Physics 464/511 Lecture D Fall, 2016

By now we have introduced the basis differential operations for vector calculus, divergence, curl, and laplacian, in terms of their expressions in cartesian coordinates, and we have also learned about vectors, and the ability to express vectors in terms of different bases. As physics is just full of partial differential equations requiring manipulations with these ideas, we need to develop good formalisms for them.

At this point there are two ways to proceed. As we are usually dealing with fields defined on three-dimensional Euclidean space, and differential operators which include the Laplacian, one approach is to discuss the coordinate systems which can be used to describe 3-D Euclidean space and in terms of which the Laplacian is tractable, which is to say, separable. In addition to the ones you are quite familiar with, cartesian, cylindrical polar, and spherical polar, there are a dozen more, such as elliptic cylindrical or oblate spheriodal, that are useful for solving problems with certain symmetries. These coordinate systems are called *curvilinear coordinates*, and this is the way the books of Arfken proceed (especially the second edition, which gives the full set).

But not all fields of physics are defined in Euclidean space. We are now increasingly aware that we really live in a curved space determined by General Relativity. This space is described as a differentiable manifold, which is like a Euclidean space (or, including time, a Minkowski space) in the infinitesimal neighborhood of every point, but which differs from it both in the notions of distance and parallelism, and in that there may not be a single set of coordinates which smoothly describe the whole space. This is how one can describe curved spaces. Dealing with manifolds is the topic of differential geometry. We will follow this path, and only after discussing this more general theory, specialize to the useful curvilinear coordinates for Euclidean space.

People often have trouble understanding what it can mean for three dimensional space to be curved, because we are so accustomed to thinking that three dimensional Euclidean space is an absolute given. We have an easier time picturing two dimensional spaces which are curved, because we can picture some of them embedded in our Euclidean three dimensional space. Consider the surface of a sphere, or, using mathematical language correctly, consider a 2-sphere S^2 which is the surface of a three-dimensional ball. We are accustomed to using coordinates θ and ϕ, or latitude and longitude, to describe points on the sphere, but notice that

- These coordinates are singular at the north and south poles, as $\theta = 0$ (also $\theta = \pi$) is a single point[1] regardless of the value of ϕ, while in general on a two-dimensional surface you must specify two coordinates to define a point.
- The distance needs to be defined, but it is certainly not $\sqrt{(\Delta\theta)^2 + (\Delta\phi)^2}$.

In fact, we need to point out that the distance between two points on the two-dimensional space does not mean the distance between them in three dimensions, but rather the length of the shortest path **in the surface** between them, and that is a complex problem. More generally, when we describe a curved space what we say must be independent of any notion of it being embedded in a larger Euclidean space. The language in which to do this is the language of manifolds and differential geometry.

1 Manifolds

An n-dimensional smooth manifold $\mathcal{M}$ is a topological[2] collection of points $\mathcal{P}$, for which in some neighborhood of an arbitrary point $\mathcal{P}$, we can parameterize the points by n coordinates[3] q^j. That is, there is a 1–1 continuous correspondance between points in an open set $\mathcal{U}$ of $\mathcal{M}$ containing $\mathcal{P}$ and a *chart* C with cartesian coordinates q^j, which respects the topology, although the distances[4] between points on $\mathcal{M}$ are not necessarily those on the chart

[1] Another disagreement between mathematician's and physicist's notation: θ is the azimuthal angle (angle from the North pole) for physicists, while that is ϕ for mathematicians. We use physicist's language here.

[2] As mentioned in the first lecture, a topological space has a notion of open sets. We will mostly consider manifolds which also have a metric, defining distances, in which case the topology is given, as for $\mathbb{R}^n$, by declaring a set open if every point in it also contains a neighborhood (or ball) that lies within the set.

[3] We are going to be talking about covariant and contravariant vectors and tensors, which we distinguish notationally by placing the indices as subscripts or as superscripts. The problem of distinguishing a superscript index from a power is not usually a problem, judging in context. The coordinates of the charts are always indexed with superscripts.

[4] So far we have not restricted ourselves to manifolds that have metrics. For example, we might want to describe a 3-dimensional manifold of visible colors. But actually we will focus on Riemannian or pseudo-Riemannian manifolds which do have metrics, but when we do, finite distances on $\mathcal{M}$ will still not equal Euclidean distances on the charts.

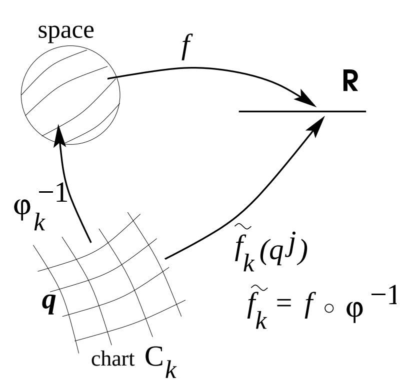
C. Thus, for example, Piscataway will appear on a chart C_1 that contains Middlesex County and a chart C_2 that contains New Jersey, and an open set on one will correspond to an open set on the other. Although Rand-McNally may tell us real distances (on $\mathcal{M}$) are proportional to distances on each chart C, that is a lie, because the surface of the Earth is not flat, and the chart is. Let us give the map from an open set $\mathcal{U}$ in $\mathcal{M}$ to C_k the name $\phi_k : \mathcal{U} \to C_k$. If two charts C_1 and C_2 both contain images of $\mathcal{U}$, the *transition function* map from C_1 to C_2 given by $\phi_2 \circ \phi_1^{-1}$ is a map from[5] $\phi_1(\mathcal{U}) \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and we insist that all our charts are such that the transition functions are differentiable (as many times as we need). Then the manifold is differentiable.

We require that every point in the manifold $\mathcal{M}$ appear on at least one chart. The set of charts for all the points $\mathcal{P} \in \mathcal{M}$ is called an *atlas*.

This is not so abstract — if you buy an atlas of the world, there are charts of each country, which agree on the topology but do not have the distances exactly correct. No single chart can cover the whole surface of the Earth, because the maps from $\mathcal{U} \subset \mathcal{M}$ into C_k have to be invertible, which excludes maps that have Hawaii on both the left and right edges.

The purpose of all this structure is so that we can discuss properties such as fields on the curved space, or manifold, in terms of things intrinsic to the space, not depending on the artificial structures like coordinate systems which we need to describe them. If we have a real scalar field defined on the manifold, $f : \mathcal{M} \to R$, for example the temperature in a continuous medium, we would like to describe the gradient of the temperature, and its properties should be describable in some way independent of which chart we use to describe its argument.

Of course the map ϕ_k into the chart C_k permits us to define a function $\tilde{f}_k(\{\mathbf{q}\})$ on the chart by $\tilde{f}_k(\{\mathbf{q}\}) = f(\phi_k^{-1}(\{\mathbf{q}\}))$. $\tilde{f}_k$ is thus a map from an open set in $\mathbb{R}^n$ into $\mathbb{R}$, so partial derivatives $\partial \tilde{f}_k / \partial q^j$ are well defined. The physics is in the function f, but that is hard to write down. The partial derivatives of $\tilde{f}_k$ are well defined, but they will, of course, depend on the parameterization $\mathbf{q}$ of C_k, and are not intrinsic properties of the field f.



[5]Of course $n = 2$ for Rand-McNally charts.

1.1 Distances and Metric

We will nearly always be interested in differentiable manifolds that have a defined notion of distance, at least for two points sufficiently close to each other. For a Riemannian manifold we require that in an infinitesimal neighborhood of any point $\mathcal{P}$ there exists a particular chart $C_{\mathcal{P}}$ with variables $\tilde{q}^j$ such that the measure is given by Euclidean rules, so that the distance between $\phi^{-1}(\tilde{\mathbf{q}})$ and $\phi^{-1}(\tilde{\mathbf{q}}+d\tilde{\mathbf{q}})$ is given by Pythagoras $(ds)^2 = \sum_j (d\tilde{q}^{\,j})^2$, to lowest order in $d\tilde{\mathbf{q}}$. We will also consider pseudo-Riemannian manifolds for which the chart is Minkowski, and $(ds)^2 = (d\tilde{q}^{\,0})^2 - \sum_{j=1}^3 (d\tilde{q}^{\,j})^2$. The statement that this is always possible, and that in this chart, locally, the laws of physics are what they would be without gravity, is the statement of the equivalence principle of general relativity.

Of course this requires a special chart for each point in space(-time), which is not very useful, and in a general chart containing the point $\mathcal{P}$, the expression in terms of the new dq^j will be more complicated. As the transition functions are differentiable, we will have $d\tilde{q}^{\,j} = \displaystyle\sum_k \frac{\partial \tilde{q}^{\,j}}{\partial q^k}\, dq^k$, so

$$(ds)^2 = \sum_\ell (d\tilde{q}^{\,\ell})^2 = \sum_{jk} g_{jk} dq^j dq^k, \quad \text{with} \quad g_{jk} = \sum_\ell \frac{\partial \tilde{q}^{\,\ell}}{\partial q^j}\frac{\partial \tilde{q}^{\,\ell}}{\partial q^k}.$$

The coefficients g_{jk} form the *metric tensor*. Note we may (and do) always take g_{jk} to be symmetric, $g_{jk} = g_{kj}$.

Thus the basic structure on the manifold is this metric, the distance between infinitesimally separated points. We can then define the length of a path as the integral, so if we have a parameterized path $\mathcal{P}(\lambda)$ which gets mapped into $q^j(\lambda)$ in a chart C, the length of the path between A and B is

$$\ell = \int_{\mathbf{q}(A)}^{\mathbf{q}(B)} \sqrt{\sum_{jk} g_{jk}(\mathbf{q})\, dq^j(\lambda) dq^k(\lambda)} = \int_{\mathbf{q}(A)}^{\mathbf{q}(B)} \sqrt{\sum_{jk} g_{jk}(\mathbf{q})\, \dot{q}^j(\lambda)\dot{q}^k(\lambda)}\; d\lambda\,,$$

where $\dot{q}^j$ means $dq^j/d\lambda$.

