Physics 250
Fall 2008
Notes 1
Manifolds, Tangent Vectors and Covectors

1. Introduction

Some of the figures for these notes are drawn by hand, and are presented separately.

Most of the "spaces" used in physical applications are technically *differentiable manifolds*, and this will be true also for most of the spaces we use in the rest of this course. After a while we will drop the qualifier "differentiable" and it will be understood that all manifolds we refer to are differentiable. We will build up the definition in steps.

A differentiable manifold is basically a topological manifold that has "coordinate systems" imposed on it. Recall that a topological manifold is a topological space that is Hausdorff and locally homeomorphic to $\mathbb{R}^n$. The number n is the *dimension* of the manifold. On a topological manifold, we can talk about the continuity of functions, for example, of functions such as $f : M \to \mathbb{R}$ (a "scalar field"), but we cannot talk about the derivatives of such functions. To talk about derivatives, we need coordinates.

2. Charts and Coordinates

Generally speaking it is impossible to cover a manifold with a single coordinate system, so we work in "patches," technically *charts*. Given a topological manifold M, a *chart* on M is a pair (U, ϕ), where $U \subset M$ is an open set and $\phi : U \to V \subset \mathbb{R}^m$ is a homeomorphism. See Fig. 1. Since ϕ is a homeomorphism, V is also open (in $\mathbb{R}^m$), and $\phi^{-1} : V \to U$ exists. If $p \in U$ is a point in the domain of ϕ, then $\phi(p) = (x^1, \ldots, x^m)$ is the set of *coordinates* of p with respect to the given chart. We use upper (contravariant) indices for the coordinates. Note that the number of coordinates m is the dimension of M (otherwise ϕ cannot be a homeomorphism).

Given a chart, it becomes possible to talk about the smoothness of a scalar field on M, that is, a function $f : M \to \mathbb{R}$. That is, we just express f as a function of the coordinates $f(x^1, \ldots, x^m)$, and declare that f is smooth if this coordinate function is smooth. Actually, this is abuse of notation, since f depends on points p of M, not on coordinate m-tuples.

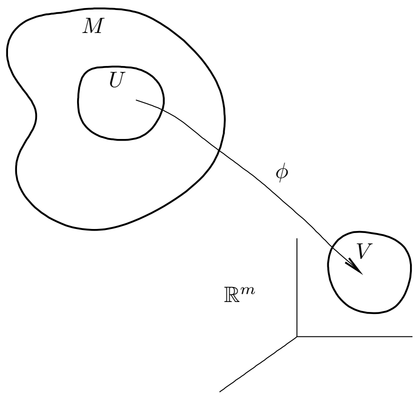

Fig. 1. A chart establishes coordinates in some open set U on a manifold M. The coordinates of $p \in U$ are $\phi(p) = (x^1, \ldots, x^m)$.

We should really write $f \circ \phi^{-1}$ instead of f above, that is, $f \circ \phi^{-1} : V \to \mathbb{R}$ is a real-valued function of a collection of real coordinates, so its partial derivatives are meaningful. Thus, smoothness is defined relative to a given chart.

3. Smoothness

In this course *smooth* will mean C^∞, that is, a smooth function is one that possesses continuous derivatives of all orders. It is convenient in differential geometry (and in this course) to assume that all functions are smooth (where this is meaningful), unless otherwise specified. Thus we will not have to qualify every theorem we encounter with warnings about how many continuous derivatives are needed for the theorem to be true (an irritating aspect of a lot of mathematical literature). We will not, however, assume that functions are analytic, since this is too strong for much of differential geometry, and is not needed in any case. (We make an exception to this when we discuss complex manifolds.)

