

Physics 570

Tangent Vectors and Differential Forms over Manifolds

Geometrical Requirements for a Local Description

I. Manifolds, at the lowest-possible degree of precision

As everyone “knows,” general relativity is associated with “curved spaces,” or, more precisely, with “curved spacetimes.” Therefore, we first need to understand what—in a relatively clear, mathematical sense—is meant by the word “curved” in this context. We do this by beginning with the idea of a *manifold*, perhaps at a similar level of detail as in Carroll’s text. However, do please read his descriptions of manifolds, which have very nice figures.

A manifold is a set of points such that every point has a neighbourhood that “looks like” some open set of some ordinary, flat, m -dimensional space, i.e., like \mathbb{R}^m , for some fixed integer m . The phrase “looks like” means that we may use, locally, a set of m coordinates just as we would in an ordinary flat space, i.e., a piece of \mathbb{R}^m , although we must surely allow the choice of coordinates to vary as we move over the manifold. It will be desirable to restrict this varying however so that the coordinates are continuous and invertible functions. In fact we also want mappings to be “sufficiently continuous” to allow their differentiation, presumably as many times as we think appropriate. Therefore, we must allow coordinates on a manifold, with some restrictions perhaps, to be as differentiable as we like. This induces us to introduce **the general notion of a ‘smooth’ manifold.**

Defn. 1. A *chart* for a manifold \mathcal{M} is a set of **invertible** mappings, $\{\varphi_1, \varphi_2, \dots\}$, and associated open neighborhoods, $\{U_1, U_2, \dots\}$, which together cover all of \mathcal{M} . Each such mapping sends its neighborhood into (an open subset of) \mathbb{R}^m ; i.e., we have the following, where we label an arbitrary point in the manifold by the symbol P :

$$\forall P \in \mathcal{M}, \exists \text{ within the chart at least one } U \subseteq \mathcal{M} \text{ and } \varphi: U \rightarrow \mathbb{R}^m \quad (0.1)$$

such that $P \in U$, and $\varphi(P) \equiv (\varphi^1(P), \varphi^2(P), \dots, \varphi^m(P)) \in \mathbb{R}^m$.

The quantities $\{\varphi^k(P)\}$, where k varies from 1 to m , are referred to as the *coordinates for P* relative to the choice of open set U . The entire set of values of this type are then *a coordinate system for U* , relative to the mapping φ .

We acquire differentiability for functions over our manifold by considering behavior in overlaps of coordinate neighborhoods. For an arbitrary point $P \in \mathcal{M}$, consider two distinct U_i 's, say U_1 and U_2 , both of which contain P . Then, the mapping $\varphi_2 \circ \varphi_1^{-1}$ sends \mathbb{R}^m to \mathbb{R}^m ; we insist that it be differentiable arbitrarily often. [These functions are called *transition functions*.] There are actually different sorts of manifolds, having differing levels of continuity for the transition functions; in our case, **we will insist that all transition functions must be C^∞ , and refer to such manifolds as “smooth.”**

Defn. 2. A set of charts that contains the largest possible number of charts for a given manifold, \mathcal{M} , consistent with the requirement that **all** transition functions be smooth, is called a (maximal) smooth *atlas* for \mathcal{M} .

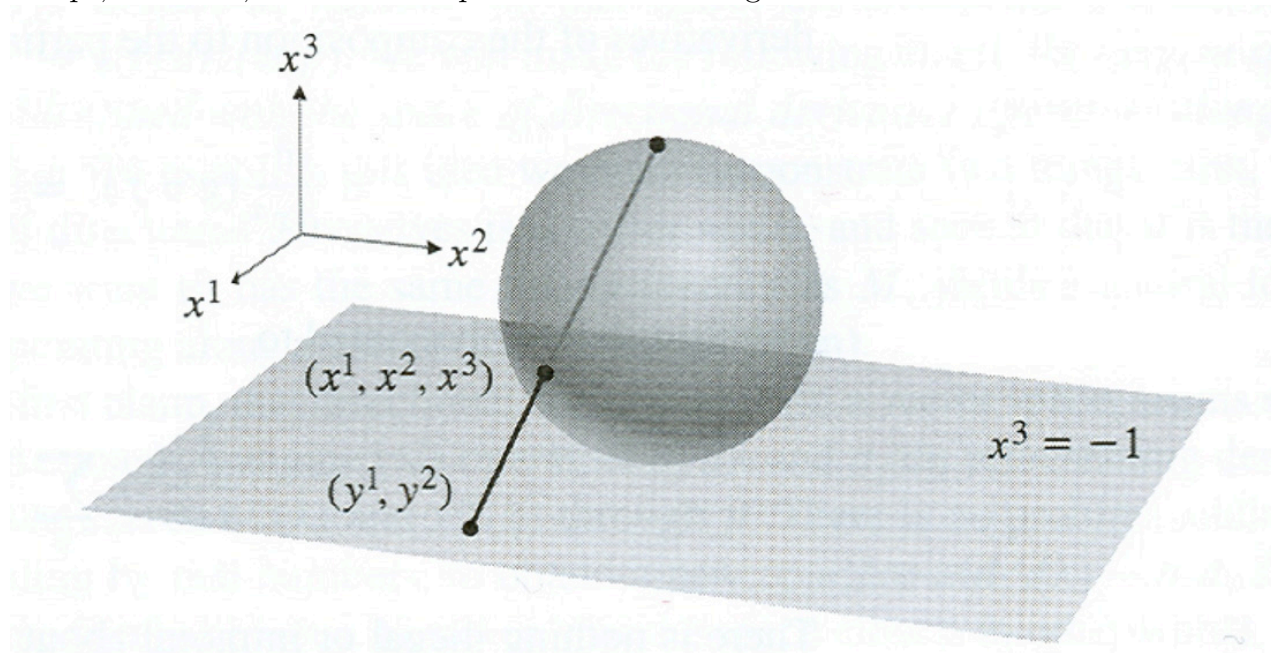
Defn. 3. Any space of points, \mathcal{M} , which admits a maximal smooth atlas, with a constant value of the dimension m , will be called a *smooth manifold*.

An obvious example to consider for a (2-dimensional) manifold is the sphere, S^2 . The usual set of labels we use for the sphere, namely $\{\theta, \varphi\}$, are good labels, but have singularities, concerning invertibility, at both the north and south poles. Therefore, one choice of charts would be to use those coordinates in a neighborhood, say, between -85° latitude and $+85^\circ$ latitude. Then, to define a ‘new’ north pole in Quito, Ecuador, say, and define a second neighborhood, done in the same way, and use coordinates $\{\theta_Q, \varphi_Q\}$ —Q for Quito. This choice has quite an awkward set of transition functions. We will now describe in some detail a simpler approach to a “good” set of charts, usually called *stereographic projection for the sphere*. If you believe that you already understand the concept of a manifold, you can skip that and go on to understand the various things that we use on a manifold to understand our physics. However, in this example we will see that the relationships between sets of coordinates on a manifold are usually complicated functions that are not at all like the linear relationships between different

choices of basis vectors in a vector space, or the linear relationships between different sets of components of vectors relative to different choices of basis vectors.

Stereographic Projection of the Sphere

Stereographic projection creates a map of the sphere onto a plane on which it sits. We place the sphere so that its South pole sits squarely on an infinite plane, which has $\{x, y\}$ coordinates with its (2-dimensional) origin at the point of contact. Then, we create a mapping between the sphere and the plane by drawing a straight line beginning at the North pole, passing through some point on the sphere, and ending where it intersects the plane. This constitutes a clearly-defined one-to-one mapping of all points on the sphere onto the plane, except, of course, for the North pole itself. See the figure below.



The figure is drawn in a 3-dimensional space, and—I am sorry—is labeled with various sets of coordinates with superscripts. For clarity of exposition, I want to use (x, y, z) as coordinates in the general 3-dimensional space we are depicting, and to choose their **origin of coordinates at the center of the sphere**. However, I want to use other symbols to label those coordinates when they apply to points on the surface of the sphere, believing that this will make things somewhat easier, or clearer; nonetheless, I have not taken the extra time to create a new figure.

Therefore, let us name the 3-dimensional coordinates that lie on the surface of the sphere as (ξ, η, ζ) , with the condition that they must satisfy

$$\xi^2 + \eta^2 + \zeta^2 = a^2 , \quad (1.1)$$

where the sphere, currently, simply has a radius that I have labeled as a .

An arbitrary line beginning at the North pole of the sphere is also shown in the figure. It passes from the North pole into the interior of the sphere, and then exits that interior at some point, which we may imagine is labeled P , and then continues until it strikes the plane, at some point which depends on P , and which we might label as Q . The purpose of this exposition is to use the (2-dimensional, **Cartesian-like**) coordinates of the point Q as well-defined coordinates for the point P itself, noticing of course that this approach will not give us any coordinates for the North pole. That is a problem that we must return to, later. We label the point P in terms of its (3-dimensional) coordinates, (ξ, η, ζ) , which lie on the surface of the sphere. We may also associate with this point the usual spherical coordinates, $\{\theta, \varphi\}$. Those coordinates of course are ill-defined in the neighborhood of both the North and the South pole, since (a) the value of φ is not defined when $\sin \theta = 0$, i.e., at the poles, and (b) the mapping is not invertible in a neighborhood of either pole.