Einstein introduced the *summation convention*: indices occurring once upstairs and once downstairs are implicitly summed over. If greek, $\sum_0^3$; if latin, $\sum_1^3$. Einstein said this was the "greatest contribution of my life". Now

$$\ell = \int_{\mathbf{q}(A)}^{\mathbf{q}(B)} \sqrt{g_{jk}(\mathbf{q})\, \dot{q}^j(\lambda)\dot{q}^k(\lambda)}\; d\lambda\,.$$

See how much that helps!

The analogue of a straight line in Euclidean space is the geodesic, or shortest path, between two points on the manifold. We may find the *geodesic equation* for it using the same variational principle we use to get the equations of motion from a Lagrangian in classical mechanics. Let's review that:

In the Lagrangian presentation of a classical mechanical system, dynamics is described by the time evolution of some coordinates of a manifold, $q^j(t)$, and is determined by the Lagrangian function of those coordinates and their first time derivatives, $\dot{q}^j$, $L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The way the Lagrangian determines the dynamical path taken by the system is by the requirement that the action $I = \int L(\mathbf{q}, \dot{\mathbf{q}}, t)\, dt$, the time integral of the Lagrangian, be an extremum under variation of the path, holding the initial and final points $\mathbf{q}(t_i)$ and $\mathbf{q}(t_f)$ fixed. The path $q^j(t)$ can be the actual motion of the system if the variation

$$\delta I = \int \left(\frac{\partial L}{\partial q^j} \delta q^j(t) + \frac{\partial L}{\partial \dot{q}^j} \frac{d}{dt} \delta q^j(t) \right) dt = 0.$$

By integrating the second term by parts, and discarding the endpoints because $\delta q^j(t_i) = \delta q^j(t_f) = 0$ is required, we find the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^j} - \frac{\partial L}{\partial q_j} = 0.$$

Now let's consider finding the shortest path between two points on a Riemannian manifold. The action is replaced by the distance, ℓ, the time by the path parameter λ, and the Lagrangian (times dt) by the infinitesimal distance

$$ds = \sqrt{g_{jk}(\mathbf{q})\, \dot{q}^j \dot{q}^k}\, d\lambda,$$

where now $\dot{q}^j$ means $dq^j/d\lambda$. So we see that the Lagrangian is

$$L = \sqrt{g_{jk}(\mathbf{q})\, \dot{q}^j \dot{q}^k}$$

so $\quad L^2 = g_{jk}(\mathbf{q})\, \dot{q}^j \dot{q}^k, \quad 2L \frac{\partial L}{\partial q^j} = \frac{\partial g_{k\ell}}{\partial q^j}(\mathbf{q})\, \dot{q}^k \dot{q}^\ell \quad$ and $\quad 2L \frac{\partial L}{\partial \dot{q}^j} = 2 g_{jk}(\mathbf{q})\, \dot{q}^k$.

Thus Euler-Lagrange gives the geodesic equation

$$\frac{d}{d\lambda} \frac{1}{L} \left(g_{jk}(\mathbf{q})\, \dot{q}^k \right) - \frac{1}{2L} \frac{\partial g_{k\ell}}{\partial q^j}(\mathbf{q})\, \dot{q}^k \dot{q}^\ell = 0. \tag{1}$$

The $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ complicates things[6]

$$\text{Now} \qquad \frac{d}{d\lambda}\frac{1}{L}\left(g_{jk}(\mathbf{q})\,\dot q^k\right) = \frac{1}{L}\left(\frac{\partial g_{jk}}{\partial q^\ell}\dot q^\ell \dot q^k + g_{jk}\ddot q^k\right) - \frac{1}{2L^3}\left(g_{jk}\dot q^k\right)\frac{dL^2}{d\lambda}$$

$$\text{and} \qquad \frac{dL^2}{d\lambda} = \frac{\partial g_{jk}}{\partial q^\ell}\dot q^j \dot q^k \dot q^\ell + 2g_{jk}\ddot q^j \dot q^k = \dot q^k\left(2g_{jk}\ddot q^j + \frac{\partial g_{jk}}{\partial q^\ell}\dot q^j \dot q^\ell\right).$$

We define the Christoffel symbol (of the first kind)

$$\Gamma_{j,k\ell} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial q^\ell} + \frac{\partial g_{j\ell}}{\partial q^k} - \frac{\partial g_{k\ell}}{\partial q^j}\right) \tag{2}$$

(which is symmetric under $k \leftrightarrow \ell$) and note we may write

$$\frac{dL^2}{d\lambda} = 2\dot q^k\left(g_{kj}\ddot q^j + \Gamma_{k,j\ell}\dot q^j \dot q^\ell\right) . \tag{3}$$