4. Atlases

Since in general M cannot be covered by a single coordinate chart, we introduce the concept of an *atlas*, which is a collection of charts $\{(U_i, \phi_i)\}$ such that the open sets U_i cover M,

$$\bigcup_i U_i = M, \tag{1}$$

that is, every point of M is contained in at least one chart, and such that a certain *compatibility condition* is satisfied for any two charts that overlap. The compatibility condition is required so that functions that are smooth in one chart remain smooth in an overlapping chart. Let (U_i, ϕ_i), (U_j, ϕ_j) be two charts on M such that $U_i \cap U_j \neq \emptyset$ (see Fig. 2). In the diagram the two "coordinate spaces" $\mathbb{R}^m$ are drawn separately for convenience (two copies of $\mathbb{R}^m$), but you can combine them if you want. A point p in the overlap region $U_i \cap U_j$ has two sets of coordinates, the "i-coordinates" $\phi_i(p) = (x^1, \ldots, x^m)$ and the "j-coordinates" $\phi_j(p) = (x'^1, \ldots, x'^m)$. The i- and j- coordinates occupy regions (open sets) $R_i = \phi_i(U_i \cap U_j)$ and $R_j = \phi_j(U_i \cap U_j)$ of the two coordinate spaces, and by mapping coordinates to points and back to coordinates again we can define a map $\psi_{ij} : R_i \to R_j$, where

$$\psi_{ij} = \phi_j \circ \phi_i^{-1}. \tag{2}$$

(Actually ψ_{ij} is this function restricted to the domain R_i.) Note that ψ_{ij}^{-1} exists since ϕ_i and ϕ_j are invertible. Then the compatibility condition is that ψ_{ij} and ψ_{ij}^{-1} be smooth, which guarantees that the smoothness of functions does not depend on the chart. An *atlas* is a collection of charts that cover M such that any two charts with nonempty overlap satisfy this compatibility condition.

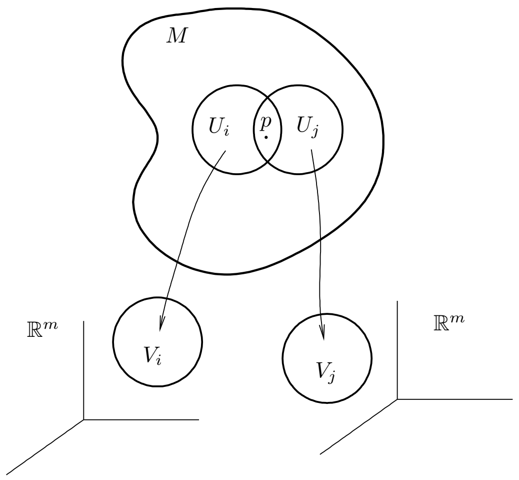

Fig. 2. Any two overlapping charts of an atlas must satisfy a compatibility condition in the overlap region.

In more ordinary language, the function ψ_{ij} gives the "new" coordinates as functions of the "old" coordinates,

$$x'^\mu = x'^\mu(x^\nu), \tag{3}$$

and ψ_{ij}^{-1} is the inverse function,

$$x^\nu = x^\nu(x'^\mu). \tag{4}$$

Thus, ψ_{ij} or ψ_{ij}^{-1} is a "coordinate transformation." The compatibility condition implies that all partial derivatives of all orders of both of these functions exist. This means in particular that the Jacobian matrix and its inverse,

$$J^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu}, \qquad (J^{-1})^\nu{}_\mu = \frac{\partial x^\nu}{\partial x'^\mu}, \tag{5}$$

exist. But since these matrices are inverses of one another, each matrix is nonsingular.

Two comments on this definition of charts and atlases. In practice, we often use coordinate systems for which the map ϕ or ϕ^{-1} does not exist everywhere, that is, there are "coordinate singularities." An example of this is ordinary spherical coordinates (θ, ϕ) at the north or south pole, where the angle ϕ is not defined, or on one meridian (Greenwich, for example) where the angle ϕ jumps discontinuously from 2π to 0. Such coordinates are excluded from the formalism presented here, or at least the domain U of definition of the coordinates must be restricted (excluding the "bad points") so that the map ϕ is a homeomorphism. If we do that, then ordinary spherical coordinates make a perfectly good chart for the sphere S^2, but the chart does not cover all of the sphere. This is an example of how in general a manifold cannot be covered by a single chart (for the sphere at least two overlapping charts are necessary).

