_1^0\mathbf{T}(\mathbf{x}_0,\theta) + A_1^1\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_1) + A_1^2\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_2) + A_1^3\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_3) = O(R)$$

The left-hand side is independent of R, so letting $R \to 0$ you obtain the identity

$$A_1^0\mathbf{T}(\mathbf{x}_0,\theta) + A_1^1\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_1) + A_1^2\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_2) + A_1^3\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_3) = 0$$

Claim:

$$\frac{A_1^i}{A_1^0} = \theta_i,\ i = 1,2,3$$

The easiest proof of this fact uses the Divergence Theorem:

Divergence Theorem For any reasonable domain Ω and vector field $\mathbf{\Phi}$ defined on Ω and continuously differentiable up to the boundary,

$$\int_\Omega dx\, \nabla\cdot\mathbf{\Phi}(\mathbf{x}) = \int_{\partial\Omega} dA(\mathbf{x})\mathbf{\Phi}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x})$$

Proof of Claim: Set $\mathbf{\Phi}(\mathbf{x}) = \mathbf{e}_i$ in the Divergence Theorem applied to $\Omega = \Gamma_1$. The left-hand side vanishes, while the right hand side is $A_0^1\theta_i + A_i^1$, which proves the Claim.

Applying this area identity to the traction identity above gives

$$\mathbf{T}(\mathbf{x}_0,\theta) = -\theta_1\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_1) - \theta_2\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_2) - \theta_3\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_3)$$

$$= S(\mathbf{x}_0)\theta$$

where $S(\mathbf{x}_0)_{j,i} = -\mathbf{T}(\mathbf{x}_0,-\mathbf{e}_i)_j$.

Since the proof has made no special assumption about the coordinate system in which traction has been described, the above identity holds for *any* coordinate system and any vector θ with positive coordinates in that system.

Consider now an arbitrary vector θ with non-zero components in some coordinate system. As mentioned at the beginning of the proof, introduce new coordinates *via* the Euclidean transformation

$$\mathbf{x}' = \mathbf{x}_0 + R(\mathbf{x} - \mathbf{x}_0),\ R = \text{diag}\left(\frac{\theta_1}{|\theta_1|}, \frac{\theta_2}{|\theta_2|}, \frac{\theta_3}{|\theta_3|}\right)$$

so that $\mathbf{x}'(\mathbf{x}_0) = \mathbf{x}_0$ and $\theta' = R^T\theta = R\theta$ has positive components. From the covariance of traction,

$$\mathbf{T}(\mathbf{x}_0, \theta) = R^T\mathbf{T}'(\mathbf{x}_0, \theta')$$

$$= -R^T(\sum_{i=1}^{3}(\theta' \cdot \mathbf{e}_i')\mathbf{T}'(\mathbf{x}_0, -\mathbf{e}_i')$$

Note that $R\mathbf{e}_i = \mathbf{e}_i'$, i.e. R carries the ith unit basis vector of the old coordinate system into the ith unit basis vector of the new coordinate system. Therefore $\theta' \cdot \mathbf{e}_i' = R\theta \cdot R\mathbf{e}_i = \theta \cdot \mathbf{e}_i = \theta_i$, so

$$\mathbf{T}(\mathbf{x}_0, \theta) = -R(\sum_{i=1}^{3}(\theta \cdot \mathbf{e}_i)R^T\mathbf{T}(\mathbf{x}_0, -R\mathbf{e}_i')$$

$$= -\sum_{i=1}^{3}\theta_i\mathbf{T}(\mathbf{x}_0, -\mathbf{e}_i) = S(\mathbf{x}_0)\theta$$

where $S(\mathbf{x}_0)$ is as defined above. That is, we have established the conclusion for all vectors θ with nonzero components.

Finally, any vector can be approximated arbitrarily closely by vectors with only nonzero components. Since both sides of

$$\mathbf{T}(\mathbf{x}_0, \theta) = S(\mathbf{x}_0)\theta$$

are continuous in θ, this identity extends by continuity to all θ. This observation completes the proof. **q.e.d.**

From the covariance property of traction, discussed in the last section, you can immediately read off

Change of Stress Tensor under change of coordinates : if coordinate systems $\mathbf{x}, \mathbf{x}'$ are related by a Euclidean change of coordinates

$$\mathbf{x}' = \mathbf{x}_0' + R(\mathbf{x} - \mathbf{x}_0)$$

then the stress tensors (describing the same state of internal forces) for the two coordinate systems are related as follows:

$$S'(\mathbf{x}') = RS(\mathbf{x})R^T$$

Another immediate consequence of Cauchy's Theorem is

Newton's Law of Action and Reaction for traction:

$$\mathbf{T}(\mathbf{x}_0, -\theta) = -\mathbf{T}(\mathbf{x}_0, \theta)$$

which follows immediately from the linearity of T. This observation also gives an alternate expression for the ith column of $S(\mathbf{x}_0)$:

$$S(\mathbf{x}_0)_{\cdot,i} = \mathbf{T}(\mathbf{x}_0, \mathbf{e}_i)$$

It also combines with the observation at the end of the last section to reveal a curious fact: if the coordinate systems $\mathbf{x}$ and $\mathbf{x}'$ are related by a Euclidean transformation with $R = -I$, i.e. the axes of the one are the axes of the other, but reversed, and $\mathbf{x}_0 = \mathbf{x}'_0$, then

$$\mathbf{T}(\mathbf{x}_0, \theta) = \mathbf{T}'(\mathbf{x}_0, \theta)$$

Note that this is a very special property of the coordinate-reversal transform, and does not hold for general Euclidean transformations.

Remember that $\mathbf{T}(\mathbf{x}_0, \theta)$ is the contact force at a point on (any!) boundary with outward normal θ. Thus the traditional interpretation of the stress tensor: $S(\mathbf{x}_0)_{i,j}$ is the ith component of force exerted at $\mathbf{x}_0$ by the material in $\{x_j > x_{0,j}\}$ on the material in $\{x_j < x_{0,j}\}$, i.e. across the j coordinate plane. Thus the diagonal components of S are often called *normal*, the off-diagonal components *tangential*.

3.4 Differential form of momentum balance

Cauchy's Theorem also gives an alternate expression for the contact force term in the momentum balance law, in terms of the stress tensor S:

$$\int_{\partial\omega} dA(\mathbf{x})\, \mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x})) = \int_{\partial\omega} dA(\mathbf{x})\, S(\mathbf{x})\mathbf{n}(\mathbf{x})$$

The ith component of this vector can be written

$$\int_{\partial\omega} dA(\mathbf{x})\, (S(\mathbf{x})_{i,\cdot})\mathbf{n}(\mathbf{x}) = \int_{\partial\omega} dA(\mathbf{x})\, (S(\mathbf{x})^T_{i,\cdot}) \cdot \mathbf{n}(\mathbf{x})$$

which has the same form as the right hand side of the Divergence Theorem.

The Divergence Theorem is so important in the remainder of these notes that I will state it again:

Divergence Theorem For any reasonable domain Ω and vector field $\mathbf{\Phi}$ defined on Ω and continuously differentiable up to the boundary,

$$\int_\Omega dx\, \nabla \cdot \mathbf{\Phi}(\mathbf{x}) = \int_{\partial\Omega} dA(x) \mathbf{\Phi}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$$

So

$$\int_{\partial\omega} dA(\mathbf{x})\, \mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x}))_{i,\cdot} = \int_\omega dx\, \nabla \cdot (S(\mathbf{x})_{1,\cdot})^T$$

Now *define* the divergence of a matrix valued function such as S by the formula

$$(\nabla \cdot S)_i = \sum_{j=1}^{3} \frac{\partial S_{ij}}{\partial x_j}$$

Then the contact force term becomes

$$\int_{\partial\omega} dA(\mathbf{x})\, \mathbf{T}(\mathbf{x}, \mathbf{n}(\mathbf{x})) = \int_\omega dx\, \nabla \cdot S(\mathbf{x})$$

This is just a special case of the

Divergence Theorem for Matrix Valued Functions: For any reasonable domain Ω and matrix valued $\mathbf{\Phi}$ defined on Ω and continuously differentiable up to the boundary,

$$\int_\Omega dx\, \nabla \cdot \mathbf{\Phi}(\mathbf{x}) = \int_{\partial\Omega} dA(\mathbf{x}) \mathbf{\Phi}(\mathbf{x}) \mathbf{n}(\mathbf{x})$$

Of course the definition of divergence for matrix valued functions is set up exactly to make this form of the divergence theorem follow from the usual one.

I have now expressed all three terms in the momentum balance law as volume integrals over the (arbitrary) part ω of the material body. The usual application of the Averaging Theorem gives the *differential form of momentum balance*:

$$\rho \frac{D\mathbf{v}}{Dt} = \nabla \cdot S + \mathbf{f}.$$

This very important equation will serve as the equation of motion for most of the continuum theories presented in the next few chapters.

3.5 Conservation of angular momentum and the symmetry of stress.

To define angular momentum, I require another relic from vector calculus, the *cross product* $\mathbf{u} \times \mathbf{w}$ of two vectors $\mathbf{u}$ and $\mathbf{w}$:

$$(\mathbf{u} \times \mathbf{w})_1 = u_2 w_3 - u_3 w_2$$

$$(\mathbf{u} \times \mathbf{w})_2 = u_3 w_1 - u_1 w_3$$
$$(\mathbf{u} \times \mathbf{w})_3 = u_1 w_2 - u_2 w_1$$

Another way to describe the cross product: form the 3×2 matrix $[\mathbf{uw}]$. Then the ith component of the cross product is the ith 2×2 minor determinant of this matrix (i.e. leave out the ith row, take the determinant of the remaining 2×2 matrix, and multiply by $(-1)^{i-1}$). Yet another way: form the skew-symmetric matrix

$$U = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}$$

Then $\mathbf{u} \times \mathbf{w} = U\mathbf{w}$.

Problem Set 3 asks you to work out a few properties of the cross product.

The *angular momentum* of a moving part ω_t is

$$a_{\omega_t}(t) = \int_{\omega_t} dx\, \rho(\mathbf{x}, t)(\mathbf{x} \times \mathbf{v}(\mathbf{x}, t))$$
$$= \int_{\omega_t} dx\, \mathbf{x} \times (\rho(\mathbf{x}, t)\mathbf{v}(\mathbf{x}, t))$$

The *external torque* is

$$\int_{\omega_t} dx\, \mathbf{x} \times \mathbf{f}(\mathbf{x}, t)$$

The *internal torque* is

$$\int_{\partial\omega_t} dA(\mathbf{x})\, \mathbf{x} \times \mathbf{T}(\mathbf{x}, t, \mathbf{n}(\mathbf{x}, t))$$

where I have made the notation reflect that both T and $\mathbf{n}$ depend explicitly on t.

The principle of *Conservation of Angular Momentum* states that the rate of change of angular momentum is the sum of all torques acting on the body, that is,

$$\frac{d}{dt}\int_{\omega_t} dx\, \rho(\mathbf{x}, t)\mathbf{x} \times \mathbf{v}(\mathbf{x}, t) = \int_{\partial\omega_t} dA(\mathbf{x})\, \mathbf{x} \times \mathbf{T}(\mathbf{x}, t, \mathbf{n}(\mathbf{x}, t)) + \int_{\omega_t} dx\, \mathbf{x} \times \mathbf{f}(\mathbf{x}, t)$$

This principle is actually independent of the Conservation of (Linear) Momentum, and in fact the two together imply the symmetry of the stress tensor. To see how this works, first analyse the left hand side of the above equation: from Problem 2 of Problem Set 2,

$$\frac{d}{dt}\int_{\omega_t} dx\, \rho(\mathbf{x}, t)\mathbf{x} \times \mathbf{v}(\mathbf{x}, t) = \int_{\omega_t} dx\, \rho(\mathbf{x}, t)\frac{D}{Dt}(\mathbf{x} \times \mathbf{v}(\mathbf{x}, t))$$
$$= \int_{\omega_t} dx\, \rho(\mathbf{x}, t)\left(\mathbf{x} \times \frac{D\mathbf{v}}{Dt}\right)$$

as follows from the vector calculus identities you will establish in Problem Set 3, especially Problem 4.

On the other hand,

$$\int_{\partial\omega_t} dA(\mathbf{x})\, \mathbf{x} \times \mathbf{T}(\mathbf{x}, t, \mathbf{n}(\mathbf{x}, t)) = \int_{\partial\omega_t} dA(\mathbf{x})\, \mathbf{x} \times S(\mathbf{x}, t)\mathbf{n}(\mathbf{x}, t)$$

$$= \int_{\omega_t} dx\, \nabla \cdot (\mathbf{x} \times S(\mathbf{x}, t))$$

$$= \int_{\omega_t} dx \sum_{j=1}^{3} \frac{\partial}{\partial x_j}(\mathbf{x} \times S(\mathbf{x}, t)_{\cdot,j})$$

$$= \int_{\omega_t} dx, \sum_{j=1}^{3} \left(\mathbf{e}_j \times S(\mathbf{x}, t)_{\cdot,j} + \mathbf{x} \times \frac{\partial S(\mathbf{x}, t)_{\cdot,j}}{\partial x_j} \right)$$

$$= \int_{\omega_t} dx\, (W(\mathbf{x}) + \mathbf{x} \times (\nabla \cdot S(\mathbf{x}, t))$$

where $W_1 = S_{32} - S_{23}$, $W_2 = S_{13} - S_{31}$, and $W_3 = S_{21} - S_{12}$.

Putting this all together, the law of angular momentum balance becomes

$$\int_{\omega_t} dx\, \mathbf{x} \times \left(\rho(\mathbf{x}, t) \frac{D\mathbf{v}}{Dt}(\mathbf{x}, t) \right) = \int_{\omega_t} dx\, (W + \mathbf{x} \times (\nabla \cdot S(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t)))$$

The Averaging Theorem, applied in the usual way, plus a little rearrangement of terms, shows that

$$W(\mathbf{x}, t) = \mathbf{x} \times \left(\rho \frac{D\mathbf{v}}{Dt}(\mathbf{x}, t) - \nabla \cdot S(\mathbf{x}, t) - \mathbf{f}(\mathbf{x}, t) \right)$$

However the differential form of the linear momentum balance law derived in the last section shows that the right hand side of the above expression vanishes. In view of the definition of W, we have proved

Theorem (Symmetry of Stress): $S = S^T$.

3.6 Problems.

Problem 3-1: **Show** that for any vectors $\mathbf{u}$ and $\mathbf{v}$, and scalar a,

- a. $\mathbf{u} \times (a\mathbf{v}) = (a\mathbf{u}) \times \mathbf{v} = a(\mathbf{u} \times \mathbf{v})$

- b. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$
- c. $\mathbf{v} \times \mathbf{v} = 0$
- d. For a matrix $A \in \mathbf{R}^{3\times 3}$, denote by $\mathbf{u} \times A$ the matrix whose jth column is $\mathbf{u} \times A_{\cdot,j}$, i.e. the columnwise cross product. **Show** that

$$\mathbf{u} \times \mathbf{v} = (\mathbf{u} \times I)\mathbf{v}$$

and that the matrix U defined in the text is in fact $\mathbf{u} \times I$.

Problem 3-2: **Assume** that $\mathbf{u}$ and $\mathbf{v}$ are continuously differentiable vector fields. **Show** that the jth column of $\nabla(\mathbf{u} \times \mathbf{v})$ is

$$\begin{aligned}\nabla(\mathbf{u} \times \mathbf{v})_{\cdot,j} &= \frac{\partial \mathbf{u}}{\partial x_j} \times \mathbf{v} + \mathbf{u} \times \frac{\partial \mathbf{v}}{\partial x_j} \\ &= (\nabla \mathbf{u})_{\cdot,j} \times \mathbf{v} + \mathbf{u} \times (\nabla \mathbf{v})_{\cdot,j}\end{aligned}$$

Problem 3-3: With notation as in Problem 2, **show** that

$$\nabla(\mathbf{x} \times \mathbf{u}) = -\mathbf{u} \times I + \mathbf{x} \times \nabla \mathbf{u}$$

Problem 3-4: In this problem, $\mathbf{v}$ denotes the velocity field, *not* an arbitrary vector field. Use the results of Problems 1 - 3 to **show** that

$$\frac{D}{Dt}(\mathbf{x} \times \mathbf{v}) = \mathbf{x} \times \frac{D\mathbf{v}}{Dt}$$

Problem 3-5: **Show** that

- a.

$$\nabla|\mathbf{x}| = \frac{\mathbf{x}}{|\mathbf{x}|}$$

- b.

$$\nabla(\nabla|\mathbf{x}|) = \frac{1}{|\mathbf{x}|}\left(I - \frac{1}{|\mathbf{x}|^2}\mathbf{x}\mathbf{x}^T\right)$$

- c. (in 3D only!)

$$\nabla \cdot \nabla \frac{1}{|\mathbf{x}|} = 0, \ \mathbf{x} \neq 0$$

- d.

$$\nabla \cdot (|\mathbf{x}|\mathbf{x}) = 4|\mathbf{x}|$$

Problem 3-6: In this problem, $\mathbf{u}$ and $\mathbf{v}$ are continuously differentiable vector fields, as in Problem 2, and $\nabla\times$ denotes the *curl operator*:

$$\nabla \times \mathbf{u} = \left(\frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}, \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}, \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right)^T .$$

Show that

- a.
$$\nabla \cdot (\mathbf{u} \times \mathbf{v}) = (\nabla \times \mathbf{u}) \cdot \mathbf{v} - (\nabla \times \mathbf{v}) \cdot \mathbf{u}$$
- b. If $\mathbf{u}$ is *twice* continuously differentiable, then
$$\nabla \times (\nabla \times \mathbf{u}) = \nabla(\nabla \cdot \mathbf{u}) - \nabla \cdot \nabla \mathbf{u}$$

Chapter 4

Ideal Fluids and the Euler Equations

4.1 Characterization of Ideal Fluids

The characteristic property of *fluids* is that *stress is scalar*, i.e.

$$S(\mathbf{x}, t) = -p(\mathbf{x}, t)I$$

where p is a scalar function, the *pressure*. The minus sign is a convention.