In order to write the relationship between these two sets of coordinates the algebra will be slightly simplified if we choose the radius, a , to have the particular value $1/2$, so that the diameter of the sphere, and in particular the diameter between the north and south poles, has the simple value of 1. We will therefore choose the diameter to have the value 1 from now on, which allows us to write the following transformation equations between the two sets of coordinates for the point P , the intersection of the line from the north pole, where it exits the interior of the sphere:

$$\begin{aligned} \xi^2 + \eta^2 + \zeta^2 &= \left(\frac{1}{2}\right)^2 , \\ \left. \begin{aligned} \theta &= \arccos(\zeta/\frac{1}{2}) , \\ \varphi &= \arctan(\eta/\xi) , \end{aligned} \right\} &\iff &\left\{ \begin{aligned} \xi &= \frac{1}{2} \sin \theta \cos \varphi , \\ \eta &= \frac{1}{2} \sin \theta \sin \varphi , \\ \zeta &= \frac{1}{2} \cos \theta . \end{aligned} \right. \end{aligned} \quad (1.2)$$

The line strikes the plane at $z = -\frac{1}{2}$ and some coordinates (x, y) , which we now want to determine, as functions of the coordinates at P . Notice that this is just an “unwrapping” and “stretching” of the sphere, minus the North pole, onto a plane. (This mapping is something like the usual Mercator projection, used commonly on flat maps of the earth.) It is also the mapping commonly used in complex variables, where it is useful to go in the other direction, i.e., to “roll up” the complex plane onto a sphere, so that the ‘point at infinity’ is more easily “seen.”

To determine the values of (x, y) for a given point P , on the sphere, we propose to create it as a parametrized curve, with parameter t , which varies from 0 to 1, as the line moves from the North pole to the (southern) plane. Of course a straight line is a single linear relation that gives the values of (x, y, z) as first-order polynomials in t , i.e., of the form $a + bt$. Since the coordinate z must vary from $z = +\frac{1}{2}$, when $t = 0$, at the North pole, to $z = -\frac{1}{2}$, when $t = 1$, when it arrives on our plane, it must be of the form $z = z(t) = \frac{1}{2}(1 - 2t)$. Since the equations for $x(t)$ and $y(t)$ should have value 0 at the North pole, where $t = 0$, it follows that $x(t)$ and $y(t)$ must simply be proportional to t , not requiring an additional constant term. We may determine the proportionality constants by noting that there is some particular value of t at which the line intersects the sphere again. Referring to that value as t_P , it is given by:

$$\zeta = \frac{1}{2}(1 - 2t_P) \implies t_P = \frac{1}{2} - \zeta = \frac{1}{2}(1 - \cos \theta) = \sin^2(\theta/2) \leq 1. \quad (1.3)$$

As the equations for $x(t)$ and $y(t)$ are simply proportional to t , and must have the values ξ and η at the point P , it follows that they are simply $x(t) = \xi(t/t_P)$ and $y(t) = \eta(t/t_P)$. Inserting the value of t_P just calculated, we have the desired parametric equations for this line:

$$\left. \begin{aligned} z &= \frac{1}{2} - t = \frac{1}{2} \cos \theta, \\ x &= \frac{2\xi}{1 - 2\zeta} t = \frac{\sin \theta \cos \varphi}{1 - \cos \theta} t = t \cot(\theta/2) \cos \varphi, \\ y &= \frac{2\eta}{1 - 2\zeta} t = \frac{\sin \theta \sin \varphi}{1 - \cos \theta} t = t \cot(\theta/2) \sin \varphi, \end{aligned} \right\} \begin{aligned} &t \in [0, 1], \\ &t_P = \frac{1}{2} - \zeta = \sin^2(\theta/2) \leq 1. \end{aligned} \quad (1.4)$$

We may note that the value of t_P is independent of φ , horizontal circles on the sphere, i.e., lines of constant θ , are mapped into circles on the plane; in particular the sphere’s equator is

mapped onto the circle in the plane of radius 1. The South pole maps into the plane's origin, i.e., $(0, 0)$. There is no point in the plane which receives the mapping of the North pole!

The desired coordinate chart for the sphere, i.e., some appropriate open subset of that sphere, that excludes at least the North pole, is mapped onto the x, y -plane, i.e., onto \mathbb{R}^2 , is effected by these formulae when $t = 1$:

$$\begin{aligned} x_1(P) &= \frac{2\xi}{1-2\zeta} = \frac{\sin\theta \cos\varphi}{1-\cos\theta} = \cot(\theta/2) \cos\varphi, \\ y_1(P) &= \frac{2\eta}{1-2\zeta} = \frac{\sin\theta \sin\varphi}{1-\cos\theta} = \cot(\theta/2) \sin\varphi, \end{aligned} \tag{1.5}$$

where the subscript 1 on the coordinates is to remind us that they are only valid in some region of the sphere, which we label U_1 , which excludes some non-trivial closed set that contains the North pole, and everywhere within U_1 , these coordinates are well-defined, invertible, continuous, etc.. For an arbitrary point on the sphere, P , we may of course still think of it in terms of the usual spherical angles, θ and φ ; however, as we have already seen they do not constitute a good, everywhere continuous and differentiable, mapping of the sphere, and we will consider examples soon where that property will be important. We also note that these coordinates have a nice simple form when expressed in terms of a single complex variable, $Z_1(P) \equiv x_1(P) + iy_1(P)$:

$$Z_1 \equiv x_1 + iy_1 = \cot(\theta/2) e^{i\varphi} \quad ,$$

along with the inverse mapping (1.6)

$$\cot(\theta/2) = |Z_1| = \sqrt{x_1^2 + y_1^2} \quad , \quad \tan\varphi = \frac{y_1}{x_1} = \tan(\text{the complex phase of } Z_1) \quad ,$$

However, since it does not include the entire sphere, in order to have this “very good” sort of coordinates everywhere on the sphere, we **must** choose a second neighborhood U_2 , which of course must include the North pole. Therefore, we do the same thing with another plane, which for simplicity we choose to be tangent to the sphere **at the North pole**. As well we now use lines of projection that begin at the South pole. Following through the same algebra

as above, we find that these coordinates, $(x_2(P), y_2(P))$ are given by

$$Z_2 \equiv x_2 + iy_2 = \tan(\theta/2) e^{i\varphi} \quad ,$$

along with the inverse mapping (1.7)

$$\tan(\theta/2) = |Z_2| = \sqrt{x_2^2 + y_2^2} \quad , \quad \tan \varphi = \frac{y_2}{x_2} \quad .$$

In the overlap of regions U_1 and U_2 , the transition functions, i.e., the relationships between the two sets of coordinates, must be differentiable, at least, and are easily found to be given by the following:

$$x_2 = \frac{x_1}{x_1^2 + y_1^2} \quad , \quad y_2 = \frac{y_1}{x_1^2 + y_1^2} \quad (1.8)$$

or, more elegantly, simply as

$$Z_2 = \frac{1}{\overline{Z_1}} = \frac{Z_1}{|Z_1|^2} \quad , \quad (1.9)$$

where the “overbar” denotes the complex conjugate. These transition functions clearly have all the required continuity properties, and are then valid for the overlap region, which may be chosen to be as large as desired, so long as it excludes a non-trivial closed set containing the North Pole and also a non-trivial closed set containing the South Pole.

II. Tangent Vectors to Curves

A curve on a manifold is a continuous selection of points on the manifold, which can most easily be thought of by letting those points vary as some (real-valued) parameter, for instance, λ , varies. (Perhaps the curve is the trajectory of an ant walking along on the surface of the manifold, and the parameter is the distance said ant has traveled.) We can then think about the vector which is tangent to this curve—at some fixed point, P , on the curve, corresponding to a particular value of $\lambda = \lambda_P$ —as the rate at which the points on the curve are varying as λ varies. This is to say that we should take the derivative of the function which defines the curve, with respect to λ , and evaluate it at λ_P . It should be clear that this tangent is a vector which lies in a flat “plane” that is tangent to the manifold at that point. The entire plane may perhaps be thought of as the plane in which the tangent vectors lie for all possible curves on the manifold at that point. One can visualize this, for instance, in our description of the

sphere as a manifold, just above. The remainder of this section is devoted to several attempts to visualize the difference between a plane tangent to a manifold at each individual point on the manifold—where tangent vectors live—and the points on the manifold itself. The points on the manifold are to be described by coordinates—local mappings to \mathbb{R}^m —while the vectors will be described by components relative to some basis, in a vector space available at each point of the manifold.

This is quite reasonably thought about by considering the electric field in some region in space, created by several different sources of charge. Each source creates its own electric field—at each point in the region—and we must add those vectors together to obtain the total field.

So let us visualize a manifold which lies in a flat space of larger dimension—like our sphere—so that we may specify the points on the sphere in a different way, namely via a vector $\vec{r}(\lambda)$ coming out from the origin. Then the tangent vector is the vector $d\vec{r}(\lambda)/d\lambda$, evaluated at λ_P , which can be seen as lying in a flat plane that is tangent to the manifold at this particular point. As other curves could be drawn through that point, one can see that an entire *tangent plane* may be described there, which intersects the sphere, i.e., the manifold, only at the one point P . This is the proper “place” to think about tangent vectors. They are not to be “drawn” on the manifold itself, which is of course curved and therefore has no place where one could draw straight lines, but rather on the flat tangent plane that touches the manifold at that point. Moreover, at some different point on the manifold, say Q , there would be a distinct flat tangent plane. Each of these distinct tangent planes is a place where we may draw vectors—in the sense of distinct line segments—and add them together, and multiply them by scalars. Therefore, each such tangent plane, at each point on the manifold, is in and of itself a vector space. The entire set of all these tangent spaces, for all points on the manifold, is called *the tangent bundle*.