A second comment is that in the applied, nongeometrical literature it is common to confuse a manifold with some standard system of coordinates on it (for example, Euler angles on $SO(3)$). Thus, coordinate singularities are confused with supposed "singularities" of the manifold itself (which do not exist). Understanding a geometrical object by means of some coordinate representation always runs the risk of confusing the idiosyncracies of the coordinate system with geometrical issues on the manifold itself that have an invariant meaning. It is part of the philosophy of geometrical methods that we think as much as possible in terms of the invariant (coordinate-independent) geometrical constructions on a manifold itself. Nevertheless, it must be noted that a differentiable manifold is one upon which coordinates can be imposed, so it is never wrong to look at something through a coordinate representation. Traditional (nongeometrical) physics literature tends to view things exclusively through coordinate representations (often examining how those representations change under a change of coordinates, in order to extract invariant meaning). The modern mathematical literature, on the other hand, sometimes goes to the other extreme, never putting anything in coordinate language no matter how convoluted it may be to say it in purely geometrical language. In this course we will try to strike a balance between these two points of view.

5. Differentiable Structure

Obviously there is an infinite number of ways of imposing an atlas on a given manifold. In order that two atlases agree as to which functions are smooth, it is necessary that the atlases satisfy a compatibility condition. This is that any two charts taken from either of the two atlases must satify the compatibility condition for charts given above. This compatibility condition for atlases is an equivalence relation, so the space of all possible atlases on a given topological manifold breaks up into equivalence classes, such that all atlases in a given equivalence class agree on which functions are smooth. An equivalence class of compatible atlases constitutes a *differentiable structure* on M, and a topological manifold plus a differentiable structure constitutes a *differentiable manifold*. Henceforth all manifolds we encounter will be assumed to be differentiable manifolds.

This definition of a differentiable manifold leaves open the possibility that a given topological manifold can possess more than one differentiable structure, and, indeed, this happens. In physical applications, however, manifolds are usually "born" with one obvious differentiable structure imposed on them (for example, they may arise as submanifolds of $\mathbb{R}^n$), and the alternative possibile differentiable structures play no role. As far as I know, there are no physical applications for such alternative differentiable structures. In this course we will always assume that there is one differentiable structure we are working with (usually an obvious one).

6. Manifolds With Boundary

What we have defined so far is a *manifold without boundary.* It is also possible to have a *manifold with boundary.* The official definition is given in the text, and will not be repeated here (it takes a little thinking to see what it means). Here we will just give two examples. The subset of $\mathbb{R}^3$ defined by

$$x^2 + y^2 + z^2 = 1 \tag{6}$$

is just the 2-sphere S^2, a manifold without boundary, $\partial S^2 = 0$. The subset of $\mathbb{R}^3$ defined by

$$x^2 + y^2 + z^2 \le 1 \tag{7}$$

is the 3-disk D^3, a manifold with boundary, $\partial D^3 = S^2$. Manifolds with boundary are more difficult to handle than manifolds without boundary, since the boundary points are exceptional and obey special rules. In the following, when we say simply "manifold," we will mean a manifold without boundary.

7. Maps Between Manifolds

Consider now mappings between (differentiable) manifolds, $f : M \to N$ (see Fig. 3). Let $\dim M = m$ and $\dim N = n$ (the manifolds do not have to have the same dimension). Let $x \in M$ and $y \in N$ be points such that $y = f(x)$, and let ϕ be a chart on M containing x and ψ a chart on N containing y. Then we will say that f is smooth at x if the coordinate representation of f (in the given charts), $\psi \circ f \circ \phi^{-1}$, is smooth. Note that in coordinates, the derivative matrix of f,

$$D^i{}_j = \frac{\partial y^i}{\partial x^j}, \tag{8}$$

is an $m \times n$ matrix (rectangular, if the dimensions of M and N are different).