Then the variational equation becomes

$$\frac{1}{L}\left(\frac{\partial g_{jk}}{\partial q^\ell}\dot q^\ell \dot q^k + g_{jk}\ddot q^k\right) - \frac{1}{L^3}\Big(g_{jk}\dot q^k \dot q^\ell\left[g_{\ell m}\ddot q^m + \Gamma_{\ell,mn}\dot q^m \dot q^n\right]\Big) - \frac{1}{2L}\frac{\partial g_{k\ell}}{\partial q^j}\dot q^k \dot q^\ell$$
$$= \frac{1}{L}\left(\delta^\ell_j - \frac{g_{jk}\dot q^k \dot q^\ell}{L^2}\right)\left[g_{\ell m}\ddot q^m + \Gamma_{\ell,mn}\dot q^m \dot q^n\right] = 0.$$

This equation has one free index (j, the others are summed over) and so appears to be n equations determining the functions $q^j(\lambda)$. But this is not correct, because if we take a linear combination of them by multiplying by $\dot q^j$ and summing, the first factor becomes

$$\dot q^j\left(\delta^\ell_j - \frac{g_{jk}\dot q^k \dot q^\ell}{L^2}\right) = \dot q^\ell - \frac{g_{jk}\dot q^k \dot q^k}{L^2}\dot q^\ell \equiv 0,$$

[6]The $1/L$ inside the derivative makes things rather messy, and we can derive Eq.(4) below with simpler algebra if we make use of the insight found laboriously below, by noting from the start that the length of any path is independent of the parameterization of it. So we could parameterize the arbitrary path with the length of that path starting from one end up to the point in question, divided by the total length of the path. Then $L = ds/d\lambda$ is constant along the path, we can pull $1/L$ out of the derivative and have $\frac{d}{d\lambda}\left(g_{jk}(\mathbf{q})\,\dot q^k\right) - \frac{1}{2}\frac{\partial g_{k\ell}}{\partial q^j}(\mathbf{q})\,\dot q^k \dot q^\ell = 0$. The scale factors out, so renaming $\dot q^k = \partial q^k/\partial s$, and using $\frac{dg_{jk}(\mathbf{q}(s))}{ds} = \frac{\partial g_{jk}}{\partial q^\ell}\dot q^\ell$, we see $g_{jk}\ddot q^k = \left(\frac{1}{2}\frac{\partial g_{k\ell}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^\ell}\right)\dot q^k \dot q^\ell = -\Gamma_{\ell,jk}\,\dot q^k \dot q^\ell$. This gives Eq.(4) upon raising the free index j.

so this linear combination vanishes without imposing any constraint on $q^j(\lambda)$. Thus we really have only $n-1$ equations, and $q^j(\lambda)$ is not determined.

There is a simple reason for this. Our parameterized path $\mathcal{P}(\lambda)$ would have the same length if it were parameterized by any other monotone parameter $\phi(\lambda)$, so $q'^j(\lambda) := q^j(\phi(\lambda))$ is the same path with a different parameterization but the same length, and clearly minimizing the length cannot determine the parameterization. In fact, we might have used this initially to simplify our calculation by insisting that we use a parameter proportional to the length s from one endpoint up to the point in question as our parameter, in which case $L = ds/d\lambda$ is constant, as we did in the footnote below Eq.(1). Then the factor in parenthesis in Eq. (3) must vanish, and the square bracket in the next equation vanishes by itself, giving n equations and determining $q^j(s)$.

We will have cause later to define a matrix $g^{j\ell}$ as the inverse of our metric $g_{\ell m}$, so $g^{j\ell} g_{\ell m} = \delta^j_m$, and to define $\Gamma^j{}_{mn} = g^{j\ell}\Gamma_{\ell,mn}$ to be the Christoffel symbol of the second kind. Multiplying the square bracket by $g^{j\ell}$ gives us the *geodesic equation*

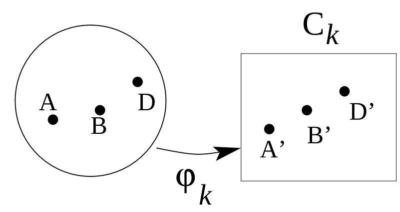
$$D^j := \ddot{q}^j + \Gamma^j{}_{mn}\dot{q}^m\dot{q}^n = 0\,, \tag{4}$$

where we are assuming the parameter is the distance s, and $\dot{q}^j = dq^j/ds$.

1.2 Vectors

We now want to include vectors. We are used to thinking of a vector as an object unchanged by translations, and being a vector in a mathematical sense, that is, taking linear combinations, *etc.*. In Euclidean space, the displacement, the difference of two points, is a vector. But vectors fields on a manifold are not so simple. What does the difference of two points in a curved space mean? I can subtract in $\mathbb{R}^n$, but not in $\mathcal{M}$. And how can I move a vector at one point in $\mathcal{M}$ to another to ask if the vectors change with position?

Consider three points, A, B, and D, in $\mathcal{M}$, which map into three points A', B', and D' in the chart $C_k \subset \mathbb{R}^n$. If $\Delta x = B' - A'$ and $D' - A' = 2\Delta x$, does that mean $D - A$ is twice $B - A$? Not at all, for such a statement depends on the chart C_k as well as any physical properties of the points A, B, and D. Then $D' - A' = 2(B' - A')$ will not be true for some other chart, or choice of coordinates. The same problem occurs in trying to discuss directions. In the chart C_k it appears $\vec{B}\,' - \vec{A}\,' = \vec{D}\,' - \vec{B}\,'$, but again this will not be true for some other chart, so it tells us nothing about the actual directions on the manifold.