As already discussed, a vector space should be equipped with a choice of basis vectors. Therefore, suppose that our manifold has some coordinate system that can be used to label the points—in the case of our sphere this is either the set $\{x_1, y_1\}$ or the alternative

$\{x_2, y_2\}$. We may then ask how the vector \vec{r} depends on these coordinates, and consider the set $\{\partial\vec{r}/\partial x_i, \partial\vec{r}/\partial y_i\}$ as a basis set for the vector space, at whichever point on the manifold that we are evaluating these derivatives. Let us change the notation slightly, and suppose now that we have a set of coordinates on our manifold, as per the definitions of a manifold, which we call φ^k , and define the desired basis vectors in the tangent space by $\tilde{e}_k \equiv \partial\vec{r}/\partial\varphi^k$, we may then talk in general about our tangent vector on a given curve as follows, using the chain rule of calculus:

$$\frac{d}{d\lambda}\vec{r}(\lambda) = \frac{\partial\vec{r}}{\partial\varphi^k} \frac{d\varphi^k}{d\lambda} = \frac{d\varphi^k}{d\lambda} \tilde{e}_k, \quad (2.1)$$

showing explicitly that this choice is indeed a basis for any vector in our tangent plane.

The only problem with the approach used above is that one would like the descriptions of a manifold to be independent of its possible existence in some larger-dimension, “flat” space, where the vector $\vec{r}(\lambda)$ could be defined. We can simply repeating the argument above, without that $\vec{r}(\lambda)$, but maintaining the dependence on λ , the curve parameter. However, first, let me also give yet some more description of how better to think about these problems, involving directional derivatives along a curve, instead of just the tangent vector.

Since a very important use of vectors is to keep track of “directions,” we will use the idea of **directional derivatives** along curves, evaluated at any given point, as our generalization for vectors on curved spacetimes. As a concrete example, consider the variations of the temperature in a room, with $T = T(x, y, z) \equiv T(\vec{r})$. We want to determine the change in the temperature as we follow along some curve whose tangent vector is given by $\vec{u} = \vec{u}(\vec{r})$. The method, from the usual vector-analysis class, is to describe this via

$$\vec{u} \cdot \nabla T(\vec{r}) = u^i \partial_i T(\vec{r}) = u^x \frac{\partial T}{\partial x} + u^y \frac{\partial T}{\partial y} + u^z \frac{\partial T}{\partial z} = \{u^i \partial_i\} T(\vec{r}) \quad . \quad (1.1)$$

Evaluated at some particular, arbitrary point, P , with coordinates \vec{r} , we say that

**this gives us the rate at which the temperature is changing as
we move away from P in the direction $\vec{u}|_P$.**

The direction \vec{u} is the tangent vector to some curve, through some portion of the room, which we can characterize as a neighborhood of some particular point P ; the notation above, namely $\vec{u}|_P$, means that we are evaluating that vector at the particular point, P .

We would like to be able to use this approach to answer questions not only about the temperature but also other quantities which vary throughout the room. It seems clear that it is the operator $\tilde{u} \equiv u^i \partial_i$, i.e., the directional derivative in the direction \vec{u} that allows this. Its components, u^i , with respect to the basis set $\{\partial_x, \partial_y, \partial_z\}$, are, **by definition**, the same as the components of \vec{u} with respect to the more standard basis set $\{\hat{x}, \hat{y}, \hat{z}\}$. A very important difference, however, between these two basis sets for vectors is that the set $\{\partial_x, \partial_y, \partial_z\}$ is composed of objects defined at a single point, that point at which they act, and thereby a more agreeable choice of basis vectors for objects on a curved space, since we need them defined at individual points, and certainly cannot draw straight line segments at a single point on the manifold itself. At the point P , those components have numerical values and the vector \tilde{u} lies in the vector space—of linear differential operators acting on functions—over that point. We may refer to this vector space as \mathcal{J}_P , which then replaces our earlier notion of directed line segments in flat spaces. At nearby points through which the curve passes, the same is true so that we can say that this actually determines a (local) “field” of vectors, each lying in its own vector space. This makes sense because we know the values in some neighborhood of points, by reason of knowing the curve to which it is tangent.

This notion of first-order differential operators, acting on locally-defined (real-valued) functions—is our generalized notion of vectors. Therefore we associate with every point P coordinates in some differentiable way, and we consider (differentiable) functions for them to act on. We will now write down more carefully some of what is meant by these terms. As well, we introduce the idea of a tangent vector field—or a “field” of tangent vectors—defined in some neighborhood of a point under consideration.

III. Tangent Vector Fields

A (continuous) curve of points on a manifold is an important notion for us. Mathematically we describe the curve as a mapping from a subset of the real numbers, say W , into our manifold:

$$\Gamma : W \subseteq \mathbb{R} \rightarrow \mathcal{M} \implies \forall \lambda \in W \subseteq \mathbb{R}, \Gamma(\lambda) \in \mathcal{M} \quad . \quad (3.1)$$

Continuity means that the coordinates of the points are continuous functions of the parameter λ , i.e., for every allowed coordinate system, where we now denote the one we have chosen by the symbols $\{x^i\}_1^m$, with m as the dimension of the manifold. Then all the coordinates, i.e., the functions $x^i[\Gamma(\lambda)]$ must be continuous functions of λ . Note that normally we will not be so “formal” with the notation and will simply write $x^i(\lambda)$ for the functional dependence of the coordinates of the points on the curve.

Following the motivational paragraphs in the previous, introductory section, **II**, we will label the tangent vector to this curve, where it passes through the point P , as $\tilde{u}|_P \equiv \frac{d}{d\lambda}|_P$, where λ is the choice that has been made for a parameter along that curve. In terms of coordinates on the manifold, this simply gives us the directional derivative operator:

$$\tilde{u}|_P \equiv \frac{d}{d\lambda}\Big|_P = \left(\frac{dx^i[\Gamma(\lambda)]}{d\lambda} \right)\Big|_P \frac{\partial}{\partial x^i} \equiv u^i(P) \partial_{x^i} \equiv u^i(P) \partial_i, \quad \text{for } P \in \mathcal{M}. \quad (3.2)$$

The $x^i[\Gamma(\lambda)]$ are a choice for coordinates of the points along the curve Γ , and we therefore see that the tangent vector has components that are the derivative of those coordinates with respect to that parameter. (The “over-tilde” will be our standard notation to indicate that the object in question is indeed a tangent vector.)

Since these tangent vectors are differential operators they act on functions defined over the manifold. It is therefore appropriate to backup our discussion for just a moment and give a definition of smooth functions over a manifold.

Defn. 5. A continuous function over \mathcal{M} is one where the coordinate form of that function is continuous for any allowed choice of coordinates. By the phrase “coordinate form” of the function, we mean the map from \mathbb{R}^m to \mathbb{R} , which we may think of as some function $f(x^i)$.

For our smooth manifolds we prefer to restrict consideration to smooth functions, i.e., to those which are C^∞ functions.

we denote the set of all C^∞ functions over the manifold by the symbol \mathcal{F} .

At a single point, $P \in \mathcal{M}$, a vector $\tilde{u}|_P$ operates on a function, $f \in \mathcal{F}$, defined in some neighborhood of P , in the expected way. The result is a number which is the value of the directional derivative, in the direction \tilde{u} , at the point P . Therefore the vector at the point P maps functions into numbers. However, if we do this in some neighborhood of P where this vector is defined, then a different number will be generated at each point, giving us therefore a new function in that neighborhood. Therefore, referring to the *tangent vector field*, defined at all neighboring points, this vector field maps functions into other functions, namely their derivatives in the direction \tilde{u} :

$$\begin{aligned} &\text{at any single point, } \tilde{u}|_P : \mathcal{F}|_{P \in U} \rightarrow \mathbb{R}, \\ &\text{in a neighborhood of a point, } \tilde{u} : \mathcal{F} \rightarrow \mathcal{F}, \\ &\tilde{u}[f] = u^j(\lambda) \frac{\partial}{\partial x^j} f(x^i) = \frac{dx^j(\lambda)}{d\lambda} \frac{\partial}{\partial x^j} f(x^i). \end{aligned} \tag{3.3}$$

If one wants to be very pedantic about the details of the mapping, one could write the same thing in very much more detail:

$$\begin{aligned} \tilde{u}[f] &= \frac{d}{d\lambda} f[\Gamma(\lambda)] = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{f[\Gamma(\lambda + \epsilon)] - f[\Gamma(\lambda)]\} = \frac{d}{d\lambda} (f \circ \Gamma) \\ &= \frac{d}{d\lambda} (f \circ \varphi^{-1} \circ \varphi \circ \Gamma) = \left(\frac{\partial}{\partial x^i} (f \circ \varphi^{-1}) \right) \left\{ \frac{d}{d\lambda} (\varphi \circ \Gamma)^i \right\} = \frac{dx^i[\Gamma(\lambda)]}{d\lambda} \frac{\partial}{\partial x^i} (f \circ \varphi^{-1})(x^j), \end{aligned} \tag{3.4}$$

where the set $\{x^j\}$ constitute the coordinate choice for a chart labeled by φ , in some neighborhood of the point P , the $x^j[\Gamma(\lambda)]$ are the coordinate presentation of the curve, so that they constitute a set of maps from $W \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ and the $(f \circ \varphi^{-1})(x^j)$ constitute the coordinate presentation of the function f , i.e., the $f \circ \varphi^{-1}$ constitute a map from $\mathbb{R}^m \rightarrow \mathbb{R}$.

Abstractly, we summarize the situation by saying that the tangent vectors **at each point, P , of the manifold** satisfy the following properties:

1. $\tilde{u}|_P : \mathcal{F}|_U \rightarrow \mathbb{R}, \quad P \in U \subseteq \mathcal{M},$
2. $\tilde{u}|_P(f + \alpha g) = \tilde{u}|_P(f) + \alpha \tilde{u}|_P(g),$ for α a constant, —**linearity**
3. $\tilde{u}|_P(fg) = f(P) \tilde{u}|_P(g) + \tilde{u}|_P(f) g(P)$ —**derivation property.**

Objects with these general properties are often referred to as derivations. It should be clear that these properties are maintained under addition and multiplication by scalars; therefore we may say that our tangent vectors form a vector space of derivations (of functions). We will refer to this vector space of tangent vectors at the point P by the symbol \mathcal{T}_P , and the entirety of all these vector spaces as the *tangent bundle*, $\mathcal{T} \equiv \mathcal{T}^1$. Our construction shows us that for each of the vector spaces \mathcal{T}_P , one may treat the partial derivative operators at P , namely $\{\partial_{i_P}\}$, as a particular choice of basis for that particular vector space over the point P .