As a special case, let $N = \mathbb{R}$. Then we have a *scalar field*, $f : M \to \mathbb{R}$. The set of all smooth scalar fields on a manifold will be denoted by

$$\mathfrak{F}(M) = \{f : M \to \mathbb{R} | f \text{ smooth}\}. \tag{9}$$

The coordinate functions $x^i : U \to \mathbb{R}$ (the i-th coordinate in a chart on M) are like scalar fields, but they are only defined locally (inside the chart U). For many purposes, they can be treated as scalar fields. This confuses people who have learned that a "scalar" is something that is invariant under changes of coordinates.

As another special case, consider $c : I \to M$, where $I = [a, b] \in \mathbb{R}$ is an interval. Then c is a *parameterized curve*, see Fig. 4.

Another example is a map $f : M \to N$ that is an isomorphism insofar as the differentiable structure is concerned. This is a mapping f that is a bijection (so f^{-1} exists), and such that both f and f^{-1} are smooth. Such a map is called a *diffeomorphism.* If there exists a diffeomorphism between two manifolds, they are said to be *diffeomorphic.* Notice that if "smooth" were replaced by "continuous", the definition would be that of a homeomorphism. Since smoothness implies continuity, a diffeomorphism is always a homeomorphism, but the converse is not true. But since M and N are homeomorphic if they are diffeomorphic, they must have the same dimension, $\dim M = \dim N$, since dimension is a topological invariant. This means that if you write f in terms of local charts on the two manifolds, the matrix $D^i{}_j$ of Eq. (8) above is a square matrix (the Jacobian). Moreover, since both f and f^{-1} are smooth, the inverse Jacobian exists, and both Jacobian matrices are nonsingular (just like in the compatibility condition between charts).

8. Tangent Vectors

Now we turn to the concept of a *tangent vector*. This is a case in which the official definition is somewhat unintuitive, but a good intuitive understanding is necessary to use differential geometry effectively. So we will begin with intuitive ideas, and build up to the formal definition.

The word "vector" has many different meanings. The original meaning of the word is something that carries something from one place to another, for example, a mosquito is a vector of disease. In $\mathbb{R}^n$, a vector is a displacement from one point to another along a straight line, and such vectors (if based at a common point such as the origin) can be added and multiplied by scalars, that is, they form a vector space. On a manifold (such as S^2, see Fig. 5), one can talk about displacements from one point p to another p_1 along some path, but such displacements cannot be added or multiplied by scalars (no such definition is useful). The exception is when the base p and tip p_1 of the displacement are close together, since a small region of the manifold can be approximated by a plane (of the same dimensionality as the manifold). If the manifold is imbedded in some higher dimensional $\mathbb{R}^n$ space, then the plane in question can be thought of as the tangent plane to the manifold at the point p. Then small vectors form something like a vector space (as long as they remain small, the tangent plane is a good approximation to the surface of the manifold).

A first, intuitive idea of a vector X tangent to a manifold M at a point p is simply a small displacement based at p, as illustrated in Fig. 5. There are two obstacles to be dealt with, however, before this can be turned into a precise definition. The first is that "small" has no precise meaning. In fact, usually when using tangent vectors there is some small elapsed parameter ("time") in the background, call it Δt, and the vector is to be seen not as one specific displacement, but rather part of a motion that takes place (going from p to p_1) in time Δt. By dividing suitable quantities by Δt, we obtain derivatives that are ordinary numbers and not infinitesimals. The motion in question can be thought of as a small part of a parameterized curve passing through p at $t = 0$.