Thus we cannot define vectors as finite differences of positions on a manifold. Still, we do have physical quantities that are vectors rather than scalars, and we need to have some way to describe them on a manifold. This is not a problem if the fields are vectors in some abstract space having nothing to do with the underlying space (or space-time) such as vectors in isotopic spin space. But most of the vectors we want to discuss, such as electric fields, are vectors in ordinary space. How should we define them?

Consider the electric field $\vec{E}(\mathcal{P})$. What does it mean? One way of thinking about it is that if we make any infinitesimal displacement $d\mathcal{P}$ to a charge Q, it will do work $Q\vec{E}\cdot d\mathcal{P}$, whatever that means. Let's consider the map of $d\mathcal{P}$ into the chart C, which will correspond to some $d\mathbf{q}$. If we consider an arbitrary infinitesimal displacement $\sum \omega_j dq^j$, the work $Q\vec{E}\cdot d\mathcal{P}$ is some real number (of Joules), so $Q\vec{E}$ is a linear function on the vector space (called the *cotangent space* $\mathcal{T}^*_{\mathcal{P}}$) spanned by dq^j, and so is part of the dual space $\mathcal{T}_{\mathcal{P}}$, called the *tangent space*. If we define the dual basis e_k to the basis dq^j, the vector $\vec{E} = \sum E^k e_k$ and $\vec{E}\cdot d\mathcal{P} = \sum E^k \omega_k$.

One way we get vector fields on a Riemannian manifold is to take the gradient of a scalar field $f : \mathcal{M} \to \mathbb{R}$,

$$\vec{V}(\mathbf{q}) = \lim_{\mathbf{q}'\to\mathbf{q}} \frac{f(\mathbf{q}') - f(\mathbf{q})}{ds(\mathbf{q}', \mathbf{q})}.$$

As the value of f at each point is well defined, independent of chart, and so is the distance ds, this is a chart-independent but abstract definition. It doesn't tell us what the components of the vector are. We could, however, describe the numerator as a differential,

$$df := f(\mathbf{q}') - f(\mathbf{q}) = \sum_j \frac{\partial \tilde{f}}{\partial q^j} dq^j$$

as a function $df : \mathcal{M} \times \mathcal{T}^*_{\mathbf{q}} \to \mathbb{R}$, where $\mathcal{T}^*_{\mathbf{q}}$ is an n-dimensional vector space with basis vectors dq^j. $\mathcal{T}^*_{\mathbf{q}}$ is called the cotangent space to $\mathcal{M}$ at the point $\mathcal{P}$ or $\mathbf{q}$. The differential df is an example of a *1-form*. It is independent of the chart, because under a change of coordinates $\partial \tilde{f}/\partial q^j$ transforms with the inverse of the matrix transforming dq^j. Thus the 1-form is intrinsic to the manifold, even though its components are chart-dependent. More generally, we may define a general 1-form on $\mathcal{M}$, $\boldsymbol{\omega}(\mathcal{P}) = \sum_j \omega_j(\mathcal{P})\, dq^j$, where the coefficients are chart-dependent and transform as do $\partial \tilde{f}/\partial q^j$, that is,

$$\omega'_j = \sum_k \frac{\partial q^k}{\partial q'^{\,j}}\, \omega_k. \tag{5}$$

This set of coefficients is called a covariant vector. Note that the 1-form itself can be considered, as we did when considering the work done by the electric field, as a displacement field, giving a displacement $d\mathbf{q}$ at each point $\mathcal{P}$ corresponding to $\omega_j(\mathcal{P})\, dq^j$ on the chart.

The mathematicians seem to like to do things differently. They like to define vectors as differential operators on scalar fields. That is, they define a vector $\mathbf{v}$ acting on a scalar field f as associated with a parametrized curve $\mathcal{P}(\lambda)$ with an image on C_k given by $q^j(\lambda)$ as $\mathbf{v}(f) = \sum_j \frac{\partial q^j}{\partial \lambda} \frac{\partial \tilde{f}_k}{\partial q^j}$. So the vector itself can be thought of as $\mathbf{v} = \sum_j \frac{\partial q^j}{\partial \lambda} \frac{\partial}{\partial q^j}$ with coefficients v^j which transform under change of chart by

$$v'^{\,j} = \frac{\partial q'^{\,j}}{\partial q^k} v^k, \tag{6}$$

which we call *contravariant*. Note the basis vectors of this space in the particular chart are the differential operators $\frac{\partial}{\partial q^k}$, which we will write more succinctly as ∂_k, so the basis vectors are

$$\mathbf{e}_k := \partial_k.$$

The space of possible vectors at the point $\mathcal{P}$ is called the tangent space $\mathcal{T}_\mathcal{P}$. Notice that $\mathbf{v}$ is defined for a given direction and "speed". Mathematicians then like to define define 1-forms as the dual to vectors, $\boldsymbol{\omega} : \mathcal{T}_\mathcal{P} \to \mathbb{R}$ by $\boldsymbol{\omega}(\mathbf{v}) = \sum_j \omega_j v^j$, which is chart-independent. Thus $\mathcal{T}^*_\mathcal{P}$ is the dual vector space to $\mathcal{T}_\mathcal{P}$, consistent with its notation.