Suppose, now, that we consider what happens when one makes a different choice of coordinates for the neighborhood. For instance we consider a transformation from coordinates x^i to some other choice y^s , where we may write either $y^s = y^s(x^i)$ or vice versa, since the transformation is required to be invertible. We want to determine how this changes the presentation of a tangent vector \tilde{u} . As the curve on \mathcal{M} , to which \tilde{u} is tangent does not change, then it is reasonable to desire that the vector itself also does not change. We refer back to Eqs. (3.2), for a description of the tangent vector \tilde{u} , tangent to a curve with parameter λ , and expressed in terms of some coordinates $\{x^i\}$, some of the content of which we re-iterate here:

$$\tilde{u} = u^i \partial_{x^i} = \frac{dx^i}{d\lambda} \frac{\partial}{\partial x^i}. \quad (3.2')$$

Then “advanced calculus” tells us how to change the partial derivatives, and the derivatives along the curve, i.e., with respect to λ , from one set of coordinates to the other, and that the transformations that perform these changes are inverse to one another:

$$\begin{aligned} \partial_i \equiv \partial_{x^i} &= \left(\frac{\partial y^s}{\partial x^i} \right) \partial_{y^s} \equiv Y_i^s \partial_{y^s}, & u^i \equiv \frac{dx^i}{d\lambda} &= \left(\frac{\partial x^i}{\partial y^s} \right) \frac{dy^s}{d\lambda} \equiv X_s^i u'^s, \\ Y_i^s X_t^i &\equiv \left(\frac{\partial y^s}{\partial x^i} \right) \left(\frac{\partial x^i}{\partial y^t} \right) = \delta_t^s, & \left(\frac{\partial x^i}{\partial y^s} \right) \left(\frac{\partial y^s}{\partial x^j} \right) &= X_s^i Y_j^s = \delta_j^i, \end{aligned} \quad (3.5)$$

where we have denoted the components of \tilde{u} with respect to these new coordinates as $u'^s \equiv dy^s/d\lambda$. As the two matrices of partial derivatives are inverse to each other, the form of \tilde{u} , i.e., the tangent vector itself, is unchanged by such a change in one's choice of coordinates—as surely seems appropriate. However, it is also clear that the **components** of the vector are changed when one goes from one choice of coordinates to another. It is worth repeating this very important transformation equation for the components of a tangent vector here:

$$u^i = X_s^i u'^s . \quad (3.6)$$

Since Eqs. (3.5) show us that the components of a tangent vector transform with a matrix that is inverse, or opposite, to the matrix that transforms the basis vectors themselves, the transformation of the components of the tangent vector is said to be *contravariant*. Therefore tangent vectors are sometimes referred to as “contravariant vectors,” even though, truthfully, it is their components that transform in this way, whenever we choose to make a change of basis vectors.

It is also worth comparing the **linear** transformation of the components of a vector, via a square matrix, with the transformation of the coordinates themselves which involve complicated functional dependencies which are seldom linear. It is for this reason that it is quite important to keep well in mind the distinction between coordinates and components of vectors—or, more generally, components of tensors—since in freshman physics these two are commonly confused.

IV. Differential Forms: an Alternative Bundle of Vector Spaces

As we have already discussed, every vector space has associated with it a *dual space* of quantities that map the vectors into whatever field of numbers over which the vectors are defined—in our case, either the real or complex numbers. In addition, these mappings are required to preserve the structure of the original vector space; i.e., the mappings should be linear and simply let scalars pass through. More precisely, if ϖ is such a mapping, \tilde{u} , \tilde{v} are arbitrary (tangent) vectors and a is a scalar, then we require that

$$\varpi(\tilde{u} + a\tilde{v}) = \varpi(\tilde{u}) + a\varpi(\tilde{v}) , \quad (4.1)$$

where we habitually use the “under-tilde” to indicate that a given symbol denotes a differential form. Since these are mappings on vectors, it is straightforward to show that they are also vectors. Therefore vector spaces always come in pairs. However, in many of the simpler cases, there is a way to identify the two spaces, so that it is often difficult to see that there are actually two. An example is given by the bra’s and ket’s in the Hilbert space for quantum mechanics. A second, well-known example is given by the *polar vectors* and *axial vectors* in classical mechanics, say. The ordinary linear momentum is a polar vector, while the angular momentum is a cross product of two polar vectors, and therefore is an axial vector. In classical mechanics vectors are often characterized by their behavior under rotations and translations. For the usual rotations, the two sorts of vectors behave exactly the same; however, for a parity transformation, polar vectors change sign while axial vectors do not. Therefore it is possible to distinguish these two sorts of vectors, but it does take a little effort. It is the generalization of this particular distinction that we are now beginning to discuss.

So we now describe in detail the dual space for tangent vectors over a manifold, which will of course be a particular case of the more general notion of differential forms. We begin by considering a function, $f \in \mathcal{F}$, defined in a neighborhood of some point $P \in \mathcal{M}$. We will define the **differential form**, df , associated with $f \in \mathcal{F}$, as a member of the vector space dual to \mathcal{T}_P by giving its action on tangent vectors. We do this first at a single point, also introducing the notation $f_{,i}$ as an abbreviation for the derivative of f with respect to the i -th coordinate, x^i :

$$df|_P(\tilde{u}|_P) \equiv \tilde{u}|_P(f) = (u^i \partial_{x^i} f)|_P \equiv (u^i f_{,i})|_P \quad , \quad (4.2)$$

and we do not bother to put an “under-tilde” on a differential form like df since its form as a differential already makes that clear. This definition does indeed satisfy the requirements of linearity, for the so-defined differential form, df , since the tangent vector \tilde{u} satisfies those requirements. The calculation takes a vector and gives a number. However by doing this in each of the vector spaces over the points of some neighborhood containing P we obtain

a function from our tangent vector. Therefore, over a neighborhood we may think of df as mapping tangent vector fields into functions over the manifold,

$$\text{i.e., } df : \mathcal{T} \rightarrow \mathcal{F}.$$

We can now proceed further along the path of determining the entirety of the dual space by next using our notions from “advanced calculus,” concerning *differentials*. Certainly we know that $df = f_x dx + f_y dy + \dots$. This shows us that the set $\{dx, dy, \dots\}$ is a reasonable choice for a basis for this vector space. Using linearity the previous equation then tells us that

$$u^i f_{,i} \equiv u^i (\partial_i)(f) = \tilde{u}(f) \equiv df(\tilde{u}) = f_{,i} dx^i(\tilde{u}) = f_{,i} dx^i(u^j \partial_j) = (f_{,i} u^j) dx^i(\partial_j). \quad (4.3)$$

First this shows us the action of the basis-quantities, $\{dx^i\}$, on vectors, which is an altogether appropriate action since the x^i are just a rather special set of functions defined over our neighborhood. Taking $\{\partial_j\}$ as a **basis for \mathcal{T}** , we see that $\{dx^i\}$ may easily be taken as a basis for this vector space of differential forms, **and** that it is a reciprocal, or dual, basis for our 1-forms relative to the vector spaces of tangent vectors, i.e., it satisfies

$$dx^i(\partial_j) = \frac{\partial x^i}{\partial x^j} = \delta_j^i. \quad (4.4)$$

This concept allows us to easily generalize the notion of the differential of a function to define the entire vector space, $\Lambda \equiv \Lambda^1$, of all differential forms of this type. We also refer to them as 1-forms, since that require exactly one tangent vector to create a number. (Soon we will generalize that notion as well.) We say that a 1-form is simply any (finite) sum of products of functions and the differentials of functions, locally-defined in each vector space over the points of some neighborhood of any point. For now we will treat the differentials of the coordinates on the manifold as the basis 1-forms for Λ^1 . As elements of this dual space are supposed to have an action on vectors, mapping them to functions—which become scalar numbers at each point on the manifold, but, of course, in a way which varies “smoothly” from one point to the next—we may use our statements about the reciprocal basis to define the general action of the dual space on the tangent vector space:

- **Dual Action:** Let $\varrho \in \Lambda^1$ and $\tilde{u} \in \mathcal{T}$ be written in terms of their (reciprocal) basis sets as

$$\varrho = \alpha_i dx^i \quad , \quad \tilde{u} = u^i \partial_i \quad . \quad (4.5a)$$

The mapping $\varrho : \mathcal{T} \rightarrow \mathcal{F}$ is defined by

$$\varrho(\tilde{u}) = \alpha_i u^i \quad , \quad (4.5b)$$

$$\text{so that} \quad df(\partial_j) = f_{,j} \quad \text{and} \quad \varrho(\partial_j) = \alpha_j \quad \text{along with} \quad dx^i(\tilde{u}) = u^i \quad . \quad (4.5c)$$

Notice the symmetry between the components of ϱ and of \tilde{u} in the expression for the action of the 1-forms on the vectors. This tells us that we can easily extend this to a symmetric sort of operation where the vectors act on the 1-forms. More precisely then we want to define a second set of behaviors for tangent vectors, which makes the relation between them and differential forms much more symmetric. The mapping $\tilde{u} : \Lambda^1 \rightarrow \mathcal{F}$ is defined by

$$\tilde{u}(\varrho) = \alpha_i u^i = \alpha_i u^i = \varrho(\tilde{u}) \quad , \quad (4.5d)$$

$$\text{so that} \quad \partial_j(\varrho) = \alpha_j \quad \text{along with} \quad \tilde{u}(dx^i) = u^i \quad . \quad (4.5e)$$

- **Behavior of Differential Forms under Coordinate Transformations**

We recall that we have insisted that tangent vectors should not change simply because we decided to make a change of coordinates on the underlying manifold, **but** that the components of those tangent vectors do change, since the change of coordinates amounts to a change of basis for the vector space, all of this described mathematically in Eqs. (3.5-6). Therefore we should have similar relationships for differential forms. We expect that the actual 1-form does not change, but that the basis set changes, since they are just differentials of the coordinates, so that the components of the 1-form relative to that basis will also change. Remembering that the two (Jacobian-type) matrices $X_s^i \equiv \partial x^i / \partial y^s$ and $Y_j^s \equiv \partial y^s / \partial x^j$ are inverse to one another, we may rewrite a portion of Eqs. (3.5) in a manner appropriate to differential forms, beginning from Eq. (4.5a):

$$\begin{aligned} \varrho &= \alpha_i dx^i = \alpha_i \frac{\partial x^i}{\partial y^s} dy^s \equiv \alpha_i X_s^i dy^s \equiv \alpha'_s dy^s \quad , \\ \implies \alpha'_s &= X_s^i \alpha_i \quad \iff \quad \alpha_i = Y_i^s \alpha'_s \quad , \end{aligned} \quad (4.6)$$

where the very last equality should be compared with Eq. (3.6), which shows that the components of differential forms transform inversely relative to the components of tangent vectors. In an analogy with the language used for tangent vectors, differential forms transform *covariantly*. This of course seems quite reasonable if we think of the action of a differential form on a tangent vector, as described in Eq. (4.5b), written below with the appropriate transformations between the two sets of coordinates, showing that the final result does not change, as it should not:

$$\mathcal{Q}(\tilde{u}) = \alpha_i u^i = \alpha'_s Y_i^s X_t^i u'^t = \alpha'_s \delta_t^s u'^t = \alpha'_s u'^s = \mathcal{Q}(\tilde{u}) . \quad (4.7)$$

- **Geometrical Description of Differential Forms**

We created tangent vectors via an altogether reasonable attempt to generalize the very intuitive notion of a vector as a directed line segment. What, then, is a reasonable geometrical picture of the elements of our vector space of differential forms? Although I do not want to belabor the creation of such a geometrical picture for 1-forms, there is a standard approach to such a description. (Much more detail, with many pictures, is given, for instance, in MTW.)

Recall the usual geometrical notion of the gradient of a function, such as the temperature in a room. The ordinary 3-dimensional vector, $\nabla T(\vec{r})$ defines a vector field throughout the room, which we usually describe by saying that it is the “direction of **greatest change**” of T , at the particular point with coordinates \vec{r} . However, we also talk about a somewhat dual construction, namely the surfaces of constant value of the temperature, T , to which the direction ∇T is perpendicular. These are surfaces, much like equipotentials for electrostatics, where the temperature doesn’t change as you move from one point to another on them. It is altogether plausible that the surface on which the function does NOT change is the one that is **perpendicular** to the direction of its greatest change!

We therefore take the analytic idea that df is the generalization of *the gradient* of f , i.e., ∇f . As in the case of tangent vectors and directional derivatives, this comparison is reasonable since the two quantities have the same components; only the basis vectors look different. Again,

the new basis vectors—for df —are defined locally, at each single point P on the manifold. Then we may say that geometrically the 1-form df corresponds to a local view of the surfaces of constant values for f . Importantly, when we are on a manifold, the surfaces of constant values for f are curved surfaces, containing many different points on the manifold. However, the 1-form df is defined, separately, at each and every point on the manifold; therefore, at a point, P , the 1-form df corresponds to the tangent planes to that surface at the point P . Therefore, in an analogous way to the partial derivatives at each point—the basis vectors for tangent vectors, lying in the space(s) of tangent vectors—these objects must live within their own vector space, the vector space of differential forms, with one such space attached at each point of our underlying manifold. Since they are *locally-defined*, over an m -dimensional manifold, surfaces of constant values for some function, say the temperature T , are $m - 1$ -dimensional surfaces—so-called hypersurfaces since they are only one less dimension than the entire space, and the simplest 1-form $dT|_P$ is an algebraic representation of the tangent planes to those (hyper)-surfaces of constant value for T , at the point P where the differential form is being evaluated.

Since the ideal of an infinitesimal, like dT , is surely a local idea, defined, within a neighborhood, but at each single point, these surfaces must live within their own vector space, Then, in particular, the hypersurfaces dx^i correspond to constant values for the coordinate function x^i , so that they are those hypersurfaces in which all the **other** coordinates are allowed to take all possible values, the value of x^i itself being fixed to one particular value. As a brief example, we look at an ordinary 3-dimensional space and the usual coordinates, $\{x, y, z\}$, so that ∂_x corresponds to the tangent vector to the x -axis, while then the 1-form dx corresponds to the set of 2-surfaces—i.e., surfaces of dimension $3 - 1$ —we would describe by saying that they are all parallel to the y, z -plane, for different, fixed values of x .

Continuing in this vein, since the 1-forms live in a vector space, we can consider the usual more general 1-form, i.e., a sum of products of scalars times basis 1-forms, is in general not simply the differential of a single function. Therefore, although they also may be thought of

as tangent planes to the manifold, over a neighborhood, there may be no particular function defined on the manifold such that they are tangent to the subspace of constant values of that function. This is the general thought pattern behind what is called Poincaré’s Lemma, discussed below, generalizing the idea of “potentials” for vector fields.

V. Extension of Differential Forms to the entire Grassmann Algebra

As it turns out, differential forms have many very useful and important properties, quite sufficient to justify the effort required to learn to manipulate and understand them. The vector space can be turned into a *differential algebra* by adding a **product** and a special differential operator into this space; the result is called a *Grassmann algebra*, following Grassmann, who was a high-school teacher in Germany in the 1870’s.

- **The Exterior Product, and p-forms**

We must begin by first defining the tensor product of two vectors. As we have already discussed, differential forms, evaluated at a point P , map tangent vectors at that point into numbers. Moreover, when we consider them in the neighborhood of that point they map tangent vectors into a number at each point of that neighborhood, i.e., a (locally-defined) function in that neighborhood, and that action is linear and continuous. Additionally, we have defined an action of tangent vectors on differential forms with the same effect. More generally, we want to think of a general sort of a tensor as linear, continuous maps of some number, s , of tangent vectors and some number, r , of 1-forms into the real numbers at a point, and therefore into functions defined in a neighborhood of any particular point. One may also say that they are contravariant of type r and covariant of type s , or simply that they are “of rank $\begin{Bmatrix} r \\ s \end{Bmatrix}$.” As an example of a very simple sort, we may look at the tensor product of two 1-forms, as some linear, continuous mapping of **two tangent vectors** into a locally-defined function:

$$\alpha \otimes \beta : \mathcal{T} \times \mathcal{T} \rightarrow \mathcal{F}. \tag{5.1}$$

The symbol \times simply indicates that we must pick two elements, one from each of the two spaces (of tangent vectors) involved; it is usually named the Cartesian product, and has no

real properties at all. On the other hand the symbol \otimes , used in this equation, is the tensor product, and insists that one preserves the linearity of the underlying vector spaces in every possible way. Examples are given by the following requirements on the tensor product, of two spaces of 1-forms, along with quite a few other ways of rewriting the same thing, moving the scalars, $\{a, b, c\}$ into and out of the various spaces, and translating addition from the smaller vector space into the larger one:

$$(a\mathcal{Q}) \otimes (b\mathcal{P} + c\mathcal{R}) = ab(\mathcal{Q} \otimes \mathcal{P}) + ac(\mathcal{Q} \otimes \mathcal{R}) = a(b\mathcal{Q}) \otimes (\mathcal{P}) + (c\mathcal{Q}) \otimes (a\mathcal{R}) . \quad (5.2)$$

Therefore, an easy way to think about the vector space of all tensor products of two 1-forms is to first define a basis for it. Let us suppose, for instance, that $\{dx^i\}_1^m$ is a basis for the first vector space (of differential forms) and that $\{dy^a\}_1^m$ is a basis for the second vector space. Then the following set of tensor products is a basis for the vector space of all such tensor products:

$$\{dx^i \otimes dy^a \mid i, a = 1, \dots, m\} \quad \text{is a basis for } \Lambda^1 \otimes \Lambda^1 , \quad (5.3)$$

i.e., an arbitrary element of $\Lambda^1 \otimes \Lambda^1$ is any finite linear combination of this set of m^2 basis elements, the coefficients being (scalar) functions over the manifold. Such a tensor, which requires 2 tangent vectors and 0 1-forms in order to create a function is referred to as a tensor of rank $\left\{ \begin{smallmatrix} 0 \\ 2 \end{smallmatrix} \right\}$. More generally, a tensor of rank $\left\{ \begin{smallmatrix} r \\ s \end{smallmatrix} \right\}$ is a member of the tensor product of r copies of \mathcal{T}^1 and s copies of Λ^1 and the natural choice of basis is

$$\{\partial_{x_1^{\mu_1}} \otimes \partial_{x_2^{\mu_2}} \otimes \dots \otimes \partial_{x_r^{\mu_r}} \otimes dx^{\lambda_1} \otimes dx^{\lambda_2} \otimes \dots \otimes dx^{\lambda_s} \mid \mu_1, \dots, \mu_r, \lambda_1, \dots, \lambda_s = 1, \dots, n\} . \quad (5.4)$$

The special case of tensors of rank $\left\{ \begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right\}$ have components which can easily be presented via a matrix. In this case, those components may be created by taking the matrix presentations of the two types of tensors, the components of the tangent vector as a column vector, i.e., a matrix with only 1 column and m rows, and the components of the differential form as a row vector, i.e., a matrix with only 1 row and m columns. One then creates the so-called outer

product (or Kronecker product) of the two of them, multiplied in the order so that one acquires an $m \times m$ matrix.

We may then next use the (**skew-symmetric**) part of a general tensor product for differential forms, which we will name the Grassmann product, also referred to, more usually, as the “*wedge*” product, since it is denoted with the symbol \wedge between the two 1-forms being multiplied together. This skew-symmetric product is of considerable interest in m -dimensional vector analysis, when m is greater than or equal to 4, and plays the same role that the usual “cross-product” plays in the more usual 3-dimensional vector algebra.

We define the wedge product of two 1-forms, say α and β , as

$$\alpha \wedge \beta \equiv \alpha \otimes \beta - \beta \otimes \alpha . \quad (5.5)$$

Then a basis for the vector space of 2-forms, Λ^2 , is just

$$\{dx^\alpha \wedge dx^\beta \mid \alpha, \beta = 1, \dots, m ; \alpha < \beta\} . \quad (5.6)$$

where the restriction on the range so that α is less than β is because $dx^2 \wedge dx^4$, say, is just the same as $-dx^4 \wedge dx^2$, since changing the order simply changes the sign, and of course the wedge product of a 1-form with itself is just identically zero!

- **p-forms**

The result of this wedge product is of course no longer a 1-form, i.e., it is no longer a linear combination of the differentials of the coordinates of \mathcal{M} , but, rather, a 2-form. Therefore, we now generalize the notion of 1-forms to allow for some additional vector spaces, whose elements we refer to as *p-forms*, where $p=1, \dots, m$. We have already defined 2-forms as the vector space of all linear combinations of multiples of the wedge product of two 1-forms, requiring this product to be linear and associative over addition. We then generalize to products of more than two at a time, again requiring associativity. For each positive integer, $p \leq m$, we may then define the vector space Λ^p as the set of all possible (finite) linear combinations of

multiples, via functions, of its basis set, which is the set of all linearly-independent, non-zero “wedge-products” of p of the $\{dx^i\}$ at once.

Of course the number of linearly-independent p -forms is reduced by the requirements of skew symmetry. In general the dimension of the space $\Lambda_{\mathcal{M}}^p$ of p -forms, i.e., the number of different ways that one may choose p distinct things from a set of m of them, is just the binomial coefficient $\binom{m}{p}$. When there are m dimensions to \mathcal{M} , then one may have spaces of 1-forms, 2-forms, 3-forms, etc., up to and including m -forms, but no higher. This is because one cannot make objects skew symmetric in the interchange of two adjacent entries in more than m things at once if you only have m things to play with! In addition, notice that there is only one independent dimension to the space of m -forms, e.g., $dx^1 \wedge dx^2 \wedge \dots \wedge dx^m$. Some particular choice of basis for that 1 dimension is usually referred to as *the volume form* for the space. If we also decide that λ^0 is to be defined as the same as the space of functions, \mathcal{F} , which also has dimension 1, we can see a duality in these various vector spaces, at least in dimensions. For example, $\binom{m}{m-1} = \binom{m}{1}$, and also $\binom{m}{m-2} = \binom{m}{2}$, etc. This is called Hodge duality, and we will return to it later, when we have allowed a metric to act in the vector spaces over our manifold, in the second set of notes concerning using vector spaces over spacetime to study gravitational physics.

Example: If the underlying space is 3-dimensional—just the usual space, \mathbb{R}^3 , with coordinates $\{x, y, z\}$, then we take $\{dx, dy, dz\}$ as a basis for the space of 1-forms, then the space of 2-forms is the vector space with basis $\{dx \wedge dy, dy \wedge dz, dz \wedge dx\}$, and the space of 3-forms has basis $\{dx \wedge dy \wedge dz\}$. (We don’t count both $dx \wedge dy$ and $dy \wedge dx$, since they are not linearly independent; one is just the negative of the other.) As it happens, the space of 2-forms is the same dimension as the space of 1-forms. If we wanted to, we could characterize all this by saying that $\Lambda^2(\mathbb{R}^3)$ is *spanned* by the set $\{dx \wedge dy, dy \wedge dz, dz \wedge dx\}$. This (interesting) equality occurs only for 3 dimensions, and is the source of the existence of the “cross-product” only in 3 dimensions.

On the other hand, in 4 dimensions, the space of 1-forms of course has 4 dimensions,

being spanned by $\{dx, dy, dz, dt\}$, while the space of 2-forms already has 6 dimensions, since it is spanned by $\{dx \wedge dy, dy \wedge dz, dz \wedge dt, dt \wedge dx, dy \wedge dt, dz \wedge dx\}$.

- **Action of p -forms, on Tangent Vectors**

As a finite linear combination of wedge products of p 1-forms at a time, a p -form may be “fed” p distinct tangent vectors, which will result in the production of a function, i.e., a number at each point. At a more sophisticated level, one could only feed some smaller number of vectors to the p -form, resulting, then, in a form of some lesser order. That is, if we gave q vectors to a p -form, where $q < p$, then the result would be a $(p - q)$ -form since it would still be capable of taking $p - q$ vectors and then giving us a function. The first thing to notice is that the skew-symmetry of a p -form arranges it so that it operates on its vectors in a skew-symmetric way. More precisely, **if the p vectors fed to the p -form are not all linearly independent, then the resulting number will simply be zero!**

Of especial interest in physics are 2-forms; therefore, we want to establish, as a formula, how one determines the action of a 2-form on a pair of vectors. A reasonable approach for the skew-symmetric action of the 2-form, $\mathcal{Q}_1 \wedge \mathcal{Q}_2$ on a pair of vectors, $\tilde{u}_1, \tilde{u}_2 \in \mathcal{T}$, is given by

$$(\mathcal{Q}_1 \wedge \mathcal{Q}_2)(\tilde{u}_1, \tilde{u}_2) = \begin{vmatrix} \mathcal{Q}_1(\tilde{u}_1) & \mathcal{Q}_1(\tilde{u}_2) \\ \mathcal{Q}_2(\tilde{u}_1) & \mathcal{Q}_2(\tilde{u}_2) \end{vmatrix} = \text{determinant}((\mathcal{Q}_i(\tilde{u}_j))) \quad . \quad (5.7)$$

Using the structure of a determinant arranges it so that the skew-symmetry is made automatic, as desired. (There are in fact other normalizations in the literature: sometimes a factor of $2!$ is used to multiply the definition above.) If, now, the 2-form given is more general, the fact that $\varpi^i \wedge \varpi^j$ satisfies the form given in Eq. (5.2) and that $\varpi^i(\tilde{u}) = u^i$ allow us to write the following, remembering that the components of a 2-form are always skew-symmetric, i.e., $\beta_{ij} = -\beta_{ji}$ in the calculation below:

$$\beta \equiv \frac{1}{2} \beta_{ij} \varpi^i \wedge \varpi^j \implies \beta(\tilde{u}, \tilde{v}) = \beta_{ij} u^i v^j \quad . \quad (5.8)$$

For a general p -form, then we acquire the similar formula:

$$\psi \equiv \frac{1}{p!} \psi_{i_1 \dots i_p} \varpi^{i_1} \wedge \dots \wedge \varpi^{i_p} \implies \psi(\tilde{u}, \dots, \tilde{w}) = \psi_{i_1 \dots i_p} u^{i_1} \dots w^{i_p} \quad . \quad (5.9)$$

Notice that I have adopted the notational convention of putting an “under-tilde” on any p -form, analogous to the method that I use to denote 1-forms. In general, we won’t usually denote the value of p , hoping that it’s reasonably obvious from context.

- **The Exterior Differential and p -forms**

Since the space of 1-forms begins with the differentials of functions, one should be able to differentiate other things as well. It is reasonable to look at the space of functions as a “zero-th” order space of forms, so that it seems that differentiation takes 0-forms into 1-forms. Therefore, we generalize the notion of differentiation so that it acts on any p -form, with the result being a $(p + 1)$ -form. Such a generalization should maintain the usual “product rule” for derivatives while also maintaining and creating the skew-symmetry of p -forms which we have just endeavored to create. More precisely, we would like it to have the following sorts of properties:

1. **the exterior derivative d is a derivation** so that it handles its arguments according to (some version of) the product rule we know from calculus, i.e., for (scalar) functions, $f, g \in \Lambda^0 \equiv \mathcal{F}$, we have $d(fg) = f dg + (df) g$,
2. to preserve skew-symmetry of p -forms, it anticommutes with 1-forms, since our product is anti-symmetric.

This motivation is sufficient to allow us to define the **exterior derivative** as a mapping on any p -form, acting something like a derivative.

$$d: \Lambda^p \mathcal{M} \longrightarrow \Lambda^{p+1} \mathcal{M} \quad \text{such that}$$

$$\begin{cases} df = f_{,i} dx^i \equiv \frac{df}{dx^i} dx^i & , \quad \text{for functions } f \in \Lambda^0 \mathcal{M}, \\ d(\phi_i dx^i) = d\phi_i \wedge dx^i \in \Lambda^2 \mathcal{M} & , \quad \text{for 1-forms } \phi = \phi_i dx^i \in \Lambda^1 \mathcal{M}, \\ d(\underline{\alpha} \wedge \underline{\beta}) = (d\underline{\alpha}) \wedge \underline{\beta} - \underline{\alpha} \wedge (d\underline{\beta}) & , \quad \text{for } \underline{\alpha}, \underline{\beta} \in \Lambda^1 \mathcal{M}. \end{cases} \quad (5.10)$$

These properties are sufficient to allow one to prove the following more general form, regarding commutativity, for its action on arbitrary p - and q -forms, where here it is useful to denote the level of the two different differential forms:

$$\text{for } \underline{\lambda} \in \Lambda^p \mathcal{M}, \quad \underline{\nu} \in \Lambda^q \mathcal{M}, \quad d(\underline{\lambda} \wedge \underline{\nu}) = (d\underline{\lambda}) \wedge \underline{\nu} + (-1)^p \underline{\lambda} \wedge (d\underline{\nu}) \quad , \quad (5.11)$$

the rationale of the above requirement being that, in the second term, the d has passed itself to the other side of exactly p 1-forms, the p basis forms of $\frac{\lambda}{p}$.

Examples:

0. Suppose, now, that $f = f(x, y, z)$ is a function over \mathbb{R}^3 , then $df \wedge dy$ is a 2-form, which, in the usual basis, can be thought of as having 2 non-zero components, i.e.,

$$\text{over } \mathbb{R}^3, \quad df \wedge dy = \frac{df}{dx} dx \wedge dy + \frac{df}{dz} dz \wedge dy \quad \equiv f_{,x} dx \wedge dy - f_{,z} dy \wedge dz \quad , \quad (5.12)$$

where we have also used above the very common notation which allows a partial derivative to be denoted by a subscript which begins with a comma!

Notice the following comparisons to ordinary 3-dimensional vector analysis.

1. We may think of the usual 3-dimensional vector, \vec{v} , as a 1-form, ζ , over a space of 3 variables, by identifying the components of the two; then the 2-form $d\zeta$ has 3 components which are just the usual components of the *curl* of \vec{v} , i.e., $\nabla \times \vec{v}$;
2. Similarly we may identify the 3 components of the 2-form β , over a space of 3 variables, with the components of a vector \vec{w} , then the 3-form $d\beta$ has just one component, which is just the *divergence* of \vec{w} , i.e., $\nabla \cdot \vec{w}$.
3. In any number of dimensions, a 2-form is specified by components which (a) have two indices, and (b) are skew-symmetric with respect to the interchange of those two indices. Therefore, we may always use matrix notation to display (or present) the components of a 2-form; in general that matrix will be skew-symmetric. A common example is the electromagnetic field tensor in 4 dimensions, which we will discuss in more detail later.

- **Poincaré’s Lemma** Because of the anti-symmetry of the wedge product, we get some slightly unexpected properties of the exterior derivative operator, which actually give it much more power than one might have expected. Firstly, notice that the skew symmetry of wedge products and the fact that partial derivatives commute, i.e., are symmetric with respect to order, make it easy to see that, for any function, f , and indeed for any p -form, ψ , we have

$$d^2 \equiv d(df) = 0, \quad d(d\psi) = 0. \quad (5.13)$$

The statement of Poincaré’s Lemma given below is a generalization of the (rather more well-known) statement from ordinary 3-dimensional vector analysis that

$\operatorname{div} \operatorname{curl} \vec{v} \equiv \nabla \cdot \nabla \times \vec{v} = 0$, i.e., identifying \vec{v} with the 1-form ζ , and utilizing the previous example, we have the equivalence of this with $dd\zeta = 0$.

To state Poincaré’s Lemma, we first need the following two definitions:

Closed Forms: A p-form, $\alpha \in \Lambda^p$ is closed $\iff d\alpha = 0 \in \Lambda^{p+1}$;

Exact Forms: A p-form β in Λ^p is exact \iff there is some (p-1)-form, $\gamma \in \Lambda^{p-1}$ such that $\beta = d\gamma$.

Poincaré’s Lemma tells us that for “nice” regions of a set of coordinates, the two sorts of forms are equivalent; i.e.,

an exact form is closed, for all values of the coordinates,

a closed form is exact, inside some sort of appropriate region.

The first half is easy to prove by a simple computation using skew-symmetry and symmetry, as suggested above. That the converse is also “true” (locally) is a moderately difficult exercise in advanced calculus; nonetheless it is true for proper sorts of neighborhoods, usually referred to as “star-shaped” neighborhoods. (This means that there is some “central” point in the neighborhood such that there is a straight line from that point to every other point in the neighborhood.) Over large regions of space, the converse may well not be true; for just which forms it is true is the subject of the mathematical theory called cohomology. That theory is used to study the global properties of manifolds.

Notice the following comparisons to ordinary 3-dimensional vector analysis, which may help us to keep the content in mind:

Example: If we think of the usual 3-dimensional vector in terms of a 1-form over a space of 3 variables, as in the examples just above, we can notice that Poincaré’s Lemma has several special cases already known to us:

1. If $\nabla \times \vec{v} = 0$, then there exists a (scalar) potential V such that $\vec{v} = \nabla V$ becomes, identifying \vec{v} and ζ , that $d\zeta = 0$ implies the existence of a 0-form $f \in \Lambda^0\mathcal{M}$ such that $\zeta = df$;
2. If $\nabla \cdot \vec{w} = 0$, then there exists a vector potential \vec{h} such that $\vec{w} = \nabla \times \vec{h}$ becomes, identifying \vec{w} and the 2-form β , that $d\beta = 0$ implies the existence of a 1-form η such that $\beta = d\eta$.

VI. Choices of Bases for our Vector Spaces

We have so far only discussed basis sets for the spaces of tangent vectors, \mathcal{T} , and differential forms, Λ^1 , made directly from the coordinates. However the discussion of transformations between basis sets, above, could easily lead one to consider other sets of bases by taking linear combinations of the old ones, as begun with a discussion in Grøn, §4.3. Quite often there may be very good **physical reasons** for making some particular choice of basis vectors for either \mathcal{T} or Λ^1 . In general, we may choose any set of basis vectors for \mathcal{T}^1 that are linearly independent and sufficiently many, i.e., m of them. Under those circumstances, let us agree to refer to a general set of basis vectors for \mathcal{T}^1 by the standard symbols $\{\tilde{e}_a\}_{a=1}^m$, while for Λ^1 we refer to an associated set of basis 1-forms by the symbols $\{\omega^b\}_{b=1}^m$, the association between the two being determined by the requirement that they remain dual to one another; i.e., they should satisfy an appropriate version of duality, analogous to Eq. (4.3):

$$\text{dual basis sets are defined by} \quad \omega^a(\tilde{e}_b) = \delta_b^a \quad . \quad (6.1a)$$

Of course there are always (local) coordinates on the manifold, as well, which can be used to form a basis; therefore, these more general ones are simply linear combinations of those, with, usually, non-constant coefficients. Therefore, let us choose to denote those coordinates as the set $\{x^i\}_1^m$, so that we may give names to those coefficients that are often useful:

$$\begin{aligned} \tilde{e}_a &\equiv E^i{}_a \partial_{x^i} , & \omega^b &\equiv W^b{}_j dx^j , \\ \delta_a^b &= \omega^b(\tilde{e}_a) = W^b{}_j E^i{}_a dx^j(\partial_{x^i}) = W^b{}_j E^i{}_a \delta_i^j = W^b{}_j E^j{}_a , \end{aligned} \quad (6.1b)$$

from which we see that the two matrices of linear combinations are inverse to each other. Again, this is a fact that is, hopefully, not too difficult to believe, with the proof just above.

We want now to consider some particular ones that are often useful. In fact their usefulness comes from the properties they have relative to a metric, i.e., a definition of scalar products, and lengths, which we have not yet found necessary to introduce. Therefore, in principle this discussion should NOT come here, but, rather, after we introduce the metric. Nonetheless, this is, otherwise, a good and reasonable place to talk about particular physical reasons for making choices about basis sets for tangent vectors, and for differential forms. Therefore, we will look at a few examples now, using the usual metric conventions that we already know about, that are appropriate for flat space, in 3-dimensions, and also in 4-dimensions.

Orthonormal Choices for Bases

When dealing with ordinary physics problems in a very simple space, such as say, this room, we traditionally use orthonormal sets of basis vectors, i.e., basis vectors which are orthogonal to one another and of unit length. It would seem reasonable that our intuition about the behavior of the components of vectors is normally built over the assumption that the basis sets are orthonormal, so that we can find all of the physics in the components themselves.

However, in general on a more sophisticated surface, it is impossible to require **both** that the basis sets be made as partial derivatives with respect to a set of coordinates, **and** that they are orthonormal with respect to some notion of a choice of metric. Many years ago it was thought that the requirement that the basis set be derivatives with respect to a set of coordinates was the most important of these requirements; such basis sets are referred to as *holonomic bases*. Therefore, the difference between covariant and contravariant components was important, and one needed to develop accurate intuitions concerning the two. However, over the last 30 years, physicists have realized that it is not actually more difficult but, instead, somewhat easier to use *non-holonomic* sets of basis vectors, i.e., basis sets for which there is no choice of coordinates such that they would be the partial derivatives with respect to those coordinates. These non-holonomic basis sets make it much simpler to provide quick physical interpretations of the results of calculations! Therefore, following Grøn, §4.6, and nearby material, we will

now begin the introduction to non-holonomic sets of basis vectors, beginning by giving a very simple example, and then criteria for knowing which kind one has.