The second obstacle is that we require a definition of a tangent vector that is intrinsic to the manifold, one that does not rely on any picture of an imbedding space with tangent planes sticking out into that space (and colliding with one another). If the manifold M (or its imbedding space) were a vector space, we could subtract points $p_1 - p$, and use that as a measure of the small displacement, but on a general manifold the subtraction of points is not meaningful. Therefore we shift attention to the functions defined on M, that is, the scalar fields, and consider the difference $f(p_1) - f(p)$, where $f : M \to \mathbb{R}$. This is meaningful, because we can subtract values of f. In this way we are led to associate with the tangent vector X a differential operator acting on scalar fields,

$$Xf = \lim_{\Delta t \to 0} \frac{f(p_1) - f(p)}{\Delta t} = \frac{df}{dt}(0), \tag{10}$$

where the d/dt is the convective derivative of f along the small segment of parameterized curve joining p and p_1. It is evaluated at $t = 0$ because we are assuming that the parameterized curve passes through p at $t = 0$. This is a second interpretation of a tangent vector at a point $p \in M$, that is, as a differential operator d/dt along a curve passing through p. By this interpretation, X is a linear mapping $X : \mathfrak{F}(M) \to \mathbb{R}$.

9. Equivalence Classes of Curves

The curve passing through p at $t = 0$ and p_1 at $t = \Delta t$ is free to do anything it wants when we move out of the neighborhood of p. In effect, we only need an infinitesimal segment of the curve to define the action of the d/dt operator. But we cannot talk precisely about an infinitesimal segment of a curve, so we talk about finite curve segments, that is mappings $c : [a, b] \to M$, $a < 0 < b$, which satisfy $c(0) = p$. We will consider any two such curves equivalent if they have the same d/dt operator at $t = 0$, that is, $c_1 \sim c_2$ if

$$\left.\frac{d(f \circ c_1)}{dt}\right|_{t=0} = \left.\frac{d(f \circ c_2)}{dt}\right|_{t=0}, \qquad \text{for all } f \in \mathfrak{F}(M). \tag{11}$$

Here we are being careful to indicate the convective derivative of a scalar field f along a curve c by $(d/dt)(f \circ c)$, since $f \circ c : [a, b] \to \mathbb{R}$ is a function whose t derivative is meaningful. (The simpler notation df/dt is abusive.) Then any curve in an equivalence class of such curves serves to define the same d/dt operator. This leads to a precise definition of a tangent vector to M at p: Such a vector X is an equivalence class of curves $X = [c]$, all passing through p at $t = 0$ and satisfying the above equivalence relation (see Fig. 6). This equivalence class is also interpreted as a map

$$X : \mathfrak{F}(M) \to \mathbb{R} : f \mapsto \frac{d(f \circ c)}{dt}(0), \tag{12}$$

for any c in the equivalence class X.

10. Tangent Vectors in Coordinates

Finally, let us introduce a chart on M containing p, and let $\{x^i\}$ be the coordinates. Then by the chain rule

$$\frac{df}{dt} = \sum_i \frac{\partial f}{\partial x^i} \frac{dx^i}{dt}, \tag{13}$$

where we revert to the simpler (but abusive) notation for the convective derivative of scalars. The convective derivative of an arbitrary scalar f is expressed in terms of the convective derivatives of the coordinates x^i. But for any scalar, $(df/dt)(0) = Xf$, where X is the given vector. Let us define

$$X^i = Xx^i = \frac{dx^i}{dt}(0), \tag{14}$$

and call the numbers $(X^1, \ldots, X^n)$ the *components* of X with respect to the coordinates $\{x^i\}$. Then

$$Xf = \sum_i X^i \frac{\partial f}{\partial x^i}, \tag{15}$$

where the partial derivatives are evaluated at p. Since f is arbitrary, we can write

$$X = \sum_i X^i \left. \frac{\partial}{\partial x^i} \right|_p, \tag{16}$$

which provides yet another interpretation of a tangent vector at $p \in M$, namely, as a *first order, linear, partial differential operator* acting at p (on scalar fields, producing a number).