It is important to keep in mind that in our definition, so far at least, the vector is defined at a particular point $\mathcal{P}$ of the manifold. Its tail is tied down, and there is no way (yet) to compare vectors defined at different points. The n-dimensional vector space spanned by $\mathbf{e}_k$ is the tangent space at the point $\mathcal{P}$, $\mathcal{T}_\mathcal{P}$. The collection of all the tangent spaces from all the points $\mathcal{P}$ of the manifold is called the *tangent bundle*, and similarly for the *cotangent bundle*.

Example: Consider a particle sensing a scalar field, for example temperature. As the particle passes the point $\mathcal{P}$, at what rate does its ambient temperature change? Consider a chart. Then $q^k(t)$ is the image of its position as a function of time t, and

$$\frac{dT}{dt} = \frac{\partial q^k}{\partial t}\frac{\partial T_C}{\partial q^k} = \frac{\partial q^k}{\partial t}\mathbf{e}_k(T)$$

is the rate of change of temperature. This is true for any other scalar field as well, so we have an operator $\mathbf{u} = \dfrac{\partial q^k}{\partial t}\dfrac{\partial}{\partial q^k}$ with

$\mathbf{u}$: scalar field $\mapsto$ time derivative of the field felt by a particle.

$\mathbf{u}$ is the velocity of the particle![7]

---

[7] In relativity, the manifold $\mathcal{M}$ is space-time, x^μ are the four coordinates in the chart, one of which is time, and a particle has a worldline which maps on the chart into $x^\mu(\tau)$, where τ is the proper time. Then

$$\frac{dT}{d\tau} = \frac{\partial x^\mu}{\partial \tau}\frac{\partial T_C}{\partial x^\mu} = \frac{\partial x^\mu}{\partial \tau}\mathbf{e}_\mu(T)$$

is the rate of change of temperature with respect to the proper time measured by the particle. Again this works for any other scalar field, so we have an operator

$\mathbf{u}$: scalar field $\to$ proper time derivative of the field felt by a particle.

$\mathbf{u}$ is the 4-velocity of the particle! Its four components on a given chart is what we call a 4-vector in courses on special relativity.

1.3 Tensor products and metric

Mathematicians tell us, and we discussed in Lecture B, that given two vector spaces U and V, we define the tensor product, so there is nothing new in things like $U \otimes V$. Physically it will only be useful to consider such tensor products of vectors defined at the same point. If e_j and e'_j are basis vectors of U and V respectively, a basis for $U \otimes V$ is $e_j \otimes e'_k$, and an arbitrary tensor in $U \otimes V$ is $\sum_{jk} v^{jk} e_j \otimes e'_k$. And of course we can tensor more than two vector spaces. We are particularly interested in tensors where each U and V is either $\mathcal{T}_\mathcal{P}$ (with basis vectors $e_j = \partial/\partial q^j$) or $\mathcal{T}^*_\mathcal{P}$ (with basis vectors $e^j = dq^j$). An arbitrary vector $\Gamma \in \mathcal{T}_\mathcal{P} \otimes \mathcal{T}^*_\mathcal{P} \otimes \mathcal{T}^*_\mathcal{P}$ will have coefficients ${\Gamma^j}_{k\ell}$. Note it is important to keep the order of the indices clear. The upper indices are called contravariant, and the lower indices covariant. Under change of chart, the coefficients transform with a product of partial derivative matrices, each according to Eq. (5) or Eq. (6).

We do have one example of such a tensor, the metric

$$(ds)^2 = g_{jk} dq^j dq^k.$$

As this is to be an intrinsic field on the manifold, the coefficients g_{jk} must transform as a covariant tensor,

$$g'_{jk} = \frac{\partial q^\ell}{\partial q'^j} \frac{\partial q^m}{\partial q'^k} \, g_{\ell m}.$$

As an element of $\mathcal{T}^*_\mathcal{P} \otimes \mathcal{T}^*_\mathcal{P}$, it is intended to give a real value when applied to two contravariant vectors, say $\mathbf{u} = u^j \partial_j$ and $\mathbf{v} = v^k \partial_k$, that is $\mathbf{g}(\mathbf{u}, \mathbf{v}) = g_{jk} u^j v^k$. But that means that associated with each contravariant vector $\mathbf{u}$ in $\mathcal{T}_\mathcal{P}$, we have a natural covariant vector $\tilde{\mathbf{u}} = \mathbf{g}(\mathbf{u}, \cdot)$ in $\mathcal{T}^*_\mathcal{P}$ defined by $\tilde{\mathbf{u}}(\mathbf{v}) = \mathbf{g}(\mathbf{u}, \mathbf{v}) = g_{jk} u^j v^k$. We write the expansion of $\tilde{\mathbf{u}}$ in terms of the basis elements dq^j of $\mathcal{T}^*_\mathcal{P}$ as $\tilde{\mathbf{u}} = u_j dq^j$. But the basis elements act on contravariant vectors by $dq^j(\mathbf{v}) = v^j$, so we must have $u_j v^j = g_{jk} u^j v^k$ for any $\mathbf{v}$, or

$$u_k = g_{jk} \, u^j, \quad \text{and also} \quad u^j = g^{jk} u_k, \quad \text{with} \quad g^{jk} g_{k\ell} = \delta^j_\ell,$$

that is, $g^{\cdot\cdot}$ is the inverse matrix to $g_{\cdot\cdot}$.