The expression of stress divergence in terms of p is simple:

$$(\nabla \cdot S)_i = \sum_{j=1}^{3} \frac{\partial}{\partial x_j}(-pI_{ij}) = (-\nabla p)_i$$

i.e.

$$\nabla \cdot S = -\nabla p$$

An *ideal fluid* is characterized by two further properties:

- admits only volume preserving deformations, i.e. is *incompressible*:

$$\nabla \cdot \mathbf{v} = 0$$

- has constant density:

$$\rho(\mathbf{x}, t) = \rho_0$$

These two hypotheses together imply the conservation of mass (verify this!), so the only independent conservation law that we have discussed so far is the conservation of momentum, which reads

$$\rho_0 \frac{D\mathbf{v}}{Dt} + \nabla p = \mathbf{f}$$

The last three equations together are *Euler's equations* for an ideal fluid.

Note that the momentum balance law can also be written

$$\rho_0 \left(\frac{\partial \mathbf{v}}{\partial t} + \nabla \mathbf{v}\, \mathbf{v} \right) + \nabla p$$

which reveals the nonlinearity of Euler's equations. In fact these are quite difficult to solve - to understand the solutions mathematically, or to construct good numerical approximations. Therefore a great deal of effort has been devoted to analysis of more tractable but still useful special cases.

4.2 Steady Flow

Motion of a material (fluid or otherwise) is *steady* if all quantities describing it are independent of time. For a fluid, this means that $\mathbf{v}$, ρ, and p are all independent of time. Euler's equations become

$$\nabla \mathbf{v}\, \mathbf{v} + \nabla p = \mathbf{f},\ \nabla \cdot \mathbf{v} = 0$$

and the first of these shows that f must be independent of time as well.

One can learn a few things about flow without actually solving the equations of motion. For example, suppose that steady flow takes place in a pipe, i.e. a region Ω for which the boundary divides into three parts, representing inlet, outlet, and sides: $\partial\Omega = \Gamma_{\text{in}} \cup \Gamma_{\text{out}} \cup \Gamma_{\text{side}}$. Since the sides are non-porous, the velocity vector must be tangent to the sides, and so $\mathbf{v} \cdot n = 0$ at every point on the side Γ_{side}.

Since ρ is independent of t, the conservation of mass in spatial terms is just

$$\nabla \cdot (\rho_0 \mathbf{v}) = 0$$

Integrating this identity over the domain of steady motion throws away lots of information, though not all: the divergence theorem gives

$$0 = \int_\Omega \nabla \cdot (\rho_0 \mathbf{v}) = \int_{\partial\Omega} \rho_0 \mathbf{v} \cdot \mathbf{n}$$

$$= \int_{\Gamma_in} \rho_0 \mathbf{v} \cdot \mathbf{n} + \int_{\Gamma_out} \rho \mathbf{v} \cdot \mathbf{n}$$

(the side integral vanishing because $\mathbf{v}$ is perpindicular to n there). These two terms give the net rate of mass flow in respectively out of the pipe (actually the first term is minus the inflow), and conservation of mass says that they must cancel: in steady flow, inflow must equal outflow.

Another interesting consequence emerges if you make another assumption that is satisfied approximately by flow of ordinary fluids through pipes: on Γ_{in}, $\mathbf{v} \simeq -v_{\text{in}} n_{\text{in}}$ and $p \simeq p_{\text{in}}$; on Γ_{out}, $\mathbf{v} \simeq -v_{\text{out}} n_{\text{out}}$ and $p \simeq p_{\text{out}}$. Here the subscripted values of $\mathbf{v}$ and p are scalar constants, independent of position. Denote by A_{in} and A_{out} the areas of the inlet and outlet respectively. Then the mass balance law reduces to

$$0 = A_{\text{out}} \rho_0 v_{\text{out}} - A_{\text{in}} \rho_0 v_{\text{in}}$$

so

$$A_{\text{out}} v_{\text{out}} = A_{\text{in}} v_{\text{in}}$$

which is a phenomenon anyone who has played with a garden hose will recognize: if you make the outlet area smaller than the inlet area, the outlet velocity will be larger than the inlet velocity.

The balance of linear momentum implies that

$$0 = \int_\Omega \rho_0 (\nabla \mathbf{v}) \mathbf{v} + \nabla p$$

This seems intractable, but the following fact gets us out of trouble:

Lemma: For a volume preserving motion, the velocity term in the material derivative is a divergence:

$$(\nabla \mathbf{v}) \mathbf{v} = \nabla \cdot (\mathbf{v} \mathbf{v}^T)$$

Proof: Leibnitz says

$$\nabla \cdot (\mathbf{v} \mathbf{v}^T) = (\nabla \mathbf{v}) \mathbf{v} + \mathbf{v} (\nabla \cdot \mathbf{v})$$

and the second summand on the right vanishes for volume preserving motion. **q.e.d.**

Thus the divergence theorem applies:

$$0 = \int_{\partial\Omega} \rho_0 (\mathbf{v} \cdot \mathbf{n}) \mathbf{v} + p \mathbf{n}$$

Note: This really is the divergence theorem, for matrix valued functions: if ϕ is a scalar function,

$$\int_\Omega \nabla \phi = \int_\Omega \nabla \cdot (\phi I) = \int_{\partial\Omega} \phi \mathbf{n}$$

Then, since the flow is tangent to the pipe walls ($\mathbf{v} \cdot \mathbf{n} = 0$)

$$0 = \left(\int_{\Gamma_{\text{in}}} + \int_{\Gamma_{\text{out}}} + \int_{\Gamma_{\text{side}}} \right) (\rho_0 (\mathbf{v} \cdot \mathbf{n}) \mathbf{v} + p\mathbf{n})$$

$$= A_{\text{out}}(p_{\text{out}} + \rho_0 v_{\text{out}}^2)\mathbf{n}_{\text{out}} + A_{\text{in}}(p_{\text{in}} - \rho_0 v_{\text{in}}^2)\mathbf{n}_{\text{in}} + \int_{\Gamma_{\text{side}}} p\mathbf{n}$$

The third term has the interpretation of total force exerted by the fluid on the pipe walls. This expression shows that you can compute this total force merely from knowing the density, pressures, and velocities at the inlet and outlet.

4.3 Circulation and Irrotational Flow

A *closed path* C in a domain $\Omega \in \mathbf{R}^3$ is the image of a continuously differentiable map $\mathbf{c} : [0, 1] \to \Omega$ for which $\mathbf{c}(0) = \mathbf{c}(1)$. The *circulation* of a vector field $\mathbf{u} : \Omega \to \mathbf{R}^3$ is the integral of the tangential component of $\mathbf{u}$ around C:

$$\int_C \mathbf{u} \cdot d\mathbf{x} = \int_0^1 dt\, \mathbf{u}(\mathbf{c}(t)) \cdot \frac{d\mathbf{c}}{dt}(t)$$

The change of coordinates rule for integrals shows that this definition is independent of orientation preserving reparametrization of C: any two parametrizations $\mathbf{c}(t)$ and $\mathbf{c}'(t')$ yield the same circulation, so long as $dt/dt' > 0$. That is, the circulation depends only on the vector field, the closed curve, and the sense in which the close curve is traversed (its orientation).

A motion is *irrotational* if the circulation of the velocity field about any closed curve in the domain of the motion is zero.

A closed curve $C \subset \Omega$ is the boundary of a smooth bounded surface D iff

1. D is the image of a continuously differentiable injective map $\mathbf{d}$ from the unit disk $U = \{(x_1, x_2)^T : x_1^2 + x_2^2 \leq 1\} \subset \mathbf{R}^2$ into Ω, and
2. $\mathbf{c}(t) = \mathbf{d}(\cos 2\pi t, \sin 2\pi t)$ is a parametrization of C.

There are two continuous unit vector fields defined on such a smooth bounded surface which are perpindicular at every point to its tangent field. Given the orientation of the bounding curve implicit in a parametrization, as in the definition, the *positive normal* $\mathbf{n}$ is the one for which the determinant

$$\det \left[\frac{d\mathbf{c}}{dt}, \frac{d^2\mathbf{c}}{dt^2}, \mathbf{n} \right]$$

is positive at every point where the first two columns do not vanish. In terms of the map mentioned in the definition, this choice of normal is a positive multiple of

$$\frac{\partial \mathbf{d}}{\partial x_1} \times \frac{\partial \mathbf{d}}{\partial x_2}$$

To interpret irrotational motion, recall *Stoke's Theorem*: if the closed curve C bounds a smooth piece of surface D, and $\mathbf{n}$ is the positive unit normal field with respect to the chosen orientation of C, then for any continuously differentiable vector field $\mathbf{u}$ defined on and near D,

$$\int_C \mathbf{u} \cdot d\mathbf{x} = \int_D dA(\mathbf{x})(\nabla \times \mathbf{u}) \cdot \mathbf{n}$$

It is an advanced calculus excercise to show how the averaging theorem implies that irrotational flow satisfies

$$\nabla \times \mathbf{v} = 0$$

This is the local or differential interpretation of zero circulation.

4.4 Potential Flow

Statement of the next major result result requires more nomenclature.

A *homotopy* $\mathbf{h}$ in a domain $\Omega \subset \mathbf{R}^3$ is a continuous map from the unit square $[0,1] \times [0,1]$ into Ω. The *end curves* of a homotopy are the images of the maps $\mathbf{h}_0, \mathbf{h}_1 : [0,1] \to \Omega$ defined by

$$\mathbf{h}_0(t) = \mathbf{h}(0,t), \;\; \mathbf{h}_1(t) = \mathbf{h}(1,t)$$

In fact, for general $s \in [0,1]$, define $\mathbf{h}_s(t) = \mathbf{h}(s,t)$. The homotopy is a *homotopy of closed curves* iff every $\mathbf{h}_s, 0 \leq s \leq 1$, is a closed curve. A homotopy is continuously differentiable iff the map $\mathbf{h}$ is continuously differentiable.

Two closed curves C_0, C_1 in Ω are *homotopic* iff there exists a continuously differentiable homotopy of closed curves $\mathbf{h}$ so that C_j is the image of $\mathbf{h}_j$, $j = 0, 1$. The domain Ω is *simply connected* if every closed curve in Ω is homotopic to the constant curve ("a point").

The intuitive meaning of "simply connected" is that every closed curve in Ω can be shrunk to a point without leaving Ω. Euclidean 3-space $\mathbf{R}^3$ is simply connected, as is Euclidean 3-space with the origin deleted. However Euclidean 3-space with the x-axis deleted is not simply connected. Another example: a cylindrical tank is simply connected, but a tank with a cylindrical hole up its axis is not, nor are donuts or coffee cups with handles.

Poincaré's Lemma: Suppose that $\Omega \subset \mathbf{R}^3$ is simply connected, and that that $\mathbf{u}$ is a continuously differentiable vector field on Ω. Then the following hypotheses are equivalent:

- $\mathbf{u}$ has zero circulation around any closed curve in Ω;
- $\nabla \times \mathbf{u} = 0$;
- $\mathbf{u} = \nabla\phi$ for a twice continuously differentiable scalar valued function ϕ on Ω.

Assume that the volume occupied by a steady irrotational flow is simply connected. Then Poincaré's Lemma implies that *the velocity is the gradient of a scalar function*, called the *flow potential* ϕ:

$$\mathbf{v} = \nabla\phi$$

This identity combined with the momentum balance law leads to *Bernoulli's Theorem*, which you will develop in Problem Set 5. Expressing the incompressibility condition in terms of the flow potential leads to the *Laplace Equation*:

$$\nabla^2\phi \equiv \nabla \cdot \nabla\phi = \sum_{i=1}^{3} \frac{\partial^2\phi}{\partial x_i^2} = 0$$

arguably the most important equation in mathematical physics.

To gain some idea of the powerful constraint the Laplace equation places on flow, suppose steady, irrotational flow takes place in a simply connected domain Ω, and that the boundary confines the fluid. Necessarily, the fluid velocity is tangent to the boundary. That is,

$$\nabla\phi(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0, \ \mathbf{x} \in \partial\Omega$$

This equation is properly known as the "no flow" boundary condition, or the (homogeneous) *Neumann boundary condition*. The pair consisting of the Laplace equation and the Neumann boundary condition is known as the *Neumann Problem* for ϕ. Use vector calculus identities from old homeworks and the divergence theorem to see that

$$\begin{aligned} 0 = \int_\Omega \phi\nabla^2\phi &= \int_\Omega (\nabla \cdot (\phi\nabla\phi) - \nabla\phi \cdot \nabla\phi) \\ &= \int_{\partial\Omega} \phi\nabla\phi \cdot \mathbf{n} - \int_\Omega |\nabla\phi|^2 \\ &= -\int_\Omega |\nabla\phi|^2 = -\int_\Omega |\mathbf{v}|^2 \end{aligned}$$

The velocity field is continuously differentiable by standing assumption, so its length squared is continuous and nonnegative. Yet another important theorem from advanced calculus: if a nonnegative continuous function integrates to zero, then it *is* zero.

So we have shown: if steady irrotational flow is confined to a simply connected domain, *then there is no flow at all!!!* The velocity field vanishes identically, from which it follows that the external force field must be conservative (see Problem 2 below) and balance the pressure gradient precisely. The flow potential is constant - one cannot say more, since *all* of the conditions constraining it involve its derivatives.

If you want steady nonzero confined flow, you will have to settle for nonzero circulation, or make it happen in a domain that's not simply connected!

In fact we have proven more: if two steady irrotational flows in the same simply connected domain, with velocity fields $\mathbf{v}_1$ and $\mathbf{v}_2$, have identical *normal components of velocity* on the boundary, then the two flows have the same velocity field throughout the domain, and their pressure fields differ by a constant. Indeed $\delta\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_1$ has zero curl and divergence also, so it is the gradient of a potential which solves the homogeneous Neumann problem. We have just seen that all such (twice continuously differentiable) potentials are constant, therefore the gradient $\delta\mathbf{v}$ vanishes, i.e. $\mathbf{v}_1 = \mathbf{v}_2$.

You can phrase this observation entirely in terms of potentials: if ϕ_1, ϕ_2 both solve the same *inhomogeneous Neumann problem*

$$\nabla^2\phi_i = f \text{ in } \Omega, \nabla\phi \cdot \mathbf{n} = g \text{ on } \partial\Omega, \ i = 1, 2$$

defined by suitable right hand sides $f : \Omega \to \mathbf{R}$ and $g : \Omega \to \mathbf{R}$, then ϕ_1 and ϕ_2 differ by a constant. That is, solutions of the Neumann problem are as unique as they can be.

Another consequence: for steady irrotational flow in a pipe, as discussed earlier, the normal components of velocity at the inlet and outlet completely determine the flow at every point in the pipe.

For very simple situations, this fact is enough to pin down the flow explicitly. Suppose for example that the pipe is straight, say with its axis parallel to the x_1 axis, and simply connected. The inlet and outlet have the same area, and the normal (x_1) component ν of velocity is constant across the inlet and outlet and the same (as we saw, this last is required by conservation of mass). Then $\mathbf{v}(\mathbf{x}) = (\nu, 0, 0)^T$ is the gradient of $\phi(\mathbf{x}) = \nu x_1$, which solves the Laplace equation and matches all of the boundary conditions. Therefore the flow potential is $\phi+$ const., hence the constant $\mathbf{v}$ must actually be the velocity field in the pipe.

Functions satisfying the Laplace equation are also called *harmonic*. Harmonic functions of two variables (say, x_1 and x_2) are quite special: they are real parts of complex analytic functions of the complex variable $x_1 + ix_2$. Such potentials give rise to flow parallel to the (x_1, x_2) plane (so called planar flow), and in special cases they can be determined explicitly by the method of *conformal mapping*, for which see practically any text on complex variable theory or elementary fluid mechanics. While few problems admit the conformal mapping approach, some that do are quite important.

4.5 Green's Functions

This section shows how we may construct solutions u of the *Poisson Problem*

$$\nabla^2 u = \phi$$

for a rather tame type of right hand side ϕ. This is the beginning of a line of reasoning that leads eventually to construction of potential flows in various circumstances; here we merely sketch the beginning of these developments.

First we need a few notions related to functions, and things more general than functions, with singularities.

A *test function* on Euclidean d-space $\mathbf{R}^d$ is an infinitely often differentiable scalar-valued function which vanishes for large enough $\|\mathbf{x}\|$. The set of test functions is conventially denoted $\mathcal{D}$; it forms a vector space.

Note that any (mixed) partial derivative of any order, or linear combination of partial derivatives, or product with an infinitely differentiable function, of a test function is a test function.

It's conventional to take for the scalar field of values the complex numbers, rather than the real numbers. All of our examples will be real-valued, but it makes very little difference to the theory developed below if complex-valued test functions are used.