Finally, let us consider what happens if we have two overlapping charts, with coordinates $\{x^i\}$ and $\{x'^i\}$. Then X has two representations,

$$X = \sum_i X^i \left. \frac{\partial}{\partial x^i} \right|_p = \sum_i X'^i \left. \frac{\partial}{\partial x'^i} \right|_p, \tag{17}$$

where

$$X'^i = Xx'^i = \frac{dx'^i}{dt}(0) = \sum_j \frac{\partial x'^i}{\partial x^j} \frac{dx^j}{dt}(0) = \sum_j \frac{\partial x'^i}{\partial x^j} X^j, \tag{18}$$

where the partial derivatives are evaluated at the point p. This is what would be called the transformation law for *contravariant vectors* in old-fashioned tensor analysis. In this sense, contravariant vectors are tangent vectors.

In Eq. (16) X is a vector and the X^i are just numbers, so the operators $(\partial/\partial x^i)|_p$ must be vectors. This is certainly correct, these are first order differential operators evaluated at p. But what is the equivalence class of curves associated with $(\partial/\partial x^i)|_p$? The answer is shown in Fig. 7, where the coordinate mesh (due to some chart) in the neighborhood of the point p is illustrated. One of the curves corresponding to $(\partial/\partial x^1)|_p$ is just the coordinate line passing through p on which all the x^i's are constant except x^1. The parameter along this line is just the coordinate x^1 itself. This curve (and any other equivalent to it in the sense of Eq. (11)) constitutes the equivalence class for $(\partial/\partial x^1)|_p$. This gives a nice interpretation to the rule of calculus, that $\partial/\partial x^1$ means holding all the x^i fixed except x^1, which is allowed to vary.

11. The Tangent Space

Equation (16) illustrates another fact, which is that the vectors $\{\partial/\partial x^i)|_p\}$ (in some chart) form a basis for the space of all tangent vectors to M at p. This space is denoted T_pM, and is called the *tangent space to M at p.* It is a real vector space that can be thought of as the space of all first order, linear differential operators evaluated at p. This is the intrinsic definition of the tangent space that we desired; notice that the tangent space at one point does not "collide" with the tangent space at any other point. Another fact which emerges from Eq. (16) is that the dimension of the tangent space is the same as that of the manifold,

$$\dim T_pM = \dim M. \tag{19}$$

This of course is what we expect, based on "tangent plane" diagrams like Fig. 5. All the different tangent spaces at different points are n-dimensional, real vector spaces ($n = \dim M$), and so are isomorphic to one another, although there is no *natural* isomorphism between them (unless we introduce additional geometrical structure).

12. Covectors

Now we turn to *covectors.* As we know, every real vector space V is associated with a dual space V^*, consisting of real-valued linear maps on that vector space. The dual space to the tangent space T_pM, $(T_pM)^*$, is usually denoted T_p^*M, and is called the *cotangent space* at $p \in M$. An element of the cotangent space is variously called a *covector*, *cotangent vector*, or *1-form.*

The most important example of a covector is one associated with a scalar field $f : M \to \mathbb{R}$, which a vector $X \in T_pM$ maps into a real number Xf by Eq. (15). In that equation, if we regard f as fixed and X as variable, we have the specification of a linear map : $T_pM \to \mathbb{R}$, associated with the scalar f, which by convention is denoted $df|_p$. That is, we define

$$df|_p : T_pM \to \mathbb{R} : X \mapsto Xf = \sum_i X^i \left.\frac{\partial f}{\partial x_i}\right|_p . \tag{20}$$

Thus, $df|_p$ is a covector at p. It is called the *differential* of the scalar f at p.