This is general — we can convert between contravariant tensor indices and covariant ones by multiplying by the metric tensor $g_{\cdot\cdot}$ or its inverse $g^{\cdot\cdot}$. In particular, we did this for the Christoffel symbol of the first kind to produce the Christoffel symbol of the second kind,

$${\Gamma^j}_{mn} := g^{jk} \Gamma_{k,mn} \, .$$

1.4 n-forms

We saw that the differential of a scalar function f gives us a 1-form df, but this is not the most general 1-form in the cotangent bundle, which is more generally

$$\boldsymbol{\omega}(\mathcal{P}) = \omega_j(\mathbf{q}) dq^j.$$

As we mentioned earlier, $\boldsymbol{\omega}$ is chart independent because the coefficients ω_j transform covariantly. If a 1-form $\boldsymbol{\omega}$ is the differential of a scalar f,

$$\omega_j = \frac{\partial f}{\partial q^j} \quad \Longrightarrow \frac{\partial \omega_j}{\partial q^k} = \frac{\partial}{\partial q^k}\frac{\partial}{\partial q^j} f = \frac{\partial}{\partial q^j}\frac{\partial}{\partial q^k} f = \frac{\partial \omega_k}{\partial q^j},$$

which will not be true for a general 1-form. In fact, we define the differential of $\boldsymbol{\omega}$ to be the 2-form

$$d\boldsymbol{\omega} = \frac{1}{2}\sum_{jk}\left(\frac{\partial \omega_k}{\partial q^j} - \frac{\partial \omega_j}{\partial q^k}\right) dq^j \wedge dq^k = \sum_{jk} \frac{\partial \omega_k}{\partial q^j} dq^j \wedge dq^k,$$

where we define[8] $dq^j \wedge dq^k = \frac{1}{2}\left(dq^j \otimes dq^k - dq^k \otimes dq^j\right)$.

Intuitively, just as dq^j can be considered one component of an infinitesimal line element, $dq^j \wedge dq^k$ can be thought of as an infinitesimal piece of a surface.

We see that $\boldsymbol{\omega}$ is *exact*, which means it is the differential of a scalar function or 0-form, only if its differential vanishes, in which case we say $\boldsymbol{\omega}$ is *closed*.

We saw that a 1-form is a natural way to describe a gradient. In Euclidean space in cartesian coordinates x^j, we saw that the components of the gradient $\vec{\nabla} f$ are just the coefficients of dx^j in the 1-form df. We can consider that the dx^j are basis vectors, and if we use other coordinates dq^j, we can still consider dq^j as basis vectors, but in general not orthonormal ones. If we associate a general 1-form $\boldsymbol{\omega} = \sum_j \omega_j dx^j$ in cartesian coordinates with a vector $\vec{V}$ with components ω_j, the components of the curl,

$$\left(\vec{\nabla} \times \vec{V}\right)_i = \epsilon_{ijk} \frac{\partial}{\partial x^j} V_k$$

give the coefficients of $\frac{1}{2} dx^j \wedge dx^k$ in $d\boldsymbol{\omega}$.

Note that if we are in a Euclidian space in other than 3 dimensions, the same differential operator could be defined, but we could not replace the two

[8] See next footnote about ambiguity on whether to include the $\frac{1}{2}$.

indices j and k by a single index i with the ϵ symbol. In three dimensions an exact 2-form is a curl, but in other dimensions we still have exact 2-forms, but they are not associated with a vector, but with an antisymmetric two-index tensor.

These are the first two steps in defining the *exterior algebra* $\mathcal{E}(\mathcal{M})$ of $\mathcal{M}$, and of the *exterior derivative* d. More generally, the wedge product of n basis elements dq^j of 1-forms is the totally antisymmetric product,

$$\boldsymbol{\omega}_1 \wedge \boldsymbol{\omega}_2 \wedge \ldots \wedge \boldsymbol{\omega}_n = \frac{1}{n!} \sum_{P \in S_n} (-1)^P \boldsymbol{\omega}_{P(1)} \otimes \boldsymbol{\omega}_{P(2)} \otimes \ldots \otimes \boldsymbol{\omega}_{P(n)}$$

where the sum is over permutations P of the n indices, and $(-1)^P$ is the sign of the permutation. The exterior derivative of an n-form[9]

$$\boldsymbol{\omega}(\{\mathbf{q}\}) = \frac{1}{n!} \sum_{j_1, j_2 \cdots j_n} \omega_{j_1, j_2 \cdots j_n}(\{\mathbf{q}\})\, dq^{j_1} \wedge dq^{j_2} \wedge \ldots \wedge dq^{j_n}$$

is the $n{+}1$-form

$$d\boldsymbol{\omega} = \frac{1}{n!} \sum_{j_0, j_1, j_2 \cdots j_n} \frac{\partial \omega_{j_1, j_2 \cdots j_n}(\{\mathbf{q}\})}{\partial q_{j_0}}\, dq^{j_0} \wedge dq^{j_1} \wedge dq^{j_2} \wedge \ldots \wedge dq^{j_n}.$$

Again, $\boldsymbol{\omega}$ is called exact if there is an $(n{-}1)$-form $\boldsymbol{\nu}$ such that $\boldsymbol{\omega} = d\boldsymbol{\nu}$, and $\boldsymbol{\omega}$ is called closed if $d\boldsymbol{\omega} = 0$. And again, exact implies closed.

In cartesian coordinates in three dimensional Euclidean space, we can associate a vector $V = v^i \hat{e}_i$ with a 1-form $\boldsymbol{\omega} = \sum_i v^i dx^i$, but we can also associate a vector with a 2-form $\boldsymbol{\omega}^{(2)} = \frac{1}{2} \sum_{jk} B_{jk}\, dx^j \wedge dx^k$ with $B_{jk} = \sum_i \epsilon_{ijk} v^i$, or $v^i = \frac{1}{2} \epsilon^{ijk} B_{jk}$. So if we have a vector $\vec{V} = v^i \hat{e}_i$ associated with $\boldsymbol{\omega}$, then $\boldsymbol{\omega}^{(2)} = d\boldsymbol{\omega} = \sum_{jk} \frac{\partial v^k}{\partial x^j}\, dx^j \wedge dx^k$ has $B_{jk} = 2 \frac{\partial v^k}{\partial x^j}$, and is associated with a vector $\vec{W} = w^i \hat{e}_i$ with $w^i = \frac{1}{2} \epsilon^{ijk} B_{jk} = \epsilon^{ijk} \frac{\partial v^k}{\partial x^j}$. So we see $\vec{W} = \vec{\nabla} \times \vec{V}$.

[9]Whether or not there is an $n!$ in the relation between the form and the coefficients seems to be unsettled. With $1/n!$ this seems to agree with Vaughn. The same is true for the wedge product, above, where I agree with Vaughn but disagree with MTW p. 99, and Wikipedia waffles on this.

Now consider the vector $\vec{U} = u^i \hat{e}_i$ associated directly with the 2-form $\boldsymbol{\omega}^{(2)} = \frac{1}{2}\sum_{jk} B_{jk}\, dx^j \wedge dx^k$ with $B_{jk} = \sum_i \epsilon_{ijk}\, u^i$. Its differential is

$$d\boldsymbol{\omega}^{(2)} = \frac{1}{2}\sum_{\ell jk} \frac{\partial B_{jk}}{\partial x^\ell}\, dx^\ell \wedge dx^j \wedge dx^k = \frac{1}{2}\sum_{i\ell jk} \epsilon_{ijk} \frac{\partial u^i}{\partial x^\ell}\, dx^\ell \wedge dx^j \wedge dx^k.$$

But there is only one 3-form in three dimensions, up to a scalar field, so we define the standard volume element in three dimensions[10]

$$\boldsymbol{\Omega} = \frac{1}{6}\epsilon_{\ell jk} dx^\ell \wedge dx^j \wedge dx^k$$

(which means $dx^\ell \wedge dx^j \wedge dx^k = \epsilon^{\ell jk}\, \boldsymbol{\Omega}$). Then we have

$$d\boldsymbol{\omega}^{(2)} = \frac{1}{2}\sum_{i\ell jk} \epsilon_{ijk}\, \epsilon^{\ell jk}\, \frac{\partial u^i}{\partial x^\ell}\, \boldsymbol{\Omega} = \sum_i \frac{\partial u^i}{\partial x^i} \boldsymbol{\Omega} = (\vec{\nabla} \cdot \vec{U})\, \boldsymbol{\Omega}.$$

So we see that the exterior derivative maps

$$\begin{array}{ccccccc} f & \underset{d}{\longrightarrow} & \vec{V} = \vec{\nabla} f & & \vec{V} & \underset{d}{\longrightarrow} & \vec{\nabla} \cdot \vec{V} \\ & & \updownarrow & & \updownarrow & & \\ & & \vec{V} & \underset{d}{\longrightarrow} & \vec{\nabla} \times \vec{V} & & \end{array}$$

scalar into its gradient, vector (as 1-form) into its curl, vector (as 2-form) into its divergence. Note that applying the exterior derivative twice in a row gives zero, $d^2 = 0$, because exact forms are closed.

Note that the n-forms are intrinsic to the manifold, so the relations between n-forms are not chart-dependent, but the relationships we have given between vectors described in our usual orthonormal notation and the n-forms requires the coordinates to be cartesian, until we covariantize their notation. In particular, we will need to covariantize ϵ.

Next time: Integration

[10] We will see why we call it that in the next lecture.