2. Differentiable Manifolds and Tensors

2.1. Definition of a Manifold

Let R^n be the set of all *n*-tuples of real numbers $\left(x^1,\cdots,x^n\right)$. A set *M* of points is a (**topological**) **manifold** if each point *P* in it has an open neighborhood *U* homeomorphic to some open set *V* in R^n. In other words, there is a bi-continuous bijection (1-1 onto map)

$$\phi : U \to V \qquad \text{by} \qquad P \mapsto \phi(P) = \left(x^1(P),\cdots,x^n(P)\right)$$

for all *P* in *M*. The *n* numbers $x^j(P)$ are called the **coordinates** of *P* and *n* is the **dimension** of *M*. Thus, the topology of *M* is the same as R^n locally.

The pair (U,ϕ) is called a **chart**, or a **local coordinate system**. An **atlas** on *M* is a set $\{(U_\alpha,\phi_\alpha)\}$ of charts so that the domains $\{U_\alpha\}$ covers *M*, i.e., every *P* is in some U_α. For reasons of compatibility, an atlas of class C^k requires the maps

$$\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha\left(U_\alpha \cap U_\beta\right) \to \phi_\beta\left(U_\alpha \cap U_\beta\right) \qquad \text{(a)}$$

to be maps of class C^k. Note that $\phi_\beta \circ \phi_\alpha^{-1}$ is a map between open sets of R^n. In fact, it represents a **coordinate transformation** for points in the overlap region $U_\alpha \cap U_\beta$ of two coordinate systems given by ϕ_α and ϕ_β. A manifold with an atlas of class C^k is said to be a C^k manifold. Those with $k > 1$ are called **differentiable manifolds**. For convenience, we shall deal only with C^∞ or C^ω manifolds.

2.2. The Sphere as a Manifold

Consider the **2-sphere** S^2 which consists of the points in R^3 that satisfy

$$\left(x^1\right)^2+\left(x^2\right)^2+\left(x^3\right)^2=a^2=const$$

Any sufficiently small neighborhood of every point *P* on S^2 has a 1-1 map onto a region in R^2. Such mappings are in general neither angle- nor length- preserving.

Consider now the map

$$f:S^2\to R^2 \qquad \text{by} \qquad (a,\theta,\phi)\mapsto\left(x^1,x^2\right)=(\theta,\phi)$$

where (a,θ,ϕ) are the spherical coordinates of a point *P* with Cartesian coordinates

$$\left(x^1,x^2,x^3\right)=\left(a\sin\theta\cos\phi,a\sin\theta\sin\phi,a\cos\phi\right)$$

For any small enough region *U* on S^2, this is a diffeomorphism so that (θ,ϕ) can serve as coordinates for *U*. In fact, as long as *U* does not contain the arc *A* given by Cartesian coordinates $(a\sin\theta,0,a\cos\theta)$ with $0\le\theta\le\pi$, (θ,ϕ) is a good coordinate system. On the other hand, for $\theta\ne 0$, there are 2 images, $(\theta,0)$ and $(\theta,2\pi)$, of $(a\sin\theta,0,a\cos\theta)$ under *f*. For $\theta=0$, the image of $(0,0,a)$ under *f* is the entire line $\left(x^1,x^2\right)=(0,\phi)$ in R^2. Thus, *f* is not defined on the arc *A* so that (θ,ϕ) cannot be a coordinate system for the entire S^2. By taking out the arc *A* from its domain, the image of *f* is a rectangular open region in R^2 given by $0<x^1<\pi$ and $0<x^2<2\pi$.

What these all mean is that S^2 cannot be covered by a single chart: at least 2 are needed. Of interest is the **stereographic** map which projects the sphere onto a plane tangent to it. Here, only a single point fails to be covered by the map.

2.3. Other Examples of Manifolds

Roughly speaking, any set that can be parametrized continuously is a manifold. The number of independent parameters required is the dimension of the manifold. Some examples of manifolds are:

1. The set of all rotations of a rigid body in 3-D space. The (continuous) parameters are the three Euler angles.
2. The set of all (pure boost) Lorentz transformations. The parameters are the three components of boost velocity.
3. The 6N-D phase space of N particles in 3-D space. The parameters are the 3N position and 3N momentum coordinates.
4. Given m dependent variables $\{y_1, \cdots, y_m\}$ and n independent ones $\{x_1, \cdots, x_n\}$, the set of all points $\{y_1, \cdots, y_m; x_1, \cdots, x_n\}$ is a (m+n)-D manifold. Solutions to some set of (algebraic of differential) equations involving these variables are surfaces in the manifold.
5. A vector space is a manifold with the topology of R^n.
6. A Lie group is a C^∞ manifold so that the group operations are C^∞ maps of the manifold into itself.

2.4. Global Considerations

Since every manifold is locally like R^n for some n, all manifolds of the same dimension are locally indistinguishable. Hence, manifolds are classifed by their global properties, e.g., S^2 and R^2. To be more precise, two manifolds are equivalent (belong to the same class) if there is a **diffeomorphism** (bi-C^∞ bijection) between them.

2.5. Curves

A **curve** is a differentiable mapping *C* from an open set of *R* into *M*, i.e.,

$$C : R \to M \qquad \text{with} \qquad \lambda \mapsto P(\lambda) = \{x^i(\lambda), i = 1, \cdots, n\}$$

where λ is the **parameter** of the curve. Differentiability here means that $x^i(\lambda)$ are differentiable functions of λ. For convenience, we shall assume each mapping represents a unique curve. Thus, curves represented by different mappings are considered as different even though they may give the same set of image points in *M*.

2.6. Functions on *M*

A function f on M is an assignment of a real number $f(P)$ to each point P in M. This is denoted by

$$f : M \to R \qquad \text{with} \qquad P \mapsto f(P)$$

If a region $U \subset M$ is mapped differentiably onto some region of R^n with $P \mapsto \{x^i, i = 1, \cdots, n\} = \{x^i\}$, we can write $f(P) = f(x^1, \cdots, x^n) = f(x^i)$ so that f is a function on R^n. If f is differentiable in R^n, we say f is **differentiable** in M.

2.7. Vectors and Vector Fields

2.7.1. Vectors as Tangents to Curves

Consider a curve $C(\lambda)$ described by $x^i = x^i(\lambda)$ in *M*. Let $f(x^i)$ be a function on *M*. For points on the curve, *f* can be taken as a function of λ through $g(\lambda) = f\left[x^i(\lambda)\right]$. Thus,

$$\frac{dg}{d\lambda} = \sum_i \frac{\partial f}{\partial x^i}\frac{dx^i}{d\lambda} \qquad (2.2)$$

Since this holds for arbitrary *g*, we can write

$$\frac{d}{d\lambda} = \sum_i \frac{dx^i}{d\lambda}\frac{\partial}{\partial x^i} \qquad (2.3)$$

Now, in Euclidean space, $\frac{dx^i}{d\lambda}$ are the components of a vector tangent to the curve *C*. Thus, if we treat $\overline{e}_i = \frac{\partial}{\partial x^i}$ as basis vectors, $\frac{d}{d\lambda}$ can be identified as the **tangent vector** to the curve *C*(λ) at point $P(\lambda)$. In fact, in differential geometry, a **vector** is defined as the tangent to some curve in the manifold.

Note that the essence of a vector is its direction, not the curve to which it is tangent. Indeed, a vector is tangent to an infinite number of curves passing through its point due to 2 different reasons. First, for a manifold of dimension higher than 1, there are infinitely many paths that can pass through a given point while being tangent to each other. Secondly, a given path represents an infinite number of curves via different parametrization.

As an example, consider the curve *C*(λ) with $x^i(\lambda) = \lambda a^i$, where a^i are constants. Obviously, *C* passes through the origin *O* when $\lambda = 0$. Its tangent is $\frac{d}{d\lambda} = \sum_i a^i \frac{\partial}{\partial x^i}$. Another curve *C'*($\mu$) with $x^i(\mu) = \mu^2 b^i + \mu a^i$ also passes through *O* when $\mu = 0$. Its tangent is $\frac{d}{d\mu} = \sum_i \left(2\mu b^i + a^i\right)\frac{\partial}{\partial x^i}$, which equals to $\frac{d}{d\lambda}$ at *O*. Hence, $\frac{d}{d\lambda}$ is tangent to 2 different curves at *O*. Next, we can re-parametrized the 1st curve as $x^i(\nu) = \left(\nu^3 + \nu\right)a^i$, which passes through *O* at $\nu = 0$ with tangent

$$\left.\frac{dx^i}{d\nu}\right|_{\nu=0} = \left.(2\nu+1)a^i\right|_{\nu=0} = a^i .$$

2.7.2. Vector Space at a Point

It is straightforward to verify that the set of all tangent vectors at a point *P* forms a vector space called the **tangent space** to *M* at *P* and denoted by T_P. For example, closure under addition and scalar multiplication is proved as follows. Let

$$\frac{d}{d\lambda} = \sum_i \frac{dx^i}{d\lambda}\frac{\partial}{\partial x^i} \qquad\qquad \frac{d}{d\mu} = \sum_i \frac{dx^i}{d\mu}\frac{\partial}{\partial x^i}$$

$$\Rightarrow \qquad a\frac{d}{d\lambda} + b\frac{d}{d\mu} = \sum_i \left(a\frac{dx^i}{d\lambda} + b\frac{dx^i}{d\mu} \right)\frac{\partial}{\partial x^i} \qquad (a, b \text{ scalars})$$

Setting

$$\frac{dx^i}{d\phi} = a\frac{dx^i}{d\lambda} + b\frac{dx^i}{d\mu}$$

we have

$$a\frac{d}{d\lambda} + b\frac{d}{d\mu} = \sum_i \frac{dx^i}{d\phi}\frac{\partial}{\partial x^i} \quad = \frac{d}{d\phi}$$

so that the linear combination of 2 tangents is another tangent. Q.E.D.

Note that the vector space defined above consists only of tangents at the same point in *M*. Thus, vectors at different points are in different vector spaces. In particular, the subtraction of 2 vectors located at different points in *M* has no geometrical meaning, which is the root cause of complications in tensor analysis on a manifold.

Finally, we define a **vector field** as a rule for assigning a vector at each point of *M*.

2.8. Basis Vectors and Basis Vector Fields

As can be seen from the definition (2.3), the dimension of T_P is the same as that of M. Given a coordinate system $\{x^i\}$ for a neighborhood U of M, we call $\left\{\frac{\partial}{\partial x^i}\right\}$ the **coordinate basis** of T_P for all points in U. In general, we shall use an overbar to denote a vector. For example, a general basis for T_P is denoted by $\{\bar{e}_i\}$. Thus,

$$\bar{V} = \sum_i V^i \frac{\partial}{\partial x^i} = \sum_j V'^j \bar{e}_j$$

where V'^j is the component of $\bar{V}$ along $\bar{e}_j$. If $\bar{V}$ is a vector field, the components $\{V^i\}$ and $\{V'^j\}$ are functions on M. In which case, $\bar{V}$ is **differentiable** if $\{V^i\}$ are.

In calling the set of vectors $\left\{\frac{\partial}{\partial x^i}\right\}$ a basis, we have implicitly assumed they are linearly independent at a given point P. We shall show below that this requires $\{x^i\}$ to be a set of **good** coordinates, i.e., it provides a 1-1 map of some neighborhood U of P onto a region of R^n.

To begin, consider a set of good coordinates $\{y^i\}$ on U. The map from $\{x^i\}$ to U can be written as

$$y^j = y^j\left(x^1, \cdots, x^n\right) \qquad j = 1, \cdots, n$$

By the inverse function theorem, this map is 1-1 and hence invertible in U iff the **Jacobian** $J = \det\left|\frac{\partial y^j}{\partial x^i}\right|$ is non-vanishing. This means the set of n vectors

$\bar{V}_{(1)} = \left(\frac{\partial y^1}{\partial x^1}, \cdots, \frac{\partial y^n}{\partial x^1}\right)$, …, $\bar{V}_{(n)} = \left(\frac{\partial y^1}{\partial x^n}, \cdots, \frac{\partial y^n}{\partial x^n}\right)$ are linearly independent. On the other hand, the chain rule gives

$$\frac{\partial}{\partial x^i} = \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$$

so that $\bar{V}_{(i)}$ is just the vector $\frac{\partial}{\partial x^i}$ expressed in the $\{y^j\}$ basis. Hence, $\left\{\frac{\partial}{\partial x^i}\right\}$ are linearly independent. Finally, since a 1-1 map of a set of good coordinates is another set of good coordinates, our assertion is proved.

2.9. Fiber Bundles

The exact definition of fibre bundles is rather involved [see, e.g., Y. Choquet-Bruhat et al, "Analysis, Manifolds and Physics", 2^{nd} ed., North-Holland (1982)] and won't be introduced until §2.11. Here, the basic concepts are presented using as example the simple variety called **tangent fibre bundles**. Loosely speaking, a tangent fibre bundle is just a manifold TM obtained by attaching a tangent space T_P to each point of a manifold M. Here, M is called the **base manifold** and T_P a **typical fibre**. By definition, $\dim(TM) = \dim(M) + \dim(T_P)$.

2.10. Examples of Fiber Bundles

1. The tangent bundle *TM*. For an *n*-D manifold *M*, its *TM* is 2*n*-D.
2. Tensor bundles with tensor spaces as fibres.
3. Isospin bundles with spacetime as manifold and isospin space as fibres.
4. Galilean spacetime with time as manifold and Euclidean space as fibres.

2.11. A Deeper Look at Fiber Bundles

2.11.1. Global Properties

1. Given 2 spaces M and N, there is a (**Cartesian**) **product space** $M\times N$ consisting of all ordered pairs (a,b) where $a \in M$ and $b \in N$.
2. If M, N are topological spaces, so is $M\times N$.
3. An open set $U \subset M$ with coordinates $\{x^i ; i = 1,\cdots,m\}$, taken together with a n open set $V \subset N$ with coordinates $\{y^i ; i = 1,\cdots,n\}$, becomes an open set $(U,V) \subset M \times N$ with $m + n$ coordinates $(x^1,\cdots,x^m, y^1,\cdots,y^n)$.
4. In the formal definition of a fibre bundle, what distinguishes the base M and the fibre T_P is the existence of a projective (many-1) mapping $\pi : T_P \to P \in M$ that maps every point in T_P to the point P in M to which T_P is attached.
5. A fibre bundle is locally a product space $U \times F$ where U is an open subset of the base manifold B and F is a typical fibre. This means a fibre bundle is **locally trivial**.
6. To be **globally trivial** means that the entire bundle is a product space $B \times F$. Or, more generally, there exists a C^∞ 1-1 map (diffeomorphism) of the bundle onto $B \times F$. With a non-zero F, the inverse of this map gives a no-where zero cross section of the bundle. Hence, a necessary condition for global triviality is the existence of a no-where zero vector field.
7. The tangent bundle TS^2 of the 2-sphere S^2 is globally non-trivial. According to the **fixed point theorem** of the sphere, every C^∞ 1-1 map (diffeomorphism) of S^2 onto itself must leave at least 1 point fixed. A no-where zero vector field will generate a map that has no fixed point. Hence, TS^2 cannot be globally trivial.
8. The tangent bundle TS^1 of the circle S^1 is the product space $S^1 \times R$, which is a cylinder. Hence, it is globally trivial.
9. The mobius strip is also a tangent bundle using S^1 as base manifold and R as the typical fibre. Locally, it is trivial but globally, not.
10. .

2.11.2. Formal Definition of A Fibre Bundle

The global properties of a fibre bundle are described by its **structure group**. In fact, a fibre bundle is formally defined as the quartet (E, B, π, G) consisting of a **base manifold** *B*, a **projection** $\pi : E \to B$, a **typical fibre** *F*, a **structural group** *G* of homeomorphisms of *F* onto itself, and a family $\{U_j\}$ of open sets covering *B*, such that

1. *E* is **locally trivial**. This means the bundle over any set U_j, i.e., the inverse image $\pi^{-1}(U_j)$, is homeomorphic to the product space $U_j \times F$. This homeomorphism $\phi_j : \pi^{-1}(U_j) \to U_j \times F$ is of the form $\phi_j(p) = (\pi(p), h_j(x))$, where $p \in E$, $x = \pi(p) \in U_j \subset B$ and $h_j(x) : F_x \to F$ which maps the fibre F_x at $x \in U_j$ to the typical fibre *F*.
2. The global properties of the bundle is given by the structural group *G* whose elements are the homeomorphisms

   $$h_k(x) \circ h_j(x)^{-1} : \; F \to F$$

   where $x \in U_j \cap U_k$.
3. The induced mappings

   $$g_{jk} : U_j \cap U_k \to G \qquad \text{by} \qquad x \mapsto g_{jk}(x) = h_j(x) \circ h_k(x)$$

   are continuous. They are called **transition functions** and satisfy

   $$g_{jk}(x)\, g_{kl}(x) = g_{jl}(x)$$

2.12. Vector Fields and Integral Curves

1. A **vector field** is a rule that selects a vector from the tangent vector space at each point of *M*.
2. Consider a vector field $V^i(P)$ for $P \in M$. Given a coordinate system $\{x^i\}$, we have $V^i(P) = v^i(x^i)$. The tangent vector to a curve $x^i(\lambda)$ is given by

$$\frac{dx^i}{d\lambda} = v^i(x^i) \qquad (2.5)$$

which is just a set of 1st order ordinary differential equations so that a unique solution always exists in some neighborhood around any given point.

3. Hence, given a vector field $v^i(x^i)$, a solution, called an **integral curve**, of (2.5) is a curve whose tangent is everywhere equal to the vector field.
4. Uniqueness of solutions means that different integral curves never cross except possibly at points where $v^i = 0$.
5. Thus, by judicial choice of initial conditions, one can find a set of integral curves that fills up *M*. Such a set of curves is called a **congruence**.

2.13. Exponentiation of the Operator *d/dλ*

Consider an analytic (C^{ω}) manifold *M*. The coordinates $x^i(\lambda)$ of points along an integral curve of $\bar{Y} = \dfrac{d}{d\lambda}$ are analytic functions of λ. Hence, 2 points with parameters λ_0 and $\lambda_0+\varepsilon$ are related by the Taylor series

$$x^i(\lambda_0+\varepsilon) = x^i(\lambda_0) + \varepsilon\left(\frac{dx^i}{d\lambda}\right)_{\lambda_0} + \frac{1}{2!}\varepsilon^2\left(\frac{d^2x^i}{d\lambda^2}\right)_{\lambda_0} + \cdots$$

$$= \left(1+\varepsilon\frac{d}{d\lambda}+\frac{1}{2!}\varepsilon^2\frac{d^2}{d\lambda^2}+\cdots\right)x^i\Big|_{\lambda_0}$$

$$\equiv \exp\left(\varepsilon\frac{d}{d\lambda}\right)x^i\Big|_{\lambda_0} \qquad (2.6)$$

The operator $\exp\left(\varepsilon\dfrac{d}{d\lambda}\right)=\exp\left(\varepsilon\bar{Y}\right)$ is called the **exponentiation** of the operator $\varepsilon\dfrac{d}{d\lambda}$. As $\varepsilon\dfrac{d}{d\lambda}$ denotes an infinitesimal 'motion' along the curve, $\exp\left(\varepsilon\dfrac{d}{d\lambda}\right)$ denotes a finite motion.

2.14. Lie Brackets and Noncoordinate Bases

2.14.1. Non-Coordinate Basis

Given a coordinate system x^i, the basis $\left\{ \frac{\partial}{\partial x^i} \right\}$ is called a **coordinate basis** for the vector fields. One important characteristics of a coordinate basis is that its members commute, i.e.,

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] = \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} - \frac{\partial}{\partial x^j}\frac{\partial}{\partial x^i} = 0$$

On the other hand, 2 arbitrary vector field need not commute. For example, let $\bar{V} = \frac{d}{d\lambda}$ and $\bar{W} = \frac{d}{d\mu}$, we have

$$\begin{aligned}
\left[\bar{V}, \bar{W} \right] &= \frac{d}{d\lambda}\frac{d}{d\mu} - \frac{d}{d\mu}\frac{d}{d\lambda} \\
&= \sum_{i,j} \left[V^i \frac{\partial}{\partial x^i}\left(W^j \frac{\partial}{\partial x^j} \right) - W^j \frac{\partial}{\partial x^j}\left(V^i \frac{\partial}{\partial x^i} \right) \right] \\
&= \sum_{i,j} \left[V^i \left(\frac{\partial W^j}{\partial x^i} \right) \frac{\partial}{\partial x^j} + V^i W^j \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} - W^j \left(\frac{\partial V^i}{\partial x^j} \right) \frac{\partial}{\partial x^i} - W^j V^i \frac{\partial}{\partial x^j}\frac{\partial}{\partial x^i} \right] \\
&= \sum_{i,j} \left[V^i \left(\frac{\partial W^j}{\partial x^i} \right) \frac{\partial}{\partial x^j} - W^j \left(\frac{\partial V^i}{\partial x^j} \right) \frac{\partial}{\partial x^i} \right] \\
&= \sum_{i,j} \left[V^i \left(\frac{\partial W^j}{\partial x^i} \right) - W^i \left(\frac{\partial V^j}{\partial x^i} \right) \right] \frac{\partial}{\partial x^j} \qquad (2.7) \\
&= \sum_j U^j \frac{\partial}{\partial x^j} \quad = \bar{U}
\end{aligned}$$

where

$$U^j = \sum_i \left[V^i \left(\frac{\partial W^j}{\partial x^i} \right) - W^i \left(\frac{\partial V^j}{\partial x^i} \right) \right]$$

Thus, the commutator of 2 vectors are in general another vector. A basis consisting of non-commutating vectors is called a **non-coordinate basis**.

2.14.2. Exercise 2.1

Consider the 'unit' basis vector fields for polar coordinates in the Euclidean plane defined by

$$\hat{\mathbf{r}} = \cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{y}} \quad = \cos\theta\,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y}$$

$$\hat{\boldsymbol{\theta}} = -\sin\theta\,\hat{\mathbf{x}} + \cos\theta\,\hat{\mathbf{y}} \quad = -\sin\theta\,\frac{\partial}{\partial x} + \cos\theta\,\frac{\partial}{\partial y}$$

Show that they form a non-coordinate basis.

2.14.3. Lie Brackets

When applied to vectors, the commutator is called a **Lie bracket**.

Consider a 2-D subspace *S* of a manifold *M* described by coordinates $\left(x^1, x^2\right)$. By definition, x^1 is constant along the lines of x^2, which are integral curves of $\dfrac{\partial}{\partial x^2}$.

This is the reason why $\left[\dfrac{\partial}{\partial x^1}, \dfrac{\partial}{\partial x^2}\right] = 0$.

Next, consider 2 other arbitrary vector fields, $\bar{V} = \dfrac{d}{d\lambda}$ and $\bar{W} = \dfrac{d}{d\mu}$, in *S* [see Fig.2.20]. In general, μ can vary along an integral curve $C(\lambda)$ of $\bar{V}$ and λ can vary along an integral curve $C(\mu)$ of $\bar{W}$. Hence, $\left[\bar{V}, \bar{W}\right] \neq 0$. As will be show below, this means λ and μ are not coordinates.

Consider a point *P* at the intersection of $C(\lambda)$ and $C(\mu)$ with respective parameters λ and μ. [see Fig.2.21] If we move along $C(\lambda)$ by $\Delta\lambda = \varepsilon$, we reach point *R* with coordinates

$$x^i(R) = \exp\left(\varepsilon \frac{d}{d\lambda}\right) x^i(P)$$

Moving further along $C(\mu)$ by $\Delta\mu = \sigma$, we reach point *A* with coordinates

$$x^i(A) = \exp\left(\sigma \frac{d}{d\mu}\right) \exp\left(\varepsilon \frac{d}{d\lambda}\right) x^i(P) \tag{2.9}$$

On the other hand, if we move first along $C(\mu)$ by $\Delta\mu = \sigma$, then along $C(\lambda)$ by $\Delta\lambda = \varepsilon$, we reach point *B* with coordinates

$$x^i(B) = \exp\left(\varepsilon \frac{d}{d\lambda}\right) \exp\left(\sigma \frac{d}{d\mu}\right) x^i(P) \tag{2.10}$$

The difference between the coordinates of *A* and *B* is

$$x^i(B) - x^i(A) = \left[\exp\left(\varepsilon \frac{d}{d\lambda}\right), \exp\left(\sigma \frac{d}{d\mu}\right)\right] x^i(P) \tag{2.11}$$

For infinitesimal ε, we have

$$\left[\exp\left(\varepsilon \frac{d}{d\lambda}\right), \exp\left(\sigma \frac{d}{d\mu}\right)\right]$$

$$= \left[1 + \varepsilon \frac{d}{d\lambda} + \frac{1}{2}\varepsilon^2 \frac{d^2}{d\lambda^2} + \cdots, 1 + \sigma \frac{d}{d\mu} + \frac{1}{2}\sigma^2 \frac{d^2}{d\mu^2} + \cdots\right]$$

$$= \varepsilon\sigma \left[\frac{d}{d\lambda}, \frac{d}{d\mu} \right] + \cdots$$

so that

$$x^i(B) - x^i(A) = \varepsilon\sigma \left[\bar{V}, \bar{W} \right] x^i(P) + \cdots \qquad (2.12)$$

Thus, the Lie bracket $\left[\bar{V}, \bar{W} \right]$ is proportional to the difference of the end points when one moves along the integral curves by the same amount but in different order. Obviously, in order for λ and μ to be coordinates, we must have $x^i(B) = x^i(A)$ or $\left[\bar{V}, \bar{W} \right] = 0$.

2.14.4. Lie Algebras

A **Lie algebra** of vector fields on a region *U* of a manifold *M* is a set *A* of vector fields on *U* such that

1. It is a vector space under addition.
2. It is closed under the Lie bracket.

Note that condition (1) means *A* is closed under linear combinations of vector fields with constant coefficients. Condition (2) means the Lie bracket of 2 vector fields is another vector field.

2.15. When is a Basis a Coordinate Basis?

In §2.14, we have shown that the necessary condition for a basis to be a **coordinate basis** is that the Lie brackets of every pair of its member vectors vanish. We shall show in the following that this condition is also sufficient.

To begin, consider a 2-D region *U* of *M*. Starting from a point *P* in *U* with coordinates $\left(x^1, x^2\right)$, we move first a parameter distance λ_1 along $\overline{A} = \frac{d}{d\lambda}$ to a point *R*, then μ_1 along $\overline{B} = \frac{d}{d\mu}$ to a point *Q*. The corresponding coordinates are

$$x^i\left(Q\right) = \exp\left(\mu_1 \frac{d}{d\mu}\right) x^i\left(R\right) \quad = \exp\left(\mu_1 \frac{d}{d\mu}\right)\exp\left(\lambda_1 \frac{d}{d\lambda}\right) x^i\left(P\right)$$

This equation defines a map $\mathbb{R}^2 \to U$ by

$$\left(\alpha, \beta\right) \mapsto x^i\left(\alpha, \beta\right) = \exp\left(\beta \frac{d}{d\mu}\right)\exp\left(\alpha \frac{d}{d\lambda}\right) x^i\left(P\right) \qquad \text{(a)}$$

Thus, $\left(\alpha, \beta\right)$ forms a coordinate system in *U* if the map (a) is 1-1, i.e., if it has an inverse. It will be shown later that this requires $\overline{A}$ and $\overline{B}$ to be linearly independent. Using

$$\frac{d}{d\alpha}\exp\left(\alpha \frac{d}{d\lambda}\right) = \frac{d}{d\alpha}\left(\sum_{n=0}^{\infty}\frac{\alpha^n}{n!}\frac{d^n}{d\lambda^n}\right) \quad = \sum_{n=1}^{\infty}\frac{\alpha^{n-1}}{\left(n-1\right)!}\frac{d^n}{d\lambda^n}$$

$$= \sum_{n=1}^{\infty}\frac{\alpha^{n-1}}{\left(n-1\right)!}\frac{d^{n-1}}{d\lambda^{n-1}}\frac{d}{d\lambda} \quad = \sum_{n=0}^{\infty}\frac{\alpha^n}{n!}\frac{d^n}{d\lambda^n}\frac{d}{d\lambda}$$

$$= \exp\left(\alpha \frac{d}{d\lambda}\right)\frac{d}{d\lambda}$$

we can differentiate (a) to get

$$\frac{\partial x^i}{\partial \alpha} = \exp\left(\beta \frac{d}{d\mu}\right)\exp\left(\alpha \frac{d}{d\lambda}\right)\left.\frac{dx^i}{d\lambda}\right|_P$$

$$\frac{\partial x^i}{\partial \beta} = \exp\left(\beta \frac{d}{d\mu}\right)\frac{d}{d\mu}\left[\exp\left(\alpha \frac{d}{d\lambda}\right) x^i\left(P\right)\right]$$

where *P* is the origin of the coordinate system $\left(\alpha, \beta\right)$. If $\left[\frac{d}{d\mu}, \frac{d}{d\lambda}\right] = 0$, the 2nd equation becomes

$$\frac{\partial x^i}{\partial \beta} = \exp\left(\beta \frac{d}{d\mu}\right)\exp\left(\alpha \frac{d}{d\lambda}\right)\left.\frac{dx^i}{d\mu}\right|_P$$

Consider now the basis vectors

$$\frac{\partial}{\partial\alpha}=\sum_i \frac{\partial x^i}{\partial\alpha}\frac{\partial}{\partial x^i} \qquad \text{and} \qquad \frac{\partial}{\partial\beta}=\sum_i \frac{\partial x^i}{\partial\beta}\frac{\partial}{\partial x^i}$$

we have

$$\frac{\partial}{\partial\alpha}=\exp\left(\beta\frac{d}{d\mu}\right)\exp\left(\alpha\frac{d}{d\lambda}\right)\sum_i \left.\frac{dx^i}{d\lambda}\right|_P \frac{\partial}{\partial x^i}$$

$$=\exp\left(\beta\frac{d}{d\mu}\right)\exp\left(\alpha\frac{d}{d\lambda}\right)\left.\frac{d}{d\lambda}\right|_P \quad =\left.\frac{d}{d\lambda}\right|_Q$$

where Q is the point with coordinates (α,β). Similarly,

$$\frac{\partial}{\partial\beta}=\left.\frac{d}{d\mu}\right|_Q$$

Since $\left\{\frac{\partial}{\partial\alpha},\frac{\partial}{\partial\beta}\right\}$ is a coordinate basis, so is $\left\{\frac{d}{d\lambda},\frac{d}{d\mu}\right\}$. Q.E.D.

To complete the proof, we need to show that the map (a) has an inverse. According to the inverse function theorem, the necessary and sufficient condition for this is

$$J=\frac{\partial\left(x^1,x^2\right)}{\partial(\alpha,\beta)} = \det\begin{vmatrix} \frac{\partial x^1}{\partial\alpha} & \frac{\partial x^2}{\partial\alpha} \\ \frac{\partial x^1}{\partial\beta} & \frac{\partial x^2}{\partial\beta} \end{vmatrix} \neq 0$$

This is the case if the vectors $\frac{\partial}{\partial\alpha}$ and $\frac{\partial}{\partial\beta}$ are linearly independent, i.e., $\bar{A}$ and $\bar{B}$ are independent.

2.16. One-Forms

A **one-form** is a linear, real- or complex- valued function of vectors.
Thus, a 1-form $\tilde{\omega}$ at P is a mapping

$$\tilde{\omega} : T_P \to \mathbb{K} \qquad \text{by} \qquad \bar{V} \mapsto \tilde{\omega}\left(\bar{V}\right)$$

such that

$$\tilde{\omega}\left(a\bar{V} + b\bar{W}\right) = a\tilde{\omega}\left(\bar{V}\right) + b\tilde{\omega}\left(\bar{W}\right) \qquad a, b \in \mathbb{K} \tag{2.15}$$

Defining the addition and scalar multiplication of 1-forms by

$$\left(a\tilde{\omega} + b\tilde{\sigma}\right)\left(\bar{V}\right) = a\tilde{\omega}\left(\bar{V}\right) + b\tilde{\sigma}\left(\bar{V}\right) \tag{2.16}$$

we see that the set of all 1-forms at P is a linear space T_P^* dual to T_P. The duality is easily seen by rewriting (2.16) as

$$\bar{V}\left(a\tilde{\omega} + b\tilde{\sigma}\right) = a\bar{V}\left(\tilde{\omega}\right) + b\bar{V}\left(\tilde{\sigma}\right) \tag{2.17a}$$

so that, combining with (2.15), we have

$$\tilde{\omega}\left(\bar{V}\right) = \bar{V}\left(\tilde{\omega}\right) \equiv \left\langle \tilde{\omega}, \bar{V} \right\rangle = \left\langle \bar{V}, \tilde{\omega} \right\rangle \tag{2.18}$$

Each expression in eq(2.18) is called the **contraction** of $\tilde{\omega}$ with $\bar{V}$. In tensor analysis, 1-forms are called **covariant** vectors.

2.17. Examples of One-Forms

1. Gradient of a function.
2. In matrix algebra, column vectors correspond to vectors and row vectors to 1-forms. Contraction corresponds to matrix multiplication with row vectors always on the left.
3. In Hilbert spaces used in quantum mechanics, the Dirac kets $|\psi\rangle$ are vectors while bras $\langle\phi|$ are 1-forms so that $\langle\phi|\psi\rangle \in \mathbb{C}$.

Note that the association of each vector $|\psi\rangle$ with another 1-form $\langle\psi|$, called its conjugate or transpose, is equivalent to introducing a **metric**, or **inner product**, to the vector space [see §2.29]

2.18. The Dirac Delta Function

As an example of **function spaces**, consider the set $C[-1,1]$ of all C^{∞} real-valued functions defined on the interval $x \in [-1,1]$. This set is a group under addition and a linear space with multiplication by real numbers as scalar multiplication. The 1-forms in its dual space are called **distributions**. One example of a distribution is the **Dirac delta function** $\delta(x)$ which is defined as the 1-form whose contraction with any C^{∞} function $f \in C[-1,1]$ is a number $f(0)$, i.e.,

$$\langle \delta(x), f(x) \rangle = f(0) \qquad (2.19)$$

Thus, a distribution is a "function" which maps functions to numbers. However, this is not what Dirac meant when he called $\delta(x)$ a delta "function". To see the distinction, we first note that for any function $g \in C[-1,1]$, one can define a 1-form $\tilde{g}$ such that its contraction with any function $f \in C[-1,1]$ is given by

$$\langle \tilde{g}, f \rangle = \int_{-1}^{1} dx\, g(x)\, f(x) \qquad (2.20)$$

The proof that $\tilde{g}$ defined this way is indeed a 1-form is straightforward. Now, what Dirac meant as a delta function is really the "function" $\delta(x)$ whose 1-form $\tilde{\delta}$ satisfies

$$\langle \tilde{\delta}, f \rangle = \int_{-1}^{1} dx\, \delta(x)\, f(x) = f(0)$$

The problem is that one cannot define $\delta(x)$ rigorously as a function $R^1 \to R^1$.

The special rule used by Dirac that $\delta(x)$ gives meaningful results only inside an integral is just an implicit statement that it is really a 1-form. For example, the "derivative" of $\delta(x)$ defined by

$$\int_{-1}^{1} dx\, \delta'(x)\, f(x) = -\int_{-1}^{1} dx\, \delta(x)\, f'(x) \quad = -f'(0)$$

is really the result of the derivative of the 1-form.

2.19. The Gradient and the Pictorial Representation of a One-Form

A **field** of 1-forms is a rule assigning a 1-form to each point in *M*. The contraction with a vector field is just the application of (2.16) to every point in *M*. Furthermore, the coefficients of linear combinations can be a function on *M*. Treating the contraction $\tilde{\omega}(\bar{V})$ as a function in *M*, the differentiability of $\tilde{\omega}$ can be defined in terms of that of $\bar{V}$ and $\tilde{\omega}(\bar{V})$. For example, if $\bar{V}$ and $\tilde{\omega}(\bar{V})$ are both C^{∞}, so is $\tilde{\omega}$. As with vector fields, there is a fibre bundle called the **cotangent bundle** TM^* and consists of *M* as base and T_P^* as typical fibre. Cross-sections of TM^* are 1-form fields.

The best known example of 1-form field is the **gradient** $\tilde{d}f$ of a function *f*. It is defined as the 1-form whose contraction with a tangent vector is the derivative of *f* along the integral curve of that tangent, i.e.,

$$\tilde{d}f\left(\frac{d}{d\lambda}\right)=\frac{df}{d\lambda} \qquad (2.21)$$

To see if $\tilde{d}f$ thus defined is indeed a 1-form, we need only check on its linearity:

$$\begin{aligned}\tilde{d}f\left(a\frac{d}{d\lambda}+b\frac{d}{d\mu}\right)&=\left(a\frac{d}{d\lambda}+b\frac{d}{d\mu}\right)f && \text{[by (2.21)]}\\ &=a\frac{df}{d\lambda}+b\frac{df}{d\mu} && \text{[linearity of differentiation]}\\ &=a\,\tilde{d}f\left(\frac{d}{d\lambda}\right)+b\,\tilde{d}f\left(\frac{d}{d\mu}\right)\end{aligned}$$

For 2 points with infinitesimal coordinate differences Δx^i, we have

$$\Delta f=\sum_i\frac{\partial f}{\partial x^i}\Delta x^i \quad =\tilde{d}f\left(\Delta\bar{x}\right)$$

Hence, a 1-form $\tilde{\omega}$ can be viewed as a series of parallel planes whose separation is inversely proportional to the magnitude of $\tilde{\omega}$. The contraction $\tilde{\omega}(\bar{V})$ is the number of planes pierced by $\bar{V}$. Note that this interpretation does not require the notion of lengths or metrics.

2.20. Basis One-Forms and Components of One-Forms

A basis $\{\bar{e}_i\}$ for T_P induces a **dual basis** $\{\tilde{\omega}^i\}$ for T_P^* so that

$$\tilde{\omega}^i\left(\bar{e}_j\right)=\delta^i_j \qquad (2.23)$$

Thus,

$$\tilde{\omega}^i\left(\bar{V}\right)=\sum_j \tilde{\omega}^i\left(V^j\bar{e}_j\right)=\sum_j V^j\delta^i_j=V^i \qquad (2.22)$$

The linear independence of $\{\tilde{\omega}^i\}$ follows immediately from (2.23). Alternatively, it can be proved as follows. Consider an arbitrary 1-form $\tilde{q}$ so that

$$\tilde{q}\left(\bar{V}\right)=\tilde{q}\left(\sum_j V^j\,\bar{e}_j\right)=\sum_j V^j\tilde{q}\left(\bar{e}_j\right)=\sum_j \tilde{\omega}^j\left(\bar{V}\right)\tilde{q}\left(\bar{e}_j\right) \qquad (2.24)$$

$$=\sum_j q_j V^j \qquad (2.27)$$

where

$$q_j=\tilde{q}\left(\bar{e}_j\right) \qquad (2.25)$$

are called the **components** of $\tilde{q}$ on the basis $\{\tilde{\omega}^i\}$ dual to $\{\bar{e}_i\}$. Note that (2.24) can be written as

$$\tilde{q}\left(\bar{V}\right)=\sum_j q_j\tilde{\omega}^j\left(\bar{V}\right)$$

Since this is valid for arbitrary $\bar{V}$, we have

$$\tilde{q}=\sum_j q_j\tilde{\omega}^j \qquad (2.26)$$

Since $\tilde{q}$ is arbitrary and the number of 1-forms in $\{\tilde{\omega}^i\}$ is equal to the dimension of T_P^*, $\{\tilde{\omega}^i\}$ is indeed a basis of T_P^*.

Extension of the foregoing to 1-form fields is straightforward.

To summarize, given a coordinate patch $\{x^i\}$ on a region $U\subset M$, there is a natural basis $\left\{\frac{\partial}{\partial x^i}\right\}$ for vector fields and $\{\tilde{d}x^i\}$ for 1-form fields with

$$\tilde{d}x^i \left(\frac{\partial}{\partial x^j} \right) = \frac{\partial x^i}{\partial x^j} = \delta^i_j \tag{2.28}$$

2.21. Index Notation

Hereafter, we shall adopt the following conventions for the use of indices:

1. Components V^j for vectors $\bar{V}$ have superscript indices.
2. Members $\bar{e}_j$ of a vector basis $\{\bar{e}_j\}$ are denoted by subscript indices.
3. Components q_j for 1-forms $\tilde{q}$ have superscript indices.
4. Members $\tilde{e}^j$ of a vector basis $\{\tilde{e}^j\}$ are denoted by superscript indices.
5. **Einstein's summation convention**: a pair of repeated super- and sub-scripts indicates summation, e.g.

$$\tilde{\omega} = \sum_j \omega_j \tilde{d}x^j \quad = \omega_j \tilde{d}x^j$$

$$\bar{V} = \sum_j V^j \frac{\partial}{\partial x^j} \quad = V^j \frac{\partial}{\partial x^j}$$

$$\tilde{\omega}(\bar{V}) = \sum_j \omega_j V^j \quad = \omega_j V^j$$

6. Note that there is no summation for $V^j W^j$ or $q_j p_j$.

2.22. Tensors and Tensor Fields

Consider a point P in M. A **tensor T** of **type** (**rank**) $\begin{pmatrix} N \\ M \end{pmatrix}$ at P is defined to be a linear function

$$\mathbf{T} : \left(T_P^*\right)^N \times \left(T_P\right)^M \to \mathbb{K}$$

by $$\left(\tilde{\omega}_{(1)}, \cdots, \tilde{\omega}_{(N)}; \bar{V}^{(1)}, \cdots \bar{V}^{(M)}\right) \mapsto \mathbf{T}\left(\tilde{\omega}_{(1)}, \cdots, \tilde{\omega}_{(N)}; \bar{V}^{(1)}, \cdots \bar{V}^{(M)}\right)$$

so that

$$\mathbf{T}\left(a\tilde{\omega}_{(1)} + b\tilde{\sigma}, \tilde{\omega}_{(1)}, \cdots, \tilde{\omega}_{(N)}; \bar{V}^{(1)}, \cdots \bar{V}^{(M)}\right)$$

$$= a\mathbf{T}\left(\tilde{\omega}_{(1)}, \cdots, \tilde{\omega}_{(N)}; \bar{V}^{(1)}, \cdots \bar{V}^{(M)}\right) + b\mathbf{T}\left(\tilde{\sigma}, \cdots, \tilde{\omega}_{(N)}; \bar{V}^{(1)}, \cdots \bar{V}^{(M)}\right) \qquad (2.29)$$

and similarly for each argument. Here, super- and sub-scripts in parentheses denote different vectors and 1-forms, respectively. Blanks can be used if one wishes to refer to a tensor without specifying its arguments, e.g., $\mathbf{S}(\ ,\ ,\ ;\ )$ denotes a tensor of type $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

A **tensor field** is a rule to assign a tensor to each point in M. Linearity (2.29) then applies to each point in M while the coefficients a and b can be functions on M. By convention, scalar functions are tensors of type $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. [see §2.28]

Note that the expression $\mathbf{T}(\tilde{\omega};\ )$ denotes a 1-form and $\mathbf{T}(\ ;\bar{V})$ a vector. Hence, $\mathbf{T}(\ ;\ )$ can be thought of as a linear vector-valued function of vectors or a linear 1-form-valued function of 1-forms.

2.23. Examples of Tensors

1. In matrix algebra, column vectors are vectors, row vectors are 1-forms, matrices are $\binom{1}{1}$ tensors, and similarity transformations of matrices are $\binom{2}{2}$ tensors.
2. In the function space $C[-1,1]$, linear differential operators are $\binom{1}{1}$ tensors since they convert vectors (functions in $C[-1,1]$) into other vectors.
3. A stress tensor gives the stress vector across a plane in a strained material. Since a plane (surface) is a 1-form, the stress tensor is a linear vector-valued function of 1-forms, i.e., a $\binom{2}{0}$ tensor.

2.24. Components of Tensors and the Outer Product

Given 2 vectors $\bar{V}$ and $\bar{W}$, we can form a $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ tensor $\bar{V} \otimes \bar{W}$ such that for any two 1-forms $\tilde{p}$ and $\tilde{q}$,

$$\bar{V} \otimes \bar{W}\left(\tilde{p}, \tilde{q}\right) = \bar{V}\left(\tilde{p}\right)\bar{W}\left(\tilde{q}\right) \tag{2.30}$$

Note that the linearity of $\bar{V} \otimes \bar{W}$ follows directly from that of the vectors. The operation $\otimes$ is called the **outer**, the **direct**, or the **tensor product**. In general, the outer product of a $\begin{pmatrix} N \\ M \end{pmatrix}$ tensor with a $\begin{pmatrix} N' \\ M' \end{pmatrix}$ tensor is a $\begin{pmatrix} N + N' \\ M + M' \end{pmatrix}$ tensor.

The **components** of a tensor are the values it takes when all the arguments are basis vectors and/or 1-forms. For example, in the bases $\left\{\tilde{\omega}^j\right\}$ and $\left\{\bar{e}_j\right\}$, the components of a type $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ tensor **S** are

$$S^{ijk}_{lm} = \mathbf{S}\left(\tilde{\omega}^i, \tilde{\omega}^j, \tilde{\omega}^k; \bar{e}_l, \bar{e}_m\right) \tag{2.31}$$

2.25. Contraction

According to §2.24, the direct product $\bar{V} \otimes \tilde{\omega}$ is a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor with components $V^i \omega_j$. The contraction $V^j \omega_j = \bar{V}(\tilde{\omega})$ is a number independent of basis (see below). Therefore, it is a **scalar function** or $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ tensor. In general, each contraction of a pair of upper and lower indices reduces the rank of the product tensor by $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

We shall show that the contraction of a $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ tensor **A** with a $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor **B** is independent of basis. To begin, let the componets of **A** and **B** be A^{ij} and B_{ij} in some basis. The contraction is then $A^{ij} B_{jk}$. Consider now the $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor defined by

$$\mathbf{C}(\tilde{\sigma};\bar{V}) = \left(A^{ij} B_{jk}\right)\sigma_i V^k \quad = C^i_k \sigma_i V^k$$

Using

$$\mathbf{A}\left(\tilde{\sigma},\tilde{\omega}^j\right) = \mathbf{A}\left(\sigma_i \tilde{\omega}^i,\tilde{\omega}^j\right) \quad = \sigma_i A^{ij}$$

and $$\mathbf{B}\left(\bar{e}_j,\bar{V}\right) = \mathbf{B}\left(\bar{e}_j, V^k \bar{e}_k\right) \quad = B_{jk} V^k$$

we have

$$\mathbf{C}(\tilde{\sigma};\bar{V}) = \mathbf{A}\left(\tilde{\sigma},\tilde{\omega}^j\right)\mathbf{B}\left(\bar{e}_j,\bar{V}\right) \quad = \mathbf{A}\left(\tilde{\sigma},\mathbf{B}\left(\bar{e}_j,\bar{V}\right)\tilde{\omega}^j\right)$$

$$= \mathbf{A}\left(\tilde{\sigma},\mathbf{B}\left(\ ,\bar{V}\right)\right)$$

Thus, **C** is independent of basis $\left\{\tilde{\omega}^i,\bar{e}_j\right\}$ used for **A** and **B**. Therefore, its components $C^i_k = A^{ij} B_{jk}$ is also independent of $\left\{\tilde{\omega}^i,\bar{e}_j\right\}$.

2.26. Basis Transformations

2.26.1. Transformation Matrix

Consider the tensors defined at a point *P* of *M*. The shift from one basis $\left\{\overline{e}_j \; ; \; j = 1, \cdots, n\right\}$ to another $\left\{\overline{e}_{j'} \; ; \; j' = 1, \cdots, n\right\}$ can be accomplished by a linear **transformation matrix** Λ so that

$$\overline{e}_{j'} = \Lambda^{j}_{j'} \, \overline{e}_j \qquad (2.32)$$

Since the mapping must be 1-1 onto, the matrix $\Lambda^{j}_{j'}$ must be non-singular. Note however that Λ is not a tensor. The natural 1-form basis with respect to $\left\{\overline{e}_j\right\}$ is given by

$$\tilde{\omega}^i \left(\overline{e}_j\right) = \delta^i_j \qquad (2.23)$$

Thus,

$$\Lambda^{j}_{j'} \, \tilde{\omega}^i \left(\overline{e}_j\right) = \tilde{\omega}^i \left(\Lambda^{j}_{j'} \, \overline{e}_j\right) \; = \tilde{\omega}^i \left(\overline{e}_{j'}\right)$$

$$= \Lambda^{j}_{j'} \delta^i_j = \Lambda^{i}_{j'} \qquad (2.33)$$

The inverse of $\Lambda^{j}_{j'}$ is denoted as $\Lambda^{j'}_{j}$ so that

$$\Lambda^{j}_{j'} \Lambda^{j'}_{i} = \delta^j_i \qquad \text{and} \qquad \Lambda^{j}_{j'} \Lambda^{k'}_{j} = \delta^{k'}_{j'} \qquad (2.34)$$

Thus, $\Lambda^{i'}_{i}$ (2.33) gives

$$\Lambda^{i'}_{i} \tilde{\omega}^i \left(\overline{e}_{j'}\right) \; = \Lambda^{i'}_{i} \Lambda^{i}_{j'} = \delta^{i'}_{j'} \qquad (2.35)$$

Therefore, the natural 1-form basis with respect to $\left\{\overline{e}_{j'}\right\}$ is

$$\tilde{\omega}^{i'} = \Lambda^{i'}_{i} \tilde{\omega}^i \qquad (2.36)$$

Note that the transformation matrix for 1-forms is the inverse of that for vectors.

Some examples of the use of the transformation matrix are

$$V^{i'} = \tilde{\omega}^{i'} \left(\overline{V}\right) \; = \Lambda^{i'}_{j} \tilde{\omega}^j \left(\overline{V}\right) \; = \Lambda^{i'}_{j} V^j \qquad (2.36)$$

$$q_{k'} = \tilde{q}\left(\overline{e}_{k'}\right) \; = \tilde{q}\left(\Lambda^{j}_{k'} \, \overline{e}_j\right) \; = \Lambda^{j}_{k'} \, \tilde{q}\left(\overline{e}_j\right) \; = \Lambda^{j}_{k'} \, q_j \qquad (2.37)$$

2.26.2. Coordinate Transformations

Consider a region $U \subset M$ with coordinate system $\left\{ x^i ; i = 1, \cdots, n \right\}$. Let $\left\{ y^{i'} ; i' = 1, \cdots, n \right\}$ be a new coordination system given by

$$y^{i'} = f^{i'}\left(x^1, \cdots, x^n \right) = f^{i'}\left(x^j \right) \qquad \text{for} \quad i' = 1, \cdots, n \tag{2.39}$$

with

$$J = \frac{\partial\left(y^{1'}, \cdots, y^{n'} \right)}{\partial\left(x^1, \cdots, x^n \right)} = \det\left| \frac{\partial y^{i'}}{\partial x^j} \right| \neq 0$$

By the chain rule,

$$\frac{\partial}{\partial y^{i'}} = \frac{\partial x^j}{\partial y^{i'}} \frac{\partial}{\partial x^j} \tag{2.40}$$

so that

$$\Lambda^j_{i'} = \frac{\partial x^j}{\partial y^{i'}} \tag{2.41}$$

Similarly,

$$\frac{\partial}{\partial x^j} = \frac{\partial y^{i'}}{\partial x^j} \frac{\partial}{\partial y^{i'}} \quad \Rightarrow \quad \Lambda^{i'}_j = \frac{\partial y^{i'}}{\partial x^j} \tag{2.42}$$

and

$$\Lambda^j_{i'} \Lambda^{i'}_k = \frac{\partial x^j}{\partial y^{i'}} \frac{\partial y^{i'}}{\partial x^k} = \frac{\partial x^j}{\partial x^k} = \delta^j_k$$

Note that the coordinate transformation (2.41-2) satisfies

$$\frac{\partial \Lambda^j_{i'}}{\partial y^{k'}} = \frac{\partial^2 x^j}{\partial y^{i'} \partial y^{k'}} = \frac{\partial \Lambda^j_{k'}}{\partial y^{i'}} \tag{2.43}$$

2.27. Tensor Operations on Components

Consider a tensor **T** with components $\left\{T^{i\cdots}_{j\cdots}\right\}$ in some basis. An operation on **T** that result in another tensor is called a **tensor operation**. Obviously, such operations are independent of basis. Some examples are:

1. Addition: $\mathbf{A},\mathbf{B} \rightarrow \mathbf{A}+\mathbf{B}$ with $\left\{A^{i\cdots}_{j\cdots}\right\},\left\{B^{i\cdots}_{j\cdots}\right\} \rightarrow \left\{A^{i\cdots}_{j\cdots} + B^{i\cdots}_{j\cdots}\right\}$.

2. Multiplication by a constant: $\mathbf{T} \rightarrow a\mathbf{T} = \left\{aT^{i\cdots}_{j\cdots}\right\}$.

3. Outer product: $\mathbf{A},\mathbf{B} \rightarrow \mathbf{A}\otimes\mathbf{B}$ with $\left\{A^{i\cdots}_{j\cdots}\right\},\left\{B^{k\cdots}_{l\cdots}\right\} \rightarrow \left\{A^{i\cdots}_{j\cdots}\ B^{k\cdots}_{l\cdots}\right\}$.

4. Contraction: $\left\{A^{i\cdots}_{jk\cdots}\right\},\left\{B^{jm\cdots}_{l\cdots}\right\} \rightarrow \left\{A^{i\cdots}_{jk\cdots}\ B^{jm\cdots}_{l\cdots}\right\}$.

2.28. Functions and Scalars

A **scalar function** is a function on *M* whose value is independent of the choice of basis. For example, the contraction $V^i\omega_i$ is a scalar but the vector field component $V^1(P)$ is not. On the other hand, we can define a scalar function *f* so that for every basis, $f(P)$ always equals to the numerical value of the 1st component $V^1(P)$ of the vector field $\bar{V}(P)$ in one particular basis. Thus, whether a function is a scalar depends on its interpretation, not its numerical value.

2.29. The Metric Tensor on a Vector Space

2.29.1. Inner Products and Metric Tensors

An **inner** (**dot**) **product** is an operation $\mathfrak{g}$ that maps 2 real vectors to a real number, i.e.,

$$\mathfrak{g}: T_P \times T_P \to \mathbb{R} \quad \text{by} \quad \bar{U}, \bar{V} \mapsto \mathfrak{g}\left(\bar{U}, \bar{V}\right) = \bar{U} \cdot \bar{V} \qquad (2.44)$$

such that

1. symmetric: $\mathfrak{g}\left(\bar{U}, \bar{V}\right) = \mathfrak{g}\left(\bar{V}, \bar{U}\right)$.

2. Given a basis $\{\bar{e}_i\}$, the matrix **g** with components

$$g_{ij} = \mathfrak{g}\left(\bar{e}_i, \bar{e}_j\right) = \bar{e}_i \cdot \bar{e}_j = g_{ji} \qquad (2.45)$$

   has an inverse.

Since **g** maps 2 vectors to a scalar, it is a $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor called the **metric tensor**. A manifold with a metric for its tangent spaces is called a **Riemannian manifold**. If there exists a basis such that

$$g_{ij} = \delta_{ij}$$

then the basis is called a **Cartesian basis**, the metric a **Euclidean metric** and the manifold a **Euclidean space**.

We mention in passing that a **metric space** is a topological space X with a **distance function** $d: X \times X \to \mathbb{R}$ by $x, y \mapsto d\left(x, y\right)$ for all *points* $x, y \in X$.

2.29.2. Canonical Forms

Under a change of basis, the metric tensor transforms as

$$g_{i'j'} = \Lambda^{i}_{i'} \Lambda^{j}_{j'} g_{ij} \qquad (2.46)$$

In matrix form, this becomes

$$\mathbf{g}' = \quad {}^{T} \qquad (2.47)$$

where $\mathbf{\Lambda}$ is a matrix with the (i, j') element equals to $\Lambda^{i}_{j'}$ and T is the transpose operation. Note that $\Lambda^{i'}_{i}$ is the (i', j) element of $\mathbf{\Lambda^{-1}}$. Since $\mathbf{\Lambda}$ is arbitrary, we can write it as

$$\mathbf{\Lambda} = \mathbf{OD} \qquad (2.48)$$

where $\mathbf{O}$ is orthogonal ($\mathbf{O}^{-1} = \mathbf{O}^{T}$) and $\mathbf{D}$ diagonal. Thus,

$$\mathbf{\Lambda}^{T} = (\mathbf{OD})^{T} = \mathbf{D}^{T}\mathbf{O}^{T} = \mathbf{DO}^{-1}$$

and (2.47) becomes

$$\mathbf{g}' = \mathbf{DO}^{-1}\,\mathbf{g}\,\mathbf{OD} \qquad (2.49)$$

Now, since $\mathbf{g}$ is real and symmetric, it is normal. Therefore, there exists an orthogonal transform that diagonalizes it. Setting $\mathbf{O}$ to be this transform, we have

$$\mathbf{O}^{-1}\,\mathbf{g}\,\mathbf{O} = \mathbf{g}_{d} = diag(g_1, \cdots, g_n)$$

where g_j are the eigenvalues of $\mathbf{g}$. Eq(2.49) thus becomes

$$\mathbf{g}' = \mathbf{D}\,\mathbf{g}_{d}\,\mathbf{D} = diag(d_1^2 g_1, \cdots, d_n^2 g_n) \qquad (2.50)$$

Since $d_j^2 \ge 0$, we can choose $d_j^2 = \frac{1}{|g_j|}$ so that $d_j^2 g_j = \mathrm{sgn}(g_j) = \pm 1$. [Note that $\mathbf{g}$ is invertible so that $g_j \ne 0$] In other words, there always exists a basis so that $\mathbf{g}$ is diagonal with elements ± 1. If $\mathbf{g} = diag(-1, \cdots, -1, 1, \cdots, 1)$, it is said to be in the **canonical form** and the corresponding basis is called **orthonormal**. The trace (sum of the diagonal elements) of the canonical metric tensor is called the **signature** of the metric. Note that the signature is invariant under any similarity transformations.

2.29.3. Characteristics of Metric Tensors

The signature is a fundamental characteristics of a metric. For example, if **g** is positive- definite, its signature is $+n$, where *n* is the dimension of the vector space, which must necessarily be Euclidean. If the eigenvalues of **g** are of mixed signs, the metric is called **indefinite**. One example is the Minkowski space with canonical form $\boldsymbol{\eta} = diag(-1,1,1,1)$ and signature +2.

Another important characteristics of the metric is the symmetry goup of the basis transformations that leave it invariant. For the Euclidean metric, this is the **orthogonal group** $O(n)$. For the Minkowski space, it is the **Lorentz group** $L(n) = O(n-1,1)$.

2.29.4. Raising and Lowering Indices

One important role of the metric tensor is to serve as the transformation between vector (upper index) and 1-form (lower index) components of a tensor. For example, for a given vector $\bar{V}$,

$$\tilde{V} = \mathfrak{g}\left(\bar{V}, \right) \qquad (2.54)$$

is a 1-form called the **dual** of $\bar{V}$. Since **g** is invertible, $\tilde{V}$ is unique. Its component is given by

$$V_i = \tilde{V}\left(\bar{e}_i\right) = \mathfrak{g}\left(\bar{V}, \bar{e}_i\right) = \mathfrak{g}\left(V^j \bar{e}_j, \bar{e}_i\right) = V^j \mathfrak{g}\left(\bar{e}_j, \bar{e}_i\right)$$

$$= V^j g_{ji} = g_{ij} V^j$$

which can be viewed as the lowering of the vector index of V^j by **g**. We now denote the components of the inverse of **g** by g^{ij} so that

$$g^{ij} g_{jk} = \delta^i_k \qquad (2.55)$$

we can raise index by

$$g^{ij} V_j = g^{ij} g_{jk} V^k = \delta^i_k V^k = V^i \qquad (2.56)$$

To summarize, we have

$$V_i = g_{ij} V^j \qquad (2.57)$$

$$V^i = g^{ij} V_j \qquad (2.58)$$

Other examples are

$$A^i_j = g_{jk} A^{ik} \qquad (2.59)$$

$$A_{ij} = g_{ik} A^k_j = g_{ik} g_{jl} A^{kl} \qquad (2.60)$$

$$A^{ij} = g^{ik} g^{jl} A_{kl} \qquad (2.61)$$

Thus, in a metric space, vectors and 1-forms are closely related. Thus, it is possible, and preferable, to refer to all tensors of type $\begin{pmatrix} N \\ M \end{pmatrix}$, where $N + M = L$, simply as tensors of order L. For example, all tensors of types $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ are order 2 tensors.

Note that in a Euclidean space with Cartesian basis, $g_{ij} = \delta_{ij}$ and $g^{ij} = \delta^{ij}$ so that $V^i = V_i$, i.e., there is no numerical difference between vectors and 1-forms. However, this is no longer the case if we shift to curvilinear coordinates.

Finally, for completeness, one should prove that g^{ij} is a $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ tensor. This is left as an exercise.

2.30. The Metric Tensor Field on a Manifold

2.30.1. Metric Tensor Field

A **metric field** is a rule to assign a metric tensor $\mathfrak{g}$ to every point *P* in *M*. Once given, it serves as metric on T_P and T_P^*. Furthermore, it also induces other geometric properties like lengths and curvatures into the manifold. Further discussion of this will be deferred to Chap 6. Here, we'll explore a few simple concepts.

First, the field $\mathfrak{g}$ should at least be continuous. This means its canonical form must be a constant everywhere since its diagonal elements are all integers that cannot change continuously. Thus, it is meaningful to speak of the signature of a metric tensor field.

In general, to make $\mathfrak{g}$ take the canonical form everywhere means using different basis transformations for different points. Moreover, this transformation field may not be coordination transformations. A notable exception is the Euclidean space $\mathbb{E}^n$ with Cartesian coordinates so that $g_{ij} = \delta_{ij}$ everywhere. However, even here only the Cartesian coordinates generate an orthonormal basis. For example, the polar coordinates (r,θ) in $\mathbb{E}^2$ generate an orthogonal basis $\left(\frac{\partial}{\partial r},\frac{\partial}{\partial \theta}\right)$ but it is not normalized. The normalized version is not a coordinate basis.

2.30.2. Locally Flat

A C^∞ metric tensor field $\mathfrak{g}$ is **locally flat** in the sense that there is always a coordinate basis $\{x^i\}$ for a neighborhood U about any point P such that

1. $g_{ij}(P) = \pm\delta_{ij}$ (canonical form at *P*)

2. $\left.\dfrac{\partial g_{ij}}{\partial x^k}\right|_P = 0$ (canonical form good approximation near *P*)

3. $\left.\dfrac{\partial^2 g_{ij}}{\partial x^k \partial x^l}\right|_P$ not necessarily zero. (canonical form not necessarily substained away from *P*)

2.30.3. Lengths

The **norm** $\|\bar{V}\|$ of a vector $\bar{V}$ is define as

$$\|\bar{V}\|^2 = \mathfrak{g}\left(\bar{V},\bar{V}\right) \quad = g_{ij}V^iV^j$$

If $\mathfrak{g}$ is positive-definite, $\|\bar{V}\|$ is real and non-negative. If $\mathfrak{g}$ is indefinite, $\|\bar{V}\|$ is called a **pseudo-norm**. A vector $\bar{V}$ is called **space-like** if $\|\bar{V}\|^2 > 0$, **time-like** if $\|\bar{V}\|^2 < 0$, and **null** if $\|\bar{V}\| = 0$. Note that null vectors are not necessarily the zero vector $V^i = 0$.

Consider a curve $C(\lambda)$ with tangent $\bar{V} = \dfrac{d}{d\lambda} \quad = \dfrac{dx^i}{d\lambda}\dfrac{\partial}{\partial x^i}$ so that

$$\mathfrak{g}\left(\bar{V},\bar{V}\right) = g_{ij}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}$$

The **length** *dl* of an infinitesimal segment on *C* is defined by

$$\left(dl\right)^2 \equiv dl^2 = \mathfrak{g}\left(\bar{V},\bar{V}\right)d\lambda^2 \qquad (2.62)$$

$$= g_{ij}dx^idx^j$$

If $\mathfrak{g}$ is positive-definite, we can write

$$dl = \sqrt{\mathfrak{g}\left(\bar{V},\bar{V}\right)}\, d\lambda \quad = \sqrt{g_{ij}dx^idx^j} \qquad (2.63)$$

where *dl* is real and positive. If $\mathfrak{g}$ is indefinite, the curves are classified by the sign of dl^2 the same way as for vectors with norm *dl*.

2.31. Special Relativity

The manifold $\mathbb{R}^4$ equipped with a metric of signature +2 is called the **Minkowski space** and serves as the spacetime manifold in special relativity. The canonical form of $\mathfrak{g}$ is $\boldsymbol{\eta} = diag(-1,1,1,1)$. The corresponding orthonormal basis can be provided by a class of global coordinate systems called **Lorentz frames**. Curves in the manifold are called **wordlines**. The arc length of an infinitesimal segment of wordline is an **event interval** and defined by

$$\Delta s^2 = \eta_{\alpha\beta} \Delta x^\alpha \Delta x^\beta = -\left(\Delta x^0\right)^2 + \left(\Delta x^1\right)^2 + \left(\Delta x^2\right)^2 + \left(\Delta x^3\right)^2 \qquad (2.66)$$

$$= -c^2 \left(\Delta t\right)^2 + \left(\Delta x\right)^2 + \left(\Delta y\right)^2 + \left(\Delta z\right)^2$$

Since $\boldsymbol{\eta}$ is indefinite, the **norm** it induces is a **pseudo-norm**. Inner product between vectors is given by

$$\bar{V} \cdot \bar{W} = \eta_{\alpha\beta} V^\alpha W^\beta \qquad (2.68)$$

In a Lorentz frame, the 1-form $\tilde{V}$ dual to $\bar{V}$ is given by

$$V_0 = \eta_{0\alpha} V^\alpha = -V^0 \qquad (2.69a)$$

$$V_i = \eta_{i\alpha} V^\alpha = V^i \qquad \text{for } i = 1,2,3 \qquad (2.69b)$$

The **vector gradient** $\bar{d}f$ of a function *f* is defined as the vector dual to the (1-form) gradient $\tilde{d}f = \frac{\partial f}{\partial x^\alpha} \tilde{d}x^\alpha$. Thus,

$$\left(\bar{d}f\right)^\alpha = \eta^{\alpha\beta} \left(\tilde{d}f\right)_\beta = \eta^{\alpha\beta} \frac{\partial f}{\partial x^\beta}$$

so that

$$\left(\bar{d}f\right)^0 = -\frac{\partial f}{\partial x^0}$$

$$\left(\bar{d}f\right)^i = \frac{\partial f}{\partial x^i} \qquad \text{for } i = 1,2,3$$

2.32. Bibliography