The most confusing thing for novices about this definition is that there is nothing small or "infinitesimal" about $df|_p$. In traditional theoretical physics the notation df usually denotes a small increment in the function f. This kind of notation was also used in mathematics in the nineteenth century and earlier, but in more modern times the meaning of the symbol df has morphed into something like that given by (20), in which $df|_p$ is an *operator* (acting on vectors), instead of a *value*. The relation between the two notations is the following. In the traditional notation, the small increment df is associated with a small change in the variables upon which f depends (this is usually implicit in the use of this notation). That is, there is a small displacement vector, call it X, in the space of variables upon which f depends (as we would say in old-fashioned language). In more modern language, would say that X is a small tangent vector at a point p to a manifold M, small because only a small movement in M results from a change $\Delta t = 1$ in the parameter of one of the curves associated with X. Then

$$(df)_{\text{traditional}} = (df|_p)(X). \tag{21}$$

In other words, the traditional interpretation of df is the (small) value of the operator $df|_p$ acting on a (small) vector. The operator $df|_p$ is not small.

13. Differentials of Coordinates

The coordinates x^i in some chart are examples of (local) scalar fields, and their differentials $dx^i|_p$ are of interest. If we allow one of these to act on an arbitary vector $X \in T_pM$, we obtain,

$$(dx^i|_p)(X) = Xx^i = X^i, \tag{22}$$

according to the definition (20) and Eq. (14). Now let $\alpha \in T_p^*M$ be an arbitary covector at p. We define the *components* of α with respect to a given chart by

$$\alpha_i = \alpha\Big(\frac{\partial}{\partial x^i}\Big|_p\Big) \tag{23}$$

(this is the normal definition of the components of a covector in V^*, you just let the covector act on a basis in V). We will frequently use Greek letters for covectors. Also, the use of a lower ("covariant") index on the components of a covector is conventional, just as we use upper ("contravariant") indices on the components of a tangent vector. Then we can evaluate the action of α on any vector $X \in T_pM$ in component language just by linearity. That is, we use Eq. (16) to write,

$$\alpha(X) = \alpha\Big(\sum_i X^i \frac{\partial}{\partial x^i}\Big|_p\Big) = \sum_i X^i \alpha\Big(\frac{\partial}{\partial x^i}\Big|_p\Big) = \sum_i X^i \alpha_i. \tag{24}$$

But by Eq. (22) this can also be written,

$$\alpha(X) = \sum_i \alpha_i (dx^i|_p)(X). \tag{25}$$

Since this is true for arbitrary X, we have

$$\alpha = \sum_i \alpha_i \, dx^i|_p, \tag{26}$$

and we see that the set $\{dx^i|_p\}$ (the differentials of the coordinates at p) forms a basis in T_p^*M. In fact, this basis in T_p^*M is dual to the basis $\{(\partial/\partial x^i)|_p\}$ in T_pM, as follows from setting $X = (\partial/\partial x^j)|_p$ in Eq. (22):

$$(dx^i|_p)\Big(\frac{\partial}{\partial x^j}\Big)\Big|_p = \frac{\partial}{\partial x^j} x^i = \delta^i_j. \tag{27}$$

Notice that the components of the differential of a function $df|_p$ are given by the "chain rule":

$$df|_p = \sum_i \frac{\partial f}{\partial x^i}\Big|_p (dx^i)|_p. \tag{28}$$

Finally, suppose we have two overlapping charts with coordinates $\{x^i\}$ and $\{x'^i\}$, so that $\alpha \in T_p^*M$ has two representations:

$$\alpha = \sum_i \alpha_i \, dx^i|_p = \sum_i \alpha'_i \, dx'^i|_p. \tag{29}$$

But by Eq. (28) we have

$$dx^i|_p = \sum_j \frac{\partial x^i}{\partial x'^j} \, dx'^j|_p, \tag{30}$$

or,

$$\alpha'_i = \sum_j \frac{\partial x^j}{\partial x'^i} \, \alpha_j, \tag{31}$$

which is the transformation law for the components of "covariant vectors" in old-fashioned tensor analysis. Thus, covectors or cotangent vectors are the same as "covariant vectors".