There is a natural notion of convergence of sequences of test functions: if $\{\phi_n\}_{n=0}^{\infty}$ is such a sequence, and 0 denotes the zero function (which plays the role of the zero vector in the vector space $\mathcal{D}$, then $\phi_n \to 0$ if and only if (i) for some $R > 0$, $\phi_n(\mathbf{x}) = 0$ if $|\mathbf{x}\| > R$, any n, and (ii)

$$\sup_{\|\mathbf{x}\| \le R} \left| \frac{\partial^{j_1+...+j_d} \phi_n}{\partial x_1^{j_1} ... \partial x_d^{j_d}} \right| \to 0$$

for each $R > 0$.

It gets tedious to write the partial derivatives symbols in this expression, which motivates the definition: for a vector $\mathbf{j} = (j_1, ..., j_d)$ of non-negative indices,

$$D^{\mathbf{j}} = \frac{\partial^{j_1+...+j_d}}{\partial x_1^{j_1} ... \partial x_d^{j_d}}$$

The set of non-negative integers is denoted $\mathbf{Z}_+$, so $\mathbf{j} \in \mathbf{Z}_+^d$. We can restate the second condition for convergence $\phi_n \to 0$ in $\mathcal{D}$ as: for any $\mathbf{j} \in \mathbf{Z}_+^d$,

$$sup_{\|\mathbf{x}\| \le R} \left| D^{\mathbf{j}} \phi_n(\mathbf{x}) \right| \to 0$$

A *distribution* is a continuous scalar-valued linear function on the vector space $\mathcal{D}$. The distributions also form a vector space in a natural way, denoted by $\mathcal{D}'$.

The value of a distribution $u \in \mathcal{D}'$ on a test function $\phi \in \mathcal{D}$ is conventionally denoted $\langle u, \phi \rangle$.

Continuity means that if $u \in \mathcal{D}'$ and $\{\phi_n\}_{n=0}^{\infty}$ is a sequence of test functions, then

$$\phi_n \to 0 \;\Rightarrow\; \langle u, \phi_n \rangle \to 0$$

Two examples of distributions are already familiar, if not under that name:

The Dirac Delta function: Usually denoted δ, defined by the rule

$$\langle \delta, \phi \rangle = \phi(0)$$

Caution: in many places the value of the delta function on a test function will be written as an integral:

$$\int dx\, \delta(\mathbf{x})\phi(\mathbf{x}) = \phi(0)$$

Since the delta function is not a function, the integral notation for its value must actually be poetry - i.e. the left hand side of the preceding equation has no meaning, apart from denoting the right hand side.

Integrable Functions An integrable scalar valued function g defines a distribution in a natural way, and it's conventional to use the same letter to denote the function and the distribution it defines:

$$\langle g, \phi \rangle = \int dx\, \bar{g}(\mathbf{x})\phi(\mathbf{x})$$

The overbar denoted complex conjugation, which of course is a no-op if g is real valued - so in that case I will drop the overbar.

We extend various common operations on functions to distributions by *transposition*. Suppose that $T : \mathcal{D} \to \mathcal{D}$ is linear and continuous. Just as for scalar-valued functions, continuity means that if $\phi_n \to 0$ in the sense defined above, then $T\phi_n \to 0$ also. Define the *transpose* $T' : \mathcal{D}' \to \mathcal{D}'$ by the rule: if $u \in \mathcal{D}', \phi \in \mathcal{D}$, then

$$\langle T'u, \phi \rangle = \langle u, T\phi \rangle$$

Differentiation, being continuous on $\mathcal{D}$ (verify this!), extends to distributions via transposition. That is if $u \in \mathcal{D}'$, then $\left(\frac{\partial}{\partial x_i}\right)' u$ is the distribution defined by

$$\left\langle \left(\frac{\partial}{\partial x_i}\right)' u, \phi \right\rangle = \left\langle u, \frac{\partial \phi}{\partial x_i} \right\rangle$$

This notion of derivative is often called the *weak* or *distribution* derivative. Note that if g and $\partial g/\partial x_i$ are integrable functions, then integration by parts shows that the (weak) (x_i partial) derivative of g is actually the distribution defined by $-\partial g/\partial x_i$. Thus the weak and usual notions of derivative coincide when both make sense.

In view of this fact, it's conventional to confuse notations, and write $-\frac{\partial u}{\partial x_i}$ instead of the more cumbersome $\left(\frac{\partial}{\partial x_k}\right)' u$ even when u is a distribution. Carrying on in this fashion, it's conventional to confused the transposed Laplacian on distributions with the usual Laplacian on functions, as the two coincide on distributions which happen to be (twice-differentiable) functions.

A *Green's Function* for the Laplace operator is a distribution $g \in \mathcal{D}$ for which

$$\nabla^2 g = \delta$$

(in which the left hand side is interpreted in the sense of weak derivative, of course). Clearly any two Green's functions, say g_1 and g_2, differ by a solution of

$$\nabla^2(g_1 - g_2) = 0$$

i.e. $g_1 - g_2$ is a "harmonic distribution". As it turns out, any harmonic distribution is actually (defined by) a harmonic function which is infinitely differentiable. We are about to see that the Green's functions themselves are definitely not so regular, as you would guess by looking at the right hand side of the defining PDE.

Claim: $g : \mathbf{R}^3 \to \mathbf{R}$ defined by

$$g(\mathbf{x}) = -\frac{1}{4\pi\|\mathbf{x}\|}$$

is a Green's function for the Laplace operator. [Note: you should be able to call on your advanced calculus to show that the right hand side of this equation defines an integrable function (its singularity at zero is integrable), so this distribution is one of those defined by an integrable function.]

To see this, we need to show that the defining equation is satisfied. Since both sides are distributions, this means showing that for any test function $\phi \in \mathcal{D}$,

$$\langle \nabla^2 g, \phi \rangle = \phi(0)$$

This means (using the weak derivative definition several times) that

$$\langle g, \nabla^2 \phi \rangle = \phi(0)$$

$$= \int dx\, g(\mathbf{x}) \nabla^2 \phi(\mathbf{x})$$

since g is actually an integrable function (hence its product with $\nabla^2\phi$ is integrable).

The integral on the right hand side of the preceding equation is of course defined as an improper integral, since the integrand is not continuous - i.e. as the limit of proper integrals

$$= \lim_{\epsilon\to 0} \int_{\|\mathbf{x}\|\geq\epsilon} dx\, g(\mathbf{x})\nabla^2\phi(\mathbf{x})$$

In this "punctured" domain of integration, the factors in the integrand are twice (actually, infinitely often) continuously differentiable, so you can integrate by parts (i.e. use the divergence theorem:

$$= \lim_{\epsilon\to 0} \int_{\|\mathbf{x}\|\geq\epsilon} dx\, \nabla\cdot(g\nabla\phi) - \nabla g\cdot\nabla\phi$$

$$= \lim_{\epsilon\to 0} \left\{ \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, g\nabla\phi\cdot\mathbf{n} - \int_{\|\mathbf{x}\|\geq\epsilon} dx\, \nabla\cdot(\nabla g\phi) + \int_{\|\mathbf{x}\|\geq\epsilon} dx\, \nabla^2 g\phi \right\}$$

$$= \lim_{\epsilon\to 0} \left\{ \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, g\nabla\phi\cdot\mathbf{n} - \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, \nabla g\cdot\mathbf{n}\phi + \int_{\|\mathbf{x}\|\geq\epsilon} dx\, \nabla^2 g\phi \right\}$$

Now we can start getting rid of stuff. According to problem 3-5(c), the last term on the right hand side drops out. Note that

$$\|\mathbf{x}\| = \epsilon \;\Rightarrow\; g(\mathbf{x}) = -\frac{1}{4\pi\epsilon}$$

and that

$$\left| \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, \nabla\phi\cdot\mathbf{n} \right| = O(\epsilon^2)$$

i.e. this integral is bounded by a multiple of the area of the sphere of radius epsilon, since the integrand is bounded. Thus

$$\left| \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, g\nabla\phi\cdot\mathbf{n} \right| = O(\epsilon)$$

so the first term goes to 0 as $\epsilon\to 0$.

Finally, note that

- $\mathbf{n}(\mathbf{x}) = -\mathbf{x}/\|\mathbf{x}\|$ (this is the *outward* unit normal of the *outside* of the ball of radius ϵ!);
- from problem 3-5(a) and a little algebra,

  $$\nabla g(\mathbf{x}) = \frac{\mathbf{x}}{4\pi\|\mathbf{x}\|^3}$$

  so

  $$\nabla g(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}) = -\frac{1}{4\pi\|\mathbf{x}\|^2} = -\frac{1}{4\pi\epsilon^2}$$

  on the boundary $\|\mathbf{x}\| = \epsilon$.;

- if $\|\mathbf{x}\| = \epsilon$, then $\phi(\mathbf{x}) = \phi(0) + O(\epsilon)$..

Putting all of this together,

$$-\int_{\|\mathbf{x}\|=\epsilon} dA(x)\, \nabla g \cdot \mathbf{n}\phi = -(\phi(0) + O(\epsilon)) \int_{\|\mathbf{x}\|=\epsilon} dA(x)\, \nabla g \cdot \mathbf{n}$$

$$= \phi(0) + O(\epsilon) \;\rightarrow\; \phi(0)$$

as $\epsilon \to 0$. **q.e.d.**

The Green's function bears the same relation to the Laplace operator that the inverse of a matrix bears to the matrix. To see this, we need some more information about transposition.

The *translation* operator $T_{\mathbf{y}} : \mathcal{D} \to \mathcal{D}$ moves its inputs by $\mathbf{y} \in \mathbf{R}^n$:

$$(T_{\mathbf{y}}\phi)(\mathbf{x}) = \phi(\mathbf{x} + \mathbf{y})$$

Claim: for any $\phi \in \mathcal{D}$, $\mathbf{y} \in \mathbf{R}^n$, $i = 1, ..., n$, and sequence of reals $h_k \to 0$,

$$\lim_{k\to\infty} \frac{1}{h_k}(T_{\mathbf{y}+h_k\mathbf{e}_i}\phi - T_{\mathbf{y}}\phi) = T_{\mathbf{y}}\frac{\partial\phi}{\partial x_i}$$

in the sense of convergence in $\mathcal{D}$.

Proof: we must show that (a) all of the functions on the left of the above limit statement vanish outside some large ball, and (b) they and their derivatives of all orders converge uniformly to the RHS. Concerning (a), suppose that ϕ vanishes outside a ball of radius R, and $|h_k| \le H$ for all k (there must be some such H, since the sequence is convergent). Then the function on the left hand side of the above equation vanishes outside the ball of radius $R + H + |\mathbf{y}|$.

According to the Mean Value Theorem, for each k there is $\hat{h}_k$ between h_k and 0 so that for any $\mathbf{x} \in \mathbf{R}^n$,

$$\left|\frac{1}{h_k}(T_{\mathbf{y}+h_k\mathbf{e}_i}\phi - T_{\mathbf{y}}\phi)(\mathbf{x}) - T_{\mathbf{y}}\frac{\partial\phi}{\partial x_i}(\mathbf{x})\right| = \left|T_{\mathbf{y}+\hat{h}_k\mathbf{e}_i}\frac{\partial\phi}{\partial x_i}(\mathbf{x}) - T_{\mathbf{y}}\frac{\partial\phi}{\partial x_i}(\mathbf{x})\right|$$

$$= \left|\phi(\mathbf{x} + \mathbf{y} + \hat{h}_k\mathbf{e}_i) - \phi(\mathbf{x} + \mathbf{y})\right|$$

Since the ϕ is uniformly continuous (being continuous and vanishing outside a bounded set), the last absolute value tends to zero uniformly in $\mathbf{x}$ as $k \to \infty$.

If $\phi \in \mathcal{D}$ and $u \in \mathcal{D}'$, then the rule

$$\mathbf{y} \mapsto \langle u, T_{\mathbf{y}}\phi\rangle = \langle T'_{\mathbf{y}}u, \phi\rangle$$

defines a function on $\mathbf{R}^d$, and that this function is infinitely often differentiable (proof: use continuity and the claim to see that any first partial exists; now apply the same reasoning repeatedly, using the fact that any mixed partial derivative of a test function is a test function.) This function has a name: it is the *crosscorrelation* of u and ϕ, written

$$u \star \phi(\mathbf{y}) = \langle u, T_{\mathbf{y}}\phi \rangle.$$

As a by-product of the induction, you get the rule: for any $\mathbf{j} \in \mathbf{Z}_+^d$,

$$D^{\mathbf{j}}u \star \phi = u \star D^{\mathbf{j}}\phi = (D^{\mathbf{j}})'u \star \phi$$

Now take for u the Green's function g derived above, a distribution defined by an integrable function. Applying the second partials in turn and adding them up (distributions are *linear* functions!), we get

$$\nabla^2 g \star \phi(\mathbf{y}) = (\nabla^2)' g \star \phi(\mathbf{y}) = \delta \star \phi(\mathbf{y})$$

$$= \langle \delta, T_{\mathbf{y}}\phi \rangle = (T_{\mathbf{y}}\phi)(0) = \phi(\mathbf{y})$$

for any $\mathbf{y} \in \mathbf{R}^d$, i.e.

$$\nabla^2(g \star \phi) = \phi.$$

That is, the crosscorrelation of the Green's function with the right hand side (assumed to be a test function) solves the Poisson equation!

Recall that the Green's function is actually a function, i.e. acts as a distribution via integration. Also, it is an even function of its argument. So

$$g \star \phi(\mathbf{y}) = \int dx g(\mathbf{x})\phi(\mathbf{x}+\mathbf{y}) = \int dx g(\mathbf{x}-\mathbf{y})\phi(\mathbf{x}) = \int dx g(\mathbf{y}-\mathbf{x})\phi(\mathbf{x})$$

The last integral defines another common type of product of functions, namely the *convolution*: the convolution of an integrable function f and a test function $\psi \in \mathcal{D}$ is the function $f * \phi$ defined by

$$\mathbf{y} \mapsto f * \phi(\mathbf{y}) = \int dx f(\mathbf{y}-\mathbf{x})\phi(\mathbf{x})$$

It's easy to verify that for integrable f, $\psi \in \mathcal{D}$,

$$f * \phi = \check{f} \star \phi = f \star \check{\phi}$$

where $\phi \mapsto \check{\phi}$ is the *parity* operator (thanks to Rani Namour for suggesting this name):

$$\check{\phi}(\mathbf{x}) = \phi(-\mathbf{x}).$$

The definition of the parity operator makes sense for any type of function - test, integrable, or neither.

This observation also motivates the extension of the notion of convolution to distributions: by definition, if $u \in \mathcal{D}'$, $\phi \in \mathcal{D}$, then $u * \phi$ is the (C^∞) function given by

$$u * \phi(\mathbf{y}) =< T'_{\mathbf{y}} u, \check{\phi} >= u \star \check{\phi}(\mathbf{y}).$$

For the solution of the Poisson problem with right-hand side $\phi \in \mathcal{D}$, we can just as well write

$$\nabla^2(g \star \phi) = \nabla^2(g * \check{\phi}) = \nabla^2(\check{g} * \phi) = \nabla^2(g * \phi)$$

since g is even (i.e. $\check{g} = g$). In other words, you get the solution of the Poisson problem by *convolving* the right hand side with the Green's function.

You should be aware of a common notational convention in the applied literature. The application of the transpose translation to the delta distribution produces another distribution, and just as with delta it's common to denote it as if it were a function:

$$\delta(\mathbf{x} - \mathbf{y})\text{“} =\text{”} (T'_{\mathbf{y}}\delta)(\mathbf{x})$$

which simply means that

$$\langle \delta(\cdot - \mathbf{y}), \phi \rangle = \phi(\mathbf{y}) \text{ “} = \int dx\, \delta(\mathbf{x} - \mathbf{y})\phi(\mathbf{x})\text{”}$$

As before, the last "equation", giving a common notation for the evaluation of this distribution, has no independent meaning.

It's common to use this notation to express the solution of the Poisson problem by the Green's function in the form

$$\nabla^2 g(\mathbf{x} - \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}).$$

4.6 Problems

Problem 1: Suppose that M is a continuously differentiable matrix-valued function on a subset of $\mathbf{R}^3$. **Show** that the divergence $\nabla \cdot M$ is the only vector valued function defined on the same domain as M and satisfying

$$(\nabla \cdot M) \cdot c = \nabla \cdot (M^T c)$$

for an arbitrary constant vector $c \in \mathbf{R}^3$.

Problem 4-2: Suppose that $\mathbf{v}(\mathbf{x}, t)$ is the velocity field of an ideal fluid moving under the influence of a *conservative* external force density $\mathbf{f}$, i.e.

$$\mathbf{f}(\mathbf{x}, t) = \rho_0 \nabla \beta(\mathbf{x}, t)$$

for some scalar field β (which, you may note, must have units of force $\times$ distance per unit mass, or energy per unit mass). Suppose that a material part of the fluid occupies the volume Ω_t at time t, and that the velocity field is tangent to the boundary at all times and places:

$$\mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}, t) = 0$$

for any point $\mathbf{x}$ on the boundary of Ω_t with normal $\mathbf{n}(\mathbf{x}, t)$.

Show that the *kinentic energy* of this part is constant in t, i.e. that

$$\frac{d}{dt} \frac{\rho_0}{2} \int_{\Omega_t} |\mathbf{v}|^2 = 0$$

Hint: You will have to remember the transport theorem, and work out for example that $\nabla |\mathbf{v}|^2 = 2(\nabla \mathbf{v})^T \mathbf{v}$.

Problem 4-3: **Prove** the following special case of *Bernoulli's theorem*: suppose that $\mathbf{v}(\mathbf{x}), p(\mathbf{x})$ are the velocity and pressure filed of a steady and irrotational motion of an ideal fluid moving under the influence of a conservative external force $\mathbf{f} = \rho_0 \nabla \beta$. **Show** that

$$\frac{|\mathbf{v}|^2}{2} + \frac{p}{\rho_0} - \beta$$

is **constant**, ie. independent of position.

Problem 4-4: Suppose that $\mathbf{v}(\mathbf{x}), p(\mathbf{x})$ are the velocity and pressure field of a steady and irrotational motion of an ideal fluid moving under the influence of no external force, i.e. in my usual notation $\mathbf{f} = 0$. Suppose that the volume occupied by the fluid excludes a domain Σ which is filled with a hard, nonfluid obstacle. **Show** that the total force exerted on the obstacle by the fluid is exactly

$$\frac{\rho_0}{2} \int_{\partial \Omega} dA(\mathbf{x}) \, |\mathbf{v}(\mathbf{x})|^2 \mathbf{n}(\mathbf{x})$$

Hint: You will need to use the result of Problem 3, and also justify several bits of calculus, for instance the identity

$$\int_{\partial \Omega} dA(\mathbf{x}) \mathbf{n}(\mathbf{x}) = 0$$

for any bounded domain Ω - this is a consequence of the divergence theorem for matrix-valued functions.

Problem 4-5: Show that

$$g(\mathbf{x}) = \frac{1}{2\pi} \log(\|\mathbf{x}\|), \ \mathbf{x} \in \mathbf{R}^2$$

is a Green's functions for the Laplacian in $\mathbf{R}^2$.

Chapter 5

Elastic Fluids and Acoustic Waves

5.1 Characterization of Elastic Fluids

While everyday fluids are mostly very difficult to compress, hence well modeled by incompressible continua, gases are notably compressible. In another respect, gases behave like (ideal) fluids: they support only small shear stress. In the idealized limit of no shear stress, gases might be modeled as compressible, or *elastic*, fluids: once again, the stress is scalar (pressure) but now motions in general do not preserve volume, so the divergence of the velocity field is no longer constrained to vanish.

This model consists in part of the laws of mass and momentum balance:

$$\frac{D\rho}{Dt} + (\nabla \cdot \mathbf{v})\rho = 0$$

$$\rho \frac{D\mathbf{v}}{Dt} + \nabla p = f$$

These conditions are not sufficient to determine the motion uniquely under any reasonable circumstances - a replacement is needed for the velocity divergence constraint. The replacement is the *constitutive law* relating density and pressure:

$$p = P(\rho)$$

This law is supposed to reflect on a macroscopic scale the molecular properties of the material. For real gases, it is observationally true that an increase in density is correlated with an increase in pressure, so assume in addition that

$$\frac{dP}{d\rho} > 0.$$

In an ideal fluid, it was not possible to compute the pressure from the density, as the density was assumed constant but the pressure was free to vary from place to place (in general). Moreover the pressure did not determine the velocity field either, as one could add a constant of arbitrary size to the pressure without violating any of the equations of motion. This freedom no longer exists for an elastic fluid.

A familiar example of an elastic fluid constitutive law is the *ideal gas law*, which relates pressure p, density ρ, temperature θ, and a constant R characteristic of the material:

$$p = R\theta\rho$$

If you assume that a volume V is filled with a material obeying the ideal gas law, with homogeneous density, temperature, and pressure, then

$$pV = R\theta\rho V = Rm\theta = \text{const}\,\theta$$

which is the famous "PV=T" of high school chemistry, apart from choice of units.

There is a Bernoulli Theorem for elastic fluids, just as for ideal fluids - see [Gurtin, 1981], pp. 130 ff. I want to emphasize two other aspects of typical gas behaviour here, both of which will be developed much more extensively in the second half of this course.

5.2 Acoustics

If you set $\rho(\mathbf{x}, t) = \rho_0$, a constant, $\mathbf{v}(\mathbf{x}, t) = 0$, and $p(\mathbf{x}, t) = p_0 = P(\rho_0)$, and assume absence of external force, you have a constant or *equilibrium* solution of the elastic fluid equations of motion: pressure and density are homogeneous and the material (gas) is not moving. Now suppose that you add a very small, possibly time-varying force density $\delta\mathbf{f}$ to the right hand side of the momentum balance law. Then presumably the motion varies by a small amount from the equilibrium solution:

$$p = p_0 + \delta p,\ \ \rho = \rho_0 + \delta\rho,\ \ \mathbf{v} = \delta\mathbf{v}$$

where anything with a δ in front of it is supposed to be small, and even to have small derivatives.

Substitute these expressions in the mass and momentum balance laws and the constitutive law:

$$\frac{\partial(\rho_0 + \delta\rho)}{\partial t} + (\nabla \cdot \delta\mathbf{v})(\rho_0 + \delta\rho) = 0$$

$$(\rho_0 + \delta\rho)\frac{\partial \delta\mathbf{v}}{\partial t} + (\nabla\delta\mathbf{v})\delta\mathbf{v} + \nabla(p_0 + \delta p) = \delta\mathbf{f}$$

$$p_0 + \delta p = P(\rho_0 + \delta\rho)$$

Presumably when all things with δs are small, the products of things with δs are much smaller. So *approximate* the balance laws by dropping these product terms, and the constitutive law by using its first order Taylor series. Make use of the equilibrium equations to eliminate the terms without δs. The result is the system

$$\frac{\partial \delta\rho}{\partial t} + \rho_0 \nabla \cdot \delta\mathbf{v} = 0$$

$$\rho_0 \frac{\partial \delta\mathbf{v}}{\partial t} + \nabla \delta p = \delta\mathbf{f}$$

$$\delta p = \frac{dP}{d\rho}(\rho_0)\delta\rho$$

Since $dP/d\rho > 0$, you can solve this equation for $\delta\rho$ in terms of δp, and substitute the result in the conservation of mass. The rate of change of pressure with respect to density is so important that it gets its own name:

$$c = \sqrt{\frac{dP}{d\rho}(\rho_0)}$$

is the *sound speed* in the fluid (the reason for the name will become apparent shortly).

With this substitution to eliminate $\delta\rho$, the remaining two equations become the *acoustic equations of motion*:

$$\frac{1}{\rho_0 c^2}\frac{\partial \delta p}{\partial t} + \nabla \cdot \delta\mathbf{v} = 0$$

$$\rho_0 \frac{\partial \delta\mathbf{v}}{\partial t} + \nabla \delta p = \delta\mathbf{f}$$

This system is equivalent to a single second order partial differential equation: differentiate the first equation with respect to time, use the second equation to eliminate the time derivative of velocity, and you get

$$\frac{1}{\rho_0 c^2}\frac{\partial^2 \delta p}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho_0}\nabla \delta p\right) = -\frac{\nabla \cdot \delta\mathbf{f}}{\rho_0}$$

Since ρ_0 is constant, this is equivalent to

$$\frac{1}{c^2}\frac{\partial^2 \delta p}{\partial t^2} - \nabla^2 \delta p = -\nabla \cdot \delta\mathbf{f}$$

which is the *acoustic wave equation*.

5.3 Problems.

Problem 5-1: [50 pts] This problem concerns small amplitude motion of an elastic fluid contained in a domain $\Omega \subset \mathbf{R}^3$. The motion is described by the perturbational fields $\delta p(\mathbf{x}, t), \delta \mathbf{v}(\mathbf{x}, t)$. These are assumed to satisfy the acoustic wave equation described in section 5.2. Define the function of time

$$E(t) = \frac{1}{2} \int_\Omega dx \left\{ \rho_0 \|\delta \mathbf{v}(\mathbf{x}, t)\|^2 + \frac{\delta p(\mathbf{x}, t)^2}{\rho_0 c_0^2} \right\}$$

Show that $E(t)$ is *independent of* t if no external force (perturbation) is acting on the elastic fluid and either

(i) $\delta \mathbf{v}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = 0, \mathbf{x} \in \partial\Omega$ (the "no-flow" condition - i.e. Ω is enclosed by an impermeable barrier), or

(ii) $\delta p(\mathbf{x}, t) = 0, \mathbf{x} \in \partial\Omega$ (the "free surface" or "pressure release" condition).

As will be established in Ch. 7, $E(t)$ is actually the *energy* in the perturbational field, and this result can be paraphrased by the statement that this energy like quantity is *conserved* by the motion, provided that either no-flow or pressure-release conditions are imposed on $\partial\Omega$.

Problem 5-2: [25 pts] Show that any function of the form

$$\delta p(\mathbf{x}, t) = f(t - \mathbf{p} \cdot \mathbf{x})$$

is a solution of the acoustic wave equation, for any *waveform* f and any *slowness vector* $\mathbf{p}$, provided that

$$\|\mathbf{p}\| = \frac{1}{c_0}$$

(the units implied by this equation explain the name of $\mathbf{p}$).

Problem 5-3: [25 pts] Suppose that $\Omega = \{\mathbf{x} : x_3 > 0\}$ is the positive-x_3 half space, and that $f_+(t)$ and $\mathbf{p}_+$ are a waveform and slowness vector respectively. Exhibit second pair $f_-(t), \mathbf{p}_-$ of waveform and slowness vector so that the sum of the corresponding solutions of the acoustic wave equation

$$\delta p(\mathbf{x}, t) = f_+(t - \mathbf{p}_+ \cdot \mathbf{x}) + f_-(t - \mathbf{p}_- \cdot \mathbf{x})$$

is also a solution of the acoustic wave equation, for which the free surface condition holds:

$$\delta p(\mathbf{x}, t) = 0, \ \mathbf{x} \in \partial\Omega$$

["Exhibit" here means: give explicit formulas for $f_-, \mathbf{p}_-$ in terms of $f_+, \mathbf{p}_+$.]

Show that the vectors $\mathbf{p}_+, \mathbf{p}_-$ make equal angles with the boundary normal $\mathbf{e}_3$. [This result is known as *Snell's law of reflection.*]

Chapter 6

Newtonian Fluids and the Navier-Stokes Equations

6.1 Basic Consitutive Assumptions

A Newtonian fluid is a material whose motion satisfy the following constraints:

- density is constant, $\rho(\mathbf{x}, t) \equiv \rho_0$;
- only volume-preserving motions occur, i.e. $\nabla \cdot \mathbf{v} \equiv 0$;
- the stress is a pressure plus a linear function of the *velocity gradient* $\nabla \mathbf{v}$, which measures the relative velocities of nearby fluid particles.

Thus Newtonian fluids are incompressible and have homogeneous density, like ideal fluids. Unlike ideal fluids, Newtonian fluids can support shear stress, when the velocity is nonuniform, i.e. changes in space.

The stress is (in part) a matrix-valued linear function of a matrix, and such a thing is a bit complex. You can express it with the help of a four-index array, C_{ijkl}:

$$S_{ij} = -p\delta_{ij} + \sum_{k,l=1}^{3} C_{ijkl} \frac{\partial v_k}{\partial x_l}$$

or more compactly

$$S = -pI + C[\nabla \mathbf{v}]$$

Together with the laws of conservation of mass and momentum and the volume-preserving condition in their familiar forms, this condition provides the ten equations one would expect to need to determine the ten unknowns $\rho, \mathbf{v}$, and S.

Stress is symmetric, hence has six independent components; the velocity gradient appears to have nine, so there would appear to be no less than fifty-four independent constants C_{ijkl} involved in the relation between stress and velocity. That there are actually not nearly so many is a consequence of a natural physical invariance property, to which we now turn.

6.2 Observer independence

The principle of *observer independence* codifies the idea that the constitutive assumptions characterizing a type of material should not depend on how it is observed, i.e. on changes of coordinates, *even time varying ones*, which are distance-preserving. This is not to say that the various component functions which go into the statements of the dynamical laws (conservation principles) will necessarily remain the same, only that so long as distances remain the same in a new coordinate system, the *form* of these equations will remain the same as in the old coordinate system.

In equations, imagine that you have two coordinate systems, $\mathbf{x}$ and $\mathbf{x}'$, related by a time-varying Euclidean transformation (which preserves distances!):

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}, t) = \mathbf{x}'_0(t) + R(t)(\mathbf{x} - \mathbf{x}_0)$$

in which $\mathbf{x}'_0 : \mathbf{R} \to \mathbf{R}^3$ and $R : \mathbf{R} \to \mathbf{R}^{3\times 3}$ are differentiable and $R(t)$ is orthogonal for each t and $\det R = 1$. As pointed out in section 1.4.2, only affine coordinate changes with orthogonal R preserve all distances. You'll see why $\det R$ must be positive shortly.

You are to think of $\mathbf{x}$ and $\mathbf{x}'$ as two different descriptions of *the same point in space*, one of whom is moving and spinning, possibly with time-dependent rates, relative to the other.

If $\mathbf{x}(\mathbf{X}, t)$ is a motion, then *the same motion*, described in the $\mathbf{x}'$ coordinate system, is

$$\mathbf{x}'(\mathbf{X}, t) = \mathbf{x}'_0(t) + R(t)(\mathbf{x}(\mathbf{X}, t) - \mathbf{x}_0) = \mathbf{x}'(\mathbf{x}(\mathbf{X}, t), t)$$

Various quantities related to the motion functions are also related, as follows.

From the chain rule,

$$\frac{\partial \mathbf{x}'}{\partial t}(\mathbf{X}, t) = \frac{d\mathbf{x}'_0}{dt}(t) + \frac{dR}{dt}(t)\mathbf{x}(\mathbf{X}, t) + R(t)\frac{\partial \mathbf{x}}{\partial t}(\mathbf{X}, t)$$

$$\nabla_{\mathbf{X}}\mathbf{x}'(\mathbf{X}, t) = \nabla_{\mathbf{x}}\mathbf{x}'(\mathbf{x}(\mathbf{X}, t), t)\nabla_{\mathbf{X}}\mathbf{x}(\mathbf{X}, t) = R(t)\nabla_{\mathbf{X}}\mathbf{x}(\mathbf{X}, t)$$

This last formula justifies the limitation $\det R > 0$: otherwise the essential property $\det \nabla \mathbf{x} > 0$ cannot hold in both coordinate systems. Write $\mathbf{X}'(\mathbf{x}', t)$ for the inverse function of $\mathbf{x}'(\mathbf{X}, t)$, i.e. $\mathbf{X}'(\mathbf{x}'(\mathbf{X}, t), t) = \mathbf{X}$ for any $\mathbf{X} \in \mathbf{R}^3$. If $\mathbf{x}'$ and $\mathbf{x}$ refer to the same *material*, i.e. if $\mathbf{x}' = \mathbf{x}'(\mathbf{x}, t)$, then necessarily $\mathbf{X}'(\mathbf{x}'(\mathbf{x}, t), t) = \mathbf{X}(\mathbf{x}, t)$, since the material coordinates simply label the material.

The material velocity field $\mathbf{v}'(\mathbf{x}', t)$ in the $\mathbf{x}'$ coordinate system, is

$$\mathbf{v}'(\mathbf{x}', t) = \frac{\partial \mathbf{x}'}{\partial t}(\mathbf{X}'(\mathbf{x}', t), t)$$

so

$$\mathbf{v}'(\mathbf{x}'(\mathbf{x}, t), t) = \frac{\partial \mathbf{x}'}{\partial t}(\mathbf{X}'(\mathbf{x}'(\mathbf{x}, t), t), t)$$

$$= \frac{\partial \mathbf{x}'}{\partial t}(\mathbf{X}(\mathbf{x}, t), t) = \frac{d\mathbf{x}'_0}{dt}(t) + \frac{dR}{dt}(t)\mathbf{x}(\mathbf{X}(\mathbf{x}, t), t) + R(t)\frac{\partial \mathbf{x}}{\partial t}(\mathbf{X}(\mathbf{x}, t), t)$$

$$= \frac{d\mathbf{x}'_0}{dt}(t) + \frac{dR}{dt}(t)\mathbf{x} + R(t)\mathbf{v}(\mathbf{x}, t)$$

So the velocity in the $\mathbf{x}'$ coordinate system differs from that in the $\mathbf{x}$ coordinate system both by the instantaneous rotation $R(t)$ and by the velocity of the $\mathbf{x}'$ coordinates relative to the $\mathbf{x}$ coordinates (the first two summands).

The velocity gradients (Jacobians) are also related:

$$\nabla_{\mathbf{x}} \mathbf{v}'(\mathbf{x}'(\mathbf{x}, t), t) = (\nabla_{\mathbf{x}'} \mathbf{v}')(\mathbf{x}'(\mathbf{x}, t), t) \nabla_{\mathbf{x}} \mathbf{x}'(\mathbf{x}, t)$$

$$= (\nabla_{\mathbf{x}'} \mathbf{v}')(\mathbf{x}'(\mathbf{x}, t), t) R(t)$$

$$= R(t) \nabla_{\mathbf{x}} \mathbf{v}(\mathbf{x}, t) + \frac{dR}{dt}(t)$$

or

$$(\nabla_{\mathbf{x}'} \mathbf{v}')(\mathbf{x}'(\mathbf{x}, t), t) = R(t) \nabla_{\mathbf{x}} \mathbf{v}(\mathbf{x}, t) R(t)^T + \frac{dR}{dt}(t) R(t)^T$$

which is the required relation.

To see how velocity divergences are related, note the following facts:

- a matrix and its transpose have the same trace;
- skew symmetric matrices Q ($Q^T = -Q$) have zero trace: $\mathrm{tr}Q = \mathrm{tr}Q^T = -\mathrm{tr}Q$, so $\mathrm{tr}Q = 0$;
- the trace of a matrix product does not depend on order: $\mathrm{tr}AB = \mathrm{tr}BA$ so long as both sides make sense and are square;
- similar matrices have the same trace: $\mathrm{tr}(RAR^{-1}) = \mathrm{tr}(AR^{-1}R) = \mathrm{tr}A$

- if $R(t)$ is differentiable and orthogonal for each t, then $\frac{dR}{dt}(t)R^T(t)$ is skew-symmetric:

$$0 = \frac{d}{dt}I = \frac{d}{dt}(R(t)R^T(t)) = \frac{dR}{dt}(t)R^T(t) + R(t)\left(\frac{dR}{dt}(t)\right)^T$$

$$= \frac{dR}{dt}(t)R^T(t) + \left(\frac{dR}{dt}(t)R^T(t)\right)^T$$

The velocity divergence is the trace of its Jacobian:

$$(\nabla_{\mathbf{x}'} \cdot \mathbf{v}')(\mathbf{x}'(\mathbf{x},t),t) = \operatorname{tr}(\nabla_{\mathbf{x}'}\mathbf{v}')(\mathbf{x}'(\mathbf{x},t),t) =$$

$$= \operatorname{tr} R(t)\nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x},t)R(t)^T + \operatorname{tr}\left(\frac{dR}{dt}(t)R(t)^T\right) = \nabla_{\mathbf{x}} \cdot \mathbf{v}(\mathbf{x},t)$$

using the just-listed results on traces.

One prominent quantity has not been mentioned so far in this discussion: the stress. We have actually already seen how the stress changes when coordinates change by a Euclidean transformation (section 1.4.2), and that analysis applies equally well when the Euclidean transformation depends on time:

$$S'(\mathbf{x}'(\mathbf{x},t),t) = R(t)S(\mathbf{x},t)R(t)^T$$

These relations quickly imply that the constitutive assumptions appearing in the fluid models of the last two sections are independent of observer (homework!). For Newtonian fluids, two critical properties of the mapping C follow from observer independence:

1. C depends only on the *symmetric part* of the velocity gradient: for any 3×3 matrix L,
$$C[L] = C\left[\frac{1}{2}(L + L^T)\right]$$
2. C is an *isotropic* transformation, i.e. for any symmetric L, orthogonal R,
$$RC[L]R^T = C[RLR^T]$$

Proof of (1) : First, observe that for any 3×3 matrix L, there is a motion $\mathbf{x}(\mathbf{x}_0,t)$ for which $\mathbf{v}(\mathbf{x},t) = L\mathbf{x}$, hence $\nabla\mathbf{v}(\mathbf{x},t) = L$: set

$$\mathbf{x}(\mathbf{X},t) = e^{tL}\mathbf{X}.$$

Note that $a(t) = \det e^{tL}$ satisfies

$$\frac{da}{dt}(t) = \det e^{tL} \left(\operatorname{tr} \left(\frac{d}{dt} e^{tL} \right) e^{-tL} \right)$$

$$= (\operatorname{tr} L) a(t)$$

and $a(0) = 1$, whence $\det e^{tL}$ is always positive. The map $\mathbf{X} \mapsto \mathbf{x}(\mathbf{X}, t)$ is plainly invertible, so in fact $\mathbf{x}(\mathbf{X}, t)$ actually defines a motion.

Note that

$$\frac{d\mathbf{x}}{dt}(\mathbf{X}, t) = Le^{tL}\mathbf{X} = L\mathbf{x}(\mathbf{X}, t)$$

whence $\mathbf{v}(\mathbf{x}, t) = L\mathbf{x}$ as claimed.

Note that if $\mathbf{x}(\mathbf{X}, t)$ as just described is to be the motion of a Newtonian fluid, it must be volume-preserving, whence $\nabla \cdot \mathbf{v} = \operatorname{tr} L = 0$. So assume from now on that L has zero trace.

The stress tensor for a Newtonian fluid under this motion must take the form

$$S(\mathbf{x}, t) = p(\mathbf{x}, t)I + C[L]$$

where $p(\mathbf{x}, t)$ is the pressure, which we leave unspecified, and $\operatorname{tr} C[L] = 0$.

For any 3×3 matrix L, motion $\mathbf{x}(\mathbf{X}, t)$ as just constructed, denote by W the skew-symmetric part of L:

$$W = \frac{1}{2}(L - L^T)$$

Set

$$R(t) = e^{-tW}, \ \mathbf{x}'(\mathbf{x}, t) = R(t)\mathbf{x}$$

Since W is skew-symmetric, $R(t)$ is orthogonal for each t.

In the $\mathbf{x}'$ coordinate system,

$$S'(\mathbf{x}', t) = p'(\mathbf{x}', t)I + C[\nabla_{\mathbf{x}'}\mathbf{v}'(\mathbf{x}', t)]$$

$$= R(t)S(\mathbf{x}, t)R(t)^T = R(t)(p(\mathbf{x}, t)I + C[L])R(t)^T$$

$$= p(\mathbf{x}, t)I + R(t)C[L]R(t)^T$$

Since S and S' are similar, they have the same trace, so $p'(\mathbf{x}', t) = p(\mathbf{x}, t)$. As shown previously,

$$\nabla_{\mathbf{x}'}\mathbf{v}'(\mathbf{x}', t) = R(t)\nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x}, t)R(t)^T + \frac{dR}{dt}(t)R(t)^T$$

$$= R(t)LR(t)^T - W$$

At $t = 0$, then,

$$S'(\mathbf{x}', 0) - p'(\mathbf{x}', 0)I = C[L - W] = C\left[\frac{1}{2}(L + L^T)\right]$$
$$= S(\mathbf{x}, 0) - p(\mathbf{x}, 0)I = C[L]$$

q.e.d.

Proof of (2) : this time take R constant, $\mathbf{x}' = R\mathbf{x}$, so that

$$\nabla_{\mathbf{x}'}\mathbf{v}'(\mathbf{x}, t) = R\nabla_{\mathbf{x}}\mathbf{v}(\mathbf{x}, t)R^T$$

From (1), we might as well confine our attention to symmetric, zero-trace matrices L: any matrix orthogonally symmetric to such a matrix is also symmetric and zero-trace. For the motion generated by such a matrix as above, and arbitrary orthogonal R,

$$S'(\mathbf{x}', t) - p'(\mathbf{x}', t)I = C[RLR^T]$$

$$= R(S(\mathbf{x}, t) - p(\mathbf{x}, t)I)R^T = RC[L]R^T$$

q.e.d.

What does isotropy mean? Recall that $C[\nabla\mathbf{v}]$ is (part of the) stress: it transforms under a coordinate change $\mathbf{x}' = \mathbf{x}'_0 + R\mathbf{x}$ (with orthogonal R) like any other stress, namely to $RC[\nabla\mathbf{v}]R^T$. That is, *if you measure this part of the stress tensor in the new coordinate system $\mathbf{x}'$, it will be related to the same part measured in the old system as indicated.* Thus you can take your *new* measurements C' and convert them back to the old measurements by the inverse similarity transform: $C = R^T C' R$. On the other hand, the velocity Jacobian $\nabla\mathbf{v}$ also transforms by similarity, $\nabla\mathbf{v}' = R\nabla\mathbf{v}R$. So the following two sequences of operations:

- measure velocity Jacobian in coordinate system $\mathbf{x}$
- compute corresponding stress
- transform stress to new coordinate system

and

- measure velocity Jacobian in coordinate system $\mathbf{x}$
- transform velocity gradient to new coordinate system $\mathbf{x}'$
- compute stress in new coordinate system

will give you *the same result.* In that sense, orientation of the coordinate system does not make any difference in the prediction of stress from velocity Jacobian.

6.3 Isotropic Transformations

The structure of isotropic transformations on the space of symmetric matrices is quite restricted. There is no reason to limit ourselves to three dimensions in studying the implications of isotropy, though the only applications in this course will be three dimensional!

Structure Theorem for Isotropic Transformations: Suppose that C is a linear transformation on the vector space of symmetric $n \times n$ matrices, and is isotropic: for any orthogonal $n \times n$ matrix R,

$$C[RLR^T] = RC[L]R^T$$

for any symmetric L. Then there exist constants μ, λ so that

$$C[L] = 2\mu L + \lambda(\operatorname{tr} L)I$$

Proof: The spectral theorem for symmetric matrices states that

$$L = \sum_{i=1}^{k} \lambda_i P_i$$

where the λ_i are the k (real) distinct eigenvalues of L and P_i corresponding orthogonal projectors onto the subspaces of eigenvectors for λ_i. The dimension of range(P_i) is the dimension of the eigenspace for λ_i, i.e. its multiplicity. If the multiplicity of λ_i is 1, then $P_i = \mathbf{u}\mathbf{u}^T$ where $\mathbf{u}$ is a unit eigenvector with eigenvalue λ_i. In general P_i is a sum of rank 1 matrices of this form, with $\mathbf{u}$ ranging over an orthonormal set of eigenvectors for λ_i.

First, we claim that every eigenvector of L is an eigenvector of $C[L]$. Let $\mathbf{u}$ be an eigenvector of L, and choose R to be the reflection across the plane orthogonal to $\mathbf{u}$, i.e.

$$R\mathbf{u} = -\mathbf{u},\ R\mathbf{v} = \mathbf{v} \text{ if } \mathbf{v} \cdot \mathbf{u} = 0$$

Since R has the same eigenvectors as L, R and L commute, hence

$$RC[L]R^T = C[RLR^T] = C[LRR^T] = C[L]$$

i.e.

$$RC[L] = C[L]R$$

Thus

$$RC[L]\mathbf{u} = C[L]R\mathbf{u} = -C[L]\mathbf{u}$$

which means that $C[L]\mathbf{u}$ is an eigenvector of R with eigenvalue -1 - which is only possible if $C[L]\mathbf{u}$ is a multiple of $\mathbf{u}$, since the -1 eigenspace of R has dimension 1. In other words, $\mathbf{u}$ is an eigenvector of $C[L]$. Since $\mathbf{u}$ was any eigenvector of L, this establishes the claim.

Since C is linear, its behaviour is therefore determined by its action on matrices of the form $L = \mathbf{u}\mathbf{u}^T$, where $\mathbf{u}$ is a unit vector. The eigenvalues of L are 1 (multiplicity 1) (eigenvectors = multiples of $\mathbf{u}$) and 0 (multiplicity $n-1$) (eigenvectors = all vectors orthogonal to $\mathbf{u}$). Thus $\mathbf{u}$ is also an eigenvector of $C[L]$, and the entire subspace of $\mathbf{R}^n$ orthogonal to $\mathbf{u}$ also consists of eigenvectors. If all vectors in a subspace are eigenvectors, then they all belong to the same eigenvalue. So

$$C[L] = \alpha(\mathbf{u})\mathbf{u}\mathbf{u}^T + \beta(\mathbf{u})(I - \mathbf{u}\mathbf{u}^T) = 2\mu(\mathbf{u})\mathbf{u}\mathbf{u}^T + \lambda(\mathbf{u})I$$

where the eigenvalues of $C[L]$ are $\alpha(\mathbf{u})$ and $\beta(\mathbf{u})$, which so far as we know at the moment depend on $\mathbf{u}$, and $\mu = \frac{1}{2}(\alpha(\mathbf{u}) - \beta(\mathbf{u}))$, $\lambda(\mathbf{u}) = \beta(\mathbf{u})$.

To see that in fact λ and μ are independent of $\mathbf{u}$, choose any orthogonal matrix R and set $\mathbf{v} = R\mathbf{u}$. Then

$$\begin{aligned} C[\mathbf{v}\mathbf{v}^T] &= 2\mu(\mathbf{v})\mathbf{v}\mathbf{v}^T + \lambda(\mathbf{v})I \\ = C[R\mathbf{u}(R\mathbf{u})^T] &= C[R\mathbf{u}\mathbf{u}^T R^T] = RC[\mathbf{u}\mathbf{u}^T]R^T \\ = R(2\mu(\mathbf{u})\mathbf{u}\mathbf{u}^T + \lambda(\mathbf{u})I)R^T &= 2\mu(\mathbf{u})\mathbf{v}\mathbf{v}^T + \lambda(\mathbf{u})I \end{aligned}$$

so

$$(2(\mu(\mathbf{v}) - \mu(\mathbf{u})))\mathbf{v}\mathbf{v}^T = -(\lambda(\mathbf{v}) - \lambda(\mathbf{u}))I$$

Since $\mathbf{v}\mathbf{v}^T$ and I are linearly independent, it follows that

$$\mu(\mathbf{v}) = \mu(\mathbf{u}),\ \lambda(\mathbf{v}) = \lambda(\mathbf{u})$$

As R ranges over all orthogonal matrices, $\mathbf{v} = R\mathbf{u}$ ranges over all unit vectors. Therefore in fact λ and μ are independent of $\mathbf{u}$.

Since $\operatorname{tr} \mathbf{u}\mathbf{u}^T = 1$ for any unit vector $\mathbf{u}$, we have actually shown that

$$C[\mathbf{u}\mathbf{u}^T] = 2\mu\mathbf{u}\mathbf{u}^T + \lambda(\operatorname{tr} \mathbf{u}\mathbf{u}^T)I$$

Since C is linear, any symmetric L is a linear combination of orthogonal projectors (spectral theorem!), and any orthogonal projector is a linear combination of rank 1 matrices of the form uu^T, the "if" part of the theorem follows.

The "only if" part is a homework excercise. q.e.d.

Corollary: Special case of symmetric zero trace matrices: Suppose that C is an isotropic linear transformation on the vector space of symmetric zero trace matrices. Then for some constant μ,

$$C[L] = 2\mu L$$

Proof: Again, the proof of the "if" part is given here. The "only if" part is actually just a special case of the "only if" part of the previous theorem.

For an arbitrary symmetric matrix L, define

$$\hat{C}[L] = C\left[L - \frac{1}{n}(\operatorname{tr} L)I\right]$$

This makes sense, since the matrix in brackets on the right has trace zero. It's very easy to verify that $\hat{C}$ is isotropic, so by the preceding theorem

$$\hat{C}[L] = 2\mu L + \lambda(\operatorname{tr} L)I$$

for suitable constants μ, λ. If L has zero trace, then $C[L] = \hat{C}[L]$, so the corollary is proved. q.e.d.

6.4 The Navier-Stokes Equations

Over the last two sections we've seen that the constitutive assumption regarding the form of stress in Newtonian fluids,

$$S = -pI + C[\nabla \mathbf{v}]$$

together with the requirement of independence of observer, mandate that for some constant μ, in fact

$$S = -pI + \mu(\nabla \mathbf{v} + (\nabla \mathbf{v})^T)$$

Note that $\nabla \cdot (\nabla \mathbf{v})^T = \nabla(\nabla \cdot \mathbf{v})$ (exercise!) and the latter vanishes because all Newtonian fluid motions are volume preserving. Also $\nabla \cdot (\nabla \mathbf{v}) = \nabla^2 \mathbf{v}$, where the right hand side means the vector of Laplacians of the velocity components. Since the constant-density and volume-preserving assumptions imply the conservation of mass, the characteristic equations of Newtonian fluid motion are thus

$$\rho_0 \frac{D\mathbf{v}}{Dt} + \nabla p - \mu \nabla^2 \mathbf{v} = \mathbf{f}$$

$$\nabla \cdot \mathbf{v} = 0$$

These are the *Navier-Stokes equations*. Given ρ_0, μ, and $\mathbf{f}$, they consitute four nonlinear partial differential equations for the four components of $p, \mathbf{v}$.

The scalar μ is the *coefficient of viscosity*: it regulates the strength of shear (viscous) internal forces in the fluid.

The number of equations appears to be adequate, so that the Navier-Stokes equations should determine a Newtonian fluid motion given its initial state. On the other hand, it is

far from obvious that these equations have solutions. It is worth emphasizing once again that one cannot appeal to physics to assert the existence of solutions: many (reasonable but somewhat arbitrary) approximations and assumptions have gone into the formulation of the Newtonian fluid model, so that it is clearly only an approximation of the actual physics of real fluids. So the fact that real fluids flow does not imply that the Navier-Stokes equations have solutions! These equations, like the other models presented in this chapter, are *models*, and nothing more: their existence is justified if they predict actual motions of real materials well under the limited range of circumstances for which they are intended. Existence of solutions is a precondition for making reasonable predictions, and must be established as a property of the model. Solutions of the Navier-Stokes equations exist for some classes of initial data and limited time intervals. Global existence of solutions is still a subject of research.

If the motion is *steady* (all time derivatives are negligible) and *slow* (velocities are small, so quadratic quantities in the velocity, notably $(\nabla \mathbf{v})\mathbf{v}$, are negligible), then the Navier-Stokes equations reduce to the *Stokes equations*:

$$\nabla p - \mu \nabla^2 \mathbf{v} = \mathbf{f}$$

$$\nabla \cdot \mathbf{v} = 0$$

Take the divergence of the first equation: since the divergence commutes with the Laplacian, all that's left in view of the second equation is

$$\nabla^2 p = \nabla \cdot \mathbf{f}$$

With $\mathbf{f}$ given, this is a *Poisson problem* for p, i.e. an inhomogeneous version of the Laplace equation.

The second chapter of these notes shows how to solve the Poisson problem for various boundary conditions modeling natural constraints on the fluid. Assuming that we can solve it, we have p. Then the Stokes equations themselves might be viewed as three more Poisson problems for the three components v_i of $\mathbf{v}$. Thus solving the Stokes equations amounts to solving four Poisson problems.

[Gurtin, 1981], sections 22-24, describes some basic properties of solutions of the Navier-Stokes equations.

6.5 Problems.

Problem 6-1: Show that the ideal fluid, elastic fluid, and Newtonian fluid models satisfy observer independence. What this means: you need to show that the *constitutive assumptions*

of each model are preserved under "change of observer", i.e. time-dependent Euclidean transformations of the form presented in the text: $\mathbf{x}'(\mathbf{x},t) = \mathbf{x}'_0(t) + R(t)(\mathbf{x} - \mathbf{x}_0)$ with $R(t)$ orthogonal for each t, $\det R(t) = 1$, and $\mathbf{x}'_0, R$ continuously differentiable as functions of t. The constitutive assumptions are:

- ideal and Newtonian fluids: the density is constant
- ideal and Newtonian fluids: the motion is volume-preserving
- ideal and elastic fluids: the stress tensor is scalar ($S = -pI$)
- elastic fluids: the pressure is a monotone increasing function of the density
- Newtonian fluids: $S = -pI + \mu(\nabla\mathbf{v} + (\nabla\mathbf{v})^T)$

You need to show that each of these assumptions is also true in the $\mathbf{x}'$ coordinate system if it is true in the $\mathbf{x}$ coordinate system. For the assumptions about the form of stress, remember that the desired result is

$$S'(\mathbf{x}'(\mathbf{x},t),t) = R(t)S(\mathbf{x},t)R(t)^T$$

In showing independence of observer for the Newtonian fluid model, you will incidently show the "only if" part of the theorem on the structure of isotropic transformations.

Problem 6-2: Once again suppose that coordinate systems $\mathbf{x}$ and $\mathbf{x}'$ are related by a Euclidean transformation: $\mathbf{x}'(\mathbf{x},t) = \mathbf{x}'_0(t) + R(t)(\mathbf{x} - \mathbf{x}_0)$, with $\mathbf{x}'_0(t)$ and $R(t)$ as before. Show that the conservation of momentum law in the $\mathbf{x}'$ coordinate system takes the same form

$$\rho' \frac{D\mathbf{v}'}{Dt} = \nabla_{\mathbf{x}'} \cdot S' + \mathbf{f}'$$

as in the $\mathbf{x}$ coordinate system:

$$\rho \frac{D\mathbf{v}}{Dt} = \nabla_{\mathbf{x}} \cdot S + \mathbf{f}$$

and state an explicit relation between the external force vector fields $\mathbf{f}'$ and $\mathbf{f}$. Observe that $\mathbf{f}'(\mathbf{x}'(\mathbf{x},t),t) = R(t)\mathbf{f}(\mathbf{x},t)$ (as one might expect from the covariance of force) only if $\frac{d\mathbf{x}'_0}{dt}, R$ are independent of t. Changes of observer satisfying this constraint are called *Galilean*.

Problem 6-3: Assume that $p_0, \mathbf{v}_0$ are a constant scalar resp. vector. Observe that they represent a slow, steady motion of a Newtonian fluid, i.e. solve the Stokes equation with vanishing external force ($\mathbf{f} = 0$). Suppose that $p(\mathbf{x},t) = p_0 + \delta p(\mathbf{x},t), \mathbf{v}(\mathbf{x},t) = \mathbf{v}_0 + \delta\mathbf{v}(\mathbf{x},t), \mathbf{f} = \delta\mathbf{f}$ represents a small perturbation of this constant solution. Following a procedure similar to

that used to obtain the acoustic equations from the elastic fluid model, derive the *linearized Navier-Stokes equations* under the assumption that all δ quantities are small:

$$\rho_0 \frac{\partial \delta \mathbf{v}}{\partial t} + \rho_0 (\nabla \delta \mathbf{v}) \mathbf{v}_0 + \nabla \delta p - \mu \nabla^2 \delta \mathbf{v} = \delta \mathbf{f}$$

$$\nabla \cdot \delta \mathbf{v} = 0$$

Chapter 7

Linear Elasticity

Stress in elastic materials is a function of the Jacobian of the motion function:

$$S(\mathbf{x}, t) = \mathcal{S}((\nabla_{\mathbf{q}}\mathbf{x})(\mathbf{q}(\mathbf{x}, t), t))$$

The function $\mathcal{S} : \mathbf{R}^{3\times 3} \to \mathbf{R}^{3\times 3}_{\text{symm}}$ is characteristic of the elastic material.

The meaning of this assumption is partly in what it rules out: in contrast to Newtonian fluids, elastic material stress does not depend on the relative velocities of nearby particles, but only on their relative displacement.

7.1 Independence of Observer

The change of variables

$$\mathbf{x}'(\mathbf{x}, t) = \mathbf{x}'_0(t) + R(t)\mathbf{x}$$

with $R(t)$ orthogonal, as in the last Chapter, gives through the chain rule

$$\nabla_{\mathbf{q}}(\mathbf{x}'(\mathbf{x}(\mathbf{q}, t), t)) = R(t)\nabla_{\mathbf{q}}\mathbf{x}(\mathbf{q}, t)$$

So the independence of observer property of stress, also explained in the last Chapter, becomes

$$S'(\mathbf{x}'(\mathbf{x}, t), t) = R(t)S(\mathbf{x}, t)R(t)^T = R(t)\mathcal{S}((\nabla_{\mathbf{q}}\mathbf{x})(\mathbf{q}(\mathbf{x}, t), t))R(t)^T$$
$$= \mathcal{S}(R(t)(\nabla_{\mathbf{q}}\mathbf{x})(\mathbf{q}(\mathbf{x}, t), t)$$

Since $\nabla\mathbf{x}$ can take any value in $\mathbf{R}^{3\times 3}$, this condition implies that

$$R\mathcal{S}(L)R^T = \mathcal{S}(RL)$$

This condition is equivalent to the independence of observer property.

One immediate consequence is that the stress is determined by the symmetric part of the Jacobian: since any matrix L has a polar decomposition $L = RQ$ with orthogonal R and symmetric Q,

$$\mathcal{S}(L) = \mathcal{S}(RQ) = R\mathcal{S}(Q)R^T$$

Thus if you know $\mathcal{S}$ on $\mathbf{R}^{3\times3}_{\text{symm}}$, you know $\mathcal{S}$ on all matrices.

Irritating Question: The independence of observer property mandated that $\det R = 1$; however that property is not assured of the orthogonal factor of the polar decomposition. Why does this fact not invalidate the claim just made (that $\mathcal{S}$ is determined by its action on symmetric matrices)? (Hint: what is the relation between $\det R$ and $\det(-R)$ if R is 3×3 (or odd $\times$ odd)?

7.2 Isotropy

An elastic material is *isotropic* if the stress produced by a motion remains the same if the material is rotated before the motion. That is,

$$\mathcal{S}(\nabla_{\mathbf{q}}\mathbf{x}(\mathbf{q}, t)) = \mathcal{S}(\nabla_{\mathbf{q}}\mathbf{x}(\mathbf{q}', t))$$

where $\mathbf{q}' = \mathbf{q}_0 + R^T(\mathbf{q} - \mathbf{q}_0)$ are the coordinates of the point at $\mathbf{q}$ in the original reference configuration after it is rotated about $\mathbf{q}_0$ by the rotation matrix R^T. Note the ordering on the right hand side: points $\mathbf{q}$ in the reference configuration are first rotated ($\mathbf{q} \to \mathbf{q}'$), then moved ($\mathbf{q}' \to \mathbf{x}$) with the *same* motion function $\mathbf{x}$. [Using R^T rather than R to denote the arbitrary rotation matrix figuring in this concept is just a notational convenience, as you will see.] According to the chain rule, $\nabla_{\mathbf{q}}\mathbf{x}(\mathbf{q}', t) = \nabla_{\mathbf{q}'}\mathbf{x}(\mathbf{q}', t)\nabla_{\mathbf{q}}\mathbf{q}' = \nabla_{\mathbf{q}'}\mathbf{x}(\mathbf{q}', t)R^T$. Set $L = \nabla\mathbf{x}(\mathbf{q}_0, t)$; since $\mathbf{q}' = \mathbf{q}_0$ when $\mathbf{q} = \mathbf{q}_0$, the equation above reads

$$\mathcal{S}(L) = \mathcal{S}(LR^T)$$

Since L may be any matrix (why?), this equation characterizes the stress function of isotropic elastic materials.

Note that independence of observer also implies that

$$R\mathcal{S}(L)R^T = \mathcal{S}(RL) = \mathcal{S}(RLR^T)$$

Slightly generalizing the definition of isotropic linear transformation given in the last Chapter, then, you could say that *isotropic materials have isotropic stress functions.*

Some materials behave in a nearly isotropic elastic fashion, at least for small deformations - some plastics, some metals, sandstones, unreinforced concrete, amorphous salt. Others do not, exhibiting considerable change in stress after rotation of the reference configuration - muscle, composites (fiberglass, carbon fiber reinforced epoxy, reinforced concrete), shales.

7.3 Linear Elasticity

As was the case for elastic fluids, the theory of elasticity simplifies considerably when motions are small enough to neglect higher order effects. For variety, this section carries out the linearization process in a slightly more formal way than did Chapter 5.

The easiest way to express the small-motion hypothesis is in terms of a *1-parameter family* of motions $\mathbf{x}_h(\mathbf{q},t) = \mathbf{q} + h\mathbf{u}(\mathbf{q},t)$. Here h is a small parameter, and $\mathbf{u}$ is the *infinitesimal deformation*. We will seek the dominant parts of all quantities by expanding them in powers of h and dropping terms of order higher than linear in h. These expansions are legitimate when the quantities being expanded have enough derivatives, i.e. are well approximated by their Taylor series. These notes make the standing assumption (throughout this chapter only!) that plenty of derivatives are always available.

To begin with,

$$\mathcal{S}(\nabla \mathbf{x}_h) = \mathcal{S}(I + h\nabla \mathbf{u})$$

$$= \mathcal{S}(I) + h\frac{\partial \mathcal{S}}{\partial L}(I)\nabla \mathbf{u} + ...$$

where the ellipses denote the remainder in this first order Taylor series, ie. "quantities of second order and higher" in h.

$\mathcal{S}(I)$ is the stress at zero displacement, called the *residual stress* in the material.

For simplicity, these notes assume that the residual stress vanishes. Not all elastic materials have vanishing residual stress: as the name suggests, prestressed concrete is an example of material deliberately configured to have nonvanishing $\mathcal{S}(I)$.

In coordinates, the linear (order 1 in h) part of the stress is

$$(\frac{\partial \mathcal{S}}{\partial L}(I)\nabla \mathbf{u})_{ij} = \sum_{k,l} C_{ijkl}\frac{\partial u_k}{\partial x_l}$$

where the four-index array C_{ijkl} is the *Hooke tensor*.

Since the stress must be symmetric (so that angular momentum balance is maintained), clearly $C_{ijkl} = C_{jikl}$, all i, j, k, l.

For brevity, we will write

$$\frac{\partial \mathcal{S}}{\partial L}(I)L = C[L]$$

so C is a linear transformation: $\mathbf{R}^{3\times 3} \to \mathbf{R}^{3\times 3}_{\text{symm}}$, with "matrix" C_{ijkl}.

Actually the linear part of the stress depends only on the symmetric part ϵ of $\nabla \mathbf{u}$. To see this, note that for any matrix L,

$$e^{hL} = I + hL + ...$$

Write $\nabla u = \epsilon + r$, where $\epsilon = \frac{1}{2}(\nabla u + (\nabla u)^T), r = \frac{1}{2}(\nabla u - (\nabla u)^T)$. Then

$$\mathcal{S}(\nabla \mathbf{x}_h) = \mathcal{S}(I + h(\epsilon + r))$$

$$= \mathcal{S}((I + hr) + h\epsilon) = \mathcal{S}((I + hr)(I + h\epsilon) - h^2 r\epsilon)$$

$$= \mathcal{S}(R(I + h\epsilon) + ...)$$

where $R = e^{hr}$ is a rotation matrix because r is skew-symmetric. It is easy to see that the higher order terms (...) can only contribute the higher order terms in the Taylor expansion of $\mathcal{S}$:

$$= \mathcal{S}(R(I + h\epsilon)) + ... = R\mathcal{S}(I + h\epsilon)R^T$$

(because of observer-independence)

$$= (I + hr + ...)(\mathcal{S}(I) + hC[\epsilon] + ...)(I + hr^T + ...) + ...$$

Since residual stress has been assumed away, this simplifies to

$$= hC[\epsilon] + ...$$

which establishes the claim, i.e. that the linear part of the stress depends only on the symmetric part of ∇u.

This fact implies in turn that $C_{ijkl} = C_{ijlk}$ (Problem 1).

This symmetric part of ∇u is called the *infinitesimal strain*: indeed, the Lagrangian strain is (Chapter 2)

$$\mathcal{L} = \frac{1}{2}((\nabla \mathbf{x})^T(\nabla \mathbf{x}) - I)$$

$$= \frac{1}{2}(h(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + h^2(\nabla \mathbf{u})^T(\nabla \mathbf{u}))$$

$$= h\epsilon + ...$$

Define a corresponding *infinitesimal stress* by assuming an expansion of S in h:

$$S = h\sigma + ...$$

Then the linearized relation between stress and motion becomes

$$\sigma = C[\epsilon], \text{ i.e. } \sigma_{ij} = \sum_{kl} C_{ijkl}\epsilon_{kl}$$

which is often called *Hooke's law*: it simply states that, in the small motion limit, stress (force) is proportional to strain (relative displacement).

To understand the dynamical laws (conservation of mass, momentum) in the small motion limit, it is necessary to compute the Taylor series of the velocity field in h:

$$\mathbf{v}(\mathbf{x},t) = \frac{\partial \mathbf{x}}{\partial t}(\mathbf{q}(\mathbf{x},t),t)$$

$$h\frac{\partial \mathbf{u}}{\partial t}(\mathbf{x} - h\mathbf{u}(\mathbf{q}(\mathbf{x},t),t))$$

$$= h\frac{\partial \mathbf{u}}{\partial t}(\mathbf{x},t) + ...$$

Note the simplicity of this result: to first order in h, the velocity is simply the time derivative of displacement!

Introduce the expansion $\rho(\mathbf{x},t) = \rho_0(\mathbf{x},t) + h\rho_1(\mathbf{x},t;h)$ of the density. Then

$$\frac{D\rho}{Dt} + (\nabla\cdot\mathbf{v})\rho = \frac{\partial\rho}{\partial t} + \mathbf{v}\cdot\nabla\rho + (\nabla\cdot\mathbf{v})\rho$$

$$= \frac{\partial\rho_0}{\partial t} + h\left(\frac{\partial\rho_1}{\partial t} + \frac{\partial\mathbf{u}}{\partial t}\cdot\nabla\rho_0 + \left(\nabla\cdot\frac{\partial\mathbf{u}}{\partial t}\right)\rho_0\right) + ...$$

Since this must hold identically in h, the coefficient of each power of h must vanish. The constant term (with respect to h) is

$$\frac{\partial\rho_0}{\partial t} = 0$$

and this is all that conservation of mass amounts to in the linear limit: as you will see shortly, only ρ_0 plays a role in the conservation of momentum. So conservation of mass $\Leftrightarrow$ mass density is time-independent to lowest order in h.

Another way to see this is to recall that the spatial description of the conservation of mass is equivalent to the material description

$$\rho(\mathbf{x},t) = \rho_0(\mathbf{q}(\mathbf{x},t))\det\nabla_{\mathbf{x}}\mathbf{q}(\mathbf{x},t)$$

Expand $\mathbf{x}$ and its Jacobian in powers of h to see that I chose the right notation for ρ_0, i.e. $\rho(\mathbf{x},t) = \rho_0(\mathbf{x}) + ...$

As for conservation of momentum,

$$\rho\frac{D\mathbf{v}}{Dt} = (\rho_0 + ...)\left(h\frac{\partial^2\mathbf{u}}{\partial t^2} + ...\right)$$

$$= h\rho_0 \frac{\partial^2 \mathbf{u}}{\partial t^2} + ... = h\nabla \cdot \sigma + ... + \mathbf{f}$$

In order for this to work, clearly the external force $\mathbf{f}$ must vanish in the small motion limit, ie. $\mathbf{f} = h\mathbf{f}_1 + ...$ Equating powers of h, obtain the *linear elastodynamic equations of motion*

$$\rho_0 \frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla \cdot \sigma + \mathbf{f}_1, \ \sigma = C[\epsilon]$$

It's customary to lose the subscripts on ρ and $\mathbf{f}$, remembering that these quantities actually represent the lowest-order terms in h in the expansion of each.

So far we have established these symmetries of the Hooke tensor:

$$C_{ijkl} = C_{jikl}$$

$$C_{ijkl} = C_{ijlk}$$

These simply express the fact that C is a linear transformation on the space of symmetric 3×3 matrices. Since the space of such matrices is 6 dimensional, the number of independent $C's$ is 36 (the number of entries in a 6×6 matrix). [Idle amusement: convince yourself that this also follows from the two equations above.]

Problems 5-10 show that *conservation of energy*, a concept introduced here for the first time in this course, implies additional symmetries: C is actually a *symmetric* linear transformation on $\mathbf{R}^{3\times3}_{\text{symm}}$.

$$C_{ijkl} = C_{klij}$$

Therefore there are actually 21 independent C_{ijkl} (without any further assumptions about the nature of the elastic material).

7.4 Isotropic Linear Elasticity

Suppose that the elastic material undergoing small motion is isotropic. Expand both sides of the equation expressing isotropy in powers of h:

$$R\mathcal{S}(\nabla \mathbf{x})R^T = hRC[\nabla \mathbf{u}]R^T + ... = hRC[\epsilon]R^T + ...$$

$$= \mathcal{S}(R\nabla \mathbf{x}R^T) = hC[R\nabla \mathbf{u}R^T] + ... = hC[R\epsilon R^T] + ...$$

Equating powers of h, and noting that ϵ can be any symmetric matrix (why?), you see that C is an *isotropic* linear transformation on $\mathbf{R}^{3\times3}_{\text{symm}}$:

$$RC[\epsilon]R^T = C[R\epsilon R^T]$$

for any rotation matrix R. We have already derived and used a structure theorem for isotropic transformations: C must take the form

$$C[\epsilon] = \lambda \mathrm{tr}\, \epsilon I + 2\mu\epsilon$$

for suitable scalars λ and μ, called *Lamé parameters* in this context.

Problems 2 - 4 in this Chapter explain some of the physical significance of λ and μ, from the point of view of *statics* - i.e. motions not dependent on t (often called simply "deformations"). The next sections give a complementary, dynamical point of view.

7.5 Homogeneous isotropic media and alternate forms of the equations of motion

Since

$$\mathrm{tr}\, \epsilon = \nabla \cdot \mathbf{u}$$

and

$$\nabla \cdot (fI) = \nabla f$$

for any scalar function f, you can rewrite

$$\nabla \cdot (\lambda \mathrm{tr}\, \epsilon I) = \nabla(\lambda \nabla \cdot \mathbf{u})$$

If the material is *homogeneous*, i.e. neither ρ, λ, nor μ depend on $\mathbf{x}$, then another simplification becomes possible: since

$$\nabla \cdot \nabla \mathbf{u} = \nabla^2 \mathbf{u}$$

and

$$\nabla \cdot (\nabla \mathbf{u})^T = \nabla(\nabla \cdot \mathbf{u})$$

Thus for *homogeneous isotropic* elastic materials, the equations of motion become

$$\rho \frac{\partial^2 u}{\partial t^2} = (\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u}$$

$$= (\lambda + 2\mu)\nabla(\nabla \cdot \mathbf{u}) - \mu \nabla \times \nabla \times \mathbf{u}$$

where in the last step we have used the ever-useful identity for vector valued functions $\mathbf{w}$:

$$\nabla(\nabla \cdot \mathbf{w}) = \nabla^2 \mathbf{w} + \nabla \times \nabla \times \mathbf{w}$$

(see Problem 4-6b).

7.6 Linear elastic waves in isotropic homogeneous materials

Motivated partly by the elastic fluid example, we hypothesize solutions of the *plane wave* form

$$\mathbf{u}(\mathbf{x}, t) = f(ct - \mathbf{x} \cdot \theta)\mathbf{w}$$

where c is a positive scalar (the *wave speed*), θ is a constant vector of unit length (the *wave vector* = normal to the equal phase surfaces), and $\mathbf{w}$ is a constant vector (the *polarization vector* = direction of particle motion). Note that

$$\nabla \cdot (f\mathbf{w}) = \nabla f \cdot \mathbf{w} = -f'\theta \cdot \mathbf{w}$$

$$\nabla(\nabla \cdot (f\mathbf{w})) = f''\theta(\theta \cdot \mathbf{w})$$

$$\nabla \times (f\mathbf{w}) = -f'\theta \times \mathbf{w}$$

$$\nabla \times \nabla \times (f\mathbf{w}) = f''\theta \times \theta \times \mathbf{w}$$

and

$$\frac{\partial^2 f\mathbf{w}}{\partial t^2} = c^2 f''\mathbf{w}$$

(where in each of these equations I have suppressed the argument $ct - \mathbf{x} \cdot \theta$ of f).

Therefore for functions of this form, the equations of motion are equivalent (after dividing through by ρ) to

$$f''(c^2\mathbf{w}) = f''(\left(\frac{\lambda + 2\mu}{\rho}\right)\theta\theta \cdot \mathbf{w} - \frac{\mu}{\rho}\theta \times \theta \times \mathbf{w})$$

If f'' does not vanish, then $\mathbf{w}$ must be an eigenvector with eigenvalue c^2 of the matrix of the linear operator appearing in the last equation. Note that

$$\mathbf{w} \mapsto \theta(\theta \cdot \mathbf{w})$$

has the eigenvalue 1 with multiplicity 1 (vectors parallel to θ) and 0 with multiplicity 2 (vectors perpindicular to θ), whereas

$$\mathbf{w} \mapsto \theta \times \theta \times \mathbf{w})$$

has eigenvalue -1 with multiplicity 2 (vectors perpindicular to θ) and 0 with multiplicity 1 (vectors parallel to θ). Since these matrices have the same eigenvectors, the eigenvalues of the linear combination of the matrices are simply the same linear combination of the eigenvalues. So we get two eigenvalues:

- $\frac{\lambda+2\mu}{\rho}$, multiplicity 1, corresponding to wave speed $c = \alpha \equiv \sqrt{\frac{\lambda+2\mu}{\rho}}$ and wave vector θ. These waves move in the same direction as the particle motion (i.e. the wave vector and the polarization vector are parallel), and are called *compressional* or *longitudinal* waves.

- $\frac{\mu}{\rho}$, multiplicity 2, corresponding to wave speed $c = \beta \equiv \sqrt{\frac{\mu}{\rho}}$ and wave vectors perpindicular to θ. These waves move in a direction perpindicular to particle motion (i.e the wave vector and the polarization vector are perpindicular) and are called *shear* or *transverse* waves.

These solutions might seem to be quite special. However arbitrary solutions may be approximated arbitrarily well by linear combinations of these (Ch. 4), so they are in some sense not at all special!

7.7 A few words about anisotropic and/or heterogeneous materials

Anisotropic (i.e. not isotropic) materials abound in nature, and in general require all 21 possible independent Hooke's tensor coefficients to characterize them. Various partial symmetries are possible: for example, shales (rocks composed of tiny platelets of mica, glued together with clay and other materials) have approximate rotational symmetry about the axis perpindicular to the bedding plane in which the platelets lie ("transverse isotropy"). Crystalline salt has the symmtries of a cubic lattice. These symmetries restrict the Hooke tensor in various complicated and interesting ways [refs].

Except where material homogeneity was explicitly assumed, nothing in the derivations in the section depends on it. So the density ρ (really ρ_0 of course) and the Hooke tensor C_{ijkl} may depend on position $\mathbf{x}$.

The significance of the Lamé parameters, being microscopic, is independent of homogeneity also. However the construction of wave solutions in the last section depended explicitly on homogeneity, so the existence of elastic waves is not at this point established for heterogeneous materials.

7.8 Problems:

Problem 7-1 . In the notes it was shown that, for an elastic material, the infinitesimal stress σ depends only on the infinitesimal strain ϵ. In terms of the Hooke's tensor C, this means that for any displacement $\mathbf{u}(\mathbf{x}, t)$,

$$\frac{\partial \mathcal{S}}{\partial L}(\nabla \mathbf{u}) \equiv C[\nabla \mathbf{u}] = C\left[\frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^T)\right]$$

Show that (1) for any matrix L, there is a motion $\mathbf{x}(\mathbf{q}, t)$ and a point $\mathbf{q}_0$ in the reference configuration so that $\nabla_{\mathbf{q}}\mathbf{x}(\mathbf{q}_0, 0) = I + L$, (2) consequently for any matrix L there is a displacement $\mathbf{u}(\mathbf{x}, t)$, a time t_0, and a point $\mathbf{x}$ in the material volume at time t_0, for which $L = \nabla\mathbf{u}(\mathbf{x}_0, t_0)$, and finally (3) therefore for any matrix L,

$$C[L] = C\left[\frac{1}{2}(L + L^T)\right]$$

whence

$$C_{ijkl} = C_{ijlk},\ i, j, k, l = 1, ..., 3$$

NOTE: In the next three problems, time plays no role, so I drop it from the notation.

Problem 7-2 (shear modulus) . Define an displacement $\mathbf{u}$ by

$$\mathbf{u}(\mathbf{x}) = (0, \gamma x_1, 0)^T,\ \gamma \in \mathbf{R}$$

Draw a picture of this deformation (see Problem 2-2 for a related picture!). **Compute** ϵ. Assume that the material is an *isotropic elastic solid*, and **show** that in that case $\sigma = 2\mu\epsilon$.

This deformation is *infinitesimal shear*. The Lamé constant μ gives the relation between stress and strain for such deformations. Put another way: μ measures the force generated by, or equivalently the resistance to, a shear motion. So μ is called the *shear modulus*.

Problem 7-3 (bulk modulus) . Define the *extension* displacement $\mathbf{u}$ of an isotropic elastic solid by

$$\mathbf{u}(\mathbf{x}) = (\gamma x_1, \gamma x_2, \gamma x_3)^T,\ \gamma \in \mathbf{R}$$

Show that $\epsilon = \gamma I$, $\sigma = 3\kappa\epsilon$ where

$$\kappa = \lambda + \frac{2}{3}\mu$$

Since κ measures the resistance of the material to extension, which changes the volume without any shear deformation, it is called the *bulk modulus* of the material.

Problem 7-4 (Poisson's ratio) . **Show** that that a *pure tension* infinitesimal stress in the x_1 direction of an isotropic elastic solid, defined by

$$\sigma = \mathrm{diag}(1, 0, 0)$$

is the result of an infinitesimal deformation whose strain takes the form

$$\epsilon = \mathrm{diag}\left(\frac{1}{E}, -\frac{\nu}{E}, -\frac{\nu}{E}\right)$$

where

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}$$

is *Young's modulus*, and

$$\nu = \frac{\lambda}{2(\lambda + \mu)}$$

is *Poisson's ratio*. That is, Young's modulus measures the increase in length resulting from a unit force along an axis, and Possion's ratio controls the decrease in diameter of the cross section perpindicular to the force axis.

Young's modulus is always positive. Almost all materials have positive Poisson's ratio, i.e. they get skinnier if you pull on them.

The next few problems lead to the relation between conservation of energy and symmetries of the Hooke tensor C_{ijkl} for linearly elastic materials. Time is back in play.

In linear elasticity, the *potential energy* of the material in $\Omega \subset \mathbf{R}^3$ is

$$PE = \frac{1}{2}\int_\Omega \operatorname{tr} \sigma\epsilon = \frac{1}{2}\int_\Omega \sum_{i,j} \sigma_{i,j}\epsilon_{i,j}$$

(note that you can write the subscript i, j instead of j, i in the last factor because ϵ is symmetric). Note that this quantity expresses the work done by the displacement (strain) against the internal forces (stress), and has the correct units for energy.

The *kinetic energy* is (as usual)

$$KE = \frac{1}{2}\int_\Omega \rho \left|\frac{\partial \mathbf{u}}{\partial t}\right|^2$$

Conservation of (total mechanical) energy in this model is expressed as

$$\frac{d}{dt}(KE + PE) = 0$$

which should hold for a domain Ω in which there is no motion near the boundary, so that energy is not being transported out of the domain. That is, assume in the following that $\mathbf{u} \equiv 0$ near $\partial\Omega$.

Problem 7-5. Use calculus and Hooke's law to **show** that

$$\frac{d}{dt}KE = \int_\Omega \frac{\partial \mathbf{u}}{\partial t}\rho\frac{\partial^2 \mathbf{u}}{\partial t^2}$$

$$\frac{d}{dt}PE = \frac{1}{2}\int_\Omega \mathrm{tr}\left(\frac{\partial \sigma}{\partial t}\epsilon + \sigma\frac{\partial \epsilon}{\partial t}\right)$$

$$= \frac{1}{2}\int_\Omega \sum_{ijkl}(C_{ijkl} + C_{klij})\frac{\partial \epsilon_{ij}}{\partial t}\epsilon_{kl}$$

$$= \frac{1}{2}\int_\Omega \sum_{ijkl}(C_{klij} - C_{ijkl})\frac{\partial \epsilon_{ij}}{\partial t}\epsilon_{kl} + \int_\Omega \mathrm{tr}\sigma\frac{\partial \epsilon}{\partial t}$$

Problem 7-6. Use the symmetry of σ and the definition of ϵ to show that

$$\mathrm{tr}\sigma\frac{\partial \epsilon}{\partial t} = \mathrm{tr}\sigma\left(\nabla\frac{\partial \mathbf{u}}{\partial t}\right)$$

Problem 7-7. Use the divergence theorem in some form to show that for any matrix valued function A, and any vector valued function $\mathbf{w}$ vanishing near the boundary $\partial\Omega$,

$$\int_\Omega \mathrm{tr}\left[(\nabla \mathbf{w})A\right] = -\int_\Omega \mathbf{w}\cdot(\nabla\cdot A)$$

Problem 7-8. Combine the results of the last several problems to conclude that

$$\frac{d}{dt}PE = \frac{1}{2}\int_\Omega \sum_{ijkl}(C_{klij} - C_{ijkl})\frac{\partial \epsilon_{ij}}{\partial t}\epsilon_{kl} - \int_\Omega \frac{\partial \mathbf{u}}{\partial t}\cdot(\nabla\cdot\sigma)$$

Problem 7-9. Apply momentum balance to conclude that

$$\frac{d}{dt}(KE + PE) = \frac{1}{2}\int_\Omega \sum_{ijkl}(C_{klij} - C_{ijkl})\frac{\partial \epsilon_{ij}}{\partial t}\epsilon_{kl}$$

Therefore energy is conserved **if** the Hooke tensor C_{ijkl} has the additional symmetry

$$C_{ijkl} = C_{klij}$$

Problem 7-10. **Show** that

$$\mathbf{u}(x,t) = (A + Bt)x$$

is a motion consistent with momentum balance, ie. satisfying the equations of motion, for **any** symmetric matrices A and B. Show that conservation of energy then implies that

$$\sum_{ijkl}(C_{klij} - C_{ijkl})A_{kl}B_{ij} = 0$$

and that since A and B are arbitrary symmetric matrices, this implies that

$$C_{ijkl} = C_{klij}.$$

Thus this symmetry is **equivalent** to conservation of energy.

Chapter 8

Thermodynamics and the Heat Equation

The goal of this chapter is to show how the heat equation arises from the first and second laws of thermodynamics and various contitutive assumptions about heat and deformation.

8.1 The First Law

The kinetic energy of a moving part ω_t is

$$KE = \frac{1}{2}\int_{\omega_t} \rho \|\mathbf{v}\|^2$$

as was discussed in Chapter 7. The potential energy however may comprise both strain energy as introduced there, even in a nonlinearized variant, and energy associated with the vibrational states of molecules. This latter motion is not represented by strain nor by any other component of continuum theory, so we must introduce a density e which is supposed to model it, together with any macroscopic (strain) contributions to energy. By convention, e has units of energy per unit mass, whence the total energy of ω_t is

$$\int \rho \left(\frac{1}{2}\|\mathbf{v}\|^2 + e\right)$$

The power supplied to ω_t is the rate of work done on it, and is supposed to consist of the work done on the material in ω_t by the material outside of it, and any work done by an otherwise unspecified body force represented by a force volume density $\mathbf{f}$. A small piece of material centered at $\mathbf{x}$ moves approximately through the displacement $\Delta t \mathbf{v}(\mathbf{x}, t)$ between times t and $t + \Delta t$, and experiences the force $\mathbf{f}(\mathbf{x}, t)\Delta V$, ΔV being the volume of the bit

of material. Thus the work done on the material due to the motion and extenal force is approximately $\mathbf{f}(\mathbf{x},t)\cdot\mathbf{v}(\mathbf{x},t)\Delta V\Delta t$. Adding up all such contributions and dividing by Δt, obtain the rate of work done on ω_t by $\mathbf{f}$:

$$\int_{\omega_t}\mathbf{f}\cdot\mathbf{v}$$

Similarly, the force on a small area ΔA at $\mathbf{x}$ on the surface of ω_t is $\Delta S(\mathbf{x})\mathbf{n}$ so by a similar argument the rate of work done on this small area is $(S\mathbf{n})\cdot\mathbf{v}$, and on the entire moving part the surface integral of this quantity..

The remaining potential energy tied up in the material is supposed to take the form of heat, for which there is both a volume source and a surface source. The surface source (i.e. heat transferred through the boundary of the moving part, due to contact with the remaining material) is supposed to be represented by an area density which is a function only of the surface normal. We supposed this for the traction (Chapter 3), and by a similar argument to that used by Cauchy conclude that the surface source of heat must take the form $-\mathbf{h}\cdot\mathbf{n}$, in which $\mathbf{h}$ is the *heat flux vectorfield*. The sign convention means that $\mathbf{h}$ represents the direction of heat flow, whence its contribution to the energy in ω_t is negative if the flux vector $\mathbf{h}$ points in the same direction as the outwoard normal. The volume source is supposed to represent processes tied to the material, for instance chemical reactions, is therefore a density r per unit mass.

The upshot of these considerations is the integral form of the principle of conservation of energy, also called the *first law of thermodynamics*: the rate of change of energy in the moving part is the sum of the four rates described. In equations,

$$\frac{d}{dt}\int_{\omega_t}\rho\left(\frac{1}{2}\|v\|^2+e\right)=\int_{\omega_t}(\mathbf{f}\cdot\mathbf{v}+\rho r)+\int_{\partial\omega_t}(S\mathbf{n}-\mathbf{h})\cdot\mathbf{v}$$

A differential form follows from Reynold's Transport Theorem, esp. problem 1-6, and the Divergence Theorem in a manner that must by now seem familiar. The only novelty is an identity for the divergence of $(S\mathbf{n})\cdot\mathbf{v}=(S\mathbf{v})\cdot\mathbf{n}$ (why?), which is

$$\nabla\cdot(S\mathbf{v})=(\nabla\cdot S)\cdot\mathbf{v}+S\cdot\nabla\mathbf{v}$$

Here the second "$\cdot$" is the *Frobenius inner product* of matrices

$$A\cdot B=\mathrm{tr}A^TB=\sum A_{ij}B_{ij}$$

We obtain

$$\rho\left(\mathbf{v}\cdot\frac{D\mathbf{v}}{Dt}+\frac{De}{Dt}\right)=(\mathbf{f}+\nabla\cdot S)\cdot\mathbf{v}+S\cdot\nabla\mathbf{v}+\rho r-\nabla\cdot\mathbf{h}$$

This simplifies still further, in view of the conservation of linear momentum, and we get the final differential form of the first law:

$$\rho\frac{De}{Dt}=S\cdot\nabla\mathbf{v}+\rho r-\nabla\cdot\mathbf{h}$$

8.2 The Second Law

Thermodynamics posits the existence of a second fundamental material field, the *entropy* of the material, denoted η. Very crudely speaking, entropy is meant to measure the relative disorder of the microscopic motion leading to the heat contribution to energy. It is supposed to have volume and surface sources in any moving part ω_t, and furthermore these are proportional to the similar quantities for heat by the reciprocal of the *temperature* θ. The difference between the first and second laws is that the rate of entropy is not supposed to be *equal* to these source contributions, but *at least* equal to it:

$$\frac{d}{dt}\int_{\omega_t}\rho\eta \geq \int_{\omega_t}\frac{\rho r}{\theta} - \int_{\partial\omega_t}\frac{\mathbf{h}\cdot\mathbf{n}}{\theta}$$

For this to make sense, the temperature must be presumed positive.

This inequality is the integral statement of the *second law of thermodynamics*, also known as the *Clausius-Duhem inequality*.

After familiar manipulations, the second law takes the differential form:

$$\rho\theta\frac{D\eta}{Dt} - \rho r + \nabla\cdot\mathbf{h} - \mathbf{h}\cdot\frac{\nabla\theta}{\theta} \geq 0$$

Two of the terms on the right hand side also appear in the first law (conservation of energy), i.e.

$$-\rho r + \nabla\cdot\mathbf{h} = S\cdot\nabla\mathbf{v} - \rho\frac{De}{Dt}$$

Combining these last two relations yields

$$\rho\left(\theta\frac{D\eta}{Dt} - \frac{De}{Dt}\right) + S\cdot\nabla\mathbf{v} - \mathbf{h}\cdot\frac{\nabla\theta}{\theta} \geq 0$$

The term in parentheses on the left is (almost) the material derivative of the *Helmholtz free energy* $\psi \equiv e - \theta\eta$:

$$\frac{D\psi}{Dt} = \frac{De}{Dt} - \theta\frac{D\eta}{Dt} - \eta\frac{D\theta}{Dt}$$

which leads to another local form of the second law:

$$-\rho\frac{D\psi}{Dt} - \rho\eta\frac{D\theta}{Dt} + S\cdot\mathbf{v} - \mathbf{h}\cdot\frac{\nabla\theta}{\theta} \geq 0$$

8.3 Stationary, undeformed heat conductor

The relations introduced so far between the various material fields of the theory - v, S, e - are insufficient to predict their evolution, i.e. their values at $t > t_0$ given their values at $t = t_0$.

A count of variables vs. equations suggests that additional relations are required, as was the case in all of the preceding chapters. Specification of these additional *consitutive laws* is a fairly complex task in general - for example, they must satisfy the principle of *independence of observer*, which strongly constrains their general form. The matter is a bit simpler in case the material undergoes no motion at all, and only the transfer of heat is to be perdicted.

"No motion" means: $\mathbf{v} \equiv 0$, material and partial (time) derivatives coincide for all fields, material and spatial descriptions coincide, and all macroscopic mechanical contributions to energy balance vanish. Conservation of energy takes the form

$$\rho \frac{\partial e}{\partial t} = \rho r - \nabla \cdot \mathbf{h}$$

and the second law may be written

$$-\rho \frac{\partial \psi}{\partial t} - \rho \eta \frac{\partial \theta}{\partial t} - \mathbf{h} \cdot \frac{\nabla \theta}{\theta} \geq 0$$

We make a major supposition: that any time "snapshot" of the fields θ, ψ, and η, for any fixed t may be reached at steady state, i.e. when these fields are in fact unchanging. This presumes that sources and sinks $(r, \mathbf{h})$ can actually be prescribed somehow to keep these fields constant, and equal to any particular values they might attain in any other way.

This is a big supposition, and it is not *a priori* obvious why it should be acceptable. It can be justified *a posteriori* for the system that will be introduced below.

"Steady state" means

$$\frac{\partial \theta}{\partial t} = \frac{\partial \psi}{\partial t} = \frac{\partial \eta}{\partial t} = 0$$

This being so, the second law boils down to

$$-\mathbf{h} \cdot \nabla \theta \geq 0$$

(recalling that θ is positive).

The important point is that *this inequality must be true regardless of whether the system is in steady state or not*, because we have assumed that any temperature field (along with the other fields) can occur in steady state if it occurs at all.

Together with the first law, this inequality implies that *heat flows from higher temperature to lower temperature regions!*.

Indeed, suppose that $\omega_t = \{\mathbf{x} : \theta(\mathbf{x}, t) \geq \theta_0\}$, and that θ is not stationary near the boundary of ω_t. Then at every point on the boundary, $\mathbf{n} = -\nabla\theta / \|\nabla\theta\|$ (note that the minus sign makes the normal outward, as is required by the divergence theorem!).Thus the

second law implies that $\mathbf{h} \cdot \mathbf{n} \geq 0$, whence *the term on the right hand side of the first law, expressing the transfer of heat through the boundary of ω_t to the remainder of the material, is non-positive.* Thus heat flows from hot places to cold places, consistent with everyday experience.

8.4 Fourier's Law and the Heat Equation

The preceding argument makes it reasonable to suppose that the heat flux $\mathbf{h}$ is a *function* of the temperature gradient, and possibly of the temperature and the position in the material: The simplest form which this dependence could take, consistent with independence of observer, was posited by Fourier:

$$\mathbf{h} = -\kappa(\theta, \mathbf{x})\nabla\theta$$

Independence of observer requires that $\kappa = kI$, in which k is the *heat conductivity* and $k(\theta, \mathbf{x}) > 0$ to ensure that $-\mathbf{h} \cdot \nabla\theta \geq 0$ as required by the second law.

Additional laws are required to determine ψ and η: the theory assumes that these are functions of θ (and possibly of position). One could permit these quantities to depend on $\nabla\theta$ as well, as does $\mathbf{h}$, but independence of observer forces that dependence to be trivial. In view of the relation between ψ, η, and e, it follows that $e = e(\theta, \mathbf{x})$, whence

$$\frac{\partial e}{\partial t} = c\frac{\partial \theta}{\partial t}$$

where $c \equiv \partial e/\partial\theta$ is the *specific heat* of the material, and is assumed positive (rate of increase of heat energy with temperature). The specific heat has units of energy per unit mass per unit temperature.

With these additional relations assumed, the first law reads

$$\rho c(\theta)\frac{\partial \theta}{\partial t} = \nabla \cdot (k(\theta)\nabla\theta) + \rho r$$

For θ varying by a small amount around an equilibrium value θ_0, first order perturbation gives a simplified, linear PDE for θ. Assuming that (i) the reference temperature is in steady state ($\partial\theta_0/\partial t \equiv 0$) and (ii) the heat conductivity $k(\theta)$ depends only weakly on θ ($\partial k/\partial\theta$ negligible) you obtain

$$\rho c_0\frac{\partial \theta}{\partial t} = \nabla \cdot (k_0\nabla\theta) + \rho r$$

for the fluctuation θ about θ_0 with $c_0 = c(\theta_0), k_0 = k(\theta_0)$. Note that these quantities may still depend on $\mathbf{x}$, which dependence accomodates variation in material properties from one part of the heat-conducting body to another.

This last equation is popularly known as *the heat equation*. Note that it takes exactly the same form as the dynamical law for density or pressure in single-phase porous flow. For a *homogeneous* body (ρ, c_0, k_0 independent of $\mathbf{x}$, it takes the even simpler form

$$\rho c_0 \frac{\partial \theta}{\partial t} = k_0 \nabla^2 \theta + \rho r$$

8.5 Problems:

Problem 9-1 . [30] Assume that $c_0 = k_0 = 1.0$ and $r \equiv 0$ (homogeneous heat equation with normalized coefficients), and that space is d-dimensional ($d = 1$, 2, or 3 are the interesting cases). Show that

$$G(\mathbf{x}, t) = \left(\frac{1}{4\pi t}\right)^{\frac{d}{2}} e^{-\frac{\|\mathbf{x}\|^2}{4t}}$$

solves

$$\frac{\partial G}{\partial t} = \nabla^2 G$$

for $t > 0$.

Problem 9-2 . [40] Show that $G(\cdot, t) \to \delta(\mathbf{x})$ as $t \to 0^+$, i.e. for any continuous function $\phi : \mathbf{R}^d \to \mathbf{R}$ vanishing outside of a large ball,

$$\lim_{t \to 0, t > 0} \int_{\mathbf{R}^d} dx' G(\mathbf{x}', t) \phi(\mathbf{x}') = \phi(\mathbf{0})$$

More generally, show that

$$\lim_{t \to 0, t > 0} \int_{\mathbf{R}^d} dx' G(\mathbf{x} - \mathbf{x}', t) \phi(\mathbf{x}') = \phi(\mathbf{x})$$

[If you show the first identity, the second follows quickly by a simple change of variable!]

You may use the following famous identity without proof, if you can see a use for it:

$$\int_{-\infty}^{\infty} dx\, e^{-x^2} = \sqrt{\pi}$$

Problem 9-3 . [30] Show how the *Gaussian kernel* G should be modified so that the properties established in problems 9-1 and 9-2 continue to hold for general positive *and constant (independent of* $\mathbf{x}, t$*)* ρ, c_0, k_0.

Problem 9-4 . [Extra Credit - 30] Show that the kernel constructed in problem 9-3 may be used to solve the inhomogeneous heat equation as stated above: that is, that

$$\theta(\mathbf{x},t) = \int_0^t dt' \int_{\mathbf{R}^n} dx' G(\mathbf{x}-\mathbf{x}', t-t')\rho r(\mathbf{x}', t') + \int_{\mathbf{R}^n} dx' G(\mathbf{x}-\mathbf{x}', t)\theta_0(\mathbf{x})$$

solves the initial value problem for the heat equation

$$\rho c_0 \frac{\partial \theta}{\partial t} = k_0 \nabla^2 \theta + \rho r,\ t > 0;$$

$$\theta(\mathbf{x},t) \to \theta_0(\mathbf{x}),\ t \to 0, t > 0$$

Chapter 9

Porous Flow

[Please see pp. 10-19 in Ewing's article "Problems arising in the modeling of processes for hydrocarbon recovery" which is Chapter 1 of *The Mathematics of Reservoir Simulation*, ed. R. Ewing, SIAM, 1983. I covered the material on single-phase flow in class; Ewing's discussion of multphase flow and chemical/compositional modeling is very readable.]

Chapter 10

Electrodynamics and Maxwell's Equations

Maxwell's equations connect the following six fields:

- electric, $\mathbf{E}(\mathbf{x}, t)$
- magnetic, $\mathbf{H}(\mathbf{x}, t)$
- electric, displacement $\mathbf{D}(\mathbf{x}, t)$
- magnetic induction, $\mathbf{B}(\mathbf{x}, t)$
- current density, $\mathbf{j}(\mathbf{x}, t)$
- charge density, $\rho(\mathbf{x}, t)$

and involve a constant c with units of velocity $\simeq 3 \times 10^5$ km/s, which will be interpreted below as the speed of light in the vacuum.

$$\nabla \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c}\mathbf{j}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho; \ \nabla \cdot \mathbf{B} = 0$$

In addition several *constitutive laws* must be specified to fully determine the evolution of these fields. As in other topics in continuum mechanics, these express properties of particular

materials. For small intensity fields in isotropic materials moving slowly relative to c, they take the form

$$\mathbf{j} = \sigma \mathbf{E};\ \mathbf{D} = \epsilon \mathbf{E};\ \mathbf{B} = \mu \mathbf{H}$$

in which the scalar fields σ, ϵ, μ are respectively the *conductivity*, the *dielectric permittivity*, and the *magnetic permeability* of the material. These satisfy the constraints $\sigma \geq 0, \epsilon, \mu \geq 1$. Electromagnetically speaking, a *vacuum* is a material in which these quantities take their extreme values: i.e. in a vacuum, $\sigma = 0, \epsilon = \mu = 1$ (in appropriate units).

"See Jackson for details."

10.1 Problems.

Problem 1. [30] Suppose that all of the Maxwell fields are independent of time - as observed in class, this *stationary* assumption decouples the Maxwell equations governing the electric fields $\mathbf{E}, \mathbf{B}$ from those governing the magnetic fields $\mathbf{H}, \mathbf{B}$. The two classes of fields remain linked however through the constitutive laws. Show that in this static situation, the *electrostatic potential* ϕ defined by

$$\mathbf{E} = -\nabla \phi$$

satisfies

$$\nabla \cdot (\sigma \nabla \phi) = 0$$

and show that this relation can be interpreted as stating *conservation of current* in electrostatics: around any closed surface $\partial\Omega$ bounding a domain Ω, the net current circulating on the surface, defined as

$$\int_{\partial\Omega} \mathbf{j} \cdot \mathbf{n},$$

vanishes.

Problem 2. (i) Show that in a vacuum containing no charges and currents ($\mathbf{j} = 0, \rho = 0$), the electric field satisfies the *homogeneous wave equation*

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \nabla^2 \mathbf{E} = 0.$$

Since we established in Ch. 5 that such equations have plane wave solutions moving with velocity c, the interpretation of c as the velocity of light (i.e. waves in the the electric field $\mathbf{E}$) in the vacuum is at least partly justified.

(ii) Show how this equation should be modified in a homogeneous dielectric material, i.e. $\sigma = 0$. Note that the constraints on ϵ, μ mentioned above imply that waves in such a material have velocity $< c$ if the material is not a vacuum.

Problem 3. [30] In class I defined a general *scalar conservation law* as a partial differential equation of the form

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{F} = 0$$

and noted that we have encountered these throughout the semester, beginning with the law of conservation of mass. The divergence theorem implies that the rate of change of total u (whatever that may be) in a volume is the integral of $\mathbf{F} \cdot \mathbf{n}$ over the boundary, whence $\mathbf{F}$ is termed a *flux*.

An interesting and much-studied special case is the *nonlinear conservation law*, in which $\mathbf{F} : \mathbf{R} \to \mathbf{R}^3$ is a function of u. [Of course it could be a linear function, but that's not so interesting - and in fact is termed a "degenerate" case!].

(i) demonstrate the vector calculus identity

$$\nabla \cdot (\mathbf{F}(\mathbf{u}(\mathbf{x}, t))) = \frac{\partial \mathbf{F}}{\partial u}(u(\mathbf{x}, t)) \cdot \nabla u(\mathbf{x}, t)$$

(ii) show that if the curve $\mathbf{x} : [0, T] \to \mathbf{R}^3$ is the solution of the ordinary differential equation

$$\frac{d\mathbf{x}}{dt}(t) = \frac{\partial \mathbf{F}}{\partial u}(u(\mathbf{x}(t), t))$$

then u is constant along this curve, i.e. $u(\mathbf{x}(t), t) = u(\mathbf{x}(0), 0)$.

(iii) deduce that any such curve (as defined in (ii)) is a straight line, and can be computed from the initial value $u(\mathbf{x}(0), 0)$.

You can interpret this as a prescription for (more or less) explicitly solving the initial value problem for a nonlinear conservation law: given $\mathbf{F}$ and the initial data $u(\mathbf{x}_0, 0)$, construct the straight line in spacetime

$$t \to \left(\frac{\partial \mathbf{F}}{\partial u}(u(\mathbf{x}_0, 0)), t \right)$$

then $u(\mathbf{x}, t) = u(\mathbf{x}_0, 0)$ at any point along this line. Such lines fill out spacetime, since you can start them anywhere on the $t = 0$ surface. Therefore you have determined u.

The fly in the ointment is that the lines may cross! Since this construction has to work if the solution u is smooth, the conclusion is that smooth solutions may not exist globally in time, even if the initial data is smooth. This observation leads to the concepts of weak solutions, shocks, and rarefaction waves, which will be discussed in CAAM 437.

Bibliography

[Chung, 2007] Chung, T. J. (2007). *General Continuum Mechanics*. Cambridge University Press, Cambridge.

[Gurtin, 1981] Gurtin, M. E. (1981). *An Introduction to Continuum Mechanics*. Academic Press, New York.

[LANG, 1993] LANG, S. (1993). *Real and Functional Analysis*. Springer Verlag.

[Marsden and Hughes, 1994] Marsden, J. E. and Hughes, T. J. R. (1994). *Mathematical Foundations of Elasticity*. Dover, New York.

[Royden, 1988] Royden, H. (1988). *Real Analysis*. Prentice Hall, Englewood Cliffs, New Jersey, 3rd edition.

[Rudin, 1986] Rudin, W. (1986). *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edition.