Example of a non-Holonomic Basis in polar coordinates, only 3 dimensions, and flat space:

We consider flat space, in 3-dimensions, but insist on using polar coordinates. Then, using coordinates $\{r, \theta, \varphi\}$, a plausible holonomic basis set for the tangent vectors would be $\{\partial_r, \partial_\theta, \partial_\varphi\}$. However, notice that these basis vectors do not even all have the same dimensions! Therefore, a set of basis vectors for the tangent space that is physically more consistent is given by the following:

$$\{\hat{e}_r \equiv \partial_r, \hat{e}_\theta \equiv \frac{1}{r}\partial_\theta, \hat{e}_\varphi \equiv \frac{1}{r\sin\theta}\partial_\varphi\} \quad \leftarrow \quad \text{a non-holonomic basis set for } \mathcal{T} \text{ over } \mathbb{R}^3. \quad (6.2)$$

All the vectors above have the same dimensions, and, as it turns out, for the usual choice of a metric on \mathbb{R}^3 , they are orthonormal, so that there is no difference between the covariant and contravariant components of a vector, thereby giving them the usual physical interpretation one has when dealing with $\{\hat{r}, \hat{\theta}, \hat{\varphi}\}$. It is for the reason that they are orthonormal, just like $\{\hat{r}, \hat{\theta}, \hat{\varphi}\}$, that I have denoted them with additional “hats” on the top of their symbols, to remind us of that.

Criteria for distinguishing holonomicity, Grøn, §4.3:

The next question that should arise is how one knows, given a particular basis set for \mathcal{T} , whether or not it is holonomic. After all, if one re-writes the holonomic set $\{\partial_r, \partial_\theta, \partial_\varphi\}$ in terms of the coordinates $\{x, y, z\}$, then it really will not appear to be holonomic. Therefore, one needs a sure test, which is given by the fact that partial derivatives with respect to the coordinates of a single coordinate system **always commute** one with another. Therefore, in general, for an arbitrary basis set $\{\tilde{e}_i\}$ of \mathcal{T} , we define the commutation coefficients:

$$[\tilde{e}_i, \tilde{e}_j] \equiv C_{ij}{}^k \tilde{e}_k \quad \text{where} \quad C_{ij}{}^k = -C_{ji}{}^k \quad . \quad (6.3)$$

Please compare Grøn, §4.3, especially his equation (4.20). However, also do notice that he defines his “structure coefficients” somewhat differently than I am doing. The quantities I am

calling C_{ij}^k he writes as C^k_{ij} . Unfortunately there is no unanimity in the literature on this notation. I will simply continue with my own symbols. We do mean the same things by the two (slightly) different notations.

Theorem: The commutation coefficients vanish \iff the basis set is holonomic, with respect to some system of coordinates.

For our example, we find that

$$\begin{aligned} [\tilde{e}_r, \tilde{e}_\theta] &= [\partial_r, \frac{1}{r}\partial_\theta] = -\frac{1}{r^2}\partial_\theta = -\frac{1}{r}\left\{\frac{1}{r}\partial_\theta\right\} = -\frac{1}{r}\tilde{e}_\theta, \quad \Rightarrow C_{r\theta}^\theta = -\frac{1}{r}, \\ [\tilde{e}_r, \tilde{e}_\varphi] &= [\partial_r, \frac{1}{r\sin\theta}\partial_\varphi] = -\frac{1}{r^2\sin\theta}\partial_\varphi = -\frac{1}{r}\left\{\frac{1}{r\sin\theta}\partial_\varphi\right\} = -\frac{1}{r}\tilde{e}_\varphi, \quad \Rightarrow C_{r\varphi}^\varphi = -\frac{1}{r}, \\ [\tilde{e}_\theta, \tilde{e}_\varphi] &= \left[\frac{1}{r}\partial_\theta, \frac{1}{r\sin\theta}\partial_\varphi\right] = -\frac{\cot\theta}{r}\frac{1}{r\sin\theta}\partial_\varphi = -\frac{\cot\theta}{r}\tilde{e}_\varphi, \quad \Rightarrow C_{\theta\varphi}^\varphi = -\frac{\cot\theta}{r}, \end{aligned}$$

with all other commutation coefficients being zero. As well, we will later find that the commutation coefficients tell us some useful and interesting things about the structure of the space on which one is studying physics.

A related question is then the structure, or behavior, of a basis for the space of 1-forms, Λ^1 . From the duality condition, one can easily see that the basis for 1-forms that is dual to a non-holonomic basis for tangent vectors would also be non-holonomic. This then means that they must also be linear combinations of the differentials of some coordinate set. Since the differentials corresponding to the exterior derivative acting on the functions that serve as coordinates are closed, it follows that a non-holonomic basis would not be *closed*, i.e., would **not** be differentials of **any other** possible system of coordinates. Therefore, a reasonable test for whether a basis set for 1-forms is holonomic would be to calculate the set of their differentials, i.e., $\{d\omega^a\}$. Such a set would of course be 2-forms, and therefore must be expressible as linear combinations of the wedge products of the basis set themselves. As it turns out, these combinations are related to the commutation coefficients already discussed:

$$d\omega^a \equiv -\frac{1}{2}C_{bc}^a \omega^b \wedge \omega^c \quad . \quad (6.4)$$

The proof of this is finally given in the next handout, on *affine connections and curvatures*, at Eq. (4.5).

Notice that the set of non-holonomic basis forms for \mathbb{R}^3 dual to the polar coordinate ones given in Eq. (6.3) is the following, where again the “hats” remind us that they are orthonormal:

$$\{\hat{\omega}^r \equiv dr, \hat{\omega}^\theta \equiv r d\theta, \hat{\omega}^\varphi \equiv r \sin \theta d\varphi\} \quad \leftarrow \text{a non-holonomic basis set for } \Lambda^1 \text{ over } \mathbb{R}^3. \quad (6.5)$$

One calculates that

$$\begin{aligned} d\hat{\omega}^r &= d dr = 0 \Rightarrow C_{ij}{}^r = 0, \\ d\hat{\omega}^\theta &= d(r d\theta) = dr \wedge d\theta = \frac{1}{r} dr \wedge r d\theta = \frac{1}{r} \hat{\omega}^r \wedge \hat{\omega}^\theta \\ &= \frac{1}{r} \left\{ \frac{1}{2} [\hat{\omega}^r \wedge \hat{\omega}^\theta - \hat{\omega}^\theta \wedge \hat{\omega}^r] \right\}, \quad \Rightarrow C_{r\theta}{}^\theta = -\frac{1}{r}, \\ d\hat{\omega}^\varphi &= d(r \sin \theta d\varphi) = \sin \theta dr \wedge d\varphi + r \cos \theta d\theta \wedge d\varphi \\ &= \frac{1}{r} \hat{\omega}^r \wedge \hat{\omega}^\varphi + \frac{\cos \theta}{r \sin \theta} \hat{\omega}^r \wedge \hat{\omega}^\varphi, \quad \Rightarrow C_{r\varphi}{}^\varphi = -\frac{1}{r}, \quad C_{\theta\varphi}{}^\varphi = -\frac{\cot \theta}{r}. \end{aligned}$$

These are in complete agreement with the ones just calculated above, using tangent vectors; again those not listed, or determined by skew-symmetry, are exactly zero.

Null tetrads in 4-dimensional spacetime:

The usual holonomic, orthonormal basis in flat, 4-dimensional spacetime is just the set $\{\partial_x, \partial_y, \partial_z, \partial_t\}$, and the corresponding basis for 1-forms, namely $\{dx, dy, dz, dt\} \equiv \{dx^\mu\}_1^4$. We may, however, use it in an interesting way to create a set of 4 basis vectors which are more agreeable for the studies of physical fields moving with the speed of light, such as electromagnetic or gravitational radiation. Motions of these fields have directions associated with them which are “null,” i.e., of zero length. Therefore it is usual in both special and general relativity to use basis vectors such as $dz + dt$ and $dz - dt$. These are both of zero length, and correspond to the tangent vectors for the paths of light rays traveling forward in time along the z -axis, incoming or outgoing, respectively. On the other hand, another linearly-independent pair of null rays does not exist; nonetheless, one should never let a simple bothersome fact like

nonexistence deter one from doing what “needs to be done.” Therefore, the standard approach to resolving this difficulty is to introduce **complex-valued**, null-length basis vectors in, for instance, the plane of the wave-front, i.e., for fixed values of z and t . In flat space, in Cartesian coordinates, this would correspond to the pair of basis vectors, $\hat{x} \pm i\hat{y}$. We now summarize the ideas defining this null tetrad: Taking $\{dx^\mu\}_1^4$ as the standard, holonomic, orthonormal tetrad, available in flat space, and $\{\varrho^\alpha\}_1^4$ as the associated null tetrad, we may write explicitly the transformation matrix between them, where we put in extra multiplicative factors of $1/\sqrt{2}$ so that the entries in the metric are just 1’s and 0’s:

$$\begin{aligned} \varrho^\alpha &= A^\alpha{}_\mu dx^\mu, \\ \begin{pmatrix} \varrho^1 \\ \varrho^2 \\ \varrho^3 \\ \varrho^4 \end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} +1 & +i & 0 & 0 \\ +1 & -i & 0 & 0 \\ 0 & 0 & +1 & -1 \\ 0 & 0 & +1 & +1 \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \\ dt \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} dx + idy \\ dx - idy \\ dz - dt \\ dz + dt \end{pmatrix}, \end{aligned} \quad (6.6)$$

where the matrix A has determinant of $-i$. I note that it is “somewhat” customary to use the symbols $\{\varrho^\alpha\}_1^4$ for the elements of a null basis, and also $\nu_{\alpha\beta}$ for the components of a metric describing a null tetrad, just as it is customary to use the symbol $\eta_{\mu\nu}$ for the components of a metric describing an orthonormal tetrad:

$$\varrho^\alpha \cdot \varrho^\beta = A^\alpha{}_\mu \eta^{\mu\nu} A^\beta{}_\nu \equiv \nu^{\alpha\beta} \implies \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (6.7)$$

There are other relative signs used in the literature for this matrix, depending, in large part, on which signature one has chosen for the metric. In particular there is a very commonly-used null-basis approach that goes by the name Newman-Penrose formalism. In that formalism the “bottom” 2×2 matrix has both the $+1$ ’s there replaced by -1 ’s, and some changes in the definitions as well, because they use the “wrong” metric.