This concludes our basic introduction to tangent vectors and covectors on a manifold. Quite a few of the formulas on the preceding pages would be unnecessary if we were completely commited to "thinking geometrically," and avoiding coordinate representations as much as possible. But the traditional way of viewing problems in physics is through coordinate representations, so it is worthwhile to explain the connections with the more modern and abstract points of view. Notice that several of the formulas above are "obvious" under an old interpretation of the symbols involved, for example, Eq. (28) is the chain rule.

14. Tangent and Cotangent Vector Fields

So far we have only been talking about tangent and cotangent vectors *at a point*, but now we shall say a few things about tangent and cotangent vector *fields*. The basic idea is simple; a tangent vector field on a manifold M is an assignment of a tangent vector at each point of M, and similarly for a covector field. Nevertheless, to avoid confusion you will often find it important to distinguish carefully between a field and a geometrical object defined at a point. We will use the same symbols (X, Y, etc.) for tangent vector fields as we use for tangent vectors at a point, and likewise the same symbols (α, β, etc.) for covector fields as for covectors at a point. Thus you must keep clear by context which is meant (any notation distinguishing the two would be awkward). When talking about fields, we drop the specification of the point at which something is evaluated, and components of vectors or covectors become functions of position (although note that the components are only defined within the domain of a given chart). For example, the expression for a vector field X or covector field α in terms of its components can be written,

$$X = \sum_i X^i(x)\, \frac{\partial}{\partial x^i}, \qquad \alpha = \sum_i \alpha_i(x)\, dx^i, \tag{32}$$

where x represents a point of M with coordinates x^i.

15. Tangent and Cotangent Bundles

Another point of view on vector and covector fields relates them to *bundles*. We define the *tangent bundle* to a manifold M, denoted TM (without any p subscript) as the set of all tangent vectors to M at all points of M:

$$TM = \bigcup_{p\in M} T_pM, \tag{33}$$

and similarly the *cotangent bundle*, denoted T^*M, is the set of all covectors at all points p of M:

$$T^*M = \bigcup_{p\in M} T_p^*M. \tag{34}$$

A vector (at a point) $X \in TM$ is a tangent vector to M at some point p, so we can define a mapping (a projection) $\pi : TM \to M$ such that $\pi(X)$ is the point of M where X is attached. In other language,

$$\pi(T_pM) = p, \tag{35}$$

showing the action of π on the subset T_pM of TM (and thereby defining π). Similarly, for the cotangent bundle, we define a projection map $\pi : T^*M \to M$ such that if α is a covector (at a point) in T^*M, then $\pi(\alpha)$ is the point at which α is attached. Equivalently,

$$\pi(T_p^*M) = p. \tag{36}$$

Of course the two π's are not the same.

In terms of bundles, we can say that a vector field is a map $X : M \to TM$, such that the vector $X(p)$ is actually attached to p, that is, such that

$$\pi(X(p)) = p. \tag{37}$$

Similarly, a covector field is a map $\alpha : M \to T^*M$, such that

$$\pi(\alpha(p)) = p. \tag{38}$$

We shall denote the set of all (smooth) vector fields on M by $\mathfrak{X}(M)$, and the set of all (smooth) covector fields by $\mathfrak{X}^*(M)$.

Recall that a vector at a point is regarded as a map $X : \mathfrak{F}(M) \to \mathbb{R}$. For a vector field X, we have vectors acting on the scalar at every point, so a vector field can be considered a map, $X : \mathfrak{F}(M) \to \mathfrak{F}(M)$. In local coordinates, this is just

$$Xf = \sum_i X^i(x) \frac{\partial f}{\partial x^i}. \tag{39}$$

Similarly, a covector at a point is a map of a vector (at the same point) to a real number, so a covector field α can be regarded as a map $\alpha : \mathfrak{X}(M) \to \mathfrak{F}(M)$, given explicitly in local coordinates by

$$\alpha(X)(x) = \sum_i \alpha_i(x)\, X^i(x). \tag{40}$$