Differential Geometry and Tensor Calculus

Naveen Durvasula

January 2019

1 Introduction

Differential geometry can be viewed as a generalization of multivariable calculus that aims to discover underlying relationships that remain no matter how we represent the space we're working with. In differential geometry, we extend calculus to *manifolds* which generalize the notions of "curve", "surface", and "volume" to higher dimensions. However, as the name suggests, we focus on the geometric nature of these objects which remain the same no matter how we define our coordinate system.

Another equally valid way to approach differential geometry is from an algebraic perspective. Key to the foundations of differential geometry are the algebraic systems which allow us to formally represent manifolds and related concepts in way that does not depend on a coordinate system.

Historically, early differential geometry arose in the mid 1700s as mathematicians found that Euclidean geometry was not the only consistent geometric system [6]. Thus, they needed a way of describing these new systems. Differential geometry arose out of that need. Initial objectives included methods for determining the local curvature of arbitrary manifolds. It has since developed into a generalized way of doing calculus.

Differential geometry serves as the underpinnings of many theoretical and applied fields. Most notably, Einstein's theory of general relativity is a result in differential geometry. However, in physics alone, differential geometry has more extensive applications, such as in the study of electromagnetism and Lagrangian/Hamiltonian mechanics. Differential geometry also served as the underpinnings for Grigori Perelman's proof of the Poincaré conjecture (one of the seven Clay millennium problems).

On the applied side, differential geometry is used to solve problems in digital signal processing, control theory, and computer vision [2]. Methods in differential geometry can also be used to image process data on curved surfaces. In chemistry and biophysics, differential geometry is used to model cell membrane structures under varying degrees of pressure. Other applications of differential geometry are in applied fields like economics, statistics, and geology [5].

2 Understanding Tensors (try 1)

Before we can understand concepts in differential geometry and tensor calculus, we must first understand both tensors and their basic algebraic operations. There are many definitions for these constructs. Some are too concrete, and as a result, they miss the true meaning. However, more big-picture definitions (see free vector space and category theoretic formulations) require too much background to understand. This text aims for a middle ground: we will define these concepts in a way that generalizes only up to the level that we use them.

To provide motivation for tensors and related concepts, we will give a series of definitions where each one is more general (and closer to the true definition) than the last. We assume that the reader has some experience with linear algebra and calculus. We begin with a surface-level definition of tensors.

Definition 2.1 (Tensors)
A **tensor** is a multidimensional array of numbers.

Although not incorrect, this definition misses the core of what a tensor truly is. However, in the field of computer science (specifically machine learning), this is the definition that is most commonly used. This definition does allow us to define (albeit crudely) the property of tensors known as *rank*.

Definition 2.2 (Rank)
The **rank** of a tensor is the number of **indices** required to reference a scalar.

Example 2.3 (Scalars)
Scalars are rank-0 tensors as no indices are required to reference the scalars contained within the object (after all, the object is simply a scalar). Examples of scalars include 1, π, e, and 4.7. We restrict our discussion of scalars to elements of $\mathbb{R}$ in this text, but in other contexts, complex numbers may be treated as scalars.

Example 2.4 (Vectors)
Vectors are rank-1 tensors as one index variable is required to reference the scalars contained within the object. Examples of vectors include $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1.2 \\ 2.4 \\ 3.6 \\ 4.8 \end{bmatrix}$, and functions like e^x (the coeffecients of e^x could be those of its Taylor expansion).

Example 2.5 (Matrices)
Matrices are rank-2 tensors as two index variables is required to reference the scalars contained within the object. Examples of matrices include $\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$, $\begin{bmatrix} 1.2 & 2.4 \\ 3.6 & 4.8 \end{bmatrix}$, and the differentiation operator.

Remark. It is important to notice the difference between rank and dimension. The dimension of a vector space is given by the size of its basis, or the number of elements that each tensor index can take on. The rank of the tensor is the "dimension" of the array given by the tensor.

3 Understanding Tensors (try 2)

As (perhaps over)emphasized earlier, this definition is very limited. The reason for this is that tensors are inherently *geometric* objects. We begin by intuitively defining tensors with a real-world example.

Example 3.1 (A Real-Life Tensor)
Pick up a pencil and point the tip at the nearest doorway. The pencil can now be considered as tensor. With your other hand, point your thumb, index, and middle fingers in three linearly independent directions (one could, in this scenario, assert their dominance as a member of the physics gang).

If we imagine the pencil as a vector in $\mathbb{R}^3$, we might describe it as a linear combination of the basis vectors given by your thumb, index, and middle fingers. However, the pencil is defined by a *geometric* relationship: it points to the nearest doorway. Thus, even if you were to change the basis by rotating your hand, the pencil tensor remains unchanged.

The above example leads us to the following better (but still intuitive) definition.

Definition 3.2 (Tensors)
A **tensor** remains invariant under a change in coordinates, and has components that change in a special way when the basis is changed.

We have yet to discuss the latter part of this definition. We do so with another example.

Example 3.3 (Demonstrating Contravariance)
Consider the vector $\mathbf{v} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$. We may alternatively write this vector as $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Now suppose that we change our basis from $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to $\{\frac{\mathbf{e}_1}{2}, \frac{\mathbf{e}_2}{2}, \frac{\mathbf{e}_3}{2}\}$. Since $\mathbf{v} = 2\frac{\mathbf{e}_1}{2} + 2\frac{\mathbf{e}_2}{2} + 2\frac{\mathbf{e}_3}{2}$, in this basis, $\mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$.

Although the basis vectors were halved in size, each of the *components* of $\mathbf{v}$ were doubled. Thus, we call this type of transformation a **contravariant** transformation.

From the above example, we learned two things: the components of a vector transform contravariantly with respect to the basis vectors, and of course, the basis vectors themselves transform **covariantly** with respect to themselves. We alter our notation to keep track of this. For all objects that transform covariantly with respect to the basis vectors, we will write their index as a subscript. Likewise, for all objects that transform contravariantly with respect to the basis vectors, we will write their index as a superscript. We would write, for example,

$$\mathbf{v} = v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + v^3\mathbf{e}_3$$

since the components v^1, v^2, and v^3 transform contravariantly, and the basis vectors $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ transform covariantly.

The notion of covariant and contravariant extends beyond simply scaling the basis: it extends to any linear transformation. The reason behind this has much to do with *reference frames*. If we were to halve the length of each of the basis vectors, in reference frame of the basis vectors, the corresponding component of a vector that remained invariant of the change would double in size. Likewise, if the entire basis were to rotate clockwise, in the reference frame of the basis, an invariant vector would rotate *counter*-clockwise.

Before we move on to our final and most complete definition of a tensor, we'll have to understand some more mathematical background.

4 Vector Spaces

In the above discussion, we've informally assumed some properties of vectors. Since vector spaces will be heavily utilized throughout this text, we give a formal definition below. We begin by defining fields.

Definition 4.1 (Field)
A **field** $\mathbb{F} \equiv (F, +, *)$ consists of a set F, and two binary operations $+ : F \times F \mapsto F$ and $* : F \times F \mapsto F$ such that the following properties hold:

- Associativity of $+$ and $*$: $a + (b + c) = (a + b) + c$ and $a(bc) = (ab)c$
- Commutativity of $+$ and $*$: $a + b = b + a$ and $ab = ba$
- Identity elements: $\exists 0, 1 \in F$ where $0 \neq 1$ such that $a + 0 = a$ and $a * 1 = a$
- Additive inverse: $\forall a \in F, \exists(-a) \in F$ such that $a + (-a) = 0$
- Multiplicative inverse: $\forall a \in F, \exists a^{-1} \in F$ such that $a(a^{-1}) = 1$
- Distributivity: $a(b + c) = ab + ac$

Any set G with a closed binary operation $+$ is called an **abelian group** if it satisfies the first four above properties.

We are now ready to define vector spaces. We establish the following notation convention: all scalars (belonging to some field) will be typeset using standard math font. Vectors will be written in bold face font.

Definition 4.2 (Vector Space)
A **vector space** $V \equiv (V, \mathbb{F}, +, *)$ is given by a set V, a **field** $\mathbb{F}$, an operation $+ : V \times V \mapsto V$, and an operation $* : F \times V \mapsto V$ such that the following properties hold:

- $(V, +)$ is an abelian group
- Scalar multiplication: $a(b\mathbf{v}) = