CHAPTER

GEOMETRY AND KINEMATICS
OF BODIES

Normally a solid continuum body occupies an open subset of three-dimensional Euclidean space (or the closure of an open subset if the boundary is counted). However, shells and rods may be modeled using surfaces and curves. “Exotic” materials, such as liquid crystals, may require higher dimensional spaces for their description. To have a unified approach, as well as for conceptual clarity, it is useful to think geometrically and to represent bodies in terms of manifolds.

This chapter concerns the description of bodies, their motions, and their configurations. The laws of dynamics which govern the motion of bodies, are discussed in the next chapter. We begin here with bodies in $\mathbb{R}^3$ and gradually work up to a description in terms of manifolds. We assume the reader is familiar with advanced calculus on $\mathbb{R}^n$. For the moment we do not distinguish between “Euclidean space” and “$\mathbb{R}^3$.”

1.1 MOTIONS OF SIMPLE BODIES

1.1 Definition A \textit{simple body} is an open set $\mathcal{B} \subset \mathbb{R}^3$. A \textit{configuration} of $\mathcal{B}$ is a mapping $\phi : \mathcal{B} \rightarrow \mathbb{R}^3$. The set of all configurations of $\mathcal{B}$ is denoted $\mathcal{C}$, or by $\mathcal{C}(\mathcal{B})$ if there is danger of confusion. Points in $\mathcal{B}$ are denoted by capital letters $X, Y, \ldots$.

A configuration represents a deformed state of the body, as in Figure 1.1.1. As the body moves, we obtain a family of configurations, depending on time $t$.

---

1 The use of the word “simple” here is not to be confused with Noll’s use of the term “simple material.” (See Truesdell and Noll [1965], p. 60 and Malvern [1969], p. 391.)
1.2 Definition A motion of a body $\mathcal{B}$ is a curve in $\mathbb{C}$; that is, a mapping $t \mapsto \phi_t \in \mathbb{C}$ of $\mathbb{R}$ to $\mathbb{C}$ (or some open interval of $\mathbb{R}$ to $\mathbb{C}$). For $t \in \mathbb{R}$ fixed, we write $\phi_t(X) = \phi(X, t)$. Likewise, if we wish to hold $X \in \mathcal{B}$ fixed, we write $\phi_X(t) = \phi(X, t)$. The map $V_t : \mathcal{B} \to \mathbb{R}^3$ defined by

$$V_t(X) = V(X, t) = \frac{\partial \phi(X, t)}{\partial t} = \frac{d}{dt} \phi_x(t)$$

(assuming the derivative exists) is called the material velocity of the motion.

If $c(t)$ is a curve in $\mathbb{R}^3$, the tangent to $c(t)$ is defined by $c'(t) = \lim_{h \to 0} (c(t + h) - c(t))/h$. If the standard Euclidean coordinates of $c(t)$ are $(c^1(t), c^2(t), c^3(t))$, then

$$c'(t) = \left( \frac{dc^1}{dt}, \frac{dc^2}{dt}, \frac{dc^3}{dt} \right).$$

To avoid confusion with other coordinate systems we shall write $\phi_t$, and so on, for the Euclidean components of $\phi$. Since $\mathbb{R}^3$ is the set of all real triples, denoted $z = (z^1, z^2, z^3)$, and for fixed $X \in \mathcal{B}$, $\phi(X, t)$ is a curve in $\mathbb{R}^3$, we get

$$V(X, t) = (V^1(x)(X, t), V^2(x)(X, t), V^3(x)(X, t))$$

$$= \left( \frac{\partial \phi^1}{\partial t}(X, t), \frac{\partial \phi^2}{\partial t}(X, t), \frac{\partial \phi^3}{\partial t}(X, t) \right).$$

We regard $V(X, t)$ as a vector based at the point $\phi(X, t)$; if $\phi(X, 0) = X$, then $V(X, t)$ is the velocity at time $t$ of the particle that started out at $X$.\footnote{If we regard the map $t \mapsto \phi_t$ as a curve in $\mathbb{C}$, then the notation $V_t = d\phi_t/dt$ is appropriate. However, we shall use $\partial \phi_t/\partial t$ for both roles for simplicity.}

1.3 Definition The material acceleration of a motion is defined by

$$A_t(X) = A(X, t) = \frac{\partial}{\partial t} V(X, t) = \frac{d}{dt} V_x(t) \quad \text{(if the derivative exists).}$$

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.1.1.png}
\caption{Figure 1.1.1}
\end{figure}
In Euclidean coordinates,

\[ A_i'(X, t) = \frac{\partial V_x'(X, t)}{\partial t} = \frac{\partial^2 \phi_x'(X, t)}{\partial t^2}. \]

**1.4 Definition** A motion \( \phi_t \) of \( \mathcal{B} \) is called *regular* or *invertible* if each \( \phi_t(\mathcal{B}) \) is open and \( \phi_t \) has an inverse \( \phi_t^{-1}: \phi_t(\mathcal{B}) \rightarrow \mathcal{B} \). A \( C^r \) *regular motion* is a \( C^r \) motion \([i.e., \phi(X, t) is a \( C^r \) function of \( (X, t) \)]\) such that \( \phi_t^{-1} \) is also \( C^r \). (The inverse function theorem, recalled later, is relevant here.)

Intuitively, a regular motion is one for which nothing “catastrophic” like ripping, pinching, or interpenetration of matter has occurred.\(^3\) Some of the commonly encountered quantities of continuum mechanics are not well defined if \( \phi \) is not regular, whereas others remain well defined. Since there are physically important cases that are not regular—such as contact problems in which \( \mathcal{B} \) may consist of two disconnected pieces that \( \phi \) brings together—it is important to differentiate between quantities that may or may not be applicable to the formulation of this class of problems.

**1.5 Definition** Let \( \phi_t \) be a \( C^1 \) regular motion of \( \mathcal{B} \). The *spatial velocity* of the motion is defined by\(^4\)

\[ \mathbf{v}_t: \phi_t(\mathcal{B}) \rightarrow \mathbb{R}^3, \quad \mathbf{v}_t = \mathbf{V}_t \circ \phi_t^{-1}. \]

If \( \phi_t \) is a \( C^2 \) regular motion, we define the *spatial acceleration* by

\[ \mathbf{a}_t: \phi_t(\mathcal{B}) \rightarrow \mathbb{R}^3, \quad \mathbf{a}_t = \mathbf{A}_t \circ \phi_t^{-1}. \]

\(^3\)Notice that if a body folds and undergoes self-contact but no interpenetration, then the motion can still be \( C^r \) regular as we have defined it. However, \( \phi_t \) cannot be extended to include the boundary of \( \mathcal{B} \) and still be regular.

\(^4\)By the local existence theory for vector fields, \( \phi_t \) will be regular exactly when \( \mathbf{v}_t \) remains defined and is at least \( C^1 \). See Section 1.6.
We do not like to use phrases like “material coordinates” or “spatial coordinates” because there are no “coordinates” involved in the general setting. Another technique for considering velocity, the “convective picture,” is described in Section 1.2.

Figure 1.1.3 shows the different domains for \( V, A \) and \( v, a \). Observe that \( v, a \) “follow the motion,” that is, they are defined on the time-dependent set \( \phi_t(\mathcal{B}) \).

![Figure 1.1.3](image)

Now we want to work out expressions for \( V, v, A, \) and \( a \) in general coordinate systems.

1.6 Definitions, Notation, and Conventions
A coordinate system \( \{x^a\} \) \((a = 1, 2, 3)\) on \( \mathbb{R}^3 \) is a \( C^\infty \) mapping \((x^1, x^2, x^3)\) of an open set \( \mathcal{U}_z \subset \mathbb{R}^3 \) to \( \mathbb{R}^3 \) such that

(i) the range is an open set \( \mathcal{U}_x \subset \mathbb{R}^3 \), and
(ii) the mapping

\[
(z^1, z^2, z^3) \mapsto (x^1(z^1, z^2, z^3), x^2(z^1, z^2, z^3), x^3(z^1, z^2, z^3))
\]

of \( \mathcal{U}_z \) to \( \mathcal{U}_x \) has a \( C^\infty \) inverse, whose components are denoted \( z^i(x^1, x^2, x^3) \).

---

\( ^5 \)We use, as far as possible, upper case letters to refer to the body \( \mathcal{B} \) and lower case for space \( \mathbb{S} \). Thus \( \{x^a\} \) is a coordinate system in space and \( \{X^a\} \) is one on the body. Another common conventions for indices are those of Truesdell and Noll: \( \{x^i\} \) for space coordinates and \( \{X^a\} \) for body coordinates. Our notation on the indices follows Rivlin [1974a] and Eringen [1975], but otherwise is largely that of Truesdell and Noll.
Coordinate lines are the curves \( c_1(t), c_2(t), c_3(t) \) whose components in Euclidean coordinates are \( z^i(c_1(t)) = z^i(t, x^2, x^3) \), where \( x^2 \) and \( x^3 \) are fixed. Similar definitions hold for \( c_2 \) and \( c_3 \). The tangents to these curves are the coordinate basis vectors; thus

\[
e_a = \frac{\partial z^i}{\partial x^a} \hat{l}_i
\]

where \( \hat{l}_i \) (\( i = 1, 2, 3 \)) are the standard basis vectors in \( \mathbb{R}^3 \). Note that \( e_a \in \mathbb{R}^3 \) and is a function of \( x^1, x^2, x^3 \); that is, \( e_a: \mathbb{U}_x \rightarrow \mathbb{R}^3 \). We always use the summation convention: summation on repeated indices is understood.

For example, spherical coordinates (where \( \mathbb{U}_z \) is \( \mathbb{R}^3 \) minus a plane) define a coordinate system in \( \mathbb{R}^3 \).

Because of condition (ii) above, the Jacobian of the transformation \( z^i \mapsto x^a(z^i) \) is nonsingular, so \( \{e_a\} \) is a basis of \( \mathbb{R}^3 \) for each \((x^1, x^2, x^3)\). Later we shall require an orientation on the vectors \( e_1, e_2, \) and \( e_3 \), but this is not needed now. Figure 1.1.4 depicts the coordinate lines and basis vectors for a general coordinate system in \( \mathbb{R}^3 \).

Figure 1.1.4

Coordinate systems on the ambient space \( \mathbb{R}^3 \) are here denoted \( \{x^a\} \) while those on \( \mathcal{B} \) are denoted \( \{X^A\} \). The corresponding Euclidean coordinates are denoted \( \{z^i\} \) and \( \{Z^I\} \). The basis vectors for the systems \( \{X^A\}, \{z^i\}, \) and \( \{Z^I\} \) are denoted, respectively, \( E_A, \hat{l}_i, \) and \( \hat{l}_I \).

1.7 Proposition Let \( \phi(X, t) \) be a \( C^1 \) motion of \( \mathcal{B} \) and let \( V \) be the material velocity. Let \( \{x^a\} \) be a coordinate system on \( \mathbb{R}^3 \) and let \( \phi^a(X, t) = x^a(\phi(X, t)) \). Then
30 GEOMETRY AND KINEMATICS OF BODIES

\[ V^a(X, t) = \frac{\partial \phi^a}{\partial t}(X, t) \]

where \( V^a(X, t) \) are the components\(^6\) of \( V \) relative to \( e_a \) at the point \( \phi(X, t) \) [with coordinates \( x^b(\phi(X, t)) \)].

**Proof** Consider \(((\partial/\partial t)\phi^a)e_a\) [where \( \phi^a \) stands for \( \phi^a(X, t) = x^a(\phi(X, t)) \)] and where \( e_a \) stands for \( e_a(x^b(\phi(X, t))) \). We have

\[
\left( \frac{\partial}{\partial t} \phi^a \right) e_a = \frac{\partial x^a}{\partial z^l} \frac{\partial \phi^l}{\partial t} e_a = \frac{\partial x^a}{\partial z^l} V^i_z \frac{\partial z^l}{\partial x^a} \delta_{ij} = \delta_{ij} V^i_z \delta_{ij} = V = V^a e_a
\]

since \( \partial x^a/\partial z^l \) and \( \partial z^l/\partial x^a \) are inverse matrices. \( \Box \)

Suppose \( \bar{x}^a \) is another coordinate system on \( \mathbb{R}^3 \) with, say, the same domain \( \mathcal{U}_z \). By composition we can form the change of coordinate functions \( \bar{x}^a(x^b) \) and \( x^a(\bar{x}^b) \); see Figure 1.1.5.

The transformation property of \( V \) is worked out next; one says \( V \) transforms "like a vector." We let \( \bar{V}^a \) denote the components\(^7\) of \( V \) in the basis \( \bar{e}_a \) associated with \( \{\bar{x}^a\} \).

---

\(^6\)If there is danger of confusion, we write \( V^a_z \) for the components of \( V \) in the \( x \)-coordinate system.

\(^7\)More consistently, we should write \( \bar{V}^a_z \) for \( \bar{V}^a \), but it is conventional in many tensor analysis texts to write \( \bar{V}^a \).
1.8 Proposition \( \bar{V}^a = (\partial \bar{x}^a/\partial x^b)V^b \), where \( \bar{V}^a \) stands for \( \bar{V}^a(X, t) \), and so forth.

Proof By the chain rule and 1.7,

\[
\bar{V}^a = \frac{\partial}{\partial t} \bar{x}^a = \frac{\partial \bar{x}^a}{\partial x^b} \frac{\partial}{\partial t} \phi^b = \frac{\partial \bar{x}^a}{\partial x^b} V^b.
\]

An alternative proof is obtained by nothing that \( \bar{e}_a = (\partial x^b/\partial \bar{x}^a) e_b \) and then equating \( V^a e_a \) with \( \bar{V}^a \bar{e}_a \).

To work out the components of the acceleration in a general coordinate system we recall a few notations from calculus on \( \mathbb{R}^n \). If \( \mathcal{U} \subset \mathbb{R}^n \) is an open set in \( \mathbb{R}^n \) and \( f \) maps \( \mathcal{U} \) to \( \mathbb{R}^m \), the derivative of \( f \) at \( x_0 \in \mathcal{U} \) is a linear transformation \( Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m \). Formally, for every \( \varepsilon > 0 \) there must be a \( \delta > 0 \) such that \( h \in \mathbb{R}^n \) and \( \|h\| < \delta \) imply

\[
\|f(x_0 + h) - f(x_0) - Df(x_0) \cdot h\| \leq \varepsilon \|h\|.
\]

Notice that only the norm and the linear structure of Euclidean space are involved. Coordinates are not mentioned. If we represent the linear transformation \( Df(x_0) \) in the standard basis, we get the matrix \( \partial f^i/\partial z^j \), the Jacobian matrix.

The usual calculus rules hold. For instance, the chain rule states that \( D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0) \); the second "o" stands for composition of linear maps—that is, matrix multiplication. The inverse function theorem\(^8\) says that if \( f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n \) is \( C^r \) (\( r \geq 1 \)) and \( Df(x_0) \) is invertible, then \( f \) has a local \( C^r \) inverse \( f^{-1} \), mapping some neighborhood of \( f(x_0) \) to some neighborhood of \( x_0 \) and \( Df^{-1}(f(x_0)) = [Df(x_0)]^{-1} \).

Returning to the context of 1.6, let us work out the derivative of \( e_a = (\partial z^i/\partial x^a) \hat{t}_i \). Clearly,

\[
\frac{\partial e_a}{\partial x^b} = \frac{\partial^2 z^i}{\partial x^a \partial x^b} \hat{t}_i = \frac{\partial^2 z^i}{\partial x^a} \frac{\partial x^c}{\partial x^b} \frac{\partial x^c}{\partial z^i} e_c.
\]

This object arises frequently so is given a name:

1.9 Definition The Christoffel symbols of the coordinate system \( \{x^a\} \) on \( \mathbb{R}^3 \) are defined by

\[
\gamma^c_{ab} = \frac{\partial^2 z^i}{\partial x^a \partial x^b} \frac{\partial x^c}{\partial z^i}
\]

which are regarded as functions of \( x^d \). The Christoffel symbols of a coordinate system \( \{X^A\} \) are denoted \( \Gamma^A_{BC} \).

Note that the \( \gamma \)'s are symmetric in the sense that \( \gamma^c_{ab} = \gamma^c_{ba} \).

---

\( ^8 \)This theorem is proved in a more general infinite-dimensional context in Section 4.1.
1.10 Proposition  Let $\phi(X, t)$ be a $C^2$ motion of $B$, and $V$ and $A$ the material velocity and acceleration. Then the components $A^a$ of $A$ in the basis $e_a$ of a coordinate system $\{x^a\}$ are given by

$$A^a = \frac{\partial V^a}{\partial t} + \gamma^a_{bc} V^b V^c.$$ 

Proof

$$A = \frac{\partial V}{\partial t} = \frac{\partial V^i}{\partial t} \hat{e}_i = \frac{\partial}{\partial t} \left( \frac{\partial z^i}{\partial x^c} V^c \right) \hat{e}_i$$

$$= \frac{\partial^2 z^i}{\partial x^b \partial x^c} \frac{\partial x^b}{\partial t} V^c \hat{e}_i + \frac{\partial z^i}{\partial x^a} \frac{\partial V^a}{\partial t} \hat{e}_i = \frac{\partial^2 z^i}{\partial x^b \partial x^c} V^b V^c \frac{\partial x^a}{\partial z^i} e_a + \frac{\partial V^a}{\partial t} e_a$$

Comparison with $A = A^a e_a$ yields the proposition.

1.11 Proposition  $A^a$ transforms as a vector; that is, $\tilde{A}^b = (\partial x^b/\partial x^a) A^a$.

1.12 Definition  Let $v$ and $w$ be two vector fields on $\mathbb{R}^3$—that is, maps of open sets in $\mathbb{R}^3$ to $\mathbb{R}^3$. Assume $v$ is $C^1$. Thus $Dv(x)$ is a linear map of $\mathbb{R}^3$ to $\mathbb{R}^3$, so $Dv(x) \cdot w(x)$ is a vector field on $\mathbb{R}^3$. It is called the covariant derivative of $v$ along $w$ and is denoted $\nabla_w v(x)$ or $w \cdot \nabla v$.

1.13 Proposition  In a coordinate system $\{x^a\}$,

$$(\nabla_w v)^a = \frac{\partial v^a}{\partial x^b} w^b + \gamma^a_{bc} w^b v^c$$

and $(\nabla_w v)^a$ transforms as a vector.

Proof  Using Euclidean coordinates, and the matrix $\partial v^i/\partial z^j$ of $Dv$ in the standard basis,

$$\nabla_w v = \frac{\partial v^i}{\partial z^j} w^j \hat{e}_i = \left[ \frac{\partial}{\partial z^j} \left( \frac{\partial z^i}{\partial x^c} v^c \right) \right] \left( w^d \frac{\partial z^j}{\partial x^d} \right) \frac{\partial x^a}{\partial z^i} e_a$$

$$= \frac{\partial^2 z^i}{\partial x^b \partial x^c} \frac{\partial x^b}{\partial t} v^c \frac{\partial z^j}{\partial x^d} \frac{\partial x^a}{\partial z^i} e_a + \frac{\partial z^i}{\partial x^a} \frac{\partial v^a}{\partial t} \frac{\partial x^b}{\partial z^j} \frac{\partial x^a}{\partial z^i} e_a$$

$$= \gamma^a_{bc} \delta^i_0 w^b v^c e_a + \delta^i_0 \frac{\partial v^a}{\partial x^b} \delta^b_0 w^d e_a = \gamma^a_{bc} v^c w^b e_a + \frac{\partial v^a}{\partial x^b} w^b e_a.$$ 

One writes $v^i_0 = \partial v^i/\partial x^b + \gamma^i_{bc} v^c$ so that $(\nabla_w v)^a = v^a_0 w^b$.

Recall that for a regular motion, $V(t) = v \circ \phi_t$, that is, $V(X, t) = v(\phi(X, t), t)$. In a coordinate system $\{x^a\}$, $V^a(X, t) = v(\phi(X, t), t)$, and so $\partial V^a/\partial t = \partial v^a/\partial t + (\partial v^a/\partial x^b) V^b$. Since $A^a = \partial V^a/\partial t + \gamma^a_{bc} V^b V^c$, we get

$$A^a(X, t) = \frac{\partial v^a}{\partial t}(\phi(X, t), t) + \frac{\partial v^a}{\partial x^b}(\phi(X, t), t)V^b(X, t) + \gamma^a_{bc}(\phi(X, t), t)V^b(X, t)V^c(X, t).$$
Substituting $A^a(X, t) = a^a(\phi(X, t), t)$ and $x = \phi(X, t)$, we get

$$a^a(x, t) = \frac{\partial v^a}{\partial t}(x, t) + \frac{\partial v^a}{\partial x^b}(x, t) v^b(x, t) + \gamma^a_{bc}(x, t) v^b(x, t) v^c(x, t).$$

The following proposition summarizes the situation:

**1.14 Proposition** For a $C^2$ motion we have

$$a^a = \frac{\partial v^a}{\partial t} + \nabla_x v^a.$$

In a coordinate system $\{x^a\}$, this reads

$$a^a = \frac{\partial v^a}{\partial t} + \frac{\partial v^a}{\partial x^b} v^b + \gamma^a_{bc} v^b v^c.$$

We call $\dot{v} = (\partial v/\partial t) + \nabla_x v$ the *material time derivative* of $v$. Thus $\dot{v} = a$.

The general definition follows.

**1.15 Definition** Let $g_t : \phi_t(\mathbb{R}) \rightarrow \mathbb{R}^3$ be a given $C^1$ mapping. We call

$$\dot{g}(x, t) = \frac{\partial}{\partial t} g(x, t) + Dg_t(x) \cdot v_t(x)$$

the *material time derivative* of $g$. Sometimes $\dot{g}$ is denoted $Dg/ Dt$.

If we define $G_t = g_t \circ \phi_t$, then an application of the chain rule gives

$$\frac{\partial}{\partial t} G(X, t) = \dot{g}(\phi_t(X, t)).$$

This formula justifies the terminology “material time derivative.” Again, $Dg_t(x) \cdot v_t(x)$ is called the *covariant derivative* of $g$ along $v_t$.

---

**Box 1.1 Summary of Important Formulas for Section 1.1**

<table>
<thead>
<tr>
<th>Motion</th>
<th>$\phi_t : \mathbb{B} \rightarrow \mathbb{R}^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_t(X) = \phi(X, t)$</td>
<td>$\phi^* = x^a \circ \phi_t$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Velocity</th>
<th>$V = \frac{\partial \phi}{\partial t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V^a = \frac{\partial \phi^a}{\partial t}$</td>
<td>$V^a = V_t^i \circ \phi_t^{-1}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Covariant Derivative</th>
<th>$Dv \cdot w = \nabla_w v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\nabla_w v)^a = \frac{\partial v^a}{\partial x^b} w^b + \gamma^a_{bc} w^b v^c$</td>
<td></td>
</tr>
</tbody>
</table>
Christoffel Symbols

\[ \gamma^a_{bc} = \frac{\partial^2 x^i}{\partial x^b \partial x^c} \frac{\partial x^a}{\partial x^i} \]

Acceleration

\[ A = \frac{\partial V}{\partial t} \]
\[ a = A_{i} \circ \phi_t^{-1} = \ddot{v} = \frac{\partial v}{\partial t} + \nabla_v v \]

\[ A^a = \frac{\partial V^a}{\partial t} + (\gamma^a_{bc} \circ \phi_t) V^b V^c \]
\[ a^a = \frac{\partial v^a}{\partial t} + \frac{\partial v^a}{\partial x^b} v^b + \gamma^a_{bc} v^b v^c \]

Coordinate Change

\[ \ddot{\phi}_t = \ddot{x} \circ \phi_t \]
\[ \ddot{v} = (\frac{\partial \ddot{x}}{\partial x^b} \circ \phi_t) V^b, \quad \ddot{a} = \frac{\partial \ddot{x}^a}{\partial x^b} v^b \]
\[ \ddot{A}^a = \frac{\partial \ddot{x}^a}{\partial x^b} \circ \phi_t A^b, \quad \ddot{\bar{a}}^a = \frac{\partial \ddot{x}^a}{\partial x^b} \bar{a}^b \]

Figure 1.1.6 goes with these formulas.

---

**Problem 1.1** Let \( \{x^a\} \) and \( \{X^A\} \) denote cylindrical coordinate systems on \( \mathbb{R}^3 \), that is,

\[ x^1 = r, \quad x^2 = \theta, \quad x^3 = z; \]
\[ X^1 = R, \quad X^2 = \Theta, \quad X^3 = Z. \]

Let \( \mathfrak{B} = \mathbb{R}^3 \) and consider a rigid counterclockwise motion about the \( z \)-axis given by

---
\[ \phi^1_t(R, \Theta, Z) = R, \quad \phi^2_\Theta(R, \Theta, Z) = \Theta + 2\pi t, \quad \text{and} \quad \phi^3_t(R, \Theta, Z) = Z \]

where \( \phi^*_t = x^a \circ \phi_t \). Compute \( \dot{V}_r, \dot{v}_r, A_r, \) and \( a_r \).

Let \( \{\tilde{x}^a\} \) denote a Cartesian coordinate system for \( \mathbb{R}^3 \) given by

\[ \tilde{x}^1 = x^1 \cos x^2, \quad \tilde{x}^2 = x^1 \sin x^2, \quad \text{and} \quad \tilde{x}^3 = x^3. \]

Determine \( \tilde{V}_r, \tilde{v}_r, \tilde{A}_r, \) and \( \tilde{a}_r \) from the change-of-coordinate formulas given in the summary, and by directly differentiating \( \dot{\phi}_t \); compare.

**Problem 1.2** For a map \( g : \phi_\alpha(\mathbb{R}) \rightarrow \mathbb{R}^3 \), work out a formula for the covariant derivative \( Dg_\alpha(x) \cdot v_\alpha(x) \) relative to a general spatial coordinate system on \( \mathbb{R}^3 \).

### 1.2 VECTOR FIELDS, ONE-FORMS, AND PULL-BACKS

We shall now start using the terminology of manifolds. We begin by giving the general definition of a manifold and then revert to our special case of open sets in \( \mathbb{R}^3 \) to allow the reader time to become acquainted with manifolds. The basic guidelines and manifold terminology we will use later in the book are given in this section.

**2.1 Definition** A smooth \( n \)-manifold (or a manifold modeled on \( \mathbb{R}^n \)) is a set \( \mathcal{M} \) such that: (1) For each \( P \in \mathcal{M} \) there is a subset \( \mathcal{U} \) of \( \mathcal{M} \) containing \( P \), and a one-to-one mapping, called a chart or coordinate system, \( \{x^a\} \) from \( \mathcal{U} \) onto an open set \( \mathbb{U} \) in \( \mathbb{R}^n \); \( x^a \) will denote the components of this mapping (\( a = 1, 2, \ldots, n \)). (2) If \( x^a \) and \( \tilde{x}^a \) are two such mappings, the change of coordinate functions \( \tilde{x}^a(x^1, \ldots, x^n) \) are \( C^\infty \).

If \( \{\tilde{x}^a\} \) maps a set \( \mathcal{U} \subset \mathcal{M} \) one-to-one onto an open set in \( \mathbb{R}^n \), and if the change of coordinate functions with the given coordinate functions are \( C^\infty \), then \( \{\tilde{x}^a\} \) will also be called a chart or coordinate system.

For instance, an open set \( \mathcal{M} \subset \mathbb{R}^n \) is a manifold. We take a single chart, \( \{Z^i\} \), to be the identity map to define the manifold structure. By allowing all possible coordinate systems that are \( C^\infty \) functions of the \( Z^i \), we enlarge our set of coordinate systems. We could also start with all of these coordinate systems at the outset, according to taste.

Thus, a manifold embodies the idea of allowing general coordinate systems (see Figure 1.1.5); it allows us to consider curved objects like surfaces (two-manifolds) in addition to open sets in Euclidean space. The manifold \( \mathcal{M} \) becomes a topological space by declaring the \( \mathcal{U} \)'s to be open sets. In our work we will

---

9One can start with a topological space \( \mathcal{M} \) or, using the differential structure, make \( \mathcal{M} \) into a topological space later. We chose the latter since it minimizes the number of necessary concepts.
consider both the body $\mathcal{B}$ and the containing space $\mathcal{S}$ to be special cases of manifolds. The descriptions of some realistic bodies require this generality, such as shells and liquid crystals (see Sections 1.5 and 2.2). However, the description of any body benefits from manifold terminology. Examples occur in the study of covariance and relativistic elasticity.

An important mathematical discovery made around the turn of the century was that one could define the tangent space to a manifold without using a containing space, as one might naively expect. Unfortunately, the abstraction necessary to do this causes confusion to those trying to learn the subject. To help guide the reader, we list three possible approaches:

(a) *Derivations.* The idea here is that in order to specify a vector tangent to $\mathcal{M}$, we can give a rule defining the derivatives of all real-valued functions in that direction; such directional derivatives are derivations (see Bishop and Goldberg [1968], pp. 47–48).

(b) *Curves.* We intuitively think of tangent vectors as velocities of curves; therefore, vectors can be specified as equivalence classes of curves, two curves being equivalent if they have the same tangent vector in some, and hence in every, chart (see Abraham and Marsden [1978], p. 43).

(c) *Local transformation properties.* We can use the transformation rules for vectors found in Section 1.1 to define a local vector in a coordinate chart as a vector in $\mathbb{R}^n$ and then use the coordinate transformation rules to define an equivalence relation on pairs of charts and local vectors. Each equivalence class obtained is a tangent vector to $\mathcal{M}$ (see Lang [1972], pp. 26, 47).

For $\mathcal{M}$ open in $\mathbb{R}^n$, the tangent space is easy to define directly:

2.2 Definition Let $\mathcal{M} \subset \mathbb{R}^n$ be an open set and let $P \in \mathcal{M}$. The tangent space to $\mathcal{M}$ at $P$ is simply the vector space $\mathbb{R}^n$ regarded as vectors emanating from $P$; this tangent space is denoted $T_P \mathcal{M}$. The tangent bundle of $\mathcal{M}$ is the product $T \mathcal{M} = \mathcal{M} \times \mathbb{R}^n$ consisting of pairs $(P, w)$ of base points $P$ and tangent vectors at $P$. The map $\pi$ (or $\pi_\mathcal{M}$ if there is danger of confusion) from $T \mathcal{M}$ to $\mathcal{M}$ mapping a tangent vector $(P, w)$ to its base point $P$ is called the projection. We may write $T_P \mathcal{M} = \{P\} \times \mathbb{R}^n$ as a set in order to keep the different tangent spaces distinguished, or denote tangent vectors by $w_P = (P, w)$ to indicate the base point $P$ which is meant.

The idea of the tangent bundle, then, is to think of a tangent vector as being a vector equipped with a base point to which it is attached (see Figure 1.2.1). Once the tangent bundle of a manifold is defined, one makes it into a manifold by introducing the coordinates of vectors as in Section 1.1. For the special case in which $\mathcal{M} = \mathcal{B}$ is an open set in $\mathbb{R}^n$ this is easy. For the rest of this section we confine ourselves to this case and we use notation adapted to it; that is, points in $\mathcal{B}$ are denoted $X$ and coordinate systems are written $\{X^A\}$. However, the reader should be prepared to apply the ideas to general manifolds after finishing the section.
2.3 Definition Let $\mathcal{B} \subset \mathbb{R}^n$ be open and $T\mathcal{B} = \mathcal{B} \times \mathbb{R}^n$ be its tangent bundle. Let $\{X^4\}$ be a coordinate system on $\mathcal{B}$. The corresponding coordinate system induced on $T\mathcal{B}$ is defined by mapping $W_x = (X, W)$ to $(X^4(X), W^4)$, where $X \in \mathcal{B}$ and $W^4 = (\partial X^4/\partial z^i)W_z$ are the components of $W$ in the coordinate system $\{X^4\}$, as explained in Section 1.1.

For $\mathcal{B} \subset \mathbb{R}^3$ open, $T\mathcal{B}$ is a six-dimensional manifold. In general if $\mathcal{B}$ is an $n$-manifold, $T\mathcal{B}$ is a $2n$-manifold.

In Euclidean space we know what is meant by a $C^r$ map. A mapping of manifolds is $C^r$ if it is $C^r$ when expressed in local coordinates.

2.4 Proposition (a) Let $\mathcal{B} \subset \mathbb{R}^n$ be open and let $f: \mathcal{B} \rightarrow \mathbb{R}$ be a $C^1$ function. Let $W_x = (X, W) \in T_x \mathcal{B}$. Let $W_x[f]$ denote the derivative of $f$ at $X$ in the direction $W_x$—that is, $W_x[f] = Df(X) \cdot W$. If $\{X^4\}$ is any coordinate system on $\mathcal{B}$, then $W_x[f] = (\partial f / \partial X^4)W^4$, where it is understood that $\partial f / \partial X^4$ is evaluated at $X$.

(b) If $c(t)$ is a $C^1$ curve in $\mathcal{B}$, $c(0) = X$, and $W_x = (X, W) = (X, c'(0))$ is the tangent to $c(t)$ at $t = 0$, then, in any coordinate system $\{X^4\}$,

$$W^4 = \frac{dc^4}{dt}, \quad \text{where} \quad c^4(t) = X^4(c(t)).$$

Proof

(a) $Df(X) \cdot W = \frac{\partial f}{\partial z^j} W^j = \frac{\partial f}{\partial X^4} \frac{\partial X^4}{\partial z^j} W^j = \frac{\partial f}{\partial X^4} W^4$.

(b) $\frac{dc^4}{dt} = \frac{\partial X^4}{\partial z^j} W^j = W^4$ (evaluated at $c(t)$).

Following standard practice, we let $c'(0)$ stand both for $(X, c'(0))$ and $c'(0) \in \mathbb{R}^n$; there is normally no danger of confusion.

The above proposition gives a correspondence between the “transformation of coordinate” definition and the other methods of defining the tangent space. Observe that the mapping $f \mapsto W_x[f]$ is a derivation; that is, it satisfies

$$W_x[f + g] = W_x[f] + W_x[g] \quad \text{(sum rule)}$$
and

\[ W_x[fg] = f W_x[g] + g W_x[f] \]  
(product rule).

In a coordinate system \( \{X^A\} \) the basis vectors \( E_A = (\partial Z^i / \partial X^A) I_j \) (see Section 1.1) are sometimes written \( \partial / \partial X^A \), since for any function \( f \), \( E_A[f] = \partial f / \partial X^A \), by (a) in 2.4 (i.e., the coordinates of \( E_A \) in the coordinate system \( \{X^P\} \) are \( \delta^A_P \)).

2.5 Definition  Let \( \mathcal{B} \) be open in \( \mathbb{R}^n \) and let \( S = \mathbb{R}^n \). If \( \phi : \mathcal{B} \rightarrow S \) is \( C^1 \), the tangent map of \( \phi \) is defined as follows:

\[ T\phi : T\mathcal{B} \rightarrow TS, \] where \( T\phi(X, W) = (\phi(X), D\phi(X) \cdot W) \).

For \( X \in \mathcal{B} \), we let \( T_X \phi \) denote the restriction of \( T\phi \) to \( T_X \mathcal{B} \), so \( T_X \phi \) becomes the linear map \( D\phi(X) \) when base points are dropped.

We notice that the following diagram commutes (see Figure 1.2.2):

![Figure 1.2.2](image)

The vector, \( T\phi \cdot W_X \) is called the push-forward of \( W_X \) by \( \phi \) and is sometimes denoted \( \phi_* W_X \). (This use of the word “push-forward” is not to be confused with its use for vector fields defined in 2.9 below.) The next proposition is the essence of the fact that \( T\phi \) makes intrinsic sense on manifolds.

2.6 Proposition  (a) If \( c(t) \) is a curve in \( \mathcal{B} \) and \( W_X = c'(0) \), then

\[ T\phi \cdot W_X = \left. \frac{d}{dt} \phi(c(t)) \right|_{t=0} \]  
(the base points \( X \) of \( W_X \) and \( \phi(X) \) of \( T\phi \cdot W_X \) being understood).

(b) If \( \{X^A\} \) is a coordinate chart on \( \mathcal{B} \) and \( \{x^a\} \) is one on \( S \), then, for \( W \in T_X \mathcal{B} \),

\[ (T\phi \cdot W_X)^a = \frac{\partial \phi^a}{\partial X^A} W^A \]

that is, in coordinate charts the matrix of \( T_X \phi \) is the Jacobian matrix of \( \phi \) evaluated at \( X \).\(^{10}\)

\(^{10}\)Sometimes \( \partial \phi^a/\partial X^A \) is written \( \partial x^a/\partial X^A \), but this can cause confusion.
Proof

(a) By the chain rule, \( \frac{d}{dt} \phi(c(t)) \bigg|_{t=0} = D\phi(X) \cdot c'(0) \).

(b) \( D\phi(X) \cdot W = \frac{\partial \phi}{\partial Z^i} W^i \frac{d}{dx} \cdot i \) (representation in the standard basis)

\[ = \frac{\partial \phi}{\partial Z^i} \left( \frac{\partial Z^i}{\partial X^a} W^a \right) \left( \frac{\partial x^a}{\partial X^b} \cdot i \right) \text{ (chain rule)} \]

\[ = \frac{\partial \phi}{\partial X^a} W^a e_a. \]

Later we shall examine the sense in which \( T\phi \) is a tensor. For now we note the following transformation rule:

\[ \frac{\partial \phi}{\partial X^a} = \frac{\partial X^b}{\partial X^a} \frac{\partial \phi}{\partial X^b}. \]

The chain rule can be expressed in terms of tangents as follows:

2.7 Proposition Let \( \phi: \mathcal{O} \rightarrow \mathcal{S} \) and \( \psi: \mathcal{S} \rightarrow \mathcal{U} \) be \( C^r \) maps of manifolds \((r \geq 1)\). Then \( \psi \circ \phi \) is a \( C^r \) map and \( T(\psi \circ \phi) = T\psi \circ T\phi \).

Proof Each side evaluated on \((X, W)\) gives \((\psi(\phi(X)), D\psi(\phi(X)) \cdot (D\phi(X) \cdot W))\) by the chain rule.

This "\( T \)" formulation keeps track of the base points automatically. The reader should draw the commutative diagram that goes with this as an exercise.

Next we formulate vector fields and the spatial and material velocities in manifold language.

2.8 Definitions If \( \mathcal{Q} \) is a manifold (e.g., either \( \mathcal{O} \) or \( \mathcal{S} \)), a vector field on \( \mathcal{Q} \) is a mapping \( \mathcal{V}: \mathcal{Q} \rightarrow T\mathcal{Q} \) such that \( \mathcal{V}(q) \in T_q \mathcal{Q} \) for all \( q \in \mathcal{Q} \).

If \( \mathcal{O} \) and \( \mathcal{S} \) are manifolds and \( \phi: \mathcal{O} \rightarrow \mathcal{S} \) is a mapping, a vector field covering \( \phi \) is a mapping \( \mathcal{V}: \mathcal{O} \rightarrow T\mathcal{S} \) such that \( \mathcal{V}(X) \in T_{\phi(X)} \mathcal{S} \) for all \( X \in \mathcal{O} \).

These diagrams commute (where \( i: \mathcal{Q} \rightarrow \mathcal{Q} \) is the identity map):

and we have the corresponding pictures shown in Figure 1.2.3.

A vector field covering the identity mapping is just a vector field. Also, if \( \mathcal{V} \) is a vector field covering an invertible map \( \phi \), then \( \mathcal{V} = \mathcal{V} \circ \phi^{-1} \) is a vector field on \( \mathcal{Q} = \phi(\mathcal{O}) \).
2.9 Definitions If $Y$ is a vector field on $\mathcal{B}$ and $\phi: \mathcal{B} \to \mathcal{S}$ is a $C^1$ mapping, then $V = T\phi \circ Y$, a vector field covering $\phi$, is called the **tilt** of $Y$ by $\phi$.

If $\phi$ is regular, then $\phi_* Y = T\phi \circ Y \circ \phi^{-1}$, a vector field on $\phi(\mathcal{B})$ is called the **push-forward** of $Y$ by $\phi$.

If $v$ is a vector field on $\phi(\mathcal{B})$ and $\phi$ is regular (i.e., $\phi(\mathcal{B})$ is open in $\mathcal{S}$ and $\phi: \mathcal{B} \to \phi(\mathcal{B})$ has a $C^1$ inverse), $\phi^* v = T(\phi^{-1}) \circ v \circ \phi$, a vector field on $\mathcal{B}$ is called the **pull-back** of $v$ by $\phi$.

If $V$ is a vector field covering $\phi$, we say that $V \circ \phi^{-1}$ (a vector field on $\phi(\mathcal{B})$) is $V$ expressed in the **spatial picture** and $T\phi^{-1} \circ V$ (a vector field on $\mathcal{B}$) is $V$ expressed in the **convected picture**.

The diagram relevant to these concepts is as follows:

![Diagram](image)

From 2.6(b) we have the coordinate description of these objects. For example, for the tilt,

$$(T\phi \circ Y(X))^a = \frac{\partial \phi^a}{\partial X^4}(X) \cdot Y^4(X),$$

and for the push-forward,

$$(\phi_* Y(X))^a = \frac{\partial \phi^a}{\partial X^4}(X) \cdot Y^4(X), \quad \text{where } x = \phi(X).$$
2.10 Definition Let $\phi_t$ be a regular motion of $\mathcal{B}$ in $\mathcal{S}$ and let $V_t$ and $v_t$ be the material and spatial velocities defined in Section 1.1. We call $\phi_t^*v_t = v_t$ the convective velocity; it is a vector field on $\mathcal{B}$ for each $t$.

The physical meaning of the convective velocity is explained by the use of “convected coordinates” as follows.

2.11 Proposition Let $\phi_t$ be a regular motion of $\mathcal{B}$ in $\mathcal{S}$, and $V_t$, $v_t$, and $\nu_t$ the material, spatial, and convective velocities, respectively. Suppose $\{X^t\}$ is a coordinate system on $\mathcal{B}$ and $\{x^a\}$ is one on $\mathcal{S}$. Let $\chi^a$ be the coordinate system on $\phi_t(\mathcal{B})$ defined by $\chi^a = X^a \circ \phi_t^{-1}$. Then the components of $\nu_t$ with respect to $X^a$ equal those of $V_t$ (or $v_t$) with respect to $\chi^a$ (evaluated at the appropriate points).

Proof Let $V^a_t$ be the components of $V$ with respect to $\chi^a$ and $v^a_t$ those of $v$ with respect to $X^a$. Thus

$$\nu = v^a E_A, \quad V = V^a A \xi_A$$

where $E_A = (\partial X^a/\partial X^b) \xi_b$, and $\xi_A = (\partial x^a/\partial \chi^A) \xi_A$. By definition, $V(X, t) = T\phi \cdot v(X, t)$, so from 2.6(b), $V = (\partial \phi^b / \partial x^a) v^b \xi_b$. However, $\phi^b_t(X, t) = (X^b \circ \phi_t^{-1})(\phi_t(X)) = X^b(X)$, so $\partial \phi^b_t / \partial X^a = \delta^b_A$, and $V = v^a \xi_A$ as required.

Problem 2.1 Prove that $\xi_A = T\phi \cdot E_A$. From this, obtain an alternative proof of 2.11.

Problem 2.2 Define the convective acceleration $\alpha_t$ by $\alpha = \phi^* a$ where $a$ is the spatial acceleration.

(a) Show that the components of $\alpha$, with respect to $\{X^t\}$ equal those of $A$, with respect to $\{x^a\}$.

(b) Find the relationship between the convected acceleration and the convected velocity in components.

The coordinates $\chi^a$ may be thought of as being convected by the motion, or being scribed on $\mathcal{B}$ and carried with it as it moves. See Figure 1.2.4. A dual procedure may also be constructed; that is, one could convect the coordinates $\{x^a\}$ on $\phi_t(\mathcal{B})$ back to $\mathcal{B}$ by composition with $\phi_t$.

The operations of pull-back and push-forward may be performed on tensor fields on a general manifold $\mathcal{Q}$. We pause here briefly to consider the situation for one-forms—that is, for “covariant” vectors as opposed to “contravariant” vectors.

2.12 Definitions Let $\mathcal{Q}$ be an $n$-manifold and $q \in \mathcal{Q}$. A one-form at $q$ is a linear mapping $\alpha_q: T_q \mathcal{Q} \to \mathbb{R}$; the vector space of one-forms at $q$ is denoted $T_q^* \mathcal{Q}$. The cotangent bundle of $\mathcal{Q}$ is the disjoint union of the sets $T_q^* \mathcal{Q}$ (made into

---

In the general case in which $\mathcal{B}$ and $\mathcal{S}$ have different dimensions, $v_t$ need not be tangent to $\phi_t(\mathcal{B})$, but only to $\mathcal{S}$. Thus $\phi_t^* v_t$ will not make sense, even if $\phi_t$ is regular. See Section 1.5 for details. Here there is no trouble since we are in the special case $\mathcal{B}$ open in $\mathbb{R}^n$ and $\mathcal{S} = \mathbb{R}^n$. 


a manifold, as was the tangent bundle). A one-form on $\mathcal{Q}$ is a map $\alpha : \mathcal{Q} \to T^*\mathcal{Q}$ such that $\alpha_q = \alpha(q) \in T_q^*\mathcal{Q}$ for all $q \in \mathcal{Q}$.

If $\{x^i\}$ are coordinates on $\mathcal{Q}$, we saw that they induce coordinates on $T\mathcal{Q}$; if $\mathcal{Q} \subset \mathbb{R}^n$ is open, $T\mathcal{Q} = \mathcal{Q} \times \mathbb{R}^n$ and $(q, v)$ is mapped to $(x'(q), v')$, where $v = v^i e_i$. For $T^*\mathcal{Q} = (\mathcal{Q} \times \mathbb{R}^n)^*$, we map $(q, \alpha)$ to $(x'(q), \alpha_j)$, where $\alpha = \alpha_j e^i$, $e^i$ being the basis dual to $e_i$; that is, $e^i(e_j) = \delta^i_j$.

2.13 Definition Let $f : \mathcal{Q} \to \mathbb{R}$ be $C^1$ so that $Tf : T\mathcal{Q} \to T\mathbb{R} = \mathbb{R} \times \mathbb{R}$. The second factor (the "vector part") is called $df$, the differential of $f$. Thus $df$ is a one-form\textsuperscript{12} on $\mathcal{Q}$.

The reader may verify the following.

2.14 Proposition If $\{x^i\}$ is a coordinate system on $\mathcal{Q}$, then $e^i = dx^i$; that is, $dx^i$ is the dual basis of $\partial / \partial x^i$; furthermore,

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

The transformation rule for one-forms is easy to work out. If $\alpha$ has components $\alpha_i$ relative to $\{x^i\}$ and $\bar{\alpha}_i$ relative to $\{\bar{x}^j\}$, then $\bar{\alpha}_i = (\partial x^j / \partial \bar{x}^i) \alpha_j$.

2.15 Definitions If $\phi : \mathcal{B} \to \mathcal{S}$ is a mapping, a one-form over $\phi$ is a mapping $\Lambda : \mathcal{B} \to T^*\mathcal{S}$ such that for $X \in \mathcal{B}$, $\Lambda_x = \Lambda(X) \in T_{\phi(x)}^*\mathcal{S}$.

If $\phi$ is $C^1$ and $\beta$ is a one-form on $\mathcal{S}$, then $\phi^*\beta$—the one-form on $\mathcal{B}$ defined by $(\phi^*\beta)_X : W_X = \beta_{\phi(x)}(T\phi \cdot W_X)$ for $X \in \mathcal{B}$ and $W_X \in T_x\mathcal{B}$—is called the pull-back of $\beta$ by $\phi$. (In contrast to vector fields, the pull-back of a one-form does not

\textsuperscript{12}Do not confuse $df$ with the gradient of $f$, introduced later, which is a vector field on $\mathcal{Q}$. In this book all tensor fields and tensor operations are denoted by boldface characters. Points, coordinates, components, mappings, and scalar fields are denoted by lightface characters.
use the inverse of $\phi$, so does not require $\phi$ to be regular.) If $\phi$ is regular, we can define the push-forward of a one-form $\gamma$ on $\mathcal{B}$ by $\phi_* \gamma = (\phi^{-1})^* \gamma$.

**Problem 2.3** Let $T^*_x \phi : T^*_x \mathcal{B} \to T^*_x \mathcal{S}$ be the dual of $T_x \phi$. Show that $(\phi^* \beta)_x = T^*_x \phi \circ \beta \circ \phi^{-1}$

The coordinate expression for pull-back follows.

**2.16 Proposition** If $\{X^a\}$ are coordinates on $\mathcal{B}$ and $\{x^a\}$ are coordinates on $\mathcal{S}$, then

$$(\phi^* \beta(X))_X = \frac{\partial \phi^a}{\partial X^a}(X) \cdot \beta_a(\phi(X)).$$

*Proof*

$$(\phi^* \beta)_x \cdot E_a = \beta^\phi(x) \cdot T \phi \cdot E_a = \beta^\phi(x) \cdot \frac{\partial \phi^a}{\partial X^a} e_a = \frac{\partial \phi^a}{\partial X^a} \beta_{\phi(x)} \cdot e_a$$

so the result follows since $\beta \cdot e_a = \beta_a$.

Using the chain rule one can easily prove the following properties of pull-back:

**2.17 Proposition**

(a) Let $\phi : \mathcal{B} \to \mathcal{S}$ be $C^1$ and $f : \mathcal{S} \to \mathbb{R}$ be $C^1$. Set $\phi^* f = f \circ \phi$. Then $\phi^* (df) = d(\phi^* f)$.

(b) If $\psi : \mathcal{S} \to \mathcal{Q}$ is $C^1$ and $\gamma$ is a one-form on $\mathcal{Q}$, then $(\psi^* \phi)^* \gamma = \phi^* (\psi^* \gamma)$

In coordinates, pull-back by $\phi$ acts the same way as if $\phi$ were a coordinate transformation. The following shows that the rule in 2.17(a) embodies the coordinate expression:

$$\phi^* \beta = \phi^* (\beta_a dX^a) = (\beta_a \circ \phi) \phi^* dX^a = (\beta_a \circ \phi) d\phi^a = (\beta_a \circ \phi) \frac{\partial \phi^a}{\partial X^b} dX^b$$

so

$$(\phi^* \beta)_a = \beta_a \circ \phi \frac{\partial \phi^a}{\partial X^b}.$$

One can think of one-forms as row vectors and vectors as column vectors in the sense of matrix algebra. The natural contraction $\beta \cdot Y$ defined by $(\beta \cdot Y)(X) = \beta_x(Y(X))$, where $\beta$ is a one-form on $\mathcal{B}$ and $Y$ a vector field, is just matrix multiplication (i.e. $\beta^\phi Y^a$). If $\beta = df$, then $\beta \cdot Y$ is just $Y[f]$ discussed earlier.

Box 2.1 below makes some further connections between "classical" and "modern" tensor analysis, and summarizes the situation.

**Problem 2.4** Consider the motion $\phi_t$ and coordinate systems of Problem 1.1.

(i) Let $w$ be a vector field on $\mathbb{R}^3$. Calculate $\phi_t^* w$.

(ii) What if $w$ is the spatial velocity?

(iii) Let $\alpha$ be a one form on $\mathbb{R}^3$. Calculate $\phi_t^* \alpha$.

(iv) Let $Y(Z^1, Z^2, Z^3) = (0, -Z^3, Z^2)$. Calculate the $\phi_t$-tilt of $Y$. 

Box 2.1 Notational Relationship between Classical Tensor Analysis and Analysis on Manifolds

Classical Tensor Analysis Let \( \{ x^a \} \) denote a (curvilinear) coordinate system defined on an open subset of \( \mathbb{R}^n \). Let \( z^i \) and \( \hat{t}_i \) denote the Cartesian coordinate functions of \( \mathbb{R}^n \) and the corresponding unit basis vectors (the standard basis of \( \mathbb{R}^n \)), respectively. We may view the \( z^i \) as functions of \( x^a \) and vice-versa. The coordinate basis vectors \( e_a \) (thought of as column vectors) corresponding to \( x^a \), are defined by \( e_a = (\partial z^i/\partial x^a) \hat{t}_i \) and are tangent to their respective coordinate curves \( x^a \). The dual basis \( e^a \) (thought of as row vectors) is defined by the inner product \( e^a \cdot e_b = \delta^a_b \). If we define the metric tensor \( g_{ab} \) by

\[
   g_{ab} = \frac{\partial z^i}{\partial x^a} \frac{\partial z^j}{\partial x^b} \delta_{ij}
\]

and let \( g^{ab} \) denote the inverse matrix of \( g_{ab} \) (so \( g^{ac} g_{cb} = \delta^a_b \)), then \( e^a = g^{ab} e_b \). (This is an easy verification; see Section 1.3 for details.) Both \( e_a \) and \( e^a \) are usually viewed as vectors although they satisfy different transformation laws; that is, if \( \tilde{x}^a \) denotes another coordinate system, then

\[
   \tilde{e}_a = \frac{\partial \tilde{x}^b}{\partial x^a} e_b \quad \text{and} \quad \tilde{e}^a = \frac{\partial \tilde{x}^a}{\partial x^b} e^b.
\]

To distinguish between the two types of vectors, the terminologies covariant and contravariant are applied to \( e_a \) and \( e^a \), respectively. (Actually, \( e^a \) is not a vector, but a one-form; the use of the metric is the source of confusion.)

Tensor Analysis on Manifolds Let \( \{ x^a \} \) denote a coordinate system defined on an open set in a manifold \( S \). The coordinate basis corresponding to \( \{ x^a \} \) is denoted \( \partial/\partial x^a \), and the dual basis is denoted \( dx^a \). The terminology “vector” is reserved for \( \partial/\partial x^a \) and the terminology “one-form” (or covector) is used for \( dx^a \). The transformation rules for \( \partial/\partial x^a \) and \( dx^a \) are suggested by the differential operator notations; that is, if \( \tilde{x}^a \) denotes another coordinate system on \( S \), then

\[
   \frac{\partial}{\partial \tilde{x}^a} = \frac{\partial x^b}{\partial \tilde{x}^a} \frac{\partial}{\partial x^b} \quad \text{and} \quad d\tilde{x}^a = \frac{\partial \tilde{x}^a}{\partial x^b} dx^b.
\]

One sees from these expressions that \( \partial/\partial x^a \) and \( dx^a \) are the analogs of \( e_a \) and \( e^a \), respectively. If we define \( e^a \) as the dual basis in the proper sense (i.e., \( e^a \) is a linear functional) without use of the metric, then \( e_a = \partial/\partial x^a \) and \( e^a = dx^a \).

Tensor analysis on manifolds is developed without assuming the structure of a “background” \( \mathbb{R}^n \). Thus Cartesian coordinate systems and standard bases are unavailable, so relations of the form \( e_a = (\partial z^i/\partial x^a) \hat{t}_i \) are not allowed. The notion of a vector can be defined operationally via the directional derivative as follows. Let \( v = v^a (\partial/\partial x^a) \) denote a vector field on \( S \). Then the directional derivative of the
function \( f: \mathbb{S} \rightarrow \mathbb{R} \) in the direction \( \mathbf{v} \), denoted \( \mathbf{v}[f] \), is given by 
\[ \mathbf{v}[f] = v^a \left( \frac{\partial f}{\partial x^a} \right) \], which is independent of the coordinate system. The analogous quantity for vectors in \( \mathbb{R}^n \) is given by
\[ \mathbf{v}[f] \overset{\text{def}}{=} \frac{d}{d\epsilon} f(z^i + \epsilon \mathbf{v}^j) \bigg|_{\epsilon=0} = v^i \frac{\partial f}{\partial z^i} = v^a \frac{\partial f}{\partial x^a}. \]
Thus the end result is the same.

The differential of \( f \), denoted \( df \), can be defined by \( df(\mathbf{v}) = \mathbf{v}[f] \).

From this, the duality relation follows immediately: 
\[ dx^a(\partial/\partial x^b) = \frac{\partial x^a}{\partial x^b} = \delta^a_b. \] This is the analog of the relationship \( e^a \cdot e_b = \delta^a_b \).

Manifolds are important in the formulation of physical theories. For example, in relativity space–time is modeled as a four-dimensional, pseudo-Riemannian manifold. There is simply no Cartesian structure to fall back on. In continuum mechanics, with a manifold as a basic notion of a body, a geometric theory of structured materials becomes a possibility. Even for a simple body, tensor analysis on manifolds clarifies the basic theory. For instance, using manifold ideas we can see clearly how to formulate the pull-back and push-forward and hence to clarify the meaning of convective velocity (and many more things that are considered in the following sections). The metric is not required to talk about \( e_a \) and \( e^a \). The confusion arises through unnecessary identifications brought about by the presence of the Cartesian structure. If introduced in the manifold context, we can see exactly how the metric is needed (see Section 1.3).

The following table summarizes the relations and notations of classical tensor analysis on \( \mathbb{R}^n \) and tensor analysis on manifolds.

<table>
<thead>
<tr>
<th>Classical Tensor Analysis</th>
<th>Tensor Analysis on Manifolds</th>
</tr>
</thead>
<tbody>
<tr>
<td>( {x^a} )</td>
<td>( {x^a} )</td>
</tr>
<tr>
<td>( e_a = \frac{\partial z^i}{\partial x^a} \hat{i}_i )</td>
<td>( \frac{\partial}{\partial x^a} = e_a )</td>
</tr>
<tr>
<td>( e^a = g^{ab} e_b )</td>
<td>( dx^a = e^a )</td>
</tr>
<tr>
<td>( e^a \cdot e_b = \delta^a_b )</td>
<td>( dx^a(\partial/\partial x^b) = \delta^a_b )</td>
</tr>
<tr>
<td>( \tilde{e}_a = \frac{\partial x^b}{\partial \tilde{x}^a} e_b )</td>
<td>( \tilde{e}_a = \frac{\partial x^a}{\partial \tilde{x}^b} \tilde{e}_b )</td>
</tr>
<tr>
<td>( \tilde{e}_a = \frac{\partial x^a}{\partial \tilde{x}^b} e_b )</td>
<td>( \tilde{e}_a = \frac{\partial x^a}{\partial \tilde{x}^b} \tilde{e}_b )</td>
</tr>
<tr>
<td>( v = v^a e_a )</td>
<td>( v = v^a(\partial/\partial x^a) )</td>
</tr>
<tr>
<td>( v^a = e^a \cdot \mathbf{v} )</td>
<td>( v^a = dx^a(\mathbf{v}) )</td>
</tr>
<tr>
<td>( \alpha = \alpha_a e^a )</td>
<td>( \alpha = \alpha_a dx^a )</td>
</tr>
<tr>
<td>( \alpha_a = e_a \cdot \alpha )</td>
<td>( \alpha_a = \alpha(\partial/\partial x^a) )</td>
</tr>
<tr>
<td>( v[f] = v^a \frac{\partial f}{\partial x^a} )</td>
<td>( v[f] = v^a \frac{\partial f}{\partial x^a} )</td>
</tr>
</tbody>
</table>
In classical tensor analysis, a vector and one-form are indistinguishable in the sense that we have the representations

\[ v = v^a e_a = v_a e^a, \quad \text{and} \quad \alpha = x_a e^a = \alpha^a e_a. \]

In tensor analysis on manifolds one avoids confusing vectors and one-forms.

**Box 2.2 Summary of Important Formulas for Section 1.2**

**Tangent Map of** \( \phi : \mathcal{B} \to \mathcal{S} \)

\[ T\phi : T\mathcal{B} \to T\mathcal{S}, \]

\[ T\phi \cdot W_x = T\phi(X, W) = (\phi(X), D\phi(X) \cdot W), \quad (T\phi \cdot W_x) = \frac{\partial \phi^a}{\partial X^A} W^A \]

**Projection**

\[ \pi_S(T\phi \cdot W_x) = \phi(X) \quad \text{(} \pi_S(T\phi \cdot W_x))^a = \phi^a(X) \]

**Tilt of a Vector Field** \( Y \)

\[ V = T\phi \circ Y \quad V^a = \frac{\partial \phi^a}{\partial X^A} Y^A \]

**Push-Forward of Vector Field** \( \phi \)

\[ \phi_\ast Y = T\phi \circ Y \circ \phi^{-1} \quad (\phi_\ast Y)^a = \left( \frac{\partial \phi^a}{\partial X^a} \circ \phi^{-1} \right) (Y^a \circ \phi^{-1}) \]

**Pull-Back of Vector Field** \( v \)

\[ \phi^\ast v = T\phi^{-1} \circ v \circ \phi \quad (\phi^\ast v)^A = \left( \frac{\partial (\phi^{-1})^A}{\partial x^a} \circ \phi \right) (v^a \circ \phi) \]

**Pull-Back of One-Form** \( \beta \)

\[ \phi^\ast \beta = (\mathcal{B} \circ \phi) \cdot T\phi \quad (\phi^\ast \beta)_a = (\beta_a \circ \phi) \frac{\partial \phi^a}{\partial X^A} \]

**Push-Forward of One-Form** \( \gamma \)

\[ \phi_\ast \gamma = (\phi^{-1})^\ast \gamma \quad (\phi_\ast \gamma)_a = (\gamma_A \circ \phi^{-1}) \frac{\partial (\phi^{-1})^A}{\partial x^a} \]

**Change of Coordinates**

\[ \frac{\partial \phi^a}{\partial X^A} = \frac{\partial \bar{x}^a}{\partial X^b} \frac{\partial X^b}{\partial X^A} \bar{\phi}_a = \frac{\partial x^b}{\partial \bar{x}^a} \beta_b \]

**Differential of a Function**

\[ df = \text{second component of } Tf \quad df = \frac{\partial f}{\partial x^a} dx^a \]

\[ dx^a = e^a, \text{ dual basis of } e_a \]
Directional Derivatives

\[ Y[f] = df \cdot Y \]

Pull-Back of a Differential

\[ \phi^* df = d(\phi^* f) \]

Chain Rules

\[ T(\psi \circ \phi) = T\psi \circ T\phi \]
\[ (\psi \circ \phi)^* = \phi^* \circ \psi^* \]

Convective Velocity

\[ \mathbf{v}_t = \phi_t^* \mathbf{v}_t \]

\[ \frac{\partial}{\partial x^A} \frac{\partial \phi^a}{\partial X^A} = \frac{\partial (f \circ \phi)}{\partial x^A} \]

GEOMETRY AND KINEMATICS OF BODIES

1.3 THE DEFORMATION GRADIENT

The derivative of the configuration of a body is called the deformation gradient. This object plays a fundamental role in the subsequent theory, so we shall devote this section to its detailed study. Several of the notions here serve as motivation for the geometric considerations of Section 1.4. Reciprocally, the general framework of that section will help give perspective and a deeper understanding of the results here.

3.1 Definition Let \( \phi : \mathcal{B} \rightarrow \mathcal{S} \) be a \( C^1 \) configuration of \( \mathcal{B} \) in \( \mathcal{S} \) (\( \mathcal{B} \) and \( \mathcal{S} \) can be general manifolds here). The tangent of \( \phi \) is denoted \( F \) and is called the deformation gradient of \( \phi \); thus \( F = T\phi \). For \( X \in \mathcal{B} \), we let \( F_X \) or \( F(X) \) denote the restriction of \( F \) to \( T_X \mathcal{B} \). Thus \( F(X) : T_X \mathcal{B} \rightarrow T_{\phi(X)} \mathcal{S} \) is a linear transformation for each \( X \in \mathcal{B} \).

In Proposition 2.6 we worked out the coordinate description of \( F \). We recall this for reference.

3.2 Proposition Let \( \{X^A\} \) and \( \{x^a\} \) denote coordinate systems on \( \mathcal{B} \) and \( \mathcal{S} \), respectively. Then the matrix of \( F(X) \) with respect to the coordinate bases \( e_a(X) \) and \( e_a(x) \) [where \( x = \phi(X) \)] is given by

\[ F^a_A(X) = \frac{\partial \phi^a}{\partial X^A}(X). \]

If we have a motion \( \phi(X, t) \), we shall write the components of the deformation gradient for each \( t \) as \( F^a_A(X, t) \), or merely as \( F^a_A \) if we are suppressing the argu-
ments. The deformation gradient $F$ is an important example of a two-point tensor. These objects will be discussed in general in Section 1.4. Notice that the coordinate expression for $F^a_\alpha$ does not involve any covariant derivatives. This is because $\phi$ is not a vector, but rather is a point mapping of $\mathfrak{B}$ to $\mathfrak{S}$. (Sometimes $x = \phi(X)$ is represented by an “arrow,” but this can be a source of confusion.)

For the remainder of this section we will assume $\mathfrak{S} = \mathbb{R}^n$ and $\mathfrak{B} \subset \mathbb{R}^n$ is a simple body. (After the reader has digested Section 1.4, $\mathfrak{B}$ and $\mathfrak{S}$ may be replaced by Riemannian manifolds.)

### 3.3 Notation

We let $\langle , \rangle_x$ denote the standard inner product in $\mathbb{R}^n$ for vectors based at $x \in \mathfrak{S}$ and similarly let $\langle , \rangle_x$ be the standard inner product in $\mathbb{R}^n$ at $X \in \mathfrak{B}$. For a vector $v \in T_x \mathfrak{S}$ we let $\|v\|_x = \langle v, v \rangle_x^{1/2}$ be the length of $v$. Similarly the length of $W \in T_X \mathfrak{B}$ is denoted $\|W\|_x$. (If there is no danger of confusion, the subscripts may be dropped.)

Let $A : T_x \mathfrak{B} \rightarrow T_x \mathfrak{S}$ be a linear transformation. Then the transpose, or adjoint of $A$, written $A^T$, is the linear transformation $A^T : T_x \mathfrak{S} \rightarrow T_x \mathfrak{B}$ such that $\langle A^T v, w \rangle_x = \langle v, A^T w \rangle_x$ for all $W \in T_x \mathfrak{B}$ and $v \in T_x \mathfrak{S}$. If $B : T_x \mathfrak{S} \rightarrow T_x \mathfrak{S}$ is a linear transformation, it is called symmetric if $B = B^T$.

In a coordinate system $\{x^a\}$ on $\mathfrak{S}$, let the metric tensor $g_{ab}$ be defined by $g_{ab}(x) = \langle e_a, e_b \rangle_x$ and similarly define $G_{AB}(X)$ on $\mathfrak{B}$ by $G_{AB}(X) = \langle E_A, E_B \rangle_X$. We let $g^{ab}$ and $G^{AB}$ denote the inverse matrices of $g_{ab}$ and $G_{AB}$; these exist since $g_{ab}$ and $G_{AB}$ are nonsingular.

In Euclidean space, $e_a = (\partial z^i/\partial x^a)\delta^i_j$, so we have the expression

$$g_{ab} = \frac{\partial z^i}{\partial x^a} \frac{\partial z^j}{\partial x^b} \delta_{ij}.$$  

Similarly,

$$G_{AB} = \frac{\partial Z^i}{\partial X^a} \frac{\partial Z^j}{\partial X^b} \delta_{ij}.$$  

### 3.4 Proposition

(i) For $v, w \in T_x \mathfrak{S}$ and a coordinate system $\{x^a\}$, we have

$$\langle v, w \rangle_x = g_{ab} v^a w^b.$$  

(ii) If $\{x^a\}$ and $\{X^A\}$ are coordinate systems on $\mathfrak{S}$ and $\mathfrak{B}$, respectively, and $\phi : \mathfrak{B} \rightarrow \mathfrak{S}$ is a $C^1$ configuration of $\mathfrak{B}$, then the matrix of $F^T$ is given by

$$\left(F^T(x)\right)_a^\alpha = g_{ab}(x) F^b_\beta(X) G^{AB}(X),$$  

where $x = \phi(X)$.  

Proof

(i) This follows from the definition of $g_{ab}$ and the expressions $v = v^a e_a$ and $w = w^b e_b$:

$$\langle v, w \rangle_x = \langle v^a e_a, w^b e_b \rangle = v^a w^b \langle e_a, e_b \rangle = v^a w^b g_{ab}.$$  

(ii) By definition,

$$\langle F^T w, W \rangle_x = \langle FW, w \rangle_x; \quad \text{that is,} \quad (F^T)^b_{a} w^b W^A_{AB} = F^a_A W^A_{AB} g_{ab}$$

for all $W \in T_x \mathcal{B}$ and $w \in T_x \mathcal{S}$, where $F^T$ and $F$ have their arguments suppressed. Since $W$ and $w$ are arbitrary, $(F^T)^b_{a} g_{ab} = F^a_A g_{ab}$. Multiplying by $G^{AC}$ and using $G_{AB} G^{AC} = \delta^C_B$ gives the result. (In order for the map $F^T: T(\phi(\mathcal{B})) \rightarrow T\mathcal{B}$ to be well defined, $\phi$ must be regular.)

Problem 3.1 Define $F^*(x): T_x^* \mathcal{S} \rightarrow T_x^* \mathcal{B}$ by $(F^*(x))_A^\beta = \beta(F(x) \cdot W)$ for $\beta \in T_x^* \mathcal{S}$ and $W \in T_x \mathcal{B}$. Show that the matrix representative of $F^*$ with respect to the dual bases $e^a$ and $E^A$ is the transpose of $F^a_A$, i.e. dropping $\beta$, $F^*(e^a) = F^a_A E^A$.

3.5 Definition The (Green) deformation tensor (also called the right Cauchy–Green tensor) $C$ is defined by:

$$C(X): T_x \mathcal{B} \rightarrow T_x^* \mathcal{B}, \quad C(X) = F(X)^T F(X)$$

Or, for short, $C = F^T F$.

If $C$ is invertible, we let $B = C^{-1}$, called the Piola deformation tensor.

The following proposition gives some of the basic properties of $C$.

3.6 Proposition Let $\phi$ be a $C^1$ configuration of $\mathcal{B}$ in $\mathcal{S}$ and $C$ the deformation tensor.

(i) If $\{x^a\}$ and $\{X^A\}$ are coordinate systems on $\mathcal{S}$ and $\mathcal{B}$, respectively, then

$$C^A_{\, B} = (F^T)^A_{\, a} F^a_{\, B} = g_{ab} G^{AC} \frac{\partial \phi^b}{\partial X^c} \frac{\partial \phi^a}{\partial X^B}.$$  

(ii) $C$ is symmetric and positive-semidefinite; that is, $\langle CW, W \rangle_x \geq 0$, and if each $F_x$ is one-to-one, then $C$ is invertible and positive-definite; that is, $\langle CW, W \rangle > 0$ if $W \neq 0$. (Note that $F$ is one-to-one if $\phi$ is regular.)

Proof

(i) follows from the definition of $C$ and 3.4(ii).

(ii) Let $W_1, W_2 \in T_x \mathcal{B}$. Then

$$\langle CW_1, W_2 \rangle_x = \langle F^T F W_1, W_2 \rangle_x = \langle FW_1, FW_2 \rangle_x = \langle W_1, F^T F W_2 \rangle_x = \langle W_1, CW_2 \rangle_x,$$

so $C$ is symmetric. Clearly, $\langle CW, W \rangle_x = \langle FW, FW \rangle_x \geq 0$, so $C$ is positive-semidefinite.
If $F$ is one-to-one and if $\langle CW, W \rangle = \langle FW, FW \rangle$ is zero, then $FW = 0$ and hence $W = 0$. In particular, $C$ is one-to-one and hence is invertible.

Symmetry of $C$ means that $C_{AB} = C_{BA}$, where $C_{AB} = G_{AC}C_{CB}$. We call $C_{AB}$ the associated components of $C$. We pause briefly to consider this notion.

3.7 Definition Let $\alpha$ be a one form on $S$, with components $\alpha_a$ in a coordinate system $\{x^a\}$; that is, $\alpha = \alpha_a e^a$. The associated vector field $\alpha^i$ is defined to have components $\alpha^a = g^{ab} \alpha_b$; that is, $\alpha^i = \alpha^a e_a = g^{ab} \alpha_b e_a$.

If $v$ is a vector field, the associated one-form $v^i$ is defined by $v^i = v_a e^a$, where $v_a = g_{ab} v^b$. If $\sigma$ is a tensor with components $\sigma_{ab}$, then the tensors with components $\sigma^b_a = \sigma_{ac} g^{cb}$, $\sigma^a_b = \sigma_{cb} g^{ca}$, and $\sigma^{ab} = \sigma_{cd} g^{ca} g^{db}$, are called associated tensors. In general, a rank $N$ tensor $t$ has $2^N$ associated tensors defined in a similar way. If we write $t^i$, we mean the tensor associated to $t$ that has all its indices lowered. Similarly, $t^i$ means $t$ with all its indices raised.

We wish to emphasize that associated tensors are different objects. Specifically, $v^a e_a$ and $v_a e^a$ are not equal. These tensors are related via the metric tensor, but they are not the same tensor. This point can become confused because of an over-reliance on the Cartesian structure. The situation will be clarified in the next section. Another thing Section 1.4 will clarify is the very definition of $C$ itself. We shall, in fact, see that $C^b = \phi^* g$, the pull-back of the metric $g$ on $S$ to $S$. In a similar way, we can push the metric $G$ on $S$ forward to $S$. This leads to a new tensor $b = c^{-1}$ with $c^i = \phi^* G$. The explicit definition follows.

3.8 Definition Let $\phi$ be a regular $C^1$ configuration of $S$ in $S$. Then the Finger deformation tensor (also called the left Cauchy–Green tensor) is defined on $\phi(S)$ by

$$b(x) = F(X)(F(X))^T : T_x \phi(S) \rightarrow T_x \phi(S)$$

[where $X = \phi^{-1}(x)$], or $b = FF^T$ for short. Also, define $c = b^{-1}$.

Note that $C$ is defined whether or not $\phi$ is regular, but $c$ and $b$ require $\phi^{-1}$ to exist to be defined. As in 3.6, one can prove the following.

3.9 Proposition If $\phi$ is $C^1$ and regular, then

(i) $b^a_b = g_{bc} G^{ab} \frac{\partial \phi^a}{\partial X^b} \frac{\partial \phi^c}{\partial X^a}$ and

(ii) $b$ is symmetric and positive-definite.

Now we shall use a result from linear algebra.

3.10 Lemma Let $U$ be a finite-dimensional inner product space and let $A : U \rightarrow U$ be a symmetric positive-definite linear transformation; that is, $A^T = A$ and $\langle Av, v \rangle > 0$ if $v \in U$ ($v \neq 0$). Then there exists a unique symmetric positive-definite linear transformation $B : U \rightarrow U$ such that $B^2 = A$. 
Let us recall the proof of existence (The summation convention is temporarily suspended): There is an orthonormal basis \( \psi_1, \ldots, \psi_n \) of eigenvectors for \( A \). Let \( \lambda_1, \ldots, \lambda_n \) be the corresponding eigenvalues. Then \( A \psi_i = \lambda_i \psi_i \), so \( \langle A \psi_i, \psi_i \rangle = \lambda_i \| \psi_i \|^2 > 0 \) and hence \( \lambda_i > 0 \). Define \( B \) by \( B \psi_i = \sqrt{\lambda_i} \psi_i \). Thus \( B^2 \psi_i \) agrees with \( A \psi_i \), so \( B^2 = A \). This shows \( B \) exists. Note that the eigenvalues of \( B \) are the square roots of those of \( A \).

3.11 Definition Let \( \phi \) be regular and let \( C \) and \( b \) be defined as above. Let \( U \) and \( V \) denote the unique symmetric, positive-definite square roots of \( C \) and \( b \), respectively. We call \( U \) and \( V \) the right and left stretch tensors, respectively. [For each \( x \in S \), \( V(x) : T_x S \to T_x S \) and for each \( X \in \mathcal{B} \), \( U(X) : T_x \mathcal{B} \to T_x \mathcal{B} \).] The eigenvalues of \( U \) are called the principal stretches.

**Warning.** Note that \( V \) also denotes the material velocity. However the meaning will be clear from the context.

3.12 Proposition Let \( \phi \) be regular. For each \( X \in \mathcal{B} \) there exists an orthogonal transformation \( R(X) : T_x S \to T_x S \) [i.e., \( R(X)^T R(X) = I \) (the identity on \( T_x S \)) and \( R(X)R(X)^T = i \) (the identity on \( T_x S \))] such that

\[
F = RU \quad \text{[that is, } F(X) = R(X) \circ U(X)\text{]}
\]

and

\[
F = VR \quad \text{[that is, } F(X) = V(X) \circ R(X)\text{]}
\]

Moreover, each of these decompositions is unique: if \( F = \tilde{R} \tilde{U} \), where \( \tilde{R} \) is orthogonal and \( \tilde{U} \) is symmetric positive-definite, then \( \tilde{R} = R \) and \( \tilde{U} = U \). We call \( R \) the rotation matrix and refer to \( F = RU \) and \( F = VR \) as the right and left polar decompositions, respectively.

**Proof.** Define \( R(X) = F(X)U(X)^{-1} \). Then \( R^T R = U^{-1} F^T F U^{-1} = U^{-1} C U^{-1} = U^{-1} U^2 U^{-1} = I \) since \( C = U^2 \) by definition of \( U \), and since \( (U^{-1})^T = U^{-1} \) as \( U \) is symmetric positive-definite. Also, \( R R^T = R(R^T R)R^T = (RR^T)^2 \), so \( RR^T = I \), since \( RR^T \) is nonsingular. Thus the right polar decomposition is established. To establish its uniqueness, let \( F = \tilde{R} \tilde{U} = RU \). Then \( F^T F = \bar{U}^T \tilde{R}^T \tilde{R} \bar{U} = \bar{U}^2 \) and \( F^T F = U^2 \). Since symmetric positive square roots are unique, \( \bar{U} = U \). Hence \( \tilde{R} = R \).

Let \( F = V \tilde{R} \) be the left polar decomposition, established in the same manner as the right polar decomposition. We prove \( R = \tilde{R} \). Indeed, \( F = VR = (\tilde{R} \tilde{R}^T) \tilde{R} \bar{U} = \tilde{R} (\tilde{R}^T \bar{U} \tilde{R}) \). This has the form of the right decomposition, so by uniqueness, \( \tilde{R} = R \) and \( U = R^T VR \).

From this proof we see that \( U \) and \( V \) are similar, and thus have the same eigenvalues. Therefore, the principal stretches can be defined in terms of either \( \sqrt{F^T F} \) or \( \sqrt{FF^T} \). The following commutative diagram summarizes the situation. Notice that, in general it does not make sense to ask that \( U = V \) since they map on different spaces. In particular, \( U \) and \( R \) do not commute. (However, if \( A \) is a symmetric transformation of an inner product space to itself, the components of its polar decomposition do commute and the left and right decompositions coincide.)
Notice that $U$ and $V$ operate within each fixed tangent space; that is, $U(X): T_x \mathcal{B} \rightarrow T_x \mathcal{B}$ and $V(x): T_x \mathcal{S} \rightarrow T_x \mathcal{S}$. On the other hand, $R$ maps $T_x \mathcal{B}$ to $T_x \mathcal{S}$; that is, it shifts the base point as well as rotating.

3.13 Algorithm for Computing the Polar Decomposition Let $X, x$ be fixed, and let $F$ be given. To compute $R$ and $U$, let $C = F^T F$, and let $\psi^1, \ldots, \psi^n$ be orthonormal eigenvectors of $C$ with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let

$$\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \quad \text{and} \quad \Psi = (\psi^1, \ldots, \psi^n)$$

so that $\Lambda = \Psi^T C \Psi$, where

$$\Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_n} \end{bmatrix} \quad (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n} \text{ are the principal stretches}),$$

and set $R = F \Lambda^{1/2}$. Use a similar procedure for the left decomposition or let $V = R U R^T$. An explicit formula for $U, R$ in the two-dimensional case is worked out in 3.15.

Observe that $b = V^2 = (R U R^T)(R U R^T) = R U^2 R^T = R C R^T$. Thus the Finger deformation tensor $b$ and the deformation tensor $C$ are conjugate under the rotation matrix.

3.14 An Example of the Polar Decomposition Let $\mathcal{B}$ be the unit circular cylinder contained in $\mathbb{R}^3$ and let $\{z^i\}$ and $\{Z^i\}$ denote coincident Cartesian coordinate systems for $\mathbb{R}^3$. Then $\mathcal{B}$ can be written as

$$\mathcal{B} = \{X | [Z^1(X)]^2 + [Z^2(X)]^2 \leq 1\}.$$

Consider the configuration $\phi: \mathcal{B} \rightarrow \mathbb{R}^3$ defined explicitly by

$$z^1(X) = \sqrt{3} Z^1(X) + Z^2(X), \quad z^2(X) = 2Z^2(X), \quad \text{and} \quad z^3(X) = Z^3(X).$$

\textsuperscript{13}The data for this example come from Jaunzemis [1967], pp. 78–79.
This configuration may be described as biaxial stretching the in $z^1$, $z^2$-plane. The boundary of $\mathcal{B}$, given by $\partial \mathcal{B} = \{X \mid [Z^1(X)]^2 + [Z^2(X)]^2 = 1\}$, is deformed under $\phi$ into an ellipse: $\partial \phi(\mathcal{B}) = \{x \mid [z^1(x)]^2 + z^1(x)z^2(x) + [z^2(x)]^2 = 3\}$. In the coordinate system $\{\xi^i\}$ defined by $\xi^1 = \sqrt{2} (z^1 + z^2)$, $\xi^2 = \sqrt{2} (z^2 - z^1)$, and $\xi^3 = z^3$, the boundary of $\phi(\mathcal{B})$ can be represented by the equation $[\xi^1(x)]^2/6 + [\xi^2(x)]^2/2 = 1$. Thus the coordinate axes $\xi^1$ and $\xi^2$, which are rotated $45^\circ$ counterclockwise with respect to $z^1$ and $z^2$, coincide with the major and minor axes of the ellipse.

The matrices pertinent to the polar decomposition are listed below.

$$F = \begin{bmatrix} \sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 3 & \sqrt{3} & 0 \\ \sqrt{3} & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Psi = \begin{bmatrix} 1/2 & \sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U = \frac{1}{2\sqrt{2}} \begin{bmatrix} 3 + \sqrt{3} & 3 - \sqrt{3} & 0 \\ 3 - \sqrt{3} & 1 + 3\sqrt{3} & 0 \\ 0 & 0 & 2\sqrt{2} \end{bmatrix}$$

$$U^{-1} = \frac{1}{4\sqrt{6}} \begin{bmatrix} 1 + 3\sqrt{3} & \sqrt{3} - 3 & 0 \\ \sqrt{3} - 3 & 3 + \sqrt{3} & 0 \\ 0 & 0 & 4\sqrt{6} \end{bmatrix}$$

$$R = \frac{1}{2\sqrt{2}} \begin{bmatrix} \sqrt{3} + 1 & \sqrt{3} - 1 & 0 \\ 1 - \sqrt{3} & \sqrt{3} + 1 & 0 \\ 0 & 0 & 2\sqrt{2} \end{bmatrix}$$

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{3} + 1 & \sqrt{3} - 1 & 0 \\ \sqrt{3} - 1 & \sqrt{3} + 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$

The physical interpretation of these results follows (see Fig. 1.3.1).

*Right decomposition* $U$ maps the unit disk in the $Z^1$, $Z^2$-plane into an ellipse with major and minor axes rotated $60^\circ$ counterclockwise with respect to the $Z^1$
and $Z^2$ axes, respectively. $R$ then rigidly rotates the elliptic cylinder $15^\circ$ clockwise about the $Z^3$-axis into its final position. 

Left decomposition $R$ maps the unit disk rigidly into itself. Then $V$ maps the unit disk into an ellipse with major and minor axes coinciding with the $\bar{z}^1$ and $\bar{z}^2$ axes, respectively.

Note that $U$ leaves three orthogonal directions unrotated (the directions are defined by its eigenvectors). Similarly, $V$ leaves three coordinate directions unrotated (these are given by its eigenvectors, which coincide with the axes of the coordinate system $\{Z_i\}$). The principal stretches—that is, $1$, $\sqrt{6}$ and $\sqrt{2}$—which are the eigenvalues of $U$ (or $V$) determine the “size” of the deformation. For example, the major axis of the deformed ellipse is $\sqrt{6}$ and minor axis is $\sqrt{2}$.

**Problem 3.2** Find the polar decompositions of

$$F = \begin{bmatrix} 1 & \kappa \\ 0 & 1 \end{bmatrix}.$$ 

Employ the substitution $\kappa = 2 \tan \alpha$ and check that

$$U = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & (1 + \sin^2 \alpha)/\cos \alpha \end{bmatrix}.$$
3.15 Proposition. In the two-dimensional case, let
\[ I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]
\[ I_c = \text{trace } C, \text{ and } II_c = \det C. \]

Then \[ U = \frac{1}{\sqrt{(I_c + 2\sqrt{II_c})}}(C - \sqrt{II_c}I) \] and \[ R = FU^{-1}. \]

Proof. By the Cayley–Hamilton theorem from linear algebra,
\[ U^2 - I_u U + II_u I = 0. \]
But \( C = U^2 \), and \( \det C = (\det U)^2 \), so \( C - I_u U + \sqrt{II_c}I = 0. \) Taking the trace gives \( I_c - (I_u)^2 + 2\sqrt{II_c} = 0 \), so \( I_u = \sqrt{I_c} + 2\sqrt{II_c}. \) Solving \( C - I_u U + \sqrt{II_c}I = 0 \) for \( U \) and substituting for \( I_u \) gives the result.

The reader might wish to re-work Problem 3.2 using 3.15.

Problem 3.3 The Cayley–Hamilton theorem for \( 3 \times 3 \) matrices \( U \) states that
\[ -U^3 + I_u U^2 - II_u U + III_u I = 0 \]
where \( I_u = \text{tr } U, \) \( II_u = \det U(\text{tr } U^{-1}) \), and \( III_u = \det U \) are the principal invariants of \( U \). Use this to work out an explicit formula for \( U \) in terms of \( C \) and its principal invariants.

Next we study how the deformation tensor and the stretch tensors measure changes in lengths and angles. Recall that if \( \sigma : [a, b] \rightarrow \mathcal{B} \) is a \( C^1 \) (or piecewise \( C^1 \)) curve, its length is given by
\[ l(\sigma) = \int_a^b \| \sigma'(\lambda) \| \, d\lambda. \]

3.16 Proposition. Let \( \sigma \) be a \( C^1 \) curve in \( \mathcal{B} \) and let \( \phi \) be a \( C^1 \) configuration of \( \mathcal{B} \) in \( S \). Let \( \tilde{\sigma} = \phi \circ \sigma \) be the image of \( \sigma \) under \( \phi \). Then the length of \( \tilde{\sigma} \) depends only on \( \sigma \) and on the stretch tensor \( U \).

Proof. From the chain rule \( \tilde{\sigma}'(\lambda) = T\phi(\sigma(\lambda))\sigma'(\lambda) = F_{\sigma(\lambda)}\sigma'(\lambda) \). Hence
\[ \| \tilde{\sigma}'(\lambda) \| = \langle F_{\sigma(\lambda)}\sigma'(\lambda), F_{\sigma(\lambda)}\sigma'(\lambda) \rangle^{1/2} \]
\[ = \langle \sigma'(\lambda), F_{\sigma(\lambda)}^T F_{\sigma(\lambda)} \sigma'(\lambda) \rangle^{1/2} \]
\[ = \langle \sigma'(\lambda), C_{\sigma(\lambda)} \sigma'(\lambda) \rangle^{1/2} \]
\[ = \langle \sigma'(\lambda), U_{\sigma(\lambda)}^2 \sigma'(\lambda) \rangle^{1/2} = \| U_{\sigma(\lambda)} \sigma'(\lambda) \|. \]
Thus
\[ l(\tilde{\sigma}) = \int_a^b \| U_{\sigma(\lambda)} \sigma'(\lambda) \| \, d\lambda. \]

We call \( \tilde{\sigma} \) the deformation of \( \sigma \) under \( \phi \) (see Figure 1.3.2).
Recall that the angle $\theta$ between two vectors $v_1, v_2 \in T_x \mathcal{S}$ is given by $\cos \theta = \langle v_1, v_2 \rangle / \| v_1 \| \cdot \| v_2 \|$, and that the angle between two $C^1$ curves is the angle between their tangent vectors at a point of intersection.

### 3.17 Proposition

Let $\sigma_1$ and $\sigma_2$ be two $C^1$ curves in $\mathcal{S}$ that intersect at $X \in \mathcal{S}$ [i.e., $\sigma_1(\lambda_1) = \sigma_2(\lambda_2) = X$]. Then the angle between the deformed curves $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$ depends only upon $\sigma_1$, $\sigma_2$, and $U_X$.

#### Proof

As in 3.16, $\tilde{\sigma}_1'(\lambda) = F_{\sigma_1(\lambda)} \sigma_1'(\lambda)$ and $\| \tilde{\sigma}_1'(\lambda) \| = \| U_{\sigma_1(\lambda)} \sigma_1'(\lambda) \|$. Now

$$\langle \tilde{\sigma}_1'(\lambda_1), \tilde{\sigma}_2'(\lambda_2) \rangle_x = \langle F_{\sigma_1(\lambda_1)} \sigma_1'(\lambda_1), F_{\sigma_2(\lambda_2)} \sigma_2'(\lambda_2) \rangle_x
= \langle R_X U_x \sigma_1'(\lambda_1), R_X U_x \sigma_2'(\lambda_2) \rangle_x = \langle U_x \sigma_1'(\lambda_1), U_x \sigma_2'(\lambda_2) \rangle_x.$$  

The angle $\theta$ between $\tilde{\sigma}_1'(\lambda_1)$ and $\tilde{\sigma}_2'(\lambda_2)$ is thus

$$\cos \theta = \frac{\langle U_x \sigma_1'(\lambda_1), U_x \sigma_2'(\lambda_2) \rangle_x}{\| U_x \sigma_1'(\lambda_1) \|_x \cdot \| U_x \sigma_2'(\lambda_2) \|_x}.$$  

Thus $U$ (or $C$) measures the distortion, as manifested by changes in lengths and angles, due to the configuration $\phi$. Note that $R$, being orthogonal, has no effect on lengths or angles.

Previously we defined the mappings $C: T\mathcal{B} \rightarrow T\mathcal{B}$ and $b: T\mathcal{S} \rightarrow T\mathcal{S}$. In the ensuing developments we deal with the associated mappings $C^\ast: T^*\mathcal{B} \rightarrow T^*\mathcal{B}$ and $b^\ast: T^*\mathcal{S} \rightarrow T^*\mathcal{S}$ which are obtained by lowering the first index of $C$ and $b$, respectively. The components of $C$ and $b$ are

$$C^A_B = G^{AC} g_{ab} F^b_c F^e_c F^e_B \quad \text{and} \quad b^a_b = G^{AB} g_{bc} F^e_A F^e_B$$

and the components of $C^\ast$ and $b^\ast$ are given by

$$C_{AB} = G_{AD} C^D_B \quad \text{and} \quad b_{ab} = g_{ac} b^c_b.$$  

(In the next section we provide an abstract treatment of raising and lowering indices.)

### 3.18 Definition

The material (or Lagrangian) strain tensor $E: T\mathcal{B} \rightarrow T\mathcal{B}$ is defined by $2E = C - I$. The associated tensor $E^\ast$ is given by $2E^\ast = C^\ast - I^\ast$. 

[Figure 1.3.2]
In components, \( 2E_{AB} = C_{AB} - \delta_{AB} \) and \( 2E_{AB} = C_{AB} - G_{AB} \). Thus \( E = 0 \) is equivalent to \( C = I \), which implies that points in \( \mathcal{B} \) experience no relative motion under \( \phi \). The reason for the factor "2" in the definition of \( E \) will be evident when we consider the linearized theory.

### 3.19 Definition

Let \( \phi \) be regular. The spatial (or Eulerian) strain tensor \( e : T\mathcal{B} \to T\mathcal{B} \) is defined by \( 2e = i - c \), where \( c = b^{-1} : T\mathcal{B} \to T\mathcal{B} \).

The associated mapping and component forms of \( e \) are given by

\[
2e^i = i^i - c^i, \quad 2e_{ab} = g_{ab} - c_{ab}, \quad \text{and} \quad 2e^a_b = \delta^a_b - c^a_b.
\]

As we shall see in Section 1.4, the various tensors defined here can be redefined in terms of pull-backs and push-forwards:

\[
\begin{align*}
C^i &= \phi^*(g) & c^i &= \phi_*(G^i) \\
B^i &= \phi^*(g^i) & b^i &= \phi_*(G^i) \\
E^i &= \phi^*(e) & e^i &= \phi_*(E^i)
\end{align*}
\]

The component forms of these definitions (with \( x = \phi(X) \), as usual) are:

\[
\begin{align*}
C_{AB}(X) &= g_{ab}(x)F^a_A(X)F^b_B(X) & c_{ab}(x) &= G_{AB}(X)(F(X)^{-1})^a_c(F(X)^{-1})^b_b \\
B^{ab}(X) &= g^{ab}(x)(F(X)^{-1})^a_c(F(X)^{-1})^b_b & b^{ab}(x) &= G^{AB}(X)F^a_A(X)F^b_B(X) \\
E^{AB}(X) &= e_{ab}(x)F^a_A(X)F^b_B(X) & e_{ab}(x) &= E_{AB}(X)(F(X)^{-1})^a_c(F(X)^{-1})^b_b.
\end{align*}
\]

The following conditions are equivalent:

\[
\begin{align*}
C &= I & c &= i \\
B &= I & b &= i \\
E &= 0 & e &= 0.
\end{align*}
\]

### Box 3.1 Shifters and Displacements

In the special case of Euclidean space, there are some additional concepts that are sometimes useful.

#### 3.20 Definition

Let \( \mathcal{B} \subset \mathbb{R}^n \) be a simple body and let \( \phi \) be a \( C^1 \) configuration. The map

\[
S : T\mathcal{B} \to T\mathcal{B}, \quad S(X, W) = (\phi(X), W)
\]

is called the shifter. If \( \phi \) is a motion, we write \( S_\phi(X, W) = (\phi(X), W) \), and let \( S_X \) be the restriction of \( S \) to \( T_x\mathcal{B} \) so that \( S_X : T_x\mathcal{B} \to T_x\mathcal{B} \).

Notice that \( S \) parallel transports vectors emanating from \( X \) to vectors emanating from \( \phi(X) \). See Figure 1.3.3.
On general Riemannian manifolds, the notion of shifter still makes sense if a motion is under consideration: $S_x(W_x)$ parallel translates $W_x$ along the curve $\phi_t(X)$. However, shifters have been used primarily in the $\mathbb{R}^n$ context.\textsuperscript{14}

\section*{3.21 Proposition}

(i) Let $\{x^a\}$ and $\{X^A\}$ be coordinate charts on $\mathcal{S}$ and $\mathcal{B}$, respectively. Then the components of $S$ are given by

$$S^a_{\,A}(X) = \frac{\partial x^a}{\partial X^A}(x) \frac{\partial Z^i}{\partial X^A}(X) \delta^i_j$$

(ii) $S$ is orthogonal: $S^T = S^{-1}$.

\textit{Proof}

(i) Since $S(X, W) = (\phi(X), W)$, $S_x$ as a linear transformation is "the identity": $S_x W^A = W^a$. But $W^A = (\partial X^A/\partial z^i) W^i$ and $W^a = (\partial x^a/\partial z^i) W^i$. Substituting these gives the result.

(ii) follows from $\langle S_x W_1, S_x W_2 \rangle_x = \langle W_1, W_2 \rangle_x$ \hfill \Box

Recall that the rotation tensor $R$ rotates as well as moves base points. Using the shifter we can break up $R$ into two parts by defining $R_1 = RS^{-1}$ and $R_2 = S^{-1}R$. Thus we have $R = R_1 S$ (shift and then rotate), and $R = SR_2$ (rotate and then shift). In general, $R_1 \neq R_2$. The diagram that goes with this follows.

\textsuperscript{14}For instance, in shell theory $\mathcal{S} = \mathbb{R}^3$, $\mathcal{B}$ is a surface, and shifters are defined as above by ordinary parallel translation in $\mathbb{R}^3$. 

Figure 1.3.3
3.22 Definition Let $\phi : \mathcal{B} \to \mathcal{S} = \mathbb{R}^n$ be a $C^1$ configuration of $\mathcal{B}$. We consider $\mathbb{R}^n$ a vector space; thus it makes sense to write $\phi(X) - X$. The vector field on $\mathcal{B}$, defined by $U(X) = (X, \phi(X) - X)$, is called the displacement. See Figure 1.3.4.

In the general case in which $\mathcal{S}$ is not a linear space, the displacement is not defined. It may be used only when the Cartesian structure of the ambient space is available.

*Warning*: Transforming $U$ as a vector to new coordinates is not compatible with transforming $\phi(X)$ and $X$ as points to the new components (i.e., the equation $U = x - X$ is not a tensorial equation).

3.23 Proposition Let $\{Z^I\}$ and $\{z^I\}$ denote collinear Cartesian coordinate systems for $\mathbb{R}^n$, and assume that the origins of these frames are connected by a vector $Y = Y^I \hat{I}_I$. Then $U^I(X) = Y^I - Z^I(X) + \delta^I_1 \phi(X)$. See Figure 1.3.5.

*Proof* If $x \in \mathbb{R}^n$, then $Z^I(x) = Y^I + \delta^I_1 z^I(x)$, by definition of the coordinates. Let the vector components of $U$ be denoted $U^I$; then $U^I(X) = \phi^I(X) - Z^I(X)$. Combining these results yields the assertion. 

![Figure 1.3.4](image)
3.24 Proposition The deformation gradient is given by $F = S(I + DU)$. (Here $DU$ is the covariant derivative of $U$; in component form, $U_{A}^{a_{B}}$.)

Proof We work first in Cartesian coordinates. Employing the result of 3.23, we find that

$$F^{i}_{j} = \delta^{i}_{j} \left( \delta^{i}_{j} + \frac{\partial U^{i}}{\partial Z^{j}} \right).$$

Using 3.21(i) and transforming to the coordinate systems $\{X^{A}\}$ and $\{x^{a}\}$ yields $F_{a_{B}}^{a} = S_{a}^{a_{B}}(\delta_{a_{B}}^{a} + U_{1a_{B}})$. 

3.25 Corollary $C = I + DU + (DU)^{T} + (DU)^{T}DU$ and $2E = DU + (DU)^{T} + (DU)^{T}DU$.

If $\phi$ is regular, then a "dual displacement" to $U$ can be defined by $u = SU \circ \phi^{-1}$. See Figure 1.3.6. Analogous to the previous development, the following formulas are easily established:

$$F^{-1} = S^{-1}(i - DU), \quad (DU)^{a_{b}} = u^{a_{b}}$$
$$c = i - DU - (DU)^{T} + (DU)^{T}(DU)$$
$$2e = DU + (DU)^{T} - (DU)^{T}(DU).$$

Figure 1.3.6
Problem 3.4 Show that $e^a = S^a_A E^A$ and $e_a = S_a^A E_A$, where $S^A_a = g_{ab} S^b_B G^{AB}$.

Problem 3.5 Show that the (unique) solution of the equations $\partial\phi_i/\partial t = v_i \circ \phi$, and $D\phi_i/Dt = 0$ with $\phi_0 = I$, $v_0 = V$ is $\phi_t = I + tV$, $v_t = S_t(V)$. Give an explicit example on $\mathbb{R}$ to show that $\phi_t$ can fail to be regular in a finite time.

Problem 3.6 Given a motion $\phi_t$ and $W \in T_x \mathcal{B}$, show that $(\partial/\partial t) S_t(W) = -\nabla_{v_t} S_t(W)$.

Suppose that a motion $\phi_t$ is under consideration. We want to calculate the rate of change of the various deformation tensors.

3.26 Definition Let $\phi_t$ be a $C^1$ motion of $\mathcal{B}$ in $\mathcal{S}$, and let $C$ be the deformation tensor. The material (or Lagrangian) rate of deformation tensor $D$ is defined by

$$2D(X, t) = \frac{\partial}{\partial t} C(X, t).$$

Note that $C(X, t): T_X \mathcal{B} \rightarrow T_X \mathcal{B}$ for each $t$, so it makes sense to differentiate $C(X, t)$ in $t$ to obtain another linear transformation. The reason for the factor “2” will become clear when we study the linearized theory; see Chapter 4.

To compute $D$ in coordinates, we introduce the covariant derivative of $V$. (It will be considered abstractly in terms of covariant differentiation of two-point tensors in Section 1.4.)

3.27 Definition The covariant derivative of the material velocity is defined in coordinates $\{x^a\}$ and $\{X^A\}$ on $\mathcal{S}$ and $\mathcal{B}$, respectively, by

$$V^a_{|\dot{b}} = \frac{\partial V^a}{\partial X^b} + \gamma^a_{bd} V^b(X, t) F^d_{\dot{b}}(X, t).$$

or, putting in the arguments,

$$V^a_{|\dot{b}}(X, t) = \frac{\partial V^a}{\partial X^b}(X, t) + \gamma^a_{bd}(\phi(X, t)) V^b(X, t) F^d_{\dot{b}}(X, t).$$

Problem 3.7 If $\mathcal{B} \subset \mathcal{S} = \mathbb{R}^n$ is a simple body, show that $V^a_{|\dot{b}}$ are simply the matrix entries of the linear map $DV(X)$ relative to the coordinate systems $\{x^a\}$ and $\{X^A\}$. (Imitate the proof of 1.13 and use the chain rule at the appropriate point.) Prove that if $\phi_t$ is regular, $V^a_{|\dot{b}}(x, t) = v^a_{|\dot{b}}(x, t) F^b_{\dot{b}}(X, t)$. 


3.28 Proposition  In coordinates, the following are the components of $D$:

$$2D^A_{\beta}(X) = G^{CA}(X)g_{ac}(x)[V^c_{\beta}(X, t)F^a_B(X, t) + V^a_B(X, t)F^c_c(X, t)].$$

The proof requires an observation about the Christoffel symbols $\gamma^a_{bc}$, which, in fact, was at the historical origins of Riemannian geometry: $\gamma^a_{bc}$ depends only on $g_{ab}$.

3.29 Lemma  We have the identities

$$\frac{\partial g_{ac}}{\partial x^d} = g_{ba}\gamma^b_{dc} + g_{bc}\gamma^b_{da}$$

and

$$2g_{ab}\gamma^a_{dc} = \left(\frac{\partial g_{cb}}{\partial x^d} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{dc}}{\partial x^b}\right).$$

Proof  The first follows at once from the explicit definitions of $g_{ab}$ and $\gamma^a_{bc}$ in 3.4 and 1.8, recalling that $(\partial z'^\beta/\partial x^\alpha)(\partial x^\alpha/\partial z^\beta) = \delta^\beta_\gamma$. The second identity follows by substituting the formula for $g_{ab}$ into its right-hand side, remembering $g_{bc}$ and $\gamma^a_{bc}$ are symmetric in $(b, c)$.

Proof of 3.28  We have $G_{AC}(X)C^A_{\beta}(X, t) = g_{ac}(\phi(X, t))F^c_c(X, t)F^a_B(X, t)$, and so, differentiating in $t$,

$$2G_{AC}D^A_{\beta}(X) = \frac{\partial g_{ac}}{\partial x^d}V^dF^c_cF^a_B + g_{ac}\frac{\partial V^c}{\partial x^c}F^a_B + g_{ae}F^c_c\frac{\partial V^a}{\partial x^b}.$$

Using the first identity from 3.29, this equals (after changing two dummy summation variables to $b$)

$$\left(g_{ab}\gamma^b_{dc}V^dF^c_c + g_{ab}\frac{\partial V^b}{\partial x^c}\right)F^a_B + \left(g_{bc}\gamma^b_{da}V^dF^a_B + g_{bc}\frac{\partial V^b}{\partial x^a}\right)F^c_c.$$

In Section 1.6, we will recognize $D^b$ as $\phi^*(L_vg)$, the pull-back of the Lie derivative of the metric.

3.30 Proposition  Let $\phi$, be a $C^2$ motion of $\mathfrak{g}$. Let $W_1, W_2 \in T_x\mathfrak{g}$ and let $w_i(t) = F(X, t)W_i$ $(i = 1, 2)$. Then

$$\frac{d}{dt}\langle w_1(t), w_2(t)\rangle_{\phi(X, t)} = 2\langle W_1, D(X, t)W_2\rangle_{x}.$$
Proof
\[
\frac{d}{dt} \langle w_1(t), w_2(t) \rangle_{\phi(x,t)} = \frac{d}{dt} \langle W_1, F(X, t)^T F(X, t) W_2 \rangle_x
\]
\[
= \frac{d}{dt} \langle W_1, C(X, t) W_2 \rangle_x = \langle W_1, 2D(X, t) W_2 \rangle_x. \]

This proposition shows that \( D(X, t) \) measures the rate of change of the inner product of vectors as they are pushed forward by the motion.
Combined with 3.28, 3.30 yields:

3.31 Corollary  In coordinates,
\[
\frac{d}{dt} \{g_{ab}(\phi(X, t)) F^a_A(X, t) W^b_B(X, t) W^c_C \}
\]
\[
= W^c_i W^B_j g_{ac}(\phi(X, t)) \{V^{1c} F^a_B + V^{1b} F^c_B \}(X, t).
\]

Problem 3.8  Work out the metric and Christoffel symbols for the spherical coordinate system \((\rho, \theta, \phi)\) on \(\mathbb{R}^3\).

Problem 3.9  Consider the motion of \(\mathbb{R}^2\) in \(\mathbb{R}^3\) given by
\[
\phi(Z^1, Z^2, t) = \begin{bmatrix} 1 & t \kappa^1 \kappa^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z^1 \\ Z^2 \end{bmatrix}
\]
(a steady shearing). Calculate the deformation tensor and the rate of deformation tensor.

<table>
<thead>
<tr>
<th>Box 3.2  Summary of Formulas for Section 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Covariant Derivatives for Simple Bodies</strong> (\mathcal{B}) in (S = \mathbb{R}^3)</td>
</tr>
<tr>
<td>(U: \mathcal{B} \rightarrow T^* S);          (DU: T^* \mathcal{B} \rightarrow T^* S)</td>
</tr>
<tr>
<td>(U^A_{1B} = \frac{\partial U^A}{\partial X^B} + \Gamma^A_{BC} U^C)</td>
</tr>
<tr>
<td>(v: S \rightarrow T S);          (DV: T S \rightarrow T S)</td>
</tr>
<tr>
<td>(v^a_{1b} = \frac{\partial v^a}{\partial x^b} + \gamma^a_{bc} v^c)</td>
</tr>
<tr>
<td>(V: \mathcal{B} \rightarrow T^* \mathcal{B});          (DV: T^* \mathcal{B} \rightarrow T^* \mathcal{B})</td>
</tr>
<tr>
<td>(V^a_{1b} = \frac{\partial V^a}{\partial X^b} + (\gamma^a_{bc} \phi) V^b F^c_B)</td>
</tr>
<tr>
<td>(V = v \circ \phi);          (DV = Dv \circ F)</td>
</tr>
<tr>
<td>(V^a_{1A} = (v^a_{1b} \circ \phi) F^b_A)</td>
</tr>
<tr>
<td><strong>Metric Tensors</strong></td>
</tr>
<tr>
<td>(G: T S \rightarrow T^* S)</td>
</tr>
<tr>
<td>(G^I = G^{-1})</td>
</tr>
<tr>
<td>(g^I = g^{-1})</td>
</tr>
</tbody>
</table>

\[G_{AB} = \frac{\partial Z^I}{\partial X^A} \frac{\partial Z^J}{\partial X^B} \delta_{IJ}\]
\[g_{ab} = \frac{\partial z^a}{\partial X^a} \frac{\partial z^b}{\partial X^b} \delta_{ij}\]
\[(G^{AB}) = (G_{AB})^{-1}\]
\[(g^{ab}) = (g_{ab})^{-1}\]
Deformation Gradient
\[ F: T\mathbb{S} \rightarrow T\mathbb{S}; \]
\[ F(X, W) = (\phi(X), D\phi(X) \cdot W) \]
\[ F^T: T\mathbb{S} \rightarrow T\mathbb{S}; \]
\[ (F^T)^A_a = \frac{\partial \phi^a}{\partial X^A} \]
\[ g_{ab}(F^b \circ \phi^{-1})(\partial^B_a \circ \phi^{-1}) \]

Polar Decomposition (\(U\) and \(V\) are the stretch tensors under this heading)
\[ F = RU = VR \]
\[ F^a_A = R^a_B U^B_A = (V^a_b \circ \phi) R^b_A \]
\[ R: T\mathbb{O} \rightarrow T\mathbb{S} \quad R^{-1} = R^T \]
\[ U: T\mathbb{S} \rightarrow T\mathbb{O} \quad (\text{symmetric positive-definite}) \]
\[ G_{AC} U^C_B = G_{BC} U^C_A \]
\[ V: T\mathbb{S} \rightarrow T\mathbb{S} \quad (\text{symmetric positive-definite}) \]
\[ g_{ac} \frac{\partial \psi}{\partial \xi^c} = g_{bc} \frac{\partial \psi}{\partial \xi^c} \]

Deformation Tensors
\[ C(X, W) = (X, D\phi(X)^T D\phi(X) W) \]
\[ C = F^T F \quad C^A_B = G^{AC}(g_{ab} \circ \phi) F^b_C F^a_B \]
\[ B = C^{-1} \quad (B^a_B) = (C^a_A)^{-1} \]
\[ 2E = C - I \quad 2E^A_B = C^A_B - \delta^A_B \]
\[ b = F F^T \quad b^{a_b} = (G^{AB} \circ \phi^{-1}) g_{bc}(F^a_A \circ \phi^{-1})(F^b_C \circ \phi^{-1}) \]
\[ c = b^{-1} \quad c^{a_b} = (b^{a_b})^{-1} \]
\[ 2e = c - i \quad 2e^{a_b} = c^{a_b} - \delta^{a_b} \]
\[ C^I = \phi^*(g) \quad C_{AB} = (g_{ab} \circ \phi) F^a_A F^b_B \]
\[ B^I = \phi^*(g^I) \quad B^{AB} = (g^{ab} \circ \phi)((F^{-1})^a_A \circ \phi)((F^{-1})^b_B \circ \phi) \]
\[ E^I = \phi^*(e) \quad E_{AB} = (e_{ab} \circ \phi) F^a_A F^b_B \]
\[ c^I = \phi^*(G) \quad c_{ab} = (G^{AB} \circ \phi^{-1})(F^{-1})^a_A (F^{-1})^b_B \]
\[ b^I = \phi^*(G^I) \quad b^{ab} = (G^{AB} \circ \phi^{-1})(F^a_A \circ \phi^{-1})(F^b_B \circ \phi^{-1}) \]
\[ e^I = \phi^*(E^I) \quad e_{ab} = (E^{AB} \circ \phi^{-1})(F^{-1})^a_A (F^{-1})^b_B \]

Shifter in \(\mathbb{R}^n\)
\[ S: T\mathbb{O} \rightarrow T\mathbb{S}; \]
\[ S(X, W) = (\phi(X), W) \]
\[ S^T: T\mathbb{S} \rightarrow T\mathbb{O}; \]
\[ S^{-1} = S^T \]
\[ S^A_a = (\frac{\partial X^a}{\partial z^I}) \delta^I_a \]
\[ (S^T)^A_a = (\frac{\partial X^a}{\partial Z^I}) \delta^I_a \]
\[ (S^{-1})^A_a = (S^T)^A_a \]

Displacements in \(\mathbb{R}^n\)
\[ U: \mathbb{O} \rightarrow T\mathbb{O}; \]
\[ U(X) = (X, \phi(X) - X) \]
\[ F = S(I + DU) \]
\[ C = I + DU + DU^T \]
\[ + DU^T DU \]
\[ 2E = DU + DU^T \]
\[ u = SU \circ \phi^{-1} \]
\[ F^{-1} = S^T (i - Du) \]
1.4 TENSORS, TWO-POINT TENSORS, AND THE COVARIANT DERIVATIVE

The previous section introduced a number of important geometric concepts. This section will examine these objects in further detail from both the abstract and the computational point of view. By this time the reader should be comfortable enough with manifolds so that we can treat the general case with no more effort than the case of open sets in \( \mathbb{R}^3 \).

Our notation throughout is as above: \( \mathcal{B} \) and \( \mathcal{S} \) are manifolds, points in \( \mathcal{B} \) are denoted by \( X \) and those in \( \mathcal{S} \) by \( x \). The tangent spaces are written \( T_X \mathcal{B} \) and \( T_x \mathcal{S} \). Coordinate systems are denoted \( \{X^i\} \) and \( \{x^a\} \) for \( \mathcal{B} \) and \( \mathcal{S} \), respectively, with corresponding bases \( E_A \) and \( e_a \) and dual bases \( E^A \) and \( e^a \). The summation convention is enforced.

4.1 Definition  A tensor of type \( \left( \frac{p}{q} \right) \) at \( X \in \mathcal{B} \) is a multilinear mapping
\[
\mathbf{T}: \underbrace{T^*_X \mathcal{B} \times \cdots \times T^*_X \mathcal{B}}_p \times \underbrace{T_X \mathcal{B} \times \cdots \times T_X \mathcal{B}}_q \rightarrow \mathbb{R}
\]
(multilinear means linear in each variable separately). The components of \( \mathbf{T} \) are defined by
\[
\mathbf{T}^{\alpha_1 \cdots \alpha_p \beta_1 \cdots \beta_q} = \mathbf{T}(E_A^{ \alpha_1}, E_A^{ \alpha_2}, \ldots, E_A^{ \alpha_p}, E_B^{ \beta_1}, E_B^{ \beta_2}, \ldots, E_B^{ \beta_q})
\]
so that
\[
\mathbf{T}(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q) = \mathbf{T}^{\alpha_1 \cdots \alpha_p \beta_1 \cdots \beta_q} \alpha_1 \cdots \alpha_p \beta_1 \cdots \beta_q
\]
where \( \alpha^i \in T^*_X \mathcal{B} \), \( \alpha^i = E_A^{ \alpha_i} \), \( W_j \in T_X \mathcal{B} \), and \( W_j = W_j^A E_A \). We say that \( \mathbf{T} \) is contravariant of rank \( p \) (\( p \) indices up) and covariant of rank \( q \) (\( q \) indices down).

A tensor field of type \( \left( \frac{p}{q} \right) \) is an assignment \( \mathbf{T}(X) \) of a tensor of type \( \left( \frac{p}{q} \right) \)
at $X$ for each $X \in \mathcal{B}$. The components are functions of $X$ and are denoted

$$\tau^{A_1\ldots A_{n-1} A_n B_1 \ldots B_n}(X).$$

We can make the space of all tensors on $\mathcal{B}$ into a vector bundle $T^p_q(\mathcal{B})$ over $\mathcal{B}$ (as with the tangent bundle) and a tensor field $T$ may be regarded as a section of this bundle; that is, $T$ assigns to each base point $X \in \mathcal{B}$ an element of the fiber of $T^p_q(\mathcal{B})$ over $X$.

A tensor field is said to be of class $C^r$ if its components are $C^r$ functions of $X^d$ in any coordinate system.

We can regard vectors as $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ tensors, $15$ one-forms as $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ tensors, and functions as $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ tensors.

We can, by a similar process, position the indices in other slots. For instance, a fourth-order tensor that is a multilinear mapping

$$T: T^p_q(\mathcal{B}) \times T^p_q(\mathcal{B}) \times T^p_q(\mathcal{B}) \times T^p_q(\mathcal{B}) \to \mathbb{R}$$

has indices positioned like this:

$$\tau^{A_1 \cdots A_n B_1 \cdots B_n}. \quad \star$$

A tensor or tensor field $T$ has a separate existence from its components $\tau^{A_1 \cdots A_n B_1 \cdots B_n}$, and it makes sense to write down $T$ without any indices, just as for vectors.

4.2 Proposition  Let $\{X^d\}$ and $\{\bar{X}^d\}$ be two coordinate systems on $\mathcal{B}$ and $T$ a tensor field of type $\left( \begin{smallmatrix} p \\ q \end{smallmatrix} \right)$. Then the components of $T$ in these two systems are related by

$$\tau^{A_1 \cdots A_n B_1 \cdots B_n} = \frac{\partial \bar{X}^{A_1}}{\partial X^{C_1}} \cdots \frac{\partial \bar{X}^{A_n}}{\partial X^{C_n}} \tau^{C_1 \cdots C_n D_1 \cdots D_n} \frac{\partial X^{D_1}}{\partial \bar{X}^{B_1}} \cdots \frac{\partial X^{D_n}}{\partial \bar{X}^{B_n}}.$$

Proof  This follows from the definition of components, multilinearity, and the formulas

$$E^A = \frac{\partial \bar{X}^A}{\partial X^B} E^B, \quad E_A = \frac{\partial X^B}{\partial \bar{X}^A} E_B. \bigstar$$

4.3 Definition  Let $T$ be a tensor of type $\left( \begin{smallmatrix} p \\ q \end{smallmatrix} \right)$ and $S$ a tensor of type $\left( \begin{smallmatrix} r \\ s \end{smallmatrix} \right)$. Then the tensor product $T \otimes S$ is the tensor of type $\left( \begin{smallmatrix} p + r \\ q + s \end{smallmatrix} \right)$ defined by

$$p \text{ copies} \quad q \text{ copies}$$

$$\tau \otimes S)(X): (T^p_q(\mathcal{B}) \times \cdots \times T^p_q(\mathcal{B})) \times (T^p_q(\mathcal{B}) \times \cdots \times T^p_q(\mathcal{B}))$$

$$r \text{ copies} \quad s \text{ copies}$$

$$\times (T^p_q(\mathcal{B}) \times \cdots \times T^p_q(\mathcal{B})) \times (T^p_q(\mathcal{B}) \times \cdots \times T^p_q(\mathcal{B})) \to \mathbb{R};$$

---

15 If $W$ is a vector at $X$, we set $W(\alpha) = \alpha(W)$ for $\alpha \in T^p_q(\mathcal{B})$. 

---
\[(T \otimes S)(X)(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q, \beta^1, \ldots, \beta^r, Y_1, \ldots, Y_s) = T(X)(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q) \cdot S(X)(\beta^1, \ldots, \beta^r, Y_1, \ldots, Y_s)\]

Thus
\[(T \otimes S)^{A_1 \cdots A_p B_1 \cdots B_q C_1 \cdots C_r D_1 \cdots D_s}(X) = T^{A_1 \cdots A_p B_1 \cdots B_q}(X) \cdot S^{C_1 \cdots C_r D_1 \cdots D_s}(X)\]

Addition and scalar multiplication of tensors are defined in the obvious way.

**Problem 4.1** In a coordinate system \(\{X^A\}\) on \(\mathfrak{B}\), show that the matrix of components of \(E_A \otimes E_B\) is a matrix of 0’s with the exception of a 1 in the \((A, B)\) slot.

The following is easily proven from the definitions.

**4.4 Proposition** The following holds for a tensor \(T\) of type \(\left(\begin{array}{c} p \\ q \end{array}\right)\):
\[T(X) = T^{A_1 \cdots A_p B_1 \cdots B_q}(X)E_{A_1} \otimes \cdots \otimes E_{A_p} \otimes E^{B_1} \otimes \cdots \otimes E^{B_q}\]

Next, we define contractions of tensors.

**4.5 Definition** The *contraction* of a one-form \(\alpha\) and a vector field \(W\) is defined by
\[\alpha \cdot W(X) = \alpha(X)(W(X)) = \alpha^A(X)W_A(X)\]

If \(T\) and \(S\) are tensors, the contraction of the \(i\)th covariant index of \(T\) and \(j\)th contravariant index of \(S\) is defined by fixing all the other slots (or indices) and contracting \(T\) in the \(i\)th covariant index as a one-form with the \(j\)th contravariant index of \(S\) as a vector field.

The operation of contracting the \(i\)th contravariant index of a tensor \(T\) with the \(j\)th covariant index of the same tensor is defined in a similar fashion. The contraction of a \(\left(\begin{array}{c} 1 \\ r \end{array}\right)\) tensor is called its *trace*. (Thus to contract two indices, one upper and one lower, we merely set the indices equal and sum.)

When two indices are simultaneously contracted, we shall use two dots. For example, if \(R\) and \(S\) are tensors of types \(\left(\begin{array}{c} 2 \\ 0 \end{array}\right)\) and \(\left(\begin{array}{c} 0 \\ 2 \end{array}\right)\), respectively, \(R \cdot S\) is the scalar defined by \(R^{AB}S_{AB}\).

The ideas of pull-back and push-forward, defined in Section 1.2 for one forms and vector fields, extend easily to tensors as follows.

**4.6 Definition** Let \(\phi: \mathfrak{B} \rightarrow \mathfrak{S}\) be a regular mapping. If \(T\) is a tensor of type \(\left(\begin{array}{c} p \\ q \end{array}\right)\), its *push-forward* \(\phi_*T\) is a tensor of type \(\left(\begin{array}{c} p \\ q \end{array}\right)\) on \(\phi(\mathfrak{B})\) defined by
\[(\phi_*T)(x)(\alpha^1, \ldots, \alpha^p, v_1, \ldots, v_q) = T(X)(\phi^*(\alpha^1), \ldots, \phi^*(\alpha^p), \phi^*(v_1), \ldots, \phi^*(v_q))\]
where $\alpha^i \in T^*_x S$, $v_i \in T_x S$, $X = \varphi^{-1}(x)$, $\phi^*(\alpha^i) \cdot v = \alpha^i \cdot (T\phi \cdot v)$ and $\phi^*(v_i) = T(\varphi^{-1}) \cdot v_i$. The pull-back of a tensor $t$ defined on $\phi(\mathcal{B})$ is given by $\phi^* t = (\phi^{-1})_* t$.

In coordinates we have the following relations which result from the definitions and the corresponding formulas for one-forms and vector fields (see 2.9 and 2.16).

**4.7 Proposition** Letting $F(X) = T_X \phi$ and $x = \phi(X)$, we have

$$(\phi_* T)^{a_1 \cdots a_b} \cdots \cdots \cdots b_k(x) = F^{a_1} A_{a_1}(X) \cdots F^{a_p} A_{a_p}(X) \cdot T^{a_1 \cdots a_p B_1 \cdots b_k}(X) \cdot (F^{-1})^{B_1} b_1(x) \cdots (F^{-1})^{B_k} b_k(x)$$

and

$$(\phi^* t)^{a_1 \cdots a_b} \cdots \cdots \cdots \cdots b_k(x) = (F^{-1})^{a_1} a_1(X) \cdots (F^{-1})^{a_p} a_p(X) \cdot (F^{-1})^{b_1} b_1(x) \cdots F^{b_k} b_k(x).$$

Before going on to two-point tensors, it is useful to consider first some of the extra structure that enters when we have a Riemannian metric.

**4.8 Definition** A Riemannian metric on $\mathcal{B}$ is a $C^\infty$ covariant two-tensor $G$ (i.e., $G$ is a tensor of type $(0, 2)$) such that for each $X \in \mathcal{B}$:

(i) $G(X)$ is symmetric; that is, for $W_1, W_2 \in T_x \mathcal{B}$, $G(X)(W_1, W_2) = G(X)(W_2, W_1)$.

(ii) $G(X)$ is positive-definite: $G(X)(W, W) > 0$ for $0 \neq W \in T_x \mathcal{B}$; in other words, $G(X)$ is an inner product on $T_x \mathcal{B}$. If there is no danger of confusion, we often write

$$G(X)(W_1, W_2) = \langle W_1, W_2 \rangle_X.$$  

(A Riemannian metric on $\mathcal{S}$ will be denoted $g$.)

Notice that symmetry of $G$ means $G_{AB} = G_{BA}$ and that $\langle W_1, W_2 \rangle_X = G_{AB}(X)W^A W^B$. Condition (ii) states that the matrix $G_{AB}$ is positive-definite.

**Warning:** In classical tensor analysis, $G$ (or $g$ if we use $\mathcal{S}$) is often used to denote $\sqrt{\det[G_{AB}]}$ (or $\sqrt{\det[g_{ab}]}$).

Next we describe associated tensors (i.e., new tensors obtained by raising and lowering indices) in greater detail than was considered in Section 1.3.

**4.9 Definition** For $X \in \mathcal{B}$, define the linear transformation $G^i(X): T_x \mathcal{B} \rightarrow T^*_x \mathcal{B}$ by $G^i(X)(W_1) = \alpha(X)$, where $\alpha(X)(W_2) = \langle W_1, W_2 \rangle_X$; that is, $G^i(X)(W_1) = \langle W_1, \cdot \rangle_X$. Since $G(X)$ is positive-definite $G^i(X)$ is invertible; its inverse is denoted $G^i(X): T^*_x \mathcal{B} \rightarrow T_x \mathcal{B}$.

---

16Regularity is not needed to pull back the covariant components. Pull-back of contravariant components and push-forward of both covariant and contravariant components require regularity. See Box 2.2.
4.10 Proposition \( G'(X) \cdot E^A = G^{AB}(X) \cdot E_B \), where \( G^{AB}(X) \) is the inverse matrix of \( G_{AB}(X) \).

Proof \( G'(X)(G^{AB}(X) \cdot E_B)(E_C) = G^{AB}(X)(E_B, E_C) = G^{AB}(X)G_{BC}(X) = \delta^A_C = E^A(E_C) \). Thus \( G'(X)(G^{AB}(X) \cdot E_B) = E^A \), so the result holds. \( \blacksquare \)

We may regard \( G' \) as a \( \begin{pmatrix} 2 \\ 0 \end{pmatrix} \) tensor on \( \mathfrak{B} \); 4.10 states that its components are \( G^{AB} \).

4.11 Definition Let \( T \) be a tensor of type \( \left( \begin{array}{c} p \\ q \end{array} \right) \) on \( \mathfrak{B} \). The \( 2^{p+q} \) associated tensors are defined by applying \( G' \) or \( G \) to each slot. For instance, lowering the first index on \( T \) gives the tensor \( T_1 \) defined by

\[
T_1: T_x(\mathfrak{B} \times T_{x_1}^p \mathfrak{B} \times \cdots \times T_{x_{p-1}}^p \mathfrak{B} \times T_{x_q}^q \mathfrak{B} \times \cdots \times T_{x}^q \mathfrak{B}) \to \mathbb{R}
\]

\[
T_1(Y, \alpha_2, \ldots, \alpha_p, W_1, \ldots, W_q) = T(G' \cdot Y, \alpha_2, \ldots, \alpha_p, W_1, \ldots, W_q).
\]

That is,

\[
(T_1)_{A_1A_2\cdots A_pB_1\cdots B_q} = G_{A_1C}T^{C A_2 \cdots A_p B_1\cdots B_q}.
\]

These indices are denoted by \( T_{A_1A_2\cdots A_pB_1\cdots B_q} \) (although they are the components of \( T_1 \), which does not equal \( T \)). Similarly, the components of the tensor \( T_2 \) obtained by raising the first covariant index of \( T \) are denoted

\[
T^{A_1\cdots A_pB_1\cdots B_q} = G^{BC}T^{A_1\cdots A_pCB_1\cdots B_q}.
\]

Invariantly,

\[
T_2(\alpha_1, \ldots, \alpha_p, \beta_1, W_2, \ldots, W_q) = T(\alpha_1, \ldots, \alpha_p, G' \cdot \beta_1, W_2, \ldots, W_q).
\]

If \( F \) is a real-valued function, the vector field \( (dF)' = \nabla F \) is called the gradient of \( F \); that is,

\[
\nabla F = G^{AB} \frac{\partial F}{\partial X^B} E_A \quad \text{or} \quad (\nabla F)^A = G^{AB} \frac{\partial F}{\partial X^B}.
\]

4.12 Definition If \( T \) and \( S \) are tensors of the same type (or same total order), their inner product is defined by

\[
\langle T, S \rangle = \text{contraction of } T' \text{ and } S' \text{ on all indices}
\]

\[
= T^{A_1\cdots A_pB_1\cdots B_q}S_{A_1\cdots A_pB_1\cdots B_q}.
\]

For instance, suppose \( T \) is a \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensor and \( S \) is a \( \left( \begin{array}{c} 2 \\ 0 \end{array} \right) \) tensor. Then \( T \cdot S \) is usually used to denote the contraction \( T_{AB}S^{BC} \) and \( T : S \) denotes the double contraction; thus \( T : S = \langle T, S \rangle \) is the trace of \( T \cdot S \).
Warning: Raising and lowering indices may not, in general, be interchanged with pull-back or push-forward. Thus, \((\phi^*t)^i \neq \phi^*(t^i)\).

**Problem 4.2** Show that \((\phi^*t)^i\) does coincide with \(\phi^*(t^i)\), provided the first \# is taken with respect to \(\phi^*g\) rather than \(G\).

**4.13 Proposition** The deformation tensor \(C\) is given by

\[
C^i = \phi^*g_i,
\]

where \(\phi: \mathcal{B} \to S\) is a configuration and \(g\) is the metric tensor on \(S\).

**Proof** From Proposition 3.6 we have

\[
(C^i)_{AB}(X) = C_{AB}(X) = g_{ab}(x)F^a_A(X)F^b_B(X),
\]

but from 4.7 these are exactly the components of \(\phi^*g\).

Note that \(\phi\) need not be regular for \(\phi^*g\) and hence for \(C\) to be defined. As an exercise we ask the reader to show, using Proposition 3.9, that

\[
b^i = \phi_*G^i \quad \text{and} \quad B^i = \phi^*g^i,
\]

but we again warn that

\[
b^i \neq \phi_*G = c^i \quad \text{and} \quad B^i \neq \phi^*g = C^i.
\]

Next we shall discuss two-point tensors. These objects play an important role in continuum mechanics; a prime example is the deformation tensor \(F^a_A\).

**4.14 Definition** A two-point tensor \(T\) of type \(\left(\begin{array}{c} q \\ p \\ m \end{array}\right)\) at \(X \in \mathcal{B}\) over a mapping \(\phi: \mathcal{B} \to S\) is a multilinear mapping

\[
T: (T_x^a \mathcal{B} \times \cdots \times T_x^a \mathcal{B}) \times (T_x \mathcal{B} \times \cdots \times T_x \mathcal{B}) \times (T_x^a S \times \cdots \times T_x^a S) \times (T_x S \times \cdots \times T_x S) \to \mathbb{R},
\]

where \(x = \phi(X)\).

The components of \(T\) are defined by

\[
T^{A_1, \ldots, A_q, B_1, \ldots, B_p, a_1, \ldots, a_q, b_1, \ldots, b_p} = T(E^{A_1}, \ldots, E^{A_q}, E_{B_1}, \ldots, E_{B_p}, e^{a_1}, \ldots, e^{a_q}, e_{b_1}, \ldots, e_{b_p})
\]

relative to coordinates \([X^a]\) on \(\mathcal{B}\) and \([x^a]\) on \(S\).

A two-point tensor field \(T\) assigns a two-point tensor \(T(X)\) to each point \(X \in \mathcal{B}\), as with ordinary tensor fields.

---

17 A main reference for two-point tensors is Ericksen [1960].

18 In manifold language, a two-point tensor is a section of the bundle \(T^p_\phi(\mathcal{B}) \otimes T^m_\phi(\phi^*(T_S))\) over \(\mathcal{B}\). Here \(\phi^*(T_S)\) denotes the pull-back bundle whose fiber over \(X \in \mathcal{B}\) is \(T^p_\phi(x)\).
The positioning of the indices can, of course, be altered just as with ordinary tensors. For example, a two-point tensor with indices $T^a_{\cdot b}$ is a multilinear mapping $T: T^*_a S \times T_X \mathfrak{S} \times T_x S \rightarrow \mathbb{R}$.

The deformation gradient $F^a_A$ may be regarded as a two-point tensor as follows:

$$F(X): T^*_a S \times T_X \mathfrak{S} \rightarrow \mathbb{R}, \quad (\alpha, W) \mapsto \alpha(T_x \phi \cdot W).$$

The components of $F$ are, of course, just $F^a_A = \partial \phi^a / \partial X^A$.

Two-point tensors are natural generalizations of vector fields and one forms over maps. Note that a two-point tensor is regarded as a function of $X \in \mathfrak{S}$, and not of $x \in S$. One can think of a two-point tensor as having two tensor "legs," one in $\mathfrak{S}$ and one in $S$. We leave it for the reader to write down the general coordinate transformation rule for two-point tensors.

The operations on tensors generalize naturally to operations on two-point tensors. These include: raising and lowering indices, tensor products, push-forward, pull-back, tilt, and contraction. For instance, suppose we have a two-point tensor $T^a_{\cdot b}$ and wish to pull back the first contravariant index $a$ from $\mathfrak{S}$ to $\mathfrak{B}$. Doing this gives a new two-point tensor $\tilde{T}$ whose components are denoted $\tilde{T}^B_{\cdot A}$ and is defined in the following ways.

Abstractly:

$$\tilde{T}(X): T^*_B \mathfrak{B} \times T_X \mathfrak{S} \times T_x S \rightarrow \mathbb{R}, \quad \tilde{T}(X)(\alpha, W, v) = T(X)(\phi \cdot \alpha, W, v),$$

In components:

$$\tilde{T}^B_{\cdot A}(X) = T^a_{\cdot b}(X)(F^{-1})^a_{\cdot b}(x) = T^a_{\cdot b}(X) \frac{\partial X^B}{\partial x^a}.$$

Notice that raising and lowering indices on the $S$ leg is done using $g$ and on the $\mathfrak{S}$ leg using $G$. For instance,

$$T_{a\cdot b}(X) = g_{ac}(x) T^c_{\cdot b}(X) \quad \text{and} \quad T^{a\cdot c}(X) = G^{AB}(X) T^a_{\cdot b}(X)$$

A number of fundamental two-point tensors will come up in our later work, so it is essential that the reader become familiar with them. (The first Piola–Kirchhoff stress tensor $P$ is an example.)

**Problem 4.3** Consider the motion of Problem 1.1. Calculate the components of the two-point tensors $F \otimes V$ and $F \otimes v$ in the given coordinates.

**Problem 4.4** (a) Let $A^a_{\cdot b}$ be a given fourth-order two-point tensor. Write down a formula for its pull-back to $\mathfrak{S}$.

(b) What is the pull-back of the fourth-order tensor $F^T \otimes F$ to $\mathfrak{S}$? Of the second-order tensor $F^T \cdot F$ (contraction on one pair of indices)? What are the associated components of $F^T \cdot F$?

Our next task is to consider some further aspects of Riemannian geometry. We want to generalize covariant differentiation from vector fields on $\mathbb{R}^n$ (Section 1.1) to general tensor fields and two-point tensor fields on Riemannian mani-
folds. We shall develop what is needed briefly and concisely since there are detailed treatments available in standard text books.

4.15 Definition Let \( W_1 \) and \( W_2 \) be vector fields on a manifold \( \mathcal{M} \). The Lie bracket of \( W_1 \) and \( W_2 \) is defined (as a derivation) by means of the commutator:

\[
[W_1, W_2][f] = W_1[W_2[f]] - W_2[W_1[f]].
\]

Recall that \( W[f] = W^A(\partial f/\partial x^A) \), the derivative of \( f \) in direction \( W \). From this and the definition, one sees that

\[
[W_1, W_2]^A = W_2^B \frac{\partial W_1^A}{\partial x^B} - W_1^B \frac{\partial W_2^A}{\partial x^B}.
\]

That is, in Euclidean notation, \((W_1 \cdot \nabla)W_2 - (W_2 \cdot \nabla)W_1\).

**Problem 4.5** Verify Jacobi's identity: \([[[X, Y], Z], Y] + [[[Z, X], Y], X] = 0\).

Before defining the covariant derivative, it is useful to have the abstract notion of a connection (or covariant derivative) at hand. We shall link it with the metric shortly.

4.16 Definition A connection on a manifold \( \mathcal{M} \) is an operation \( \nabla \) that associates to each pair of vector fields \( W, Y \) on \( \mathcal{M} \) a third vector field denoted \( \nabla_W Y \) and called the covariant derivative of \( Y \) along \( W \), such that:

(i) \( \nabla_W Y \) is linear in each of \( W \) and \( Y \);
(ii) \( \nabla_{fW} Y = f \nabla_W Y \) for scalar functions \( f \); and
(iii) \( \nabla_W(fY) = f \nabla_W Y + (W[f])Y \).

These conditions are reasonable requirements for an operation that is supposed to differentiate \( Y \) in the direction \( W \). Note that (iii) is analogous to the product rule for differentiation.

4.17 Definition The Christoffel symbols of a connection \( \nabla \) on \( \mathcal{M} \) are defined on a coordinate system \( \{x^A\} \) by \( \Gamma^A_{BC}(X)E_A(X) = (\nabla_{E_B}E_C)(X) \). (Consistent with our previous conventions, we denote the Christoffel symbols on \( \mathcal{B} \) by \( \Gamma^A_{BC} \) and those on \( \mathcal{S} \) by \( \gamma^A_{BC} \).

**Problem 4.6** If \( \{\tilde{x}^A\} \) is another coordinate system on \( \mathcal{M} \), show that

\[
\Gamma^A_{BC} = \frac{\partial \tilde{x}^A}{\partial x^D} \Gamma^D_{EF} \frac{\partial x^E}{\partial \tilde{x}^C} \frac{\partial x^F}{\partial \tilde{x}^B} + \frac{\partial \tilde{x}^A}{\partial x^D} \frac{\partial^2 x^D}{\partial \tilde{x}^B \partial \tilde{x}^C}.
\]

Conclude that \( \Gamma^A_{BC} \) are not the components of a tensor.
4.18 Proposition  In coordinates, we have
\[(\nabla_w Y)^A = \frac{\partial Y^A}{\partial x^B} w^B + \Gamma^A_{BC} w^B y^C.\]

Proof  By the properties (i)-(iii) of a connection:
\[
\nabla_w Y = \nabla_{w(x)} (Y^C E_C) = w^B \nabla_{E_B} (Y^C E_C) = w^B (Y^C \nabla_{E_B} E_C) + E_B [Y^A] E_A = \Gamma^A_{BC} w^B Y^C E_A + w^B \frac{\partial Y^A}{\partial x^B} E_A.  \]

From this proposition, observe that \( \nabla_w Y(X) \) depends only on the point value of \( W \) at \( X \), and not on its derivatives. Thus if \( \sigma \) is a curve, \( \nabla_{\sigma} Y \) makes sense. In fact, this is an important object that we now identify.

4.19 Definition  We call \( \nabla_{\sigma} Y \) the material derivative of \( Y \) along \( \sigma \), and write it \( D Y/dt \), or if \( \sigma(t) = \phi_X(t) \) [where \( \phi(X, t) \) is a given motion], by \( \dot{Y} \).

Using coordinates, 4.18, and the chain rule,
\[
\left( \frac{dY}{dt} \right)^A = \frac{d}{dt} [Y^A(\sigma(t))] + \Gamma^A_{BC}(\sigma(t)) Y^B(\sigma(t)) \partial C. \]

If \( Y \) is time dependent, set \( D Y/dt = \partial Y/\partial t + \nabla_{\sigma} Y \). In particular, for a motion \( \phi_t \) of \( \mathcal{S} \) in \( \mathcal{S} \) we see that the acceleration is
\[
a = \frac{\partial v}{\partial t} + \nabla_{\sigma} v = \dot{v}. \]

Another consequence of the fact that \( \nabla_w Y(X) \) depends only on the point value of \( W \) at \( X \) is that \( \nabla Y \) defines a \((\tilde{\sim})\) tensor with components \( (\nabla Y)^{A}_{B} = a^{A}_{B}/a^{B}_{C} + r_{zC} Y_{C}. \)

There are several ways of defining the covariant derivative of a tensor. One may proceed directly (see, Abraham and Marsden [1978]), using the fact that we desire \( \nabla(T \otimes S) = \nabla T \otimes S + T \otimes \nabla S \), or one may use parallel translation. Since we shall need this concept, we choose the latter.

4.20 Definition  A vector field \( Y \) defined along a curve \( \sigma(t) \) is called parallel if \( \nabla_{\sigma} Y = 0 \). A vector \( Y_0 \) at \( \sigma(t_0) \) is called parallel transported or translated along \( \sigma \) if it is extended to a parallel vector field \( Y \) along \( \sigma \).

A vector can be parallel transported uniquely since, given \( \sigma(t) \), \( D Y/dt = 0 \) is a linear differential equation for \( Y(\sigma(t)) \), which has a unique solution with initial data \( Y_0 \) (as long as \( \sigma \) is, say, continuous).

The intuitive idea is that \( \nabla \) reflects the local geometry in such a way that an observer carrying an arrow along a curve in what seems to be a parallel manner,
is, in fact, parallel transporting the arrow. On a sphere, for instance, a parallel transported vector is not parallel transported in space. See Figure 1.4.1.

The following proposition demonstrates a basic link between parallel transport and the covariant derivative.

4.21 Proposition  Let $\sigma(t)$ be a given curve in $\mathcal{M}$, and $Y$ a vector field defined along $\sigma$. Let $S_{t,s} : T_{\sigma(s)}\mathcal{M} \rightarrow T_{\sigma(t)}\mathcal{M}$ denote parallel translation from $\sigma(s)$ to $\sigma(t)$ along $\sigma$. (We also call $S_{t,s}$ the shifter.) Then

$$
\nabla_{\sigma(t)}Y(\sigma(t)) = \left. \frac{d}{ds} (S_{t,s} Y(\sigma(s))) \right|_{s=t}.
$$

(Note that $S_{t,s} Y(\sigma(s)) \in T_{\sigma(t)}\mathcal{M}$ for all $s$.)

Proof  By construction, $(S_{s,t} Y_0)^A = (S_{s,s})^A_B Y_0^B = Y_t^A$ satisfies

$$
\frac{d}{ds} (S_{s,t})^A_B Y_0^B + \Gamma^A_{BC} Y_0^C = 0.
$$

Thus, since $S_{s,s}$ is the identity at $s = t_0$,

$$
\left. \frac{d}{ds} (S_{s,t})^A_B \right|_{s=t_0} = -\Gamma^A_{BC} Y^C.
$$

From uniqueness of solution of differential equations, $S_{s,s} \circ S_{t,s} = S_{t,s}$, and so $S_{t,s} = S_{s,t}^{-1}$. Thus

$$
\left. \left( \frac{d}{ds} S_{s,t} Y(\sigma(s)) \right|_{s=t} \right)^A = \left( \left. \frac{d}{ds} S_{s,s}^{-1} Y(\sigma(s)) \right|_{s=t} \right)^A
$$

$$
= -\left. \frac{d}{ds} (S_{s,t})^A_B \right|_{s=t} Y^B(\sigma(t)) + \frac{d}{ds} Y^A(\sigma(s)) \left|_{s=t} \right.
$$

$$
= \Gamma^A_{BC} Y^C + \frac{d}{dt} Y^A(\sigma(t)) = (\nabla_{\sigma(t)} Y(\sigma(t)))^A.
$$

4.22 Definition  A connection is called torsion free if

$$
\nabla_W Y - \nabla_Y W = [W, Y]
$$
for all vector fields $W$ and $Y$. The torsion of a connection is defined by
\[ \text{Tor}(W, Y) = \nabla_W Y - \nabla_Y W - [W, Y] \]
and is a $\left( \frac{1}{2} \right)$ tensor; thus a connection is torsion free when its torsion tensor is zero.

**4.23 Proposition** A connection is torsion free if and only if $\Gamma^A_{BC} = \Gamma^A_{CB}$—that is, the Christoffel symbols are symmetric.

**Proof** Write out $\nabla_W Y - \nabla_Y W - [W, Y]$ in coordinates, using the earlier coordinate expressions to see that $(\text{Tor})^A_{BC} = \Gamma^A_{BC} - \Gamma^A_{CB}$. $lacksquare$

Now we are ready to find the connection associated with a given metric $G$ on $\mathcal{M}$. The following is sometimes called the “Fundamental Theorem of Riemannian Geometry,” and is based on the early pioneering work of Levi-Civita.

**4.24 Theorem** Let $\mathcal{M}$ be a manifold with a Riemannian metric $G$. Then there is a unique connection on $\mathcal{M}$ that is torsion free and for which parallel translation preserves inner products (that is, $S_{is}$ is an orthogonal transformation).

**Proof** First of all, suppose that such a connection exists. Let $W_1(t)$ and $W_2(t)$ be parallel vector fields along a curve $\sigma(t)$. Therefore, by hypothesis,
\[
0 = \frac{d}{dt} \langle W_1(t), W_2(t) \rangle = \frac{d}{dt} G_{AC}(\sigma(t)) W_1^A(\sigma(t)) W_2^C(\sigma(t))
\]
\[
= \frac{\partial G_{AC}}{\partial X^D} \sigma^D W_1^A W_2^C - G_{AC} \Gamma^{A}_{BD} W_1^B \sigma^D W_2^C - G_{AC} \Gamma^{C}_{BD} W_1^A \sigma^D W_2^B
\]
since
\[
\frac{d}{dt} W_i^A(t) + \Gamma^{A}_{BC} W_i^B \sigma^C = 0,
\]
that is, $W_i$ is parallel along $\sigma(t)$. If we take the particular case in which $W_1(0) = E_A$ and $W_2(0) = E_C$ and evaluate at $t = 0$, we get
\[
0 = \frac{\partial G_{AC}}{\partial X^D} \sigma^D - G_{BC} \Gamma^{A}_{BD} \sigma^D - G_{BA} \Gamma^{B}_{CD} \sigma^D.
\]
Since $\sigma^D$ is arbitrary,
\[
\frac{\partial G_{AC}}{\partial X^D} = G_{BC} \Gamma^{A}_{BD} + G_{BA} \Gamma^{B}_{CD}.
\]
Note that this is the same as the identity obtained in Euclidean space (see Lemma 3.29). By the same algebra, remembering symmetry of $G_{AB}$ and $\Gamma^{A}_{BC}$, we get
\[
2G_{AB} \Gamma^{A}_{DC} = \left( \frac{\partial G_{CB}}{\partial X^D} + \frac{\partial G_{DB}}{\partial X^C} - \frac{\partial G_{DC}}{\partial X^B} \right).
\]
This formula shows that the Christoffel symbols $\Gamma^{d}_{bc}$ are uniquely determined. Hence the connection is unique.

For the converse, we can define the Christoffel symbols in any coordinate chart by the preceding formula. Clearly they are symmetric. Then we define $\nabla_{w}Y$ in terms of the $\Gamma^{d}_{bc}$ and check that it is a well-defined vector field (transforms like a vector). Finally, reversing the above computations shows that parallel translation preserves the inner product. 

**Problem 4.7** Show that the coordinate-free analogs of the identities in the above proof are

$$\frac{d}{dt}\langle W_{1}(t), W_{2}(t)\rangle = \left\langle \frac{DW_{1}}{dt}, W_{2}\right\rangle + \left\langle W_{1}, \frac{DW_{2}}{dt}\right\rangle$$

and

$$2\langle W_{1}, \nabla_{w_{1}}W_{2}\rangle = W_{3}[\langle W_{1}, W_{2}\rangle] + \langle W_{3}, [W_{1}, W_{2}]\rangle$$

$$+ W_{2}[\langle W_{3}, W_{1}\rangle] - \langle W_{2}, [W_{3}, W_{1}]\rangle$$

$$- W_{1}[\langle W_{2}, W_{3}\rangle] - \langle W_{1}, [W_{2}, W_{3}]\rangle.$$ 

Given a Riemannian manifold, we shall use the connection above (called the Riemannian or Levi-Civita connection) unless special exception is made.

**Problem 4.8**

(a) Work out $\Gamma^{d}_{bc}$ for $M$ the unit sphere $S^{2}$ in $\mathbb{R}^{3}$ with respect to spherical coordinates.

(b) A geodesic is a curve $\sigma(t)$ such that $\dot{\sigma}$ is parallel along $\sigma$. Show that the geodesics on $S^{2}$ are great circles.

(c) Describe parallel translation of a vector around a circle of constant latitude on $S^{2}$ relative to (i) the Riemannian manifold $\mathbb{R}^{3}$ and (ii) the Riemannian manifold $S^{2}$.

4.25 Definition Let $S_{t,s}$ be parallel translation along a curve $\sigma(t)$. Extend $S_{t,s}$ to all tensor fields along $\sigma(t)$ in the same manner as push-forward was extended. Explicitly: If $\alpha$ is a one-form at $\sigma(s)$, let

$$(S_{t,s}\alpha)(W) = \alpha(S_{s,t}(W)), \quad W \in T_{\sigma(t)}\mathcal{M}$$

(see Figure 1.4.2). If $T$ is a tensor of type $(p, q)$ at $\sigma(s)$, let

$$(S_{t,s}T)(\alpha_{1}, \ldots, \alpha_{p}, W_{1}, \ldots, W_{q}) = T(S_{s,t}\alpha_{1}, \ldots, S_{s,t}\alpha_{p}, S_{s,t}W_{1}, \ldots, S_{s,t}W_{q}).$$

If $W = \dot{\sigma}(t)$, then the **covariant derivative of $T$ along $W$ at $X = \sigma(t)$** is defined by

$$(\nabla_{w}T)(X) = \left. \frac{d}{ds}\{S_{t,s}T(\sigma(s))\} \right|_{s=t}.$$ 

The covariant derivative $\nabla T$ thereby defines a $(p, q+1)$ tensor.
Notice that $S_{s,t}T(\sigma(s))$ is a tensor at $\sigma(t)$, so the $d/ds$ derivative occurs within a fixed linear space, just as with vectors in 4.21.

The coordinate expression for covariant derivative may be worked out exactly as for vector fields. That computation leads to the following:

**4.26 Proposition** Let $T$ be a tensor of type $\left( \frac{p}{q} \right)$ on $\mathbb{M}$. Writing $(\nabla T)^{A\ldots D}_{E\ldots G | K}$, we have

$$T^{A\ldots D}_{E\ldots G | K} = \frac{\partial T^{A\ldots D}_{E\ldots G}}{\partial x^K} + T^{LB\ldots D}_{E\ldots G} \Gamma^A_{LK} + (\text{all upper indices})$$

and

$$(\nabla_{\dot{\gamma}} T)^{A\ldots D}_{E\ldots G} = T^{A\ldots D}_{E\ldots G | K} \dot{W}^K + (\text{all lower indices}).$$

Notice that we use a vertical bar to designate the coordinate expression for covariant differentiation.

**Problem 4.9** (i) Prove that $\nabla G = 0$ by:
   
   (a) a coordinate calculation;
   (b) using the definition and the fact that $S_{s,t}$ preserves inner products.

(ii) Prove the product rule: $\nabla (T \otimes S) = \nabla T \otimes S + T \otimes \nabla S$. What does this assert in coordinates? For $T = T^A E_A$?

(iii) If $\phi: \mathcal{B} \rightarrow S$ is a regular mapping, show that $\phi^*(\nabla t) = \nabla^*(\phi^*t)$, where $\nabla$ is the covariant derivative with respect to $g$ on $S$, and $\nabla^*$ is the covariant derivative with respect to $\phi^*g = C^i_\ell(g)$.

(iv) Consult problem 2.2 and write the convected acceleration as a covariant derivative of the convected velocity using the metric $C^i_\ell = \varphi^*_* (g)$. 

Figure 1.4.2
4.27 Definition The divergence of a tensor field $T$ of type $\left(\begin{array}{c} p \\ q \end{array}\right)$ is a tensor of type $\left(\begin{array}{c} p - 1 \\ q \end{array}\right)$ obtained by contracting the last contravariant and covariant indices of $\nabla T$:

$$(\text{DIV} \ T)^A_{B\ldots C_p\ldots H} = T^A_{B\ldots C_p\ldots H|C}.$$ 

(The divergence of a tensor $t$ on $\mathcal{S}$ is denoted $\text{div} \ t$.)

For vector fields, the following formula is easy to check:

$$\text{DIV} \ W = W^A_{A\ldots A} = \frac{1}{\sqrt{\det G}} \frac{\partial}{\partial X^A}(\sqrt{\det G} \ W^A).$$

We shall return to this formula from another point of view in Section 1.7.

Next we consider covariant differentiation of two-point tensors.

4.28 Definition Let $T$ be a two-point tensor of type $\left(\begin{array}{c} p \\ q \end{array}\right)$ over the map $\phi: \mathcal{B} \rightarrow \mathcal{S}$; let $\sigma(t)$ be a curve in $\mathcal{B}$ and $W = \dot{\sigma}(t)$. Let $\tilde{\sigma}(t) = \phi(\sigma(t))$ be the image of $\sigma$ under $\phi$. Let $S_{s,t}$ denote parallel translation along $\sigma$ in $\mathcal{B}$ and $s_{s,t}$ denote parallel translation along $\tilde{\sigma}$ in $\mathcal{S}$.

Define the two-point tensor $\nabla_w T$ at $X = \sigma(t)$, a tensor of type $\left(\begin{array}{c} p \\ q + 1 \end{array}\right)$, as follows:

$$\nabla_w T = \left. \frac{d}{ds} (S_{s,t} T(\sigma(s))) \right|_{s=t},$$

where

$$(S_{s,t} T)(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q; \beta^1, \ldots, \beta^q, v_1, \ldots, v_m) = T(S_{s,t} \alpha^1, \ldots, S_{s,t} \alpha^p, S_{s,t} W_1, \ldots, S_{s,t} W_q;$$

$$(s_{s,t} \beta^1, \ldots, s_{s,t} \beta^q, s_{s,t} v_1, \ldots, s_{s,t} v_m),$$

where $\alpha^i \in T^*_x \mathcal{B}$, $W_i \in T_x \mathcal{B}$, $\beta^j \in T^*_x \mathcal{S}$, $v_j \in T_x \mathcal{S}$, $X = \sigma(t)$, and $x = \phi(X)$.

This defines, therefore, a two-point tensor $\nabla T$ of type $\left(\begin{array}{c} p \\ q + 1 \end{array}\right)$.

To help understand this definition, we work out the components of $\nabla_w V$ in case $V$ is a vector field over $\phi$. By definition,

$$\nabla_w V = \left. \frac{d}{ds} (S_{s,t} V(\sigma(s))) \right|_{s=t} \quad \text{where} \quad \dot{\sigma} = W.$$
(See Figure 1.4.3.) Going to a chart, as in 4.21, we get

\[ (\nabla_w V)^a = \frac{d}{dt} V^a(\sigma(t)) + \gamma_{bc}^a V^b \frac{d}{dt} \sigma^c. \]

[Note that \((d/ds)(s_{t,s})_{ab} = -\gamma_{bc} \mathbf{\sigma}^c\) since \(s_{t,s}\) is parallel translation along \(\mathbf{\sigma}\).]

Using the chain rule and the notation \(F^a_A = \partial \phi^a / \partial X^A\),

\[ (\nabla W V)^a = \frac{\partial V^a}{\partial X^A} W^A + \gamma_{bc}^a V^b F^c_A W^A, \]

that is, \((\nabla V)^a_A = V^a_A = \frac{\partial V^a}{\partial X^A} + \gamma_{bc}^a V^b F^c_A\).

The reader will recognize this expression from 3.27. A similar computation proves the following:

**4.29 Proposition**  If \(\mathbf{T}\) is a two-point tensor of type \((p \ l \ q \ m)\), then

\[
\begin{align*}
T^{AB\ldots D}_{E\ldots H} & \mathcal{a}^{d}\ldots e^{h}\mathcal{K} \\
= & \frac{\partial}{\partial X^k} T^{AB\ldots D}_{E\ldots H} \mathcal{a}^{d}\ldots e^{h} \\
& + T^{L_{E_{\ldots H}} a^{d}\ldots e^{h}}_{E_{\ldots H}} + (\text{all upper large indices}) \\
& - T^{A_{L_{E_{\ldots H}} a^{d}\ldots e^{h}}} + (\text{all lower large indices}) \\
& + T^{k_{E_{\ldots H}} a^{d}\ldots e^{h}}_{k_{E_{\ldots H}}} + (\text{all upper small indices}) \\
& - T^{A_{k_{E_{\ldots H}} a^{d}\ldots e^{h}}} + (\text{all lower small indices}).
\end{align*}
\]
The divergence of a two-point tensor is defined in the same way as for a tensor:

\[
(DIV \mathbf{T})^{AB\cdots C}_{\mathbf{E}\cdots H} = g_{\mathbf{E}A} g_{\mathbf{H}B} T^{\mathbf{A}B\cdots \mathbf{C}D}_{\mathbf{E}\cdots \mathbf{H}^{\cdots} g\cdots h | D^*}
\]

**Problem 4.10** Show that, if \( T^{Aa} = g^{ab} T^a_b \) (associated tensors), then

\[
T^{Aa} |B = g^{ab} T^a_b |B.
\]

What is the general statement?

We conclude this section with a brief discussion of curvature.\(^{19}\)

**4.30 Definition** The curvature tensor \( \mathbf{R} \) on a Riemannian manifold \( \mathcal{M} \) (with metric \( G \)) is a \((1, 3)\) tensor defined as follows:

\[
\mathbf{R} : T_x^* \mathcal{M} \times T_x \mathcal{M} \times T_x \mathcal{M} \times T_x \mathcal{M} \rightarrow \mathbf{R},
\]

where

\[
\mathbf{R}(\alpha, (W_1, W_2, W_3)) = \alpha(\nabla_{W_1} \nabla_{W_2} W_3 - \nabla_{W_2} \nabla_{W_1} W_3 - \nabla_{[W_1, W_2]} W_3),
\]

\( \alpha \in T_x^* \mathcal{M} \), and \( W_i(X) \in T_x \mathcal{M} \).

One has to verify that \( \mathbf{R} \) indeed depends only on the point values of \( W_i \). This can be done by computing the components of \( \mathbf{R} \). One gets:

\[
R^A_{BCD} = \frac{\partial \Gamma^A_{DB}}{\partial X^C} - \frac{\partial \Gamma^A_{CB}}{\partial X^D} + \Gamma^A_{CE} \Gamma^E_{DB} - \Gamma^A_{DE} \Gamma^E_{CB}.
\]

One can substitute the expression for \( \Gamma^A_{BC} \) in terms of \( G_{AB} \) to express \( R^A_{BCD} \) totally in terms of \( G_{AB} \) and \( \partial G_{AB} / \partial X^C \partial X^D \).

The contraction \( R_{BD} = R^A_{BAD} \) is called the Ricci curvature and its trace \( R^A_A = R \) is called the scalar curvature.

The geometric significance of the curvature can be illustrated by an example. Let \( \Phi : \mathbb{R}^2 \rightarrow \mathcal{M} \) be a surface in \( \mathcal{M} \). Then \( R^A_{BCD} \) measures the lack of commutativity of second covariant derivatives along the \( x \) and \( y \) directions:

\[
\frac{D}{dx} \frac{D}{dy} W - \frac{D}{dy} \frac{D}{dx} W = \mathbf{R} \cdot \left( \frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, W \right)
\]

where \( (x, y) \in \mathbb{R}^2 \) are identified with points on the surface via \( \Phi \), so \( D/dx \) denotes the covariant derivative along the \( x \)-coordinate curve, and \( W \) is a vector field along the surface.

\(^{19}\)One has to be careful of sign conventions in this topic. The frontispiece of Misner–Thorne–Wheeler [1973] summarizes the situation.
Problem 4.11 Show that on a Riemannian manifold $\mathcal{M}$,

$$W^A_{\mid B\mid C} - W^A_{\mid C\mid B} = R^A_{\ BCD} W^D$$

for a vector field $W$, by a coordinate calculation.

A manifold with $R = 0$ is called flat. For instance, $\mathbb{R}^3$ is flat: $R^A_{\ BCD} = 0$ in any coordinate system since it is obviously zero in Cartesian coordinates (note $\Gamma^B_{\ BC} \neq 0$ in general curvilinear coordinates on $\mathbb{R}^n$).

In books on Riemannian geometry it is proven that a manifold $\mathcal{M}$ is flat if and only if there are coordinates about each point in which $G_{AB} = \delta_{AB}$: that is, $\mathcal{M}$ is locally Euclidean, both topologically and metrically.\(^\natural\)

We conclude with the following, which are sometimes called the compatibility conditions. They restrict the motion $\phi$ of a simple body in terms of its deformation tensor.

4.31 Proposition Let $\mathcal{B}$ be open in $\mathbb{R}^n$ and $S = \mathbb{R}^n$. Let $\phi: \mathcal{B} \rightarrow S$ be a regular configuration. Let $K^A_{\ BCD}$ be the "curvature tensor" obtained by using the deformation tensor $C_{AB}$ in place of $G_{AB}$. Then $K^A_{\ BCD} = 0$.

Proof Combining Problem 4.9(iii) with the definition of curvature, we see that $K$ is the pull-back of $R$, the curvature tensor of $\mathbb{R}^3$. But $R = 0$, so $K = 0$ as well.

There is a related question that is of some interest. This is, given a tensor $C_{AB}$ that is symmetric and positive-definite, when is $C_{AB}$ the deformation tensor of a configuration? The following answers the question locally. (The global question is presumably hard.)

4.32 Proposition Let $\mathcal{B} \subset \mathbb{R}^n$ be open and $S = \mathbb{R}^n$. Let $C_{AB}$ be a given positive-definite symmetric two-tensor whose curvature tensor vanishes: $K^A_{\ BCD} = 0$. Then given any point $x_0 \in \mathcal{B}$, there is a neighborhood $\mathcal{U}$ of $x_0$ and a regular map $\phi: \mathcal{U} \rightarrow \mathbb{R}^n$ whose deformation tensor is $C_{AB}$.

Proof The hypotheses say that the Riemannian manifold $(\mathcal{B}, C_{AB})$ is flat. Thus, in a neighborhood $\mathcal{U}$ of $x_0$, there is a coordinate system $\chi: \mathcal{U} \rightarrow \mathbb{R}^n$ in which the $C_{AB}$'s are constants.\(^\natural\) By following $\chi$ with a linear transformation $A$, we can bring $C_{AB}$ into diagonal form $\delta_{AB}$. Let $\phi = A \circ \chi$. Then $(\phi \circ C)_{ab}$ is the Euclidean metric; that is, $C_{AB}$ is the deformation tensor of $\phi$. \(\blacksquare\)

\(^\natural\) If a metric is flat, one shows that exponential or "canonical" coordinates (which make $\Gamma^4_{\ BC} = 0$ at one point) do the job. See footnote 21.

\(^\natural\) This is a theorem of classical Riemannian geometry proved in, for example, the book of Eisenhart [1926]. The map $\chi$ is given by the exponential map.
Box 4.1 Summary of Important Formulas for Section 1.4

Tensors of type \((p, q)\) on \(\mathcal{M}\)

\[
\mathbf{T} : T^*_\mathcal{M} \times \cdots \times T^*_\mathcal{M} \times T_\mathcal{M} \times \cdots \times T_\mathcal{M} \to \mathbb{R},
\]

\(p\) indices
\[
\mathbf{T}^{AB\cdots D}_{EF\cdots H} = \mathbf{T}(E^A, \ldots, E^D, E_E \cdots E_H)
\]

\(q\) indices

Coordinate Transformations

Tensor Product

\[
\mathbf{T} \otimes \mathbf{S}(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q, \beta^1, \ldots, \beta^p, Y_1, \ldots, Y_q)
= \mathbf{T}(\alpha^1, \ldots, \alpha^p, W_1, \ldots, W_q) \mathbf{S}(\beta^1, \ldots, \beta^p, Y_1, \ldots, Y_q)
\]

Contraction

Contraction of \(\mathbf{T}\) in its last covariant and contravariant slots

\[
\text{(Contraction of } \mathbf{T} \text{ and } \mathbf{S} \text{) } = \mathbf{T} \cdot \mathbf{S}
\]

Double contraction of \(\mathbf{T}\) and \(\mathbf{S}\)

\[
\text{(Double contraction of } \mathbf{T} \text{ and } \mathbf{S} \text{) } = \mathbf{T} : \mathbf{S}
\]

Inner product of \(\mathbf{T}\) and \(\mathbf{S}\)

\[
\langle \mathbf{T}, \mathbf{S} \rangle
\]

Push-Forward

\[
(\phi_* \mathbf{T})(\alpha^1, \ldots, \alpha^p, v_1, \ldots, v_q)
= \mathbf{T}(\phi^*(\alpha^1), \ldots, \phi^*(\alpha^p), \phi^*(v_1), \ldots, \phi^*(v_q))
\]

Pull-Back

\[
\phi^* \mathbf{t} = \phi^{-1}_* \mathbf{t}
\]

Riemannian Metric

\[
G(X)(W_1, W_2) = \left\langle W_1, W_2 \right\rangle_X
\]

Inner product of \(W_1\) and \(W_2\)

\[
\left\langle W_1, W_2 \right\rangle_X = G_{AB}(X)W_1^A W_2^B
\]

\[
\frac{\partial X^A}{\partial \tilde{X}^I}, \ldots, \frac{\partial X^D}{\partial \tilde{X}^I}, \frac{\partial X^M}{\partial \tilde{X}^E}, \ldots, \frac{\partial X^P}{\partial \tilde{X}^H}
\]

\[
(T \otimes S)^{A\cdots D}_{E\cdots H} = T^{A\cdots D}_{E\cdots H} S^{I\cdots L}_{M\cdots P} = T^{A\cdots D}_{E\cdots H} S^{I\cdots L}_{M\cdots P}
\]

\[
(\phi_* \mathbf{T})^{A\cdots D}_{E\cdots H} = F^a_A \cdots F^d_D T^{A\cdots D}_{E\cdots H}(F^{-1})^e_E \cdots (F^{-1})^h_H
\]

\[
(\phi^* \mathbf{t})^{A\cdots D}_{E\cdots H} = (F^{-1})_a^A \cdots (F^{-1})_d^D \mathbf{t}^{A\cdots D}_{E\cdots H} F^e_E \cdots F^h_H
\]

\[
\langle W_1, W_2 \rangle_X = G_{AB}(X)W_1^A W_2^B
\]
Associated Tensors
Raise or lower indices by $G^i$ or $G^i$
\[ T^i = (T with all indices raised: T_{AB...DE...H}^i) \]
\[ T^i = (T with all indices lowered: T_{AB...DE...H}^i) \]

Two-Point Tensors of Type \( \left( \frac{p}{q} \frac{l}{m} \right) \) Over \( \phi: \mathcal{B} \rightarrow \mathcal{S} \)
\[ \mathbf{T}: \frac{T^p_A \otimes \cdots \otimes T^p_A}{p} \times \frac{T^q_B \otimes \cdots \otimes T^q_B}{q} \times \frac{T^{*S}_x \otimes \cdots \otimes T^{*S}_x}{l} \times \frac{T^{*S}_x \otimes \cdots \otimes T^{*S}_x}{m} \rightarrow \mathbb{R} \]

\[ T^{A...D_{E...H}}_{\sigma...\theta} = T(E^A, \ldots, E^D, E_E, \ldots, E_H, e^a, \ldots, e^d, e^e, \ldots, e_h) \]

Lie Bracket
\[ [W_1, W_2][f] = W_1[W_2[f]] - W_2[W_1[f]] \]

Christoffel Symbols
Of a connection \( \nabla \) ...
Of a metric \( G_{AB} \) ...

Covariant Derivative of a Tensor
\[ \nabla_W T = \left. \frac{d}{ds} S_{i,s} T(\sigma(s)) \right|_{s=t} \]
\[ W = \dot{\sigma}(t), \]
\( S_{i,s} = \) parallel translation (shifter)

Covariant Derivative along a Curve \( \sigma(t) \)
\[ \frac{dT}{dt} = \nabla_W T, \quad W = \dot{\sigma}(t) \]

\[ \frac{dT}{dt} \bigg|_{A...D_{E...H}}(\sigma(t)) + \frac{d}{dt} \Gamma^{L}_{EK} W^K + \frac{d}{dt} \Gamma^{L}_{EF} W^K (all upper indices) \]
\[ - \frac{d}{dt} \Gamma^{L}_{EF} W^K (all lower indices) \]
Covariant Derivative of a Two-Point Tensor $T$ of Type $\left( \begin{array}{c} p \\ q \\ m \end{array} \right)$

$$V_w T = \frac{d}{ds} S_{i,s} T(\sigma(s)) \bigg|_{s=t}$$

$S_{i,s} = \text{shifter}$

$$T^A\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h} = \frac{\partial}{\partial X^k} T^A\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h}$$

$$+ T^L\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h} \Gamma^A_{LK} + \text{(all upper large indices)}$$

$$- T^A\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h} \Gamma^L_{EK} - \text{(all lower large indices)}$$

$$+ T^A\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h} \gamma^a_{kl} F^l_{j} + \text{(all upper small indices)}$$

$$- T^A\cdots_D {^E}_{\cdots} ^{a\cdots d}_{e\cdots h} \gamma^a_{kl} F^l_{j} - \text{(all lower small indices)}$$

Divergence

DIV $T = \text{contraction of } V T \text{ on last contravariant index}$

Curvature Tensor

$$R(\alpha, W_1, W_2, W_3) = \alpha(V_{W_1} V_{W_2} W_3 - V_{W_2} V_{W_1} W_3 - V_{[W_1, W_2]} W_3)$$

Compatibility Conditions

Curvature of $C^s = \phi \ast g$ is zero.

$K^A_{BCD} = 0$, where $K^A_{BCD}$ is the curvature of $C_{AB}$. 
1.5 CONSERVATION OF MASS

We shall use conservation of mass to motivate the geometric ideas in Sections 1.6 and 1.7 and the general balance principles in Section 2.1. We begin with regular motions of simple bodies and then, as a supplement, treat the case of (thin) shells; that is, when $\mathcal{B}$ is a two-manifold, $\mathcal{S} = \mathbb{R}^3$ and the motion consists of embeddings.

We shall begin right away by assuming the existence of a mass density function $\rho$. We could alternatively assume the existence of a mass measure $m$ and then, assuming $m$ is sufficiently regular, derive $\rho$ by writing $m = \rho \, dv$ (we may use either the Radon–Nikodym theorem or differential forms to do this). We shall bypass the measure theoretic approach, although it is based on the more primitive physical concept of measuring the masses of portions of a body. The reasons are twofold: first, it plays no role in the rest of the book, and second, it is "obvious" how to bridge the gap for those who know measure theory.\textsuperscript{22}

5.1 Definition Let $\mathcal{B} \subset \mathbb{R}^n$ be a simple body and let $\phi(X, t)$ be a motion of $\mathcal{B}$. A function $\rho(x, t)$ is said to obey conservation of mass provided that for all open sets $\mathcal{U} \subset \mathcal{B}$ with piecewise $C^1$ boundary (such $\mathcal{U}$ hereafter will be called nice),

\[
\frac{d}{dt} \int_{\phi_t(\mathcal{U})} \rho(x, t) \, dv = 0
\]

where $dv$ denotes the standard Euclidean volume element in $\mathbb{R}^n$.

We call $\rho(x, t)$ the mass density (in spatial coordinates) and call $\int_0^t \rho(x, t) \, dv$ the mass of the set $\mathcal{U}$.

Conservation of mass thus states that the mass of any nice material region $\mathcal{U}$ is constant in time. Let us assume that $\phi_t$ is a regular $C^1$ motion from now on, and $\phi_0 =$ identity; that is, the body is in the reference configuration at $t = 0$.

5.2 Definition Let $J(X, t)$, called the Jacobian, denote the determinant of the linear transformation $F(X, t) = D\phi(X)$.

Let us work out $J(X, t)$ in general coordinates.

5.3 Proposition In (positively oriented)\textsuperscript{23} coordinates $\{X^a\}$ on $\mathcal{B}$ and $\{x^a\}$ on $\mathcal{S}$ we have

\[
J(X, t) = \frac{\partial(\phi^1, \ldots, \phi^n)}{\partial(X^1, \ldots, X^n)} \frac{\sqrt{\det g_{ab}(x)}}{\sqrt{\det G_{AB}(X)}},
\]

\textsuperscript{22}See, for example, Truesdell [1977].

\textsuperscript{23}Coordinates $\{X^a\}$ in $\mathbb{R}^3$ are called positively oriented when $\det(\partial X^a/\partial Z^b) > 0$. 

85
where \[ x = \phi(X, t) \quad \text{and} \quad \frac{\partial (\phi^1, \ldots, \phi^n)}{\partial (X^1, \ldots, X^n)} = \det \left( \frac{\partial \phi^i}{\partial X^a} \right). \]

**Proof** Using Cartesian coordinates, we get
\[ J(X, t) = \det \left( \frac{\partial \phi^i}{\partial Z^j} \right) = \det \left[ \left( \begin{array}{ccc} \frac{\partial \phi^1}{\partial X^a} & \cdots & \frac{\partial \phi^n}{\partial X^a} \\ \frac{\partial Z^1}{\partial x^1} & \cdots & \frac{\partial Z^n}{\partial x^1} \end{array} \right) \left( \begin{array}{ccc} \frac{\partial Z^1}{\partial x^a} & \cdots & \frac{\partial Z^n}{\partial x^a} \end{array} \right) \right] = \det \left( \frac{\partial \phi^a}{\partial X^a} \right) \det \left( \frac{\partial X^a}{\partial Z^j} \right). \]

But
\[ \det(g_{ab}) = \det \left( \frac{\partial z^i}{\partial x^a}, \frac{\partial z^j}{\partial x^b}, \delta_{ij} \right) = \left[ \det \left( \frac{\partial z^i}{\partial x^a} \right) \right]^2, \]
with a similar formula for \( \det(G_{AB}) \). Hence the result follows. \( \blacksquare \)

**Warning.** \( J \) is a scalar function, invariant under coordinate transformations. Note that \( J \neq \det(F_{ab}) \); the quantity \( \det(F_{ab}) \) picks up determinant factors under coordinate changes and is sometimes called a “modular” tensor. The factor \( (\det g_{ab}/\det G_{AB})^{1/2} \) in the formula for \( J \) corrects this, so \( J \) is a scalar.

For what follows, we shall need to note that \( J(X, t) > 0 \). Indeed, \( J(X, 0) = 1 \), and since \( \phi \), is regular, \( J(X, t) \neq 0 \). Hence \( J(X, t) \) is positive by the intermediate value theorem.

5.4 **Proposition**

\[ \frac{\partial J}{\partial t} = (\text{div} \, v)J = (v^a \sub a)J. \]

**Proof** We shall prove this by a direct calculation. For a proof without coordinates, see Section 1.7. The reader can easily prove the following fact from matrix algebra: if \( (a_{ij}(t)) \) is a time-dependent matrix, then
\[ \frac{d}{dt} \det(a_{ij}(t)) = \frac{da_{ij}}{dt}(\text{Cof})^i_j \]

(sum on both \( i \) and \( j \)), where \( (\text{Cof})^i_j \) is the \( (i, j) \)th cofactor of \( (a_{ij}) \). Therefore,
\[ \frac{\partial}{\partial t} \ J(X, t) = \left( \frac{\partial}{\partial t} \left( \frac{\partial \phi^a}{\partial X^4} \right) (\text{Cof})^a_4 \right) \sqrt{\det \frac{g_{cd}}{G_{CD}}} + \left( \frac{\partial (\phi^1, \ldots, \phi^n)}{\partial (X^1, \ldots, X^n)} \right) \left( \frac{\partial}{\partial X^4} \right) \sqrt{\det \frac{g_{cd}}{G_{CD}}} \]
\[ = \frac{\partial v^a}{\partial x^b} \frac{\partial \phi^b}{\partial X^4} (\text{Cof})^a_4 \left( \sqrt{\det \frac{g_{cd}}{G_{CD}}} \right) \frac{1}{\sqrt{\det G_{CD}}} \frac{\partial}{\partial x^a} \left( \sqrt{\det g_{cd}} \right) v^a. \]

The first sum is a sum of determinants with the \( a \)th row of \( \partial \phi^a/\partial X^4 \) replaced by \( (\partial v^a/\partial x^b)(\partial \phi^b/\partial X^4) \). Since repeated rows yield a zero determinant, only the term with \( b = a \) survives. Thus
\[
\frac{\partial}{\partial t} J(X, t) = \frac{\partial \nu^a}{\partial x^a} J(X, t) + J(X, t) \frac{1}{\sqrt{\det g_{cd}}} \frac{\partial}{\partial x^a} (\sqrt{\det g_{cd}} \nu^a)
\]

\[
= J(X, t) \frac{1}{\sqrt{\det g_{cd}}} \frac{\partial}{\partial x^a} (\sqrt{\det g_{cd}} \nu^a)
\]

\[
= J(X, t) \nu^a |_{\nu_a}.
\]

This proof will be thoroughly understood if the reader will repeat it explicitly for \( \mathbb{R}^3 \), using Cartesian coordinates and the formula

\[
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \frac{d}{dt} = \begin{vmatrix} \dot{a}_{11} & \dot{a}_{12} & \dot{a}_{13} \\ \dot{a}_{21} & \dot{a}_{22} & \dot{a}_{23} \\ \dot{a}_{31} & \dot{a}_{32} & \dot{a}_{33} \end{vmatrix} a_{11} a_{12} a_{13} + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} \dot{a}_{21} & \dot{a}_{22} & \dot{a}_{23} \\ \dot{a}_{31} & \dot{a}_{32} & \dot{a}_{33} \end{vmatrix} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \dot{a}_{31} \dot{a}_{32} \dot{a}_{33}.
\]

5.5 Definition A motion \( \phi_t \) is called volume preserving (also called isochoric or incompressible) if \( \text{volume}[\phi_t(\mathcal{U})] = \text{volume}[\mathcal{U}] \) for every nice region \( \mathcal{U} \subset \mathcal{B} \).

5.6 Proposition Assume \( \phi_t \) is a \( C^1 \) regular motion. Then the following are equivalent:

(i) \( \phi_t \) is volume preserving;
(ii) \( J(X, t) = 1 \); and
(iii) \( \text{div} \nu = 0 \).

Proof (i) is equivalent to (ii) by the change of variables formula, and (ii) is equivalent to (iii) by 5.4 and \( J(X, 0) = 1 \).

Next we use 5.4 to establish the basic equation of continuity. We write \( \rho_{\text{ref}}(X) = \rho(X, 0) \), the mass density in the reference configuration (assumed to occur at \( t = 0 \)). Recall that \( \dot{\rho} = \frac{\partial \rho}{\partial t} + \rho \cdot \nu = \frac{\partial \rho}{\partial t} + \nu[\rho] \) is the material derivative of \( \rho \).

5.7 Theorem Assume \( \phi_t \) is a \( C^1 \) regular motion and that \( \rho(x, t) \) is a \( CI \) function. Then the following are equivalent:

(i) Conservation of Mass;
(ii) \( \rho(x, t) J(X, t) = \rho_{\text{ref}}(X) \) [where \( x = \phi(X, t) \)]; and
(iii) the equation of continuity,

\[
\dot{\rho} + \rho \text{div} \nu = 0;
\]

that is, \( \frac{\partial \rho}{\partial t} + \text{div}(\rho \nu) = 0 \).

Proof Assume (i); that is, \( \int_{\phi_t(\mathcal{U})} \rho(x, t) \, dv = \int_{\mathcal{U}} \rho_{\text{ref}}(X) \, dV \). By the change of variables formula this is equivalent to \( \int_{\mathcal{U}} \rho(x, t) J(X, t) \, dV = \int_{\mathcal{U}} \rho_{\text{ref}}(X) \, dV \). Since
88 GEOMETRY AND KINEMATICS OF BODIES

CH. 1

is arbitrary, this is equivalent to (ii). (The proof employs an easy calculus exercise: if \( f(x) \) is a continuous function and \( \int_D f(x) \, dv = 0 \) for every ball \( D \), then \( f = 0 \).) By 5.4, \( (\partial/\partial t)(\rho(x, t)J(X, t) = \dot{\rho}J + \rho \dot{J} = \rho \dot{J} + (\rho \text{ div } v)J \), so (ii) is equivalent to (iii).

The process of passing from (i) to (iii) is called localization. In the volume preserving case, the equation of continuity reduces to \( \dot{\rho} = 0 \); that is, \( \rho(x, t) = \rho(X, 0) \).

**Problem 5.1** Omitting the assumption that \( \phi_0 = \text{identity} \), show that 5.7 remains true for some function \( \rho_{\text{Ref}} \) on \( \mathcal{B} \).

Next we investigate the geometry behind conservation of mass a little further. We want to make \( dv \) and \( dV \), the volume elements on \( S \) and \( \mathcal{B} \), into respectable tensors, \( dv \) and \( dV \) respectively.

**Warning.** These notations should not be confused with the spatial and material velocities.

5.8 **Definition** The *volume form*, regarded as an antisymmetric \( (0 \, n) \) tensor, is defined by \( dv(w_1, \ldots, w_n) = \sqrt{\det \langle w_i, w_j \rangle} \) for \( w_1, \ldots, w_n \) positively oriented; that is, \( \det \langle w_i, w_j \rangle > 0 \) and is extended to all \( w_1, \ldots, w_n \) by skew symmetry and multilinearity.

Recall that the volume of the parallelopiped spanned by \( n \) vectors \( w_1, \ldots, w_n \) in \( \mathbb{R}^n \) is \( \sqrt{\det \langle w_i, w_j \rangle} \). The tensor \( dv \) is completely antisymmetric to reflect the same property of the determinant: it vanishes if any \( w_i = 0 \) and therefore if \( dv \) is to be linear in each \( w_i \), it must change sign if two \( w_i \)'s are interchanged.

In terms of a general (oriented) coordinate system in \( \mathbb{R}^3 \), 5.8 reads

\[
dv = \sqrt{\det g_{ab} (dx^1 \otimes dx^2 \otimes dx^3)} \text{antisymmetrized} = \sqrt{\det g_{ab}} \, dx^1 \wedge dx^2 \wedge dx^3.
\]

(Section 1.7 elaborates on this notation.)

5.9 **Definition** The *mass form* is defined by \( m = \rho \, dv \) (each of \( m \) and \( \rho \) depends on \( (x, t) \)). Also, let \( m_{\text{Ref}} = \rho_{\text{Ref}} \, dV \), the mass form of the reference configuration.

Using the mass form we can express conservation of mass entirely in terms of pull-backs.

5.10 **Proposition** Conservation of mass is equivalent to

\[
\phi^* m = m_{\text{Ref}}.
\]

**Proof** \( \phi^* m = \phi^* (\rho \, dv) = (\rho \circ \phi) \phi^* dv \). (The pull-back of a product is the product of the pull-backs.) We claim that \( \phi^* dv = J \, dV \), from which it follows
that $\phi^*m = m_{\text{Ref}}$ is equivalent to $\rho(x, t)J(X, t) = \rho_{\text{Ref}}(X)$—that is, conservation of mass. To prove the claim we use the definitions of pull-back and $d\nu$:

$$(\phi^*d\nu)(W_1, W_2, W_3) = d\nu(D\phi_t \cdot W_1, D\phi_t \cdot W_2, D\phi_t \cdot W_3)$$

$$= [\det(D\phi_t \cdot W_1, D\phi_t \cdot W_2, D\phi_t \cdot W_3)]^{1/2}$$

$$= \left[\det\left( g^{ab} \frac{\partial \phi^a}{\partial X^A} W^A_1, \frac{\partial \phi^b}{\partial X^B} W^B_2 \right) \right]^{1/2}$$

$$= \sqrt{\det g^{ab}} \frac{\partial(\phi^1, \phi^2, \phi^3)}{\partial(x^1, x^2, x^3)} \sqrt{\frac{\det(G_{AB} W^A_1 W^B_2)}{\det(G_{AB})}}$$

$$= J(t, X) dV(W_1, W_2, W_3).$$

These ideas are concisely treated using differential forms and the Lie derivative in Section 1.7.

**Problem 5.2** Write out the equation of continuity in spherical coordinates in $\mathbb{R}^3$.

---

**Box 5.1 Conservation of Mass for Shells and the Second Fundamental Form**

The expressions for $d\nu$ and $dV$ make sense on an arbitrary Riemannian manifold as does the definition of conservation of mass. The equivalence with $\rho J = \rho_{\text{Ref}}$ and $\phi^*m = m_{\text{Ref}}$ remains valid. However, the equation of continuity requires revision; indeed, it does not make sense as stated, since div $\nu$ is not always defined. In fact, $\nu$ is a vector tangent to $S$ defined at points of $\phi,(\mathcal{B})$, so if $\mathcal{B}$ and $S$ have different dimensions, div $\nu$ does not make sense.

By definition, a shell is a body $\mathcal{B}$ that is a two-manifold in $\mathbb{R}^3 = S$. Physically, it is a sheet whose thickness is being ignored. The correct form of the equation of continuity depends on the second fundamental form of an embedded hypersurface (in our case, the shell), so we consider first some relevant geometry.

**5.11 Definition** Let $\mathcal{M}$ and $\mathcal{N}$ be Riemannian manifolds, $\mathcal{M} \subset \mathcal{N}$ with $\dim \mathcal{N} = \dim \mathcal{M} + 1$. The Riemannian structure on $\mathcal{M}$ is assumed to equal that obtained from $\mathcal{N}$. The second fundamental form of $\mathcal{M}$ is the $\left( \begin{array}{c} 0 \\ 2 \end{array} \right)$ tensor $k$ on $\mathcal{M}$ such that

$$k(W_1, W_2) = \langle \nabla_W n, W \rangle_x,$$

---

$^{24}$The standard reference for shells is Naghdi [1972].
where \( \nabla \) is the covariant derivative on \( \mathcal{M} \), \( W_1, W_2 \in T_x \mathcal{M} \), and \( n \) is the unit outward normal of \( \mathcal{M} \). (We assume \( \mathcal{M} \) is oriented; that is, it has a well defined "outward" normal.)

5.12 Proposition We have

(i) \( k(W_1, W_2) = -\langle \nabla_{W_1}W_2, n \rangle \), (Weingarten equation)

(ii) \( k \) is a symmetric tensor, and

(iii) in coordinates \( (x^1, \ldots, x^{n+1}) \) for which \( x^{n+1} \) is the outward normal direction to \( \mathcal{M} \), and such that \( e_{n+1} = n \),

\[
k_{ab} = -\gamma_{ab}^{n+1} = \frac{1}{2} \frac{\partial g_{ab}}{\partial x^{n+1}} (1 \leq a, b \leq n).
\]

Proof

(i) Since \( W_2 \) is parallel to \( \mathcal{M} \), \( \langle W_2, n \rangle = 0 \).

Hence

\[
0 = \nabla_{W_1} \langle W_2, n \rangle = \langle \nabla_{W_1}W_2, n \rangle + \langle W_2, \nabla_{W_1}n \rangle.
\]

(ii) follows from (iii), so we prove (iii); indeed, by (i),

\[
k_{ab} = -\langle \nabla_{e_a}e_b, n \rangle = -\gamma_{ab}^{n+1} \delta_{n+1} = -\gamma_{ab}^{n+1}.
\]

Using the explicit formula for \( \gamma_{bc}^{n+1} \) in terms of \( g_{ab} \) (see the proof of 4.24 or 3.29) yields the last equality.

For a surface in space, \( \kappa = \text{tr} k \), the trace of \( k \), is called the mean curvature (sum of the principal curvatures; the eigenvalues of \( k \)) and \( \det k \) (product of the principal curvatures) is called the Gaussian curvature. [Aside: The second fundamental form provides the link between the connections and curvatures on \( \mathcal{M} \) and \( \mathcal{N} \), through the Gauss–Codazzi equations:

\[
(i) \quad (\mathcal{N}) R_{abcd} = (\mathcal{M}) R_{abcd} + k_{ba} k_{ac} - k_{ad} k_{bc} (1 \leq a, b, c, d \leq n),
\]

\[
(ii) \quad (\mathcal{N}) \nabla_{W_1} W_2 - (\mathcal{M}) \nabla_{W_1} W_2 = -k(W_1, W_2)n.
\]

These are not hard to prove: for these and related formulas, see any book in Riemannian geometry, such as Yano [1970].]

Now we return to our main concern—conservation of mass. As in 5.7, the equation of continuity boils down to a computation of \( \partial J/\partial t \). The definition of \( J \) in general is \( J dV = \phi^* dv \), where \( dv \) is the volume element on \( \phi_i(\mathcal{B}) \); 5.3 remains valid. Also, \( \phi^* dv = \) volume element of \( (\phi^* g = C^i) \) from the definition of pull-back and 5.8. Write \( \mu(\phi^* g) \) for the volume element of \( \phi^* g \). By the computation in 5.4,

\[
\frac{d}{dt} \det(g_{ab}(t)) = \text{tr} \left( \frac{\partial g_{ab}}{\partial t} \right) \det(g_{ab}).
\]

Thus

\[
\frac{\partial J}{\partial t} dV = \frac{1}{2} \mu(\phi^* g) \text{tr} \left( \frac{\partial}{\partial t} \phi^* g \right) = \frac{1}{2} \mu \left( \text{tr} \left( \frac{\partial}{\partial t} \phi^* g \right) \right) dv.
\]
This calculation proves the following abstract result:

**5.13 Proposition** Let $\mathcal{B}$ and $\mathcal{S}$ be general Riemannian manifolds and $\phi_t$ an embedded $C^1$ motion of $\mathcal{B}$ in $\mathcal{S}$. Then

$$\frac{\partial J}{\partial t} = \frac{1}{2} \text{tr}_c \left( \frac{\partial}{\partial t} \phi_t^* g \right) J = D_{AB} B^A B^B,$$

where the trace is taken using the metric $\phi_t^* g$ and $B^i = (C^i)^{-1}$.

To apply this to shells, let $v$ be the spatial velocity and $n$ the unit normal. Decompose $v$ into a component parallel to $\phi_t(\mathcal{B})$ and a component normal to it (see Figure 1.5.1):

$$v = v_p + v_n n.$$

We must compute $(\partial/\partial t)\phi_t^* g$, that is, $(\partial/\partial t)C_{AB}$. Write $2D^B = (\partial/\partial t)\phi_t^* g$ as before.

**5.14 Proposition** For shells, the rate of deformation tensor is

$$2D^B = \phi_t^*((\nabla v)^t + [(\nabla v)]^t) + 2v_n k,$$

that is,

$$2D_{AB} = (v_{a;\mu}^b + v_{b;\mu}^a) F^\mu_A F^\mu_B + 2v_n k_{ab} F^a_A F^b_B,$$

where $k$ is the second fundamental form.

Proof. $C_{AB} = g_{ab} F^a_A F^b_B$, so, as in 3.28,

$$2D_{AB} = \left( \frac{\partial g_{ab}}{\partial x^\alpha} v^\alpha + g_{cb} \frac{\partial v^c}{\partial x^a} + g_{ac} \frac{\partial v^c}{\partial x^b} \right) F^a_A F^b_B$$

$$= (v_{a;\mu}^b + v_{b;\mu}^a) F^\mu_A F^\mu_B.$$
But $\nabla v = \nabla (v_i + v_n n) = \nabla v_i + \nabla v_n \otimes n + v_n \nabla n$. The covariant derivative here is on $\mathcal{S}$. Since $n$ is orthogonal to $\phi_i(\mathcal{B})$, the second term pulls back to zero. Since $(\nabla n)^i = k$, and $k$ is symmetric, we get the result.

We shall write $\tilde{V}$ for the covariant derivative on $\phi_i(\mathcal{B})$ to avoid confusion with the covariant derivative $V$ on $\mathcal{S}$. However, $\nabla v_i$ and $\tilde{\nabla} v_i$ differ by $k(v_i, n)$, which is normal to $\phi_i(\mathcal{B})$ and hence has zero pull-back to $\mathcal{B}$. Thus $\phi_i^*(\nabla v_i) = \phi_i^*(\tilde{\nabla} v_i)$. Since the divergence of $v_i$ on $\phi_i(\mathcal{B})$ is given by $\overline{\text{div}}v_i = \text{tr}(\nabla v_i)$, 5.14 yields

$$\text{tr}_c D = \phi_i^*(\overline{\text{div}}v_i) + \phi_i^*(v_n \text{tr} k) = (\overline{\text{div}}v_i) \circ \phi_i + (v_n \text{tr} k) \circ \phi_i.$$ 

Substituting in 5.13 using $\rho J = \rho_{\text{Ref}}$ gives the following:

**5.15 Theorem** Conservation of mass for shells is equivalent to the equation of continuity for shells:

$$\dot{\rho} + \rho \overline{\text{div}}v + \rho v_n \text{tr} k = 0$$

where $\dot{\rho} = \partial \rho / \partial t + \nabla \rho \cdot v$, $k$ is the second fundamental form of the surface $\phi_i(\mathcal{B})$, and $\text{tr} k$ is its mean curvature.

**Problem 5.3** Suppose $\phi_i(\mathcal{B})$ is a sphere in $\mathbb{R}^3$ with radius $r(t)$. Given $\rho_{\text{Ref}}$, derive a formula for $\rho(x, t)$ by elementary considerations. Verify that 5.15 holds by explicitly computing $k$ for a sphere of radius $r(t)$.

**Ans**: $k_{ab} = \begin{bmatrix} r & 0 \\ 0 & r \sin^2 \phi \end{bmatrix}$ in spherical coordinates.

**Problem 5.4** Derive the equation of continuity for a wire—that is, when $\mathcal{B} \subset \mathbb{R}^3$ is a 1-manifold. You will need the Frenet formulas from a geometry text.

**Problem 5.5** If $\mathcal{B} \subset \mathcal{S}$ is open, but the metric on $\mathcal{S}$ is time dependent, show that the equation of continuity becomes

$$\dot{\rho} + \rho \overline{\text{div}}v + \frac{1}{2} \text{tr} \left( \frac{\partial g}{\partial t} \right) = 0.$$
Jacobian $J(X, t)$
\[ \phi^* \, dv = J \, dV \]

Volume Element
\[ dv(w_1, \ldots, w_n) = \sqrt{\det(w_i, w_j)} \]

Mass Form
\[ m = \rho \, dv \]

Conservation of Mass
\[ \phi_i^* m = m_{\text{Ref}} \]

Rate of Change of Jacobian
Simple Bodies:
\[ \frac{\partial J}{\partial t} = (\text{div} \, v)J \]
Shells:
\[ \frac{\partial J}{\partial t} = \frac{1}{2} \text{tr}_C \left( \frac{\partial}{\partial t} \phi^* g \right) J = (\text{div} \, v)J + v_n \text{tr} \, k \]

Second Fundamental Form:
\[ k(W_1, W_2) = \langle \nabla_{W_1} n, W_2 \rangle = -\langle \nabla_{W_1} W_2, n \rangle \]

Equation of Continuity
Simple Bodies:
\[ \dot{\rho} + \rho \, \text{div} \, v = 0 \]
Shells:
\[ \dot{\rho} + \rho (\text{div} \, v) + v_n \text{tr} \, k = 0 \]

\[ J = \frac{\partial (\phi^1, \ldots, \phi^n)}{\partial (X^1, \ldots, X^n)} \sqrt{\det G_{AB}}, \]
where \[ \frac{\partial (\phi^1, \ldots, \phi^n)}{\partial (X^1, \ldots, X^n)} = \det \left( \frac{\partial \phi^a}{\partial X^A} \right) \]

\[ dv = \sqrt{\det g_{ab}} \, dx^1 \wedge \cdots \wedge dx^n \]

\[ m = \rho \sqrt{\det g_{ab}} \, dx^1 \wedge \cdots \wedge dx^n \]

\[ (\rho \circ \phi)J = \rho_{\text{Ref}} \]

1.6 FLOWS AND LIE DERIVATIVES

We shall begin by defining the notion of a tangent vector to a curve $c$ within the context of manifold theory. Heretofore the tangent has been denoted $dc/dt$. Strictly speaking, this is not an appropriate notation for a tangent vector in a
manifold since it omits the base point of the vector. If \( S \) is an \( n \)-dimensional differentiable manifold and \( c: \mathcal{D} \rightarrow S \) is a smooth curve, where \( \mathcal{D} \) is an open interval of the real line, then in a coordinate chart \( \{x^a\} \), \( dc/dt = (dc^1/dt, \ldots, dc^n/dt) \), the "vector part" of the tangent vector. To develop the appropriate notion of tangent vector we apply the tangent map \( T \) to \( c \). (The interval \( \mathcal{D} \) may be viewed as a differentiable manifold and thus \( Tc: T\mathcal{D} \rightarrow T\mathcal{S} \).) By definition of the tangent of a map, \( Tc(r, s) = (c(r), (dc/dt)(r)s) \), where \((r, s) \in \mathcal{D} \times \mathbb{R} = T\mathcal{D} \) is a tangent vector to \( \mathcal{D} \). Since we ultimately want the vector part of the tangent vector to be simply \( dc/dt \), we evaluate \( Tc \) on the unit vector based at \( r: Tc(r, 1) = (c(r), (dc/dt)(r)) \). Thus \( Tc(\cdot, 1): \mathcal{D} \rightarrow T\mathcal{S} \) is what we meant when we discussed the tangent vector to the curve \( c \). The use of \( dc/dt \) in place of \( Tc(\cdot, 1) \) is a frequent abuse of notation in manifold theory, which we shall follow.

6.1 Definition  Let \( w: \mathcal{U} \rightarrow T\mathcal{S} \) be a vector field, where \( \mathcal{U} \) is an open subset of \( \mathcal{S} \). A curve \( c: \mathcal{D} \rightarrow \mathcal{S} \), where \( \mathcal{D} \) is an open interval, is called an integral curve of \( w \) if for every \( r \in \mathcal{D} \)

\[
\frac{dc}{dt}(r) = w(c(r)).
\]

6.2 Proposition  Let \( w \) be a \( C^k \) vector field on \( \mathcal{U} \) (open in \( \mathcal{S} \)), where \( k \geq 1 \). Then for each \( x \in \mathcal{U} \) there exists an \( \epsilon > 0 \) and a \( C^{k+1} \) integral curve \( c_x \) of \( w \) defined on \((-\epsilon, \epsilon)\) such that \( c_x(0) = x \). Furthermore, \( c_x \) is unique in the sense that if \( \tilde{c}_x \) satisfies the same conditions, but is defined on an interval \((-\tilde{\epsilon}, \tilde{\epsilon})\) \((\tilde{\epsilon} > 0)\), then \( c_x = \tilde{c}_x \) on the intersection of the intervals.

Proof  Let \( \{x^a\} \) be a coordinate system about a point \( p \in \mathcal{S} \). Then \( c_x^a = x^a(c(r)) \), so by the chain rule,

\[
\frac{dc_x^a}{dt}(r) = dx^a(Tc_x(r, 1)) = dx^a(w(c_x(r))) = w^a(c_x(r)),
\]

for each \( r \in (-\epsilon, \epsilon) \). Also, \( c_x^a(0) = x^a(c_x(0)) = x^a(p) \). This is a system of \( n \) ordinary differential equations with given initial conditions. Since \( w \) is \( C^k \), so are the \( w^a \)'s (this is what \( w \) being \( C^k \) means), and thus by the fundamental existence theorem of ordinary differential equations, there exist \( C^{k+1} \) functions \( c_x^a \) satisfying the differential equations and initial conditions. These functions are unique in the sense spelled out in the hypotheses.

6.3 Generalization  If \( w \) depends upon the "time variable" \( r \) itself—that is, if \( w: \mathcal{U} \times (-\epsilon, \epsilon) \rightarrow T\mathcal{S} \)—we define an integral curve by

\[
\frac{dc}{dt} = w(c(t), t).
\]

\[\text{25 See any of the standard references on analysis or ordinary differential equations, for example, Marsden [1974a] or Hartman [1973].}\]
If $w$ depends upon $t$ in a $C^k$ manner, then the conclusions of the previous proposition still hold.

**6.4 Example** Let $v : \phi_t (\mathcal{B}) \times (-\epsilon, \epsilon) \rightarrow T\mathcal{S}$ denote the spatial (Eulerian) velocity vector field of a regular motion $\phi_t$ with $\phi_0 = \text{identity}$. An integral curve of $v$ through $X \in \mathcal{B} \subset \mathcal{S}$ is given by the motion; that is, $\phi_x(t)$ is an integral curve of $v$:

$$\frac{d\phi_x}{dt} = T\phi_x(t, 1) = \left(\phi_x(t), \frac{d}{dt} \phi_x(t)\right) = v(\phi_x(t), t)$$

and $\phi_x(0) = X$.

**6.5 Definition** Let $w : \mathcal{S} \times \mathcal{S} \rightarrow T\mathcal{S}$ be a $C^k$ vector field with $\mathcal{S}$ and $\mathcal{S}$ as before. The collection of maps $\psi_{t,s}$ such that for each $s$ and $x$, $t \mapsto \psi_{t,s}(x)$ is an integral curve of $w$ and $\psi_{s,s}(x) = x$ is called the flow or evolution operator of $w$.

The flow may be thought of as the totality of integral curves of $w$. The basic existence theory implies that the flow of a $C^k$ vector field is jointly $C^k$ where defined. Normally one takes $\psi_{t,s}(x)$ to mean the maximally extended flow—that is, with the largest domain possible.

**6.6 Example** Let $v$ be the spatial velocity vector field of a motion $\phi_t$. Then the collection $\{\phi_{t,s} | \phi_{t,s} = \phi_{t,s} \circ \phi_{s}^{-1} : \phi_{t,s} (\mathcal{B}) \rightarrow \phi_{s} (\mathcal{B})\}$ is the flow of $v$. It is clear from this definition that $\phi_{t,s} \circ \phi_{s,r} = \phi_{t,r}$ and that $\phi_{t,t} = \text{identity}$, for all $r, s, t \in \mathbb{R}$ for which the flow is defined. These are properties of flows in general that may be verified using uniqueness of integral curves.

**6.7 Remark** If the vector field $w$ is time independent, then $\psi_{t,s}$ depends only on the difference $t - s$. In this case, one writes $\psi_{t-s} = \psi_{t,s}$, and the flow constitutes a local one-parameter group—that is, $\psi_{t} \circ \psi_{s} = \psi_{t+s}$.

The tangent map of $\psi_{t,s}$ is defined in the usual way, as are the push-forward and pull-back induced by $\psi_{t,s}$, by merely holding $t$ and $s$ fixed and regarding $\psi_{t,s}$ as a mapping of $\mathcal{S}$ to $\mathcal{S}$.

**6.8 Definition** Let $w$ be a $C^1$ (time-dependent) vector field on $\mathcal{S}$ and let $\psi_{t,s}$ denote its flow. If $t$ is a $C^1$ (possibly time-dependent) tensor field on $\mathcal{S}$, then the Lie derivative\(^{26}\) of $t$ with respect to $w$ is defined by

$$L_w t = \left(\frac{d}{dt} \psi_{t,s}^* t\right)_{|s=s}.$$

---

\(^{26}\)The first uses of the Lie derivative in mechanics that we know of are Cartan [1922] and Siebodzinski [1931]. In continuum mechanics the first comprehensive reference we know of is Guo Zhong-Heng [1963].
If we hold \( t \) fixed in \( t\), we obtain the autonomous Lie derivative:

\[
\mathcal{L}_w t = \frac{d}{dt}(\psi_{t,s}^*(t))_{|_{t=s}}.
\]

Thus \( L_w t = \partial t/\partial t + \mathcal{L}_w t \).

**6.9 Remark** Note that the Lie derivative does not depend on a metric or a connection on \( S \).

The differentiation \( d/dt \) in the definition of the Lie derivative makes sense since for \( s \) fixed, the pull-back of \( t \) is a curve of tensors in a fixed linear space.

Before enumerating general properties of the Lie derivative, we compute it in coordinates for a few simple cases.

**6.10 Examples**

1. Consider the Lie derivative of a function \( f \) on \( S \). In this case, \( (\psi_{t,s}^*(x), t) \), and, therefore,

\[
(L_w f)(x, t) = \left. \left( \frac{\partial f}{\partial t} (\psi_{t,s}^*(x), t) + \frac{\partial f}{\partial x^a} (\psi_{t,s}^*(x), t) \frac{\partial \psi_{t,s}^a}{\partial t} (x) \right) \right|_{t=s}.
\]

Since \( \psi_{t,s}(x) = x \) and \( \frac{\partial \psi_{t,s}^a}{\partial t} (x) = \frac{d}{dt} \psi_{t,s}^a(t) = w^a(\psi_{x}(t), t) \), we have

\[
(L_w f)(x, t) = \frac{\partial f}{\partial t} (x, t) + \frac{\partial f}{\partial x^a} (x, t)w^a(x, t),
\]

that is, \( L_w f = df/\partial t + w[f] \).

2. Let \( \alpha \) be a one-form. Then

\[
(\psi_{t,s}^*(\alpha))(x) = \alpha_0(\psi_{t,s}(x), t) \frac{\partial \psi_{t,s}^a}{\partial x^b}(x)e^b(x).
\]

Thus

\[
(L_w \alpha)(x, t) = \left[ \left( \frac{\partial}{\partial t} \alpha_a + \alpha_\bullet \frac{\partial \psi_{t,s}^a}{\partial t} + \frac{\partial \psi_{t,s}^a}{\partial x^b} \right) \frac{\partial \psi_{t,s}^a}{\partial x^b} + \alpha_\bullet \frac{\partial \psi_{t,s}^a}{\partial x^b} \right] e^b.
\]

Note that from \( \psi_{t,s}(x) = x \), we have \( (\partial \psi_{t,s}^a/\partial x^b)(x) = (\partial x^a/\partial x^b)(x) = \delta^a_b \). We also need to note that if \( w \) is a \( \mathcal{C}^1 \) vector field,

\[
\frac{\partial^2 \psi_{t,s}^a}{\partial t \partial x^b} = \frac{\partial}{\partial x^b} \frac{\partial \psi_{t,s}^a}{\partial t} = \frac{\partial w^a}{\partial x^b}.
\]

[This uses a general fact: if \( f(x, y) \) is a \( \mathcal{C}^1 \) function of two variables and \( \partial f/\partial x \) is also \( \mathcal{C}^1 \), then the mixed partial derivatives exist and commute, and \( \partial^2 f/\partial x \partial y \) is continuous.\(^2\)]

Combining these results enables us to write

\[
L_w \alpha = \left( \frac{\partial \alpha_a}{\partial t} + \alpha_\bullet \frac{\partial w^a}{\partial x^b} + \alpha_\bullet \frac{\partial w^a}{\partial x^b} \right) e^a.
\]

Note that throughout, the basis one-form \( e^a \) is fixed at \( x \) and thus is un-

\(^2\)For a proof, see, for example, Apostol [1974].
affected by the differentiation. (This is another way of understanding that a metric structure of $S$ is not involved in the notion of Lie derivative.) Also observe that $L_w \alpha$ depends on the values of $w$ in a neighborhood of the point in question since $L_w \alpha$ involves partial derivatives of the components of $w$. This is to be contrasted with covariant differentiation, for which $\nabla_w \alpha$ depends solely on the value of $w$ at the point in question.

(3) Finally, we consider the case of a vector field $v$. Here

$$(\psi_{t,s}^* v_t)(x) = v^a(\psi_{t,s}(x), t) \frac{\partial(\psi_{t,s}^{-1})^b}{\partial x^a}(\psi_{t,s}(x)) e_b(x).$$

To compute the Lie derivative we use the formula

$$\frac{\partial}{\partial t} \left( \frac{\partial(\psi_{t,s}^{-1})^b}{\partial x^a}(\psi_{t,s}(x)) \right) \bigg|_{t=s} = -\frac{\partial w^b}{\partial x^a}$$

which is obtained in the same way as computing the derivative of the inverse of a matrix depending upon a parameter. Thus

$$L_w v = \left( \frac{\partial v^a}{\partial t} + \frac{\partial v^a}{\partial x^b} w^b - v^b \frac{\partial w^a}{\partial x^b} \right) e_a$$

or, in coordinate-free notation,

$$L_w v = \frac{\partial v}{\partial t} + [w, v]$$

where $[ , ]$ is the Lie bracket of vector fields (see Section 1.2).

A general coordinate expression can be given for the Lie derivative of a tensor of arbitrary type, namely,

$$(L_w t)^{ab\ldots c}_{de\ldots f} = \frac{\partial}{\partial t} t^{ab\ldots c}_{de\ldots f} + \frac{\partial}{\partial x^g} t^{ab\ldots c}_{de\ldots f} w^g$$

$$- t^{gb\ldots c}_{de\ldots f} \frac{\partial w^a}{\partial x^g} - (\text{all upper indices})$$

$$+ t^{ab\ldots c}_{ge\ldots f} \frac{\partial w^g}{\partial x^d} + (\text{all lower indices}).$$

This follows by using the computations in the preceding example applied to each index.

The following proposition enables us to express the Lie derivative in terms of the covariant derivative.

6.11 Proposition For a torsion-free connection the partial derivatives of the preceding general formula may be replaced by covariance derivatives; that is,

$$(L_w t)^{ab\ldots c}_{de\ldots f} = \frac{\partial}{\partial t} t^{gb\ldots c}_{de\ldots f} + t^{ab\ldots c}_{de\ldots f} w^g$$

$$- t^{gb\ldots c}_{de\ldots f} w^a_{lg} - (\text{all upper indices})$$

$$+ t^{ab\ldots c}_{ge\ldots f} w^g_{ld} + (\text{all lower indices}).$$
Recall that the torsion, Tor, of a connection $\nabla$ is the $(0, 2)$ tensor defined by $\text{Tor}(v, w) = \nabla_v w - \nabla_w v - [v, w]$. Torsion-free simply means that $\text{Tor} = 0$, from which it follows that the Christoffel symbols $\gamma^a_{bc}$ are symmetric; that is, $\gamma^a_{bc} = \gamma^a_{cb}$. With this fact the proof of the proposition is a simple manipulative exercise, which we omit.

6.12 Corollary  
(i) Let $v$ be the spatial velocity vector field of a motion $\phi_t$, and $g$ denote a Riemannian metric on $\mathbb{S}$. Then $\frac{1}{2} L_v g = d$, the spatial rate-of-deformation tensor.

(ii) Let $f$ be a function on $\mathbb{S}$. Then the material derivative is $\dot{f} = L_v f$. In particular, $L_v J = \dot{J} = (\text{div } v) J$, where $J$ is the Jacobian.

Proof  From the general formula of Proposition 6.11,

$$(L_v g)_{ab} = \frac{\partial g_{ab}}{\partial t} + g_{ac} \omega^c + g_{cb} \omega^c_{la} + g_{ac} \omega^c_{lb}.$$  
(The vertical bar is understood to signify covariant differentiation with respect to the Riemannian connection.) Since $g_{abc} = 0$, we have $(L_v g)_{ab} = v_{bia} + v_{aib}$, from which the result (i) is immediate; (ii) follows similarly.

6.13 Proposition  Let $\mu = \mu(g)$ denote the volume form of $g$ on $\mathbb{S}$. [Recall that $\mu(g) = (\det g_{ab})^{1/2} dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$, where the wedge notation refers to the skew symmetrization of $dx^1 \otimes dx^2 \otimes \cdots \otimes dx^n$.] Then $\nabla_v \mu = 0$ and $L_v \mu = (\text{div } v) \mu$.

This proposition is another way of formulating $J = (\text{div } v) J$. We return to this point and give more details in Section 1.7.

We now enumerate some general properties of the Lie derivative.

6.14 Proposition

(i) If $v$ and $t$ are $C^k$ $(k \geq 1)$, then $L_v t$ is a $C^{k-1}$ tensor field of the same type as $t$.

(ii) Symmetries and skew symmetries of $v$ are also possessed by $L_v t$.

(iii) Let $t_1$ and $t_2$ be tensor fields of the same type. Then $L_v (t_1 + t_2) = L_v t_1 + L_v t_2$.

(iv) $L_v$ is a derivation: $L_v (t_1 \otimes t_2) = L_v t_1 \otimes t_2 + t_1 \otimes L_v t_2$.

(v) Let $w_1$ and $w_2$ be vector fields. Then $\mathcal{L}_{w_1 + w_2} = \mathcal{L}_{w_1} + \mathcal{L}_{w_2}$.

(vi) $L_v$ commutes with contractions.

(vii) $L_v w = \partial w / \partial t$, that is, $\mathcal{L}_v w = 0$.

(viii) Let $f$ be a $C^2$ function and let $df$ denote the differential of $f$; that is, $df = \left( \partial f / \partial x^a \right) dx^a$. Then $L_v df = (\text{div } v) df$.

(ix) If $t$ is a general tensor and $\psi$ is a diffeomorphism, then $\psi^*(L_v t) = \mathcal{L}_{\psi^* v} \psi^* t$.

(x) The following holds: $(d/dt)(\psi^* t_i) = \psi^* (L_v t_i)$.
The proofs are relatively simple verifications, using the definitions or coordinate formulas, so we leave them as exercises.

6.15 **Definition** A map \( \psi : S \rightarrow S \) is called an *isometry* of a metric \( g \), if \( \psi^*g = g \). A vector field \( w \) is a *Killing vector field*\(^{28}\) (or *infinitesimal isometry*) if each map \( \psi_{t,s} \) of the flow of \( w \) is an isometry is \( S \).

6.16 **Proposition** If \( w \) is a Killing vector field, then \( L_w g = 0 \), and conversely.

**Proof** If \( w \) is a Killing field, then
\[
L_w g = \left( \frac{d}{dt} \psi_{t,s}^* g \right) \bigg|_{t=s} = 0
\]
since \( \psi_{t,s} \) is an isometry. The converse follows from 6.14(x).

6.17 **Remark** A Killing vector field \( w \) therefore satisfies *Killing’s Equation*
\[
w_{ab} + w_{ba} = 0 \quad \text{(its covariant derivative is skew symmetric).}
\]

**Problem 6.1** Let \( S = \mathbb{R}^3 \) with the usual metric. Show that \( S \) possesses six linearly independent Killing vector fields: namely, \( \partial/\partial z^i \) (\( i = 1, 2, 3 \)) and \( z^i \partial/\partial z^j - z^j \partial/\partial z^i \) (\( i, j = 1, 2, 3; i \neq j \)). Show that the most general form of a time-dependent isometry corresponding to these Killing vectors is
\[
\psi^i(t) = c^i(t) + Q^i_j(t)z^j(x),
\]
where the \( c^i \)'s are components of a \( C^\infty \) curve \( c : \mathbb{R} \rightarrow S \) and the \( Q^i_j \)'s are the components of a \( C^\infty \) curve of orthogonal matrices.

**Problem 6.2** Consider the vector fields \( t = xi - xyj \) and \( w = x^2i - yj \) in the plane. Calculate \( \mathcal{L}_w t \) using 6.11. Also calculate the flow of \( w \) explicitly and verify that Definition 6.8 holds.

**Problem 6.3** Define \( \mathcal{L}_w V \) for a connection \( V \) [see Problem 4.8(iii)]. Show that \( \mathcal{L}_w V = \nabla \cdot w + w \cdot R \); that is, \( (\mathcal{L}_w V)^a_{b,c} = w^a_{\ b;\ c} + w^d R^a_{\ d\ b\ c} \). (This can be used as a basis for the introduction of curvature.)

---

\(^{28}\)After W. Killing [1892].
with $\partial/\partial t$ and $\nabla$—do not commute with Lie differentiation. This fact is important in understanding this topic. Let $\sigma_1 = \sigma^{ab} e_a \otimes e_b$ be a given tensor field on $\mathcal{S}$ and let $\sigma_2 = \sigma^a_b e^a \otimes e_b$, $\sigma_3 = \sigma^{a}_b e_a \otimes e^b$, and $\sigma_4 = \sigma_{ab} e^a \otimes e^b$, denote its associated tensors (e.g., $\sigma^{ab} = g^{ac} \sigma_c^b$, etc.). Then the following relations hold:

$$(L_v \sigma_1)^{ab} = \dot{\sigma}^{ab} - \sigma^{cb} v^a c - \sigma^{ac} v^b |_c$$

$$g^{ac} (L_v \sigma_2)_c^b = \dot{\sigma}^{ab} - \sigma^{ad} v^b |_d + \sigma^{db} v^a |_d$$

$$(L_v \sigma_3)^a_c g^{cb} = \dot{\sigma}^{ab} - \sigma^{db} v^a |_d + \sigma^{ad} v^b |_d$$

$$g^{ac} (L_v \sigma_4)_c^b = \dot{\sigma}^{ab} + \sigma^{cb} v^a |_b + \sigma^{ac} v^b |_b$$

and

$$(L_v (\sigma_1 \otimes \mu))^{ab} = ((L_v \sigma_1)^{ab} + \sigma^{ab} \text{div } v) \mu$$

where $\dot{\sigma}^{ab} = \partial \sigma^{ab} / \partial t + \sigma^{ab} v c$. The tensor $L_v \sigma_1$ has been associated with the name Oldroyd (see Oldroyd [1950]) and $L_v (\sigma_1 \otimes \mu)$ with the name Truesdell (see Truesdell [1955a, b]). We see that all of these tensors are different manifestations of the Lie derivative of $\sigma$.

Any linear combination of the preceding formulas also qualifies as an “objective flux”; for example,

$$\tfrac{1}{2} ((L_v \sigma_3)_c^a g^{cb} + g^{ac} (L_v \sigma_2)_c^b) = \dot{\sigma}^{ab} + \sigma^{ad} \omega^b_d - \sigma^{db} \omega^a_d$$

where $\omega^a_d$ are associated components of the spin $\omega_{ab} = \tfrac{1}{2} (v^a_{;b} - v^b_{;a})$. This tensor is associated with the name Jaumann (see Jaumann [1911]). We note in passing that, like the Lie derivative in general, the right-hand sides may be expressed without using covariant derivatives. For example,

$$(L_v \sigma_1)^{ab} = \frac{\partial \sigma^{ab}}{\partial t} + \frac{\partial \sigma^{ab}}{\partial x^c} v^c - \sigma^{cb} \frac{\partial v^a}{\partial x^c} - \sigma^{ac} \frac{\partial v^b}{\partial x^c}.$$

And if $\sigma^{ab} = \sigma^{ba}$, then

$$(L_v \sigma_1)^{ab} = \frac{\partial \sigma^{ab}}{\partial t} + \frac{\partial \sigma^{ab}}{\partial x^c} v^c - 2 \times \text{symmetric part of } (\sigma^{ac} \frac{\partial v^b}{\partial x^c}).$$

(One seems never to see this in practice, but it could yield savings in numerical computations in that it is unnecessary to compute the Christoffel symbols.)

We conclude this box with a general discussion of what we mean by objective.

**6.18 Definition** Let $t$ be a tensor field (or tensor density) on a manifold $\mathcal{S}$ and $\psi$ a diffeomorphism of $\mathcal{S}$ to another manifold $\mathcal{S}'$. We say that $t' = \psi_\ast t$ is the objective transformation of $t$ (i.e., $t$ transforms in the usual way under the map $\psi$).

---

*Some of these results were known to Sedov around 1960.*
6.19 Theorem Let $\phi_1$ be a regular motion of $\mathcal{B}$ in $\mathcal{S}$ with spatial velocity field $v$. Let $\xi_1$ be a motion of $\mathcal{S}$ in $\mathcal{S}'$, and let $\phi_i' = \xi_i \circ \phi_i$ be the superposed motion of $\mathcal{B}$ in $\mathcal{S}'$.

Let $t$ be a given time-dependent tensor field on $\mathcal{S}$ and let $t' = \xi_* t$, that is, transform $t$ objectively.

Let $v'$ be the velocity field of $\phi_i'$. Then

$$L_v t' = \xi_* (L_v t).$$

Thus, "objective tensors (or tensor densities) have objective Lie derivative." This is remarkable since $v$ itself is not objective as we shall see immediately in the proof.

**Proof** We first note that $v' = w_i + \xi_i \ast v_i$ where $w_i$ is the spatial velocity of $\xi_i$. This follows by differentiating $\phi_i(X) = \xi_i(\phi_i(X))$ in $t$.

Now we compute

$$L_v t' = L_{w + \xi_* v} (\xi_* t) = \xi_* (L_v t) + \frac{\partial}{\partial t} (\xi_* t)$$

$$= \xi_* (L_v t) + \xi_* (\xi_* t) + \frac{\partial}{\partial t} (\xi_* t)$$

$$= \xi_* (L_v t) + \frac{d}{dt} (\xi_i \circ \xi^{-1}(\xi_i \circ \xi^{-1} t))|_{t=s}$$

(see 6.6)

$$= \xi_* (L_v t) + \frac{d}{dt} \xi_{s*} t_i|_{t=s}$$

$$= \xi_* \left( L_v t + \frac{d}{dt} t_i|_{t=s} \right) = \xi_* (L_v t).$$

In order to master the proof fully, a coordinate computation may also be done. We do so in case $t$ is a vector field $t$. Let $\{x^a\}$ be coordinates on $\mathcal{S}$, $\{\xi^a\}$ be coordinates on $\mathcal{S}'$, and write $\xi^a$ for the coordinates of $\xi$. From $\phi_i = \xi^a(\phi_i(X, t))$ we get $v'^a = w^a + (\partial \xi^a/\partial x^a)v^a$. Since $t^a$ is objective, $t'^a = (\partial \xi^a/\partial x^a)t^a$. From our coordinate formulas for the Lie derivative,

$$(L_v t')^a = \frac{\partial t'^a}{\partial t} + \frac{\partial t'^a}{\partial \xi^b} v'^b - t'^b \frac{\partial v'^a}{\partial \xi^b}.$$
The second term is there because \( t' \) has an explicit time dependence through \( \xi^{-1}_t \). All terms cancel except those for \( (\partial \xi^p / \partial x^a) (L_v t) \) — that is, \( \xi_v(L_v t) \) — so we get the result. The reader may do the same computation for one-forms or two-tensors.

As a corollary, all the “objective fluxes” discussed earlier are objective tensors with this proviso: if the metric tensor \( g_{ab} \) or \( g^{ab} \) appears explicitly on the left hand side and the rates are to transform like tensors with the same \( g_{ab} \) resulting after the transformation, \( \zeta \) must be an isometry at the point of interest. In this sense, there are two levels of objectivity, objectivity with respect to diffeomorphisms, and objectivity with respect to isometries. The Oldroyd and Truesdell rates are objective with respect to diffeomorphisms while the remaining rates discussed above are objective with respect to isometries. Rates which are objective with respect to diffeomorphisms are called covariant. This subject of covariance is taken up in Sections 2-4 and 3-3.

---

Box 6.2 Summary of Important Formulas for Section 1.6

**Tangent Vector to a Curve**
\[ c: S \to S \]
\[ \frac{dc}{dt} = Tc(t, 1) \in T_{c(t)} S \]

**Integral Curve of \( w \)**
\[ \frac{dc}{dt} = w(c(t)) \]

**Flow (Evolution operator) \( \psi_{t,s} \) of \( w \)**
\[ \frac{d\psi_{t,s}(x)}{dt} = w(\psi_{t,s}(x)) \]
\[ \psi_{s,s}(x) = x \]

\[ (\xi_v(L_v t)) = \xi_v(t) \]

\( (d\xi^p / d\xi^a) (L_v t) \) = \( x \)
Flow Associated with a Motion $\phi_t$

$\phi_{t,s} = \phi_{t} \circ \phi_{s}^{-1}$ = flow of the spatial velocity $v_t$

Lie Derivative of a Time-Dependent Tensor $t$

$$L_w t = \frac{\partial t}{\partial t} + \mathbf{L}_w t$$

$$\mathbf{L}_w t = \left. \frac{d}{dt} (\psi_{t,s}^*, t_s) \right|_{t = s},$$

where $\psi_{t,s}$ is the flow of $w$

$(L_w t)^{ab \cdots c}_{de \cdots f}$

$$= \frac{\partial}{\partial t} t^{ab \cdots c}_{de \cdots f} + \frac{\partial}{\partial x^g} t^{ab \cdots c}_{de \cdots f} \frac{\partial w^g}{\partial x^d}$$

$$- (\text{all upper indices})$$

$$+ t^{ab \cdots c}_{de \cdots f} \frac{\partial w^g}{\partial x^d}$$

$$+ (\text{all lower indices})$$

$$= \frac{\partial}{\partial t} t^{ab \cdots c}_{de \cdots f} + t^{ab \cdots c}_{de \cdots f} \frac{\partial w^g}{\partial x^d}$$

$$- (\text{all upper indices})$$

$$+ t^{ab \cdots c}_{de \cdots f} \frac{\partial w^g}{\partial x^d}$$

$$+ (\text{all lower indices})$$

(Using any torsion-free connection).

For a vector field, $L_w v = \frac{\partial v}{\partial t} + [w, v] = \frac{\partial v}{\partial t} + (v \cdot \nabla v)$.

For a differential form $\alpha$, $L_w \alpha = d_i \alpha + i_v \alpha$ (see Section 1.7).

Rate of Deformation Tensor

$$d = \frac{1}{2} L_v g$$

$$d_{ab} = \frac{1}{2} (v_{abl} + v_{bla}) = \frac{1}{2} (L_v g)_{ab}$$

Material Derivative

$$\dot{f} = L_v f$$

$$\dot{f} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^a} v^a$$

Jacobic and Lie Derivatives

$L_v J = J = (\text{div } v) J$

$J = (v^a) J$

$J = \text{Jacobian determinant of } \phi_t$

$J = \text{det}(F^a_s)$

Properties of Lie Differentiation

$L_v (t_1 + t_2) = L_v t_1 + L_v t_2$

$L_v (t_1 \times t_2) = L_v t_1 \times t_2 + t_1 \times L_v t_2$

$L_v df = dL_v f$

$$\mathbf{L}_{w_1 + w_2} t = \mathbf{L}_{w_1} t + \mathbf{L}_{w_2} t$$

$$\psi^*(\mathbf{L}_w t) = \mathbf{L}_{w^*} \psi^* t$$

$$\frac{d}{dt} (\psi_{t*, t}) = \psi_{t_1, t}(L_v t)$$

Killing's Equations

$L_v g = 0$

$w_{abl} + w_{bla} = 0$
(The flow of \( w \) consists of isometries if and only if Killing's equations hold.)

**Objective Tensors**

Transformation rule

\[ t' = \xi_\ast t \]

\[(t')^{\alpha_1 \ldots \alpha_p}_{\mu_1 \ldots \mu_q} = \frac{\partial \xi^\alpha}{\partial x^\mu} \ldots \frac{\partial \xi^\nu}{\partial x^\mu} t^{ab \ldots c}_{ef \ldots g} \frac{\partial x^e}{\partial \xi^\alpha} \ldots \frac{\partial x^g}{\partial \xi^\nu} \]

**Lie Derivatives**

\[ L_\nu t' = \xi_\ast L_\nu t \]

(The Lie derivative of an objective tensor is objective.)

1.7 DIFFERENTIAL FORMS AND THE PIOLA TRANSFORMATION

Skew-symmetric covariant tensors are called differential forms. They have a rich algebraic and differential structure with many applications to the physical sciences.\(^{30}\) We shall consider some of these here. One of the principal applications is to the Piola transformation—a fundamental operation relating the material and spatial descriptions of a continuous medium. We shall consider a few other applications as well, to Hamiltonian systems and variational principles; these are treated with more specific reference to continuum mechanics in Chapter 5. We shall also consider volume elements, integration, Stokes’ theorem and Gauss’ theorem in the language of differential forms.

7.1 **Definition** A \( k \)-form on a manifold \( \mathcal{M} \) is a \[ \left( \begin{array}{c} 0 \\ k \end{array} \right) \] tensor \( \alpha \) on \( \mathcal{M} \) that is skew symmetric; that is, for \( X \in \mathcal{M} \),

\[ \alpha_X : T_X \mathcal{M} \times \cdots \times T_X \mathcal{M} \to \mathbb{R} \]

\( k \) copies

is a multilinear mapping and

\[ \alpha_X(W_{\pi(1)}, W_{\pi(2)}, \ldots, W_{\pi(k)}) = (\text{sgn} \ \pi) \alpha_X(W_1, \ldots, W_k) \]

for any \( W_1, \ldots, W_k \in T_X \mathcal{M} \) and any permutation \( \pi \) on \( \{1, \ldots, k\} \), where \( \text{sgn} \ \pi \) is the sign of \( \pi \) (+1 or −1 according to whether \( \pi \) is an even or odd permutation). One also refers to a \( k \)-form as a **differential \( k \)-form** or simply as a **differential form**. We write \( \alpha(X) = \alpha_X \) for the value of \( \alpha \) at \( X \in \mathcal{M} \), and leave off the argument \( X \) when convenient.

In terms of components, skew symmetry means the following: if any two indices of the components \( \alpha_{A_1 \ldots A_k} \) are interchanged, then \( \alpha_{A_1 \ldots A_k} \) changes sign.

---

\(^{30}\)See, for example, Cartan [1922], Gallissot [1958], Flanders [1963], and Abraham and Marsden [1978].
If we skew-symmetrize the tensor product, we get the following basic algebraic operation on differential forms.

**7.2 Definition** Let $\alpha$ be a $k$-form and $\beta$ an $l$-form. Define the $(k + l)$-form $\alpha \wedge \beta$, called their wedge or exterior product, by

\[
(\alpha \wedge \beta)(W_1, \ldots, W_k, W_{k+1}, \ldots, W_{k+l}) = \\
\frac{1}{k!l!} \sum_{\pi} (\text{sgn } \pi) \alpha(W_{\pi(1)}, \ldots, W_{\pi(k)}) \beta(W_{\pi(k+1)}, \ldots, W_{\pi(k+l)}).
\]

There are several possible choices of normalization factor. The one chosen here gives the equivalent formula

\[
(\alpha \wedge \beta)(W_1, \ldots, W_k, W_{k+1}, \ldots, W_{k+l}) = \\
\sum' (\text{sign } \pi) \alpha(W_{\pi(1)}, \ldots, W_{\pi(k)}) \beta(W_{\pi(k+1)}, \ldots, W_{\pi(k+l)})
\]

where $\sum'$ denotes the sum over permutations satisfying $\pi(1) < \ldots < \pi(k)$ and $\pi(k + 1) < \ldots < \pi(k + l)$.

The wedge product $\wedge$ is associative, $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$, and satisfies the following commutation relation: $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$. These are algebraic verifications that we shall omit. (If you get stuck, consult, for example, Abraham and Marsden [1978].) Some algebraic relationships which help in understanding the wedge product are contained in the following problem.

**Problem 7.1**

(i) If $\alpha$ is a two-form and $\beta$ a one-form, show that

\[
(\alpha \wedge \beta)(W_1, W_2, W_3) = \alpha(W_1, W_2)\beta(W_3) - \alpha(W_1, W_3)\beta(W_2) + \alpha(W_2, W_3)\beta(W_1).
\]

(ii) If $\alpha$, $\beta$, and $\gamma$ are one-forms,

\[
(\alpha \wedge \beta \wedge \gamma)(W_1, W_2, W_3) = \alpha(W_1)\beta(W_2)\gamma(W_3) + \alpha(W_2)\beta(W_3)\gamma(W_1) + \alpha(W_3)\beta(W_1)\gamma(W_2) - \alpha(W_1)\beta(W_3)\gamma(W_2) - \alpha(W_2)\beta(W_1)\gamma(W_3) - \alpha(W_3)\beta(W_2)\gamma(W_1).
\]

(iii) In coordinates, if $\alpha$ is a $k$-form, show that

\[
\alpha = \frac{1}{k!} \alpha_{A_1 \ldots A_k} dX^{A_1} \wedge \cdots \wedge dX^{A_k}
\]

\[
= \sum_{A_1 < \cdots < A_k} \alpha_{A_1 \ldots A_k} dX^{A_1} \wedge \cdots \wedge dX^{A_k}
\]

\[
\text{31 Some books (e.g. Kobayashi and Nomizu [1963]) use the factor } 1/(k + l)! \text{ here. This convention leads to awkward factors in subsequent formulas however.}
\]
where the summation is over all indices $A_1, \ldots, A_k$ satisfying $A_1 < \cdots < A_k$.

(iv) In $\mathbb{R}^3$, if $\alpha$ and $\beta$ are one-forms, show that the coefficients of $\alpha \wedge \beta$ in the standard basis are components of the cross product of $\alpha$ and $\beta$.

(v) For mappings $\phi : \mathcal{M} \to \mathcal{N}$ and $\psi : \mathcal{L} \to \mathcal{M}$, and differential forms $\alpha$ and $\beta$ on $\mathcal{K}$, verify that $\phi^*(\alpha \otimes \beta) = \phi^*\alpha \otimes \phi^*\beta$, $\phi^*(\alpha \wedge \beta) = \phi^*\alpha \wedge \phi^*\beta$, and $(\phi \circ \psi)^*\alpha = \psi^*(\phi^*\alpha)$.

7.3 Definition If $W$ is a vector field on $\mathcal{M}$ and $\alpha$ is a $k$-form, the contraction of $W$ with the first index of $\alpha$ is called the interior product and is denoted by $i_W \alpha$ or $W \lrcorner \alpha$.

Thus $i_W \alpha$ is a $(k-1)$-form given by

$$
(i_W \alpha)(W_2, \ldots, W_k) = \alpha(W, W_2, \ldots, W_k), \quad \text{or} \quad (i_W \alpha)_{A_1 \ldots A_k} = W^{A_1} \alpha_{A_1 \ldots A_k}.
$$

Some properties of this contraction are given in the next problem:

**Problem 7.2** (i) On $\mathbb{R}^3$ and in Euclidean coordinates, let $dv$ be the volume element ($dv = dx \wedge dy \wedge dz$), and let $\alpha = \alpha_x dx + \alpha_y dy + \alpha_z dz$ and $\beta = \beta_x dx + \beta_y dy + \beta_z dz$. Let $v = (\alpha_x \beta_z - \alpha_z \beta_x) i - (\alpha_x \beta_y - \alpha_y \beta_x) j + (\alpha_y \beta_z - \alpha_z \beta_y) k$, be the cross product. Show that $\alpha \wedge \beta = i_v dv$.

(ii) Prove that if $\alpha$ is a $k$-form, $i_W (\alpha \wedge \beta) = (i_W \alpha) \wedge \beta + (-1)^k \alpha \wedge i_W \beta$.

(iii) Show that

$$
i_W \alpha = \frac{1}{(k-1)!} \omega^{a_1 \ldots a_k} \alpha_{a_1 \ldots a_k} dx^{a_1} \wedge \cdots \wedge dx^{a_k}.
$$

The three most important types of differentiation in tensor analysis are: covariant differentiation $\nabla$, Lie differentiation $\mathcal{L}_v$, and exterior differentiation $d$. We have already met the first two and now turn to $d$. The exterior derivative generalizes the notion of gradient, divergence, and curl to differential forms. We want it to capture the identities $\text{div} \, \text{curl} = 0$ and $\text{curl} \, \text{grad} = 0$ as well as $d$ being a derivative.

7.4 Theorem Given a manifold $\mathcal{M}$, there is a unique linear operator $d$ taking (smooth) $k$-forms $\alpha$ on $\mathcal{M}$ to (smooth) $(k+1)$-forms $d\alpha$ on $\mathcal{M}$ such that:

(i) $d(d\alpha) = 0$;

(ii) $d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta$;

(iii) for functions $f$, $df$ coincides with the differential of $f$ as defined in Section 1.2; and

(iv) if $U \subset \mathcal{M}$ is open, then $d(\alpha|U) = (d\alpha)|U$.
Proof First we show uniqueness by using these properties to derive a formula for $d\alpha$. Let

$$\alpha = \frac{1}{k!} \alpha_{A_1 \ldots A_k} dX^{A_1} \wedge \cdots \wedge dX^{A_k}.$$  

(we can pass to a local chart by virtue of (iv)). Then since $d dX^A = 0$ we get

$$d d\alpha = \frac{1}{k!} \frac{\partial^2 \alpha_{A_1 \ldots A_k}}{\partial X^B \partial X^C} dX^C \wedge dX^B \wedge dX^{A_1} \wedge \cdots \wedge dX^{A_k}.$$  

Thus, if there is such a $d$, it is unique. To show that $d$ exists, define it by this formula. One has to show that $d\alpha$ is a tensor (i.e., transforms properly) and satisfies (i), (ii), and (iii). These are all straightforward. For example, to prove (i), we use the above formula for $d$ twice:

$$d d\alpha = \frac{1}{k!} \frac{\partial^2 \alpha_{A_1 \ldots A_k}}{\partial X^C \partial X^B} dX^C \wedge dX^B \wedge dX^{A_1} \wedge \cdots \wedge dX^{A_k}.$$  

But

$$\frac{\partial^2 \alpha_{A_1 \ldots A_k}}{\partial X^C \partial X^B}$$

is symmetric in $B$ and $C$ (by the equality of mixed partial derivatives) and $dX^C \wedge dX^B$ is skew symmetric. Therefore, the sum vanishes.

There is a useful coordinate-free formula for $d$ (due to Palais [1954]): If $W_0, W_1, \ldots, W_k$ are vector fields, then

$$d\alpha(W_0, \ldots, W_k) = \sum_{i=0}^{k} (-1)^i W_i [\alpha(W_0, \ldots, W_{i-1}, W_{i+1}, \ldots, W_k)]$$

$$+ \sum_{i<j} (-1)^{i+j} \alpha([W_i, W_j], W_0, \ldots, W_k).$$

This can be verified using the coordinate expressions of each side; we omit the details.

**Problem 7.3** In Euclidean coordinates $(x, y, z)$ on $\mathbb{R}^3$, establish:

(i) $\text{grad} f = (df)'$;

(ii) $\text{div} \, \mathbf{v}$ is such that $d(i_v \, dv) = (\text{div} \, \mathbf{v}) \, dv$, where $dv = dx \wedge dy \wedge dz$; and

(iii) $d(v^x) = i_{\mathbf{v} \times x} \, dv$.

Use $dd = 0$ to show that $\text{curl} \, \text{grad} = 0$ and $\text{div} \, \text{curl} = 0$.  

**Problem 7.4** On $\mathbb{R}^4$ with coordinates $(t, x, y, z)$ let $F = E^t \wedge dt + i_B(dx \wedge dy \wedge dz)$ (the Faraday two-form), and $*F = -B^t \wedge dt + i_B(dx \wedge dy \wedge dz)$ (the Maxwell two-form). Show that Maxwell's equations may be written $dF = 0$, $d(*F) = 4\pi*J$, where $*J = i_J(dt \wedge dx \wedge dy \wedge dz)$ and $J^t = \rho \, dt + j_a \, dx^a$.

**7.5 Proposition** If $\phi : \mathcal{M} \to \mathcal{N}$ is a (smooth) mapping and $\alpha$ is a $k$-form on $\mathcal{N}$, then $\phi^* d\alpha = d\phi^* \alpha$; that is, pull-back commutes with exterior differentiation.

**Proof** First we verify this for functions $f$. But $\phi^* f = f \circ \phi$ and so, by the chain rule, $d(\phi^* f) = d(f \circ \phi) = df \circ T\phi = \phi^* df$. In general, let

$$\alpha = \frac{1}{k!} \alpha_{a_1 \ldots a_k} \, dx^{a_1} \wedge \ldots \wedge dx^{a_k}.$$ 

Since $\phi^* (\alpha \wedge \beta) = \phi^* \alpha \wedge \phi^* \beta$ and $d\phi^* f = \phi^* df$,

$$\phi^* \alpha = \frac{1}{k!} \alpha_{a_1 \ldots a_k} \circ \phi \, d\phi^{a_1} \wedge \ldots \wedge d\phi^{a_k}$$

(where $\phi^a$ stands for $x^a \circ \phi$). Using properties of $d$,

$$d\phi^* \alpha = \frac{1}{k!} \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} \frac{\partial \phi^b}{\partial X^A} \, dX^A \wedge d\phi^{a_1} \wedge \ldots \wedge d\phi^{a_k}$$

$$= \frac{1}{k!} \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} \phi^a (dx^b) \wedge d\phi^{a_1} \wedge \ldots \wedge d\phi^{a_k}$$

$$= \phi^* \left( \frac{1}{k!} \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} \, dx^b \wedge d\phi^{a_1} \wedge \ldots \wedge d\phi^{a_k} \right) = \phi^* d\alpha.$$

Next we shall establish the basic link between the exterior derivative $d$ and the Lie derivative $\mathcal{L}$.

**7.6 Theorem** Let $\alpha$ be a $k$-form on $\mathcal{N}$ and $w$ a vector field on $\mathcal{N}$. Then

$$\mathcal{L}_w \alpha = di_w \alpha + i_w d\alpha,$$

that is, $\mathcal{L}_w = di_w + i_w d$.

This formula is one of several "magic formulas" of Cartan [1922].

**Proof** From Section 1.6 we have

$$(\mathcal{L}_w \alpha)_{a_1 \ldots a_k} = \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} w^b + \alpha_{[a_1 \ldots a_k} \frac{\partial w^{a_k]}}{\partial x^a} + \text{(all lower indices)}.$$ 

On the other hand, from our formulas for $d$ and $i$,

$$(di_w \alpha + i_w d\alpha) = d\left( \frac{1}{(k-1)!} w^b \alpha_{a_1 \ldots a_k} \, dx^{a_1} \wedge \ldots \wedge dx^{a_k} \right)$$

$$+ \frac{1}{k!} \left( w^b \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} \, dx^{a_1} \wedge \ldots \wedge dx^{a_k} \right.$$  

$$- w^{a_1} \frac{\partial \alpha_{a_2 \ldots a_k}}{\partial x^b} \, dx^b \wedge dx^{a_1} \wedge \ldots \wedge dx^{a_k}$$

$$+ w^{a_2} \frac{\partial \alpha_{a_1 \ldots a_k}}{\partial x^b} \, dx^b \wedge dx^{a_1} \wedge \ldots \wedge dx^{a_k} - \ldots + \ldots \right).$$
In this expression the derivative from $\alpha$ in the first term cancels the last group of $k$ terms (which are all equal). Therefore, this simplifies to

$$\frac{1}{(k - 1)!} \frac{\partial w}{\partial x^a} \alpha_{ba_2 \cdots a_k} \, dx^a_1 \wedge \cdots \wedge dx^a_k$$

$$+ \frac{1}{k!} \frac{\partial \alpha_{ba_2 \cdots a_k}}{\partial x^b} \, dx^a_1 \wedge \cdots \wedge dx^a_k.$$ 

Since the expression for $(\mathcal{L}_a \alpha)_{a_2 \cdots a_k}$ is skew symmetric in $a_1, \ldots, a_k$, 

$$\mathcal{L}_a \alpha = \frac{1}{k!} w^b \frac{\partial \alpha_{ba_2 \cdots a_k}}{\partial x^b} \, dx^a_1 \wedge \cdots \wedge dx^a_k$$

$$+ \frac{1}{(k - 1)!} \frac{\partial \alpha_{ba_2 \cdots a_k}}{\partial x^b} \, dx^a_1 \wedge \cdots \wedge dx^a_k$$

since all the lower index terms are equal. Thus the two expressions agree.

7.7 Corollary \( \mathcal{L}_a d = d \mathcal{L}_a. \)

Proof

$$\mathcal{L}_a d \alpha = d \mathcal{L}_a (d \alpha) + i_a d (d \alpha) = d (i_a d \alpha + d i_a \alpha) = d \mathcal{L}_a \alpha.$$

Box 7.1 Summary of Identities Relating $d$, $i_\nu$, and $\mathcal{L}_\nu$ for Differential Forms

1. (a) $d \circ d = 0$, \hspace{1em} (b) $i_\nu i_\mu = 0$
2. (a) \(d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta, \) $\alpha$ a $k$-form
   \hspace{1em} (b) \(i_\nu(\alpha \wedge \beta) = i_\nu \alpha \wedge \beta + (-1)^k \alpha \wedge i_\nu \beta$
   \hspace{1em} (c) \(\mathcal{L}_\nu (\alpha \wedge \beta) = \mathcal{L}_\nu \alpha \wedge \beta + \alpha \wedge \mathcal{L}_\nu \beta$
3. (a) \(\phi^*(\alpha \wedge \beta) = \phi^* \alpha \wedge \phi^* \beta$
   \hspace{1em} (b) \(\phi^* d\alpha = d \phi^* \alpha$
   \hspace{1em} (c) \(\phi^* i_\nu \alpha = i_{\phi^* \nu} \phi^* \alpha \) (if $\phi$ is regular)
   \hspace{1em} (d) \(\phi^* \mathcal{L}_\nu \alpha = \mathcal{L}_{\phi^* \nu} \phi^* \alpha \) (if $\phi$ is regular)
4. (a) \(d \alpha (w_0, w_1, \ldots, w_k) = \sum_{i=0}^k (-1)^i w_i [\alpha (w_{i-1}, \ldots, w_k)] \)
   \hspace{1em} \text{and in a chart,}$w_i$ missing
   \hspace{1em} \(+ \sum_{i<j} (-1)^{i+j} \alpha ([w_i, w_j], w_{i}, \ldots, w_k)$
   \hspace{1em} $w_j$ missing
(b) \(d \alpha (w_0, \ldots, w_k) = \sum_{i=0}^k (-1)^i D \alpha_\nu \cdot \underbrace{w_i}_{w_i \text{ missing}} (w_{i-1}, \ldots, w_k)$
5. \( (\mathcal{L}_w \alpha)(w_1, \ldots, w_k) = w[\alpha(w_1, \ldots, w_k)] - \sum_{i=1}^{k} \alpha(w_1, \ldots, [w, w_i], \ldots, w_k) \)

\([w, w_i]\) is in the \(i\)th slot of \(\alpha\).

6. (a) \( \mathcal{L}_w = i_w d + di_w \)
(b) \( \mathcal{L}_w d = d \mathcal{L}_w \)
(c) \( \mathcal{L}_{[w, v]} \alpha = \mathcal{L}_w \mathcal{L}_v \alpha - \mathcal{L}_v \mathcal{L}_w \alpha \)
(d) \( i_{[w, v]} \alpha = \mathcal{L}_w i_v \alpha - i_v \mathcal{L}_w \alpha \)
(e) \( \mathcal{L}_{f \cdot w} \alpha = f \mathcal{L}_w \alpha + df \wedge i_w \alpha \)

In the boxes that follow we present two applications of differential forms which are of interest in mechanics.

**Box 7.2 The Poincaré Lemma and Variational Principles**

A differential form \(\alpha\) is called *closed* if \(d\alpha = 0\) and *exact* if \(\alpha = df\) for a form \(f\). Since \(d^2 = 0\), every exact form is closed. The converse is not true. (For example, the form \(\alpha = x dy - y dx\) restricted to the unit circle is closed but is not exact, for its line integral—in the sense of advanced calculus—is \(2\pi \neq 0\). Locally \(\alpha = d\theta\), where \(\theta\) is angular measure.)

7.8 Poincaré Lemma Let \(\alpha\) be a closed \(k\)-form on a manifold \(\mathcal{M}\). Then in some neighborhood of each point, \(\alpha\) is exact.

*Proof* (cf. Moser [1965]) Using coordinates, it suffices to prove the lemma on a ball \(U\) centered at the origin in \(\mathbb{R}^n\). Consider the radial motion given by \(\phi_t(X) = tX\). For \(t > 0\), \(\phi_t\) is a regular mapping of \(\mathbb{R}^n\) to itself. The velocity field of \(\phi_t\) is \(v_t(x) = x/t\). (\(x\) means the vector from the origin to the point \(x\).) From \(\mathcal{L}_v = i_v d + di_v\) and the flow definition of Lie derivatives,

\[
\frac{d}{dt} \phi_t^* \alpha = \phi_t^* \mathcal{L}_v \alpha = \phi_t^* (di_v \alpha + i_v d\alpha) = \phi_t^* (di_v \alpha) = \phi_t^* (d(\phi_t^* i_v \alpha) = d(\phi_t^* i_v \alpha) \quad \text{(by 7.5)}.
\]

Integrating from \(t = 0^+\) to \(t = 1\), noting \(\phi_1 = \text{Identity}\), gives \(\alpha = d \int_0^1 (\phi_t^* i_v \alpha) \, dt\), so we can choose \(\beta = \int_0^1 \phi_t^* i_v \alpha \, dt\). Explicitly,

\[
\beta_{\alpha}(w_1, \ldots, w_{k-1}) = \int_0^1 t^{k-1} \alpha_{tx}(x, w_1, \ldots, w_{k-1}) \, dt.
\]
It follows that in $\mathbb{R}^3$, if $\nabla \times \mathbf{v} = 0$, then $\mathbf{v} = \nabla f$ for some $f$ and that if $\text{div} \mathbf{v} = 0$, then $\mathbf{v} = \nabla \times \mathbf{w}$ for some $\mathbf{w}$; these are well-known results in vector calculus.

**Note.** We proved the Poincaré lemma for finite dimensional manifolds; essentially the same proof works in infinite dimensional spaces as well.

We now apply the Poincaré lemma to a problem in the calculus of variations. (We return to this topic in Chapter 5.) Let us first recall how the Euler–Lagrange equations for a Lagrangian $\mathcal{L}(q, \dot{q})$ may be regarded as the equations for a critical point. Let $\mathcal{Q}$ denote the space of all paths $q(t) \in \mathbb{R}^n$ with $0 \leq t \leq T$ and $q(0)$ and $q(T)$ fixed at specified values. Then $\mathcal{Q}$ is an infinite-dimensional space, but let us apply calculus to it in any case. (This points out the need to generalize our ideas to infinite dimensions which we do in Chapter 4.) The tangent space to $\mathcal{Q}$ is obtained by differentiating a curve $q_1$ in $\mathcal{Q}$ and thus consists of all vector functions $h(\cdot)$, which are zero at $t = 0$ and $t = T$. Define

$$L : \mathcal{Q} \to \mathbb{R} \ \text{by} \ \ L(q) = \int_0^T \mathcal{L}(q(t), \dot{q}(t)) \, dt.$$  

**7.9 Proposition** A curve $q(\cdot) \in \mathcal{Q}$ is a critical point of $L$; that is, $DL(q) = 0$ if and only if the Euler–Lagrange equations hold:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \frac{\partial \mathcal{L}}{\partial q^i} = 0 \quad (i = 1, \ldots, n).$$

(All functions appearing are assumed continuous.)

**Proof** Let us differentiate $L$ in the direction $h$ using the chain rule:

$$DL(q) \cdot h = \frac{d}{d\lambda} L(q + \lambda h)_{|\lambda = 0}$$

$$= \frac{d}{d\lambda} \int_0^T \mathcal{L}(q(t) + \lambda h(t), \dot{q}(t) + \lambda \dot{h}(t)) \, dt \bigg|_{\lambda = 0}$$

$$= \int_0^T \left( \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \cdot h^i + \frac{\partial \mathcal{L}}{\partial q^i} \cdot \dot{h}^i \right) \, dt.$$

Since $h$ vanishes at $t = 0$ and $t = T$, we can integrate by parts to get

$$DL(q) \cdot h = \int_0^T \left( \frac{\partial \mathcal{L}}{\partial q^i} \cdot \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right) \cdot h(t) \, dt.$$

Thus if the Euler–Lagrange equations hold, obviously $DL(q) = 0$. Conversely, if $DL(q) \cdot h = 0$ for all $h$ and if $\frac{\partial \mathcal{L}}{\partial \dot{q}^i} - (d/dt)(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}) \neq 0$, then choose a neighborhood about a point $t_0$ where $\frac{\partial \mathcal{L}}{\partial \dot{q}^i} - (d/dt)(\frac{\partial \mathcal{L}}{\partial \dot{q}^i})$ is nowhere zero (by continuity) and choose a parallel vector $h(t) \neq 0$ in this neighborhood and zero outside; then $DL(q) \cdot h \neq 0$. Therefore, if $DL(q) = 0$, the Euler–Lagrange equations must hold. 

Abstracting this, let $\mathcal{X}$ be a Banach space, let $\langle \cdot, \cdot \rangle$ be a bilinear form (e.g., an inner product) on $\mathcal{X}$, and let $A: \mathcal{X} \to \mathcal{X}$ be a given (nonlinear) operator.

**7.10 Definition** We say $A$ is a potential operator if there is a function $L: \mathcal{X} \to \mathbb{R}$ such that $dL(x) \cdot v = \langle A(x), v \rangle$ for all $x$ in $\mathcal{X}$ and $v \in \mathcal{X}$.

In view of 7.9, $A(x) = 0$ represents the Euler–Lagrange equations for $x \in \mathcal{X}$, in abstract form. The next theorem is due to Vainberg [1964] (although the present proof is due to the authors—see Hughes and Marsden [1977]).

**7.11 Theorem** A given operator $A$ is a potential operator if and only if for each $x \in \mathcal{X}$, $v_1$ and $v_2 \in \mathcal{X}$,

$$ \langle DA(x) \cdot v_1, v_2 \rangle = \langle DA(x) \cdot v_2, v_1 \rangle. $$

If $\langle \cdot, \cdot \rangle$ is symmetric, this is equivalent to saying $DA(x)$ is a symmetric linear operator on $\mathcal{X}$.

**Proof** Consider the one-form $\alpha(x) \cdot v = \langle A(x), v \rangle$. By definition, $A$ is a potential operator if and only if $\alpha$ is exact. By the Poincaré lemma, this is the case if and only if $d\alpha = 0$. But by Formula 4(b) in Box 7.1,

$$ d\alpha(x) \cdot (v_1, v_2) = \langle DA(x) \cdot v_1, v_2 \rangle - \langle DA(x) \cdot v_2, v_1 \rangle, $$

so the result follows immediately.

It may be instructive for the reader to write out an explicit proof using $L(x) = \int_0^1 \langle A(tx), x \rangle dt$.

This result gives necessary and sufficient conditions for a given set of equations to be the Euler–Lagrange equations for some Lagrangian.\(^{32}\)

---

**Box 7.3 A Geometric Formulation of Hamiltonian Mechanics**

In Chapter 5 we study elastodynamics as an infinite-dimensional Hamiltonian system. Here we indicate briefly how differential forms and Lie derivatives can be used in classical mechanics (see Arnold [1978] and Abraham and Marsden [1978] for further details).

---

\(^{32}\)To ensure that $L$ comes from a Lagrangian density requires the further assumption that the operator $A$ is a local operator. Then $L$ will be a density and $A$ will be the usual Euler–Lagrange operator if $\langle \cdot, \cdot \rangle$ is the $L^2$ inner product. For related work, see Lawruk and Tulczyjew [1977] and Takens [1977].
Consider a given Hamiltonian function \( H(q, p) \) for \( q \in \mathbb{R}^n, p \in \mathbb{R}^n \) (i.e., \( p = p_i \, dx^i \) is to be regarded as a one-form), and the associated Hamiltonian equations

\[
\begin{align*}
\frac{dq_i}{dt} &= \frac{\partial H}{\partial p_i}, \\
\frac{dp_i}{dt} &= -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \ldots, n)
\end{align*}
\]

(Recall that if \( H = (1/2m)\langle p, p \rangle + V(q) \), these reduce to Newton’s second law: \( dq^i/dt = p_i/m \) and \( m(d^2q^i/dt^2) = -\partial V/\partial q^i \).

Let \( X_H \) be the vector field for Hamilton’s equations,

\[
X_H = \left( \frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i} \right).
\]

Then \((q(t), p(t))\) is an integral curve of \( X_H \) if and only if Hamilton’s equations hold.

Let \( \omega = dq^i \wedge dp_i \) denote the fundamental two-form. Observe that it has the matrix of components given by the skew-symmetric matrix

\[
\begin{bmatrix}
0 & I \\
-I & 0
\end{bmatrix}
\]

We have the following key identity: \( i_{x_H} \omega = dH \). (Proof)

\[
i_{x_H} \omega = i_{x_H}(dq^i \wedge dp_i)
= (i_{x_H} dq^i) \wedge dp_i - dq^i \wedge i_{x_H} dp_i \quad \text{[see identity 2(b), Box 7.1]}
= \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q^i} dq^i = dH.
\]

Note that \( \omega \) is closed: \( d\omega = 0 \). In fact, \( \omega \) is exact: \( \omega = d\theta \), where \( \theta = -p^i \, dq^i \).

Now we can easily prove a number of key results about Hamiltonian systems:

**Conservation of energy** If \( F_t \) is the flow of \( X_H \) (see Section 1.6), then \( H \circ F_t = H \). (Proof)

\[
\frac{d}{dt} H \circ F_t = \frac{d}{dt} F_t^* H = F_t^* (i_{x_H} dH) = F_t^* (i_{x_H} \omega) = 0,
\]

thus \( H \circ F_t = H \circ F_0 = H \).

Each \( F_t \) is a canonical transformation; that is, \( F_t^* \omega = \omega \). (Proof)

\[
\frac{d}{dt} F_t^* \omega = F_t^* (i_{x_H} dH) = F_t^* (i_{x_H} \omega + i_{x_H} d\omega)
= F_t^* (i_{x_H} \omega) \quad \text{(since } d\omega = 0 \text{)} = F_t^* dH = 0.
\]

**Liouville’s Theorem** \( F_t \) preserves the phase volume:

\[
\mu = dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n.
\]
By computation, one finds that
\[
\omega \wedge \omega \wedge \cdots \wedge \omega = n! (-1)^{[n/2]} \mu.
\]
\[
\text{n times}
\]
Here \([n/2]\) denotes the largest integer \(\leq n/2\). Since \(F_t^* \omega = \omega\), and \(F_t^* (\omega \wedge \omega) = F_t^* \omega \wedge F_t^* \omega = \omega \wedge \omega\), we get \(F_t^* \mu = \mu\).

One can establish other basic facts like conservation laws in the presence of a symmetry group (Noether's theorem) in a similar way. We shall return to this in Chapter 5.

Next we summarize, without proofs, a few key facts about integration on manifolds. These are familiar results in \(\mathbb{R}^n\) which are converted to the language of manifolds.\(^{33}\) Let \(\mathcal{M}\) be a manifold, possibly with a (piecewise smooth) boundary. We assume \(\mathcal{M}\) is oriented; that is, \(\mathcal{M}\) has a covering by charts such that the Jacobian of the change of coordinates between any two of these coordinate systems is positive.

Let \(\mu\) be an \(n\)-form on \(\mathcal{M}\) (\(n\) is the dimension of \(\mathcal{M}\)). In any coordinate chart \(\{X^a\}\), \(\mu\) has the form \(\mu = f dX^1 \wedge \cdots \wedge dX^n\). If we integrate \(f\) in this chart and do the same for a covering of \(\mathcal{M}\) by charts not counting overlaps twice, we get a well-defined number, \(\int_{\mathcal{M}} \mu\) (the change of variables formula and transformation properties of \(\mu\) guarantee that this number is independent of the way in which \(\mathcal{M}\) is sliced up).

7.12 Theorem (Change of Variables) If \(\phi: \mathcal{M} \rightarrow \mathcal{N}\) is a regular \(C^1\) mapping that is orientation preserving, and \(\mu\) is an \(n\)-form on \(\phi(\mathcal{M})\), then

\[
\int_{\mathcal{M}} \phi^* \mu = \int_{\phi(\mathcal{M})} \mu.
\]

In Section 1.4 we described how to obtain a volume element on an oriented Riemannian manifold \((\mathcal{M}, G)\): \(dV = \sqrt{\det G_{AB}} dX^1 \wedge \cdots \wedge dX^n\), and the Jacobian of \(\phi\) by \(\phi^* d\nu = J dV\). Using this notation, 7.12 can be written

\[
\int_{\mathcal{M}} (f \circ \phi) J dV = \int_{\phi(\mathcal{M})} f d\nu
\]

for a scalar function \(f\) on \(\phi(\mathcal{M})\) by writing \(\nu = f d\nu\). Under an integral sign we shall usually write \(dV\) instead of \(d\nu\) in accordance with usage in integration theory.

\(^{33}\)The proofs of the following theorems are omitted; they may be found in one of the standard references (e.g., Spivak [1975], Lang [1972], or Abraham and Marsden [1978]).
7.13 Theorem (Stokes’ Theorem) If $\partial \mathcal{M}$ is positively oriented\(^{34}\) and $\alpha$ is an $(n - 1)$-form on $\mathcal{M}$, then

$$\int_{\partial \mathcal{M}} d\alpha = \int_{\mathcal{M}} \alpha.$$

This version of Stokes’ theorem includes as special cases, the usual theorems of Green, Gauss, and Stokes. We shall obtain the divergence theorem on a Riemannian manifold as a special case after a few preparatory results.

7.14 Proposition If $W$ is a vector field on $\mathcal{M}$, then $\mathcal{L}_W dV = (\text{DIV } W) dV$.

Proof Observe that $dV$ is closed (any $n$-form on an $n$-manifold is closed). Thus by the magic formula,

$$\mathcal{L}_W dV = d(i_W dV) = d(i_W \sqrt{\det G_{AB}} dX^1 \wedge \cdots \wedge dX^n)$$

$$= d[\sqrt{\det G_{AB}} (W^1 dX^2 \wedge \cdots \wedge dX^n - W^2 dX^1 \wedge dX^3 \wedge \cdots \wedge dX^n + \cdots)]$$

$$= \frac{\partial}{\partial X^c} (\sqrt{\det G_{AB}} W^c) dX^1 \wedge \cdots \wedge dX^n$$

$$= \frac{1}{\sqrt{\det G_{AB}}} \frac{\partial}{\partial X^c} (\sqrt{\det G_{AB}} W^c) dV = (\text{div } W) dV.$$

This gives an easy proof of the formula for the rate of change of the Jacobian, which was proved directly in Section 1.5.

7.15 Corollary $\frac{\partial}{\partial t} J = J \cdot (\text{div } v) \circ \phi$.

Proof $\frac{\partial}{\partial t} (\phi^* v) dV = \phi^* (\text{DIV } v) dV = \phi^* (\text{div } v) dV = (\text{div } v \circ \phi) \phi^* dV$

$$= J (\text{div } v \circ \phi) dV.$$

To get the divergence theorem we need one more important observation.

7.16 Proposition Let $N$ be the unit outward normal to $\partial \mathcal{M}$ and $W$ a vector field on $\mathcal{M}$. Then on $\partial \mathcal{M}$, $W \cdot N dA = i_W dV$, where $dA$ is the “area” element on $\partial \mathcal{M}$ (i.e., $dA$ is the “$dV$” for the $(n - 1)$-manifold $\partial \mathcal{M}$).

Proof Let $\{X^1\}$ be coordinates for $\mathcal{M}$ in which $\partial \mathcal{M}$ is the plane $X^1 = 0$ and for which $N = (1, 0, 0, \ldots)$, and $\mathcal{M}$ is described by $X^1 < 0$. We can arrange for $N$ to be normal to this plane at any particular point by a linear transformation

\(^{34}\)In an oriented coordinate chart $\{X^1, \ldots, X^n\}$ for $\mathcal{M}$ in which $\partial \mathcal{M}$ is the plane $X^1 = 0$, $\mathcal{M}$ must be represented as the half-space $X^1 < 0$. This agrees with the usual choice of normals in Green’s, Gauss’, and Stokes’ theorems in vector calculus.
of coordinates; that is \( G_{1A} = 0 \) \((A = 2, \ldots, n)\) and \( G_{11} = 1 \). The metric tensor on \( \partial\mathcal{M} \) is \( G_{AB} \) \((A, B = 2, \ldots, n)\). As above,
\[
i_w dV = i_w \sqrt{\text{det} G_{AB}} \, dX^1 \wedge \cdots \wedge dX^n
\]
\[
= \sqrt{\text{det} G_{AB}} \sum_{i=1}^{n} (-1)^{i-1} W^i \frac{dX^1 \wedge \cdots \wedge dX^n}{dX^i} \text{missing}
\]
If we evaluate this expression at a point satisfying \( X^1 = 0 \), we get only the first term in this sum:
\[
i_w dV = \sqrt{\text{det} G_{AB}} W^1 dX^2 \wedge \cdots \wedge dX^n = (W \cdot N) \, dA.
\]

7.17 Theorem \((\text{Divergence Theorem})\) If \( W \) is a vector field on \( \mathcal{M} \), then

\[
\int_{\mathcal{M}} \text{DIV} \, W \, dV = \int_{\partial \mathcal{M}} W \cdot N \, dA.
\]

Proof Let \( \alpha = i_w dV \). Then \( \int_{\partial \mathcal{M}} \alpha = \int_{\partial \mathcal{M}} W \cdot N \, dA \) by 7.16. Also,
\[
d\alpha = di_w dV = (\text{DIV} \, W) dV \text{by 7.14}. \]
Hence the result follows from Stokes' theorem.

Problem 7.5 (i) Explicitly recover the classical Stokes' theorem for oriented surfaces from 7.13.

(ii) Show that, on a Riemannian manifold \((\mathcal{M}, g)\),
\[
\omega_{ab} = \frac{1}{2} (v_a |_b - v_b |_a) = \frac{1}{2} \left( \frac{\partial v_a}{\partial x^b} - \frac{\partial v_b}{\partial x^a} \right)
\]
are the components of the spin \( \omega = \frac{1}{2} d(v) \). Formulate a Stokes' theorem for a general oriented surface in a Riemannian manifold.

Problem 7.6 Set \( S^{AB} \) be a 2-tensor on \( \mathcal{M} \) and \( \text{DIV} \, S = S^{AB}_{\,;B} \), its divergence. Let \( \alpha_A \) be a one-form on \( \mathcal{M} \). Prove the following integration by parts formula (in index notation):
\[
\int_{\mathcal{M}} \alpha_A S^{AB}_{\,;B} \, dV = -\int_{\mathcal{M}} \alpha_A S^{AB} \, dV + \int_{\partial \mathcal{M}} \alpha_A S^{AB} N^B \, dA.
\]

Now we turn our attention to the Piola transformation. This concept is of fundamental importance in the subsequent chapters. The Piola transformation is analogous to pull-back, except that there is a Jacobian present as well. This is indicative that volume or area forms are being transformed. We begin by defining the Piola transform of vector fields, and revert to \( \mathcal{B}, \mathcal{S} \) notation in place of \( \mathcal{M} \) and \( \mathcal{N} \) because we have configurations of bodies in mind.

7.18 Definition Let \( y \) be a vector field on \( \mathcal{S} \) and \( \phi: \mathcal{B} \rightarrow \mathcal{S} \) a regular (orientation preserving) \( C^1 \) mapping. The Piola transform of \( y \) is given by
\[
Y = J\phi^* y
\]
where $J$ is the Jacobian of $\phi$. In coordinates, $Y^a = J(F^{-1})_b^a y^b$, where

$$J = \frac{\sqrt{\det g_{ab}}}{\sqrt{\det G_{AB}}} \partial(\phi^1, \ldots, \phi^n)$$

and $F^a_A = \frac{\partial \phi^a}{\partial X^A}$.

We can phrase this in another useful way:

**7.19 Proposition**  $Y$ is the Piola transform of $y$ if and only if $\phi^*(i_y dv) = i_Y dV$.

*Proof* Notice that $(n-1)$-forms and vector fields are in one-to-one correspondence by way of the mapping $Y \mapsto *Y$, where $*Y = i_Y dV$. (Above we calculated $i_Y dV$ in coordinates.) From Box 7.1, $\phi^*(i_y dv) = i_{\phi^*} \phi^* dv = i_{\phi^*} \phi^* dV = i_{\phi^*} y^b dV$ so the assertion follows.

**Note.** $Y$ is an honest vector field, while $[\partial(\phi^1, \ldots, \phi^n)/\partial(X^1, \ldots, X^n)](F^{-1})_b^a y^b$ is not. The metric factors are important, even for curvilinear coordinates in $\mathbb{R}^3$, such as spherical coordinates.

**7.20 Theorem** (Piola Identity) If $Y$ is the Piola transform of $y$, then

$$\text{DIV } Y = J^a(v) \phi^a$$

We shall give two proofs of this important result.

**First proof** Let $\mathcal{U} \subset \mathbb{R}^n$ be a nice open set and $\partial \mathcal{U}$ its boundary. By the change of variables theorem and 7.19,

$$\int_{\partial \mathcal{U}} i_Y dV = \int_{\phi(\mathcal{U})} i_y dv.$$  

From 7.16 and 7.17,

$$\int_{\mathcal{U}} \text{DIV } Y dV = \int_{\phi(\mathcal{U})} \text{div } y dv = \int_{\mathcal{U}} J(\text{div } y \circ \phi) dV.$$  

Since $\mathcal{U}$ is arbitrary, the assertion follows.

**Second Proof** We compute directly, using differential forms: $(\text{DIV } Y) dV = \mathcal{E}_Y dV = d(i_Y dV)$ by the general formula $\mathcal{E}_Y \alpha = i_Y d\alpha + d(i_Y \alpha)$ and the fact that $d$ of an $n$-form is zero. Thus $(\text{DIV } Y) dV = d(i_Y dV) = d(\phi^*(i_Y dv)) = \phi^* d(i_y dv)$ (the operations of pull-back and $d$ commute) $= \phi^*(\text{div } y dv) = J(\text{div } y \circ \phi) dV$ (definition of $J$), and so $\text{DIV } Y = J(\text{div } y \circ \phi)$.

From 7.16 we get another important way of expressing the Piola transformation: $Y \circ \mathcal{N} da = y \cdot n da$, where $da$ is the area element on $\phi(\mathcal{U})$; it is related to $dA$ according to $(da)_b = J(F^{-1})_a^b dA$. See Figure 1.7.1. This equation shows how the area elements on $\partial \mathcal{U}$ and $\partial \phi(\mathcal{U})$ are related.

Since $Y^a = (F^{-1})_b^a y^b$, the Piola identity may be read this way:

$$\text{DIV}(JF^{-1}) = 0,$$

where

$$\text{(DIV}(JF^{-1}))_a = (J(F^{-1})_a^b y^b)_{|A} = \frac{1}{\sqrt{\det G_{BC}}} \partial \frac{\partial}{\partial X^A} (\sqrt{\det G_{BC}} J(F^{-1})_a^b).$$
We can also make a Piola transformation on any index of a tensor. For example, let $\sigma^{ab}$ be a given two tensor. If we make a Piola transformation on the last index, we get a two-point tensor $P$ with components $P_{ab} = J(F^{-1})^b_c \sigma^{ab}$. The Piola identity, then, tells us that $\text{DIV} \ P = J(\text{div} \ \sigma) \circ \phi$; that is, $P_{ab \mid b} = J\sigma^{ab \mid b}$.

**Problem 7.7** If $y \cdot n$ is interpreted as a flux per unit area on $\phi(U)$, show that one can interpret $Y \cdot N$ as the corresponding flux per unit of underformed area. Use this to give a physical interpretation of the Piola identity.

### Box 7.4 Summary of Important Formulas for Section 1.7

(See Box 7.1 for the key identities for differential forms.)

**Definition of a $k$-form on $\mathbb{R}$**

$\alpha$ is a $\begin{pmatrix} 0 \\ k \end{pmatrix}$ tensor that is skew symmetric.

$$\alpha(W_{\pi(1)}, \ldots, W_{\pi(k)}) = \text{sgn} \ \pi \alpha(W_1, \ldots, W_k)$$

**Wedge Product**

$$(\alpha \wedge \beta)(W_1, \ldots, W_{k+1}) = \frac{1}{k!} \sum \text{sgn} \ \pi \alpha \otimes \beta$$

$$(W_{\pi(1)}, \ldots, W_{\pi(k+1)})$$

**Interior Product**

$$(\iota_W \alpha)(W_2, \ldots, W_k) = \alpha(W, W_2, \ldots, W_k)$$

**Exterior Derivative $d$**

Characterized by:

(i) $dd\alpha = 0$

(ii) $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta$

$\alpha_{A_1 \ldots A_k}$ changes sign if any two indices are switched.

$$(\alpha \wedge \beta)_{A_1 \ldots A_k B_1 \ldots B_1}$$

is the complete anti-symmetrization of $\alpha_{A_1 \ldots A_k} \beta_{B_1 \ldots B_1}$.

$$(i_W \alpha)_{A_1 \ldots A_k} = W^A \alpha_{A_1 \ldots A_k}$$

If $\alpha = \alpha_{A_1 \ldots A_k} dX^{A_1} \wedge \ldots \wedge dX^{A_k}$, then $d\alpha = \frac{\partial \alpha_{A_1 \ldots A_k}}{\partial X^A} dX^A \wedge dX^{A_1} \wedge \ldots \wedge dX^{A_k}$.
(iii) $df$ = differential of the function $f$

(iv) $d$ is local

**Poincaré Lemma**

If $d\alpha = 0$, then locally there is a $\beta$ such that $\alpha = d\beta$.

**Euler–Lagrange Equations of a Variational Principle**

$q(t) \in \mathbb{R}^n$ satisfies $(d/dt)(\partial \mathcal{L}/\partial \dot{q}^i) - \partial \mathcal{L}/\partial q^i = 0$ ($i = 1, \ldots, n$) if and only if $q(\cdot)$ is a critical point of $\mathcal{L}(q, \dot{q})dt$ subject to the condition that $q(0), q(1)$ be fixed.

**Inverse Problem for a Variational Principle**

A nonlinear operator $A$ on a Banach space $X$ is a potential operator; that is, $\langle A(x), v \rangle = dL(x) \cdot v$ for some $L(x)$ if and only if $\langle DA(x) \cdot v_1, v_2 \rangle$ is symmetric in $v_1$ and $v_2$ (i.e., $DA(x)$ is a symmetric operator).

**Hamilton’s Equations**

Hamiltonian vector field $X_H$:

$$i_{X_H} \omega = dH$$

$$X_H = \left( \frac{\partial H}{\partial p_i}, -\frac{\partial H}{\partial q^i} \right)$$

**Change of Variables**

$$\int_{\Omega} (f \circ \phi)^* \mu = \int_{\phi(\Omega)} \mu$$

**Stokes’ Theorem**

$$\int_{\partial \Omega} d\alpha = \int_{\Omega} \alpha$$

$$\int_{\partial \Omega} \frac{\partial \alpha_{A_1, \ldots, A_n}}{\partial x^B} dx^B \wedge dx^{A_1} \wedge \cdots \wedge dx^{A_{n-1}}$$

**Divergence Theorem**

$$\int_{\Omega} (\text{DIV } W) \, dv = \int_{\partial \Omega} W^A N_A \, dA$$

**Piola Transformation**

$$Y = J\phi^* y \quad \text{or} \quad \int_{\phi \Omega} dV = \phi^*(\int_{\Omega} dV)$$

**Piola Identity**

$$\text{DIV } Y = J(\text{div } y \circ \phi)$$

$$Y^A_{1, A} = J y^a_{1, a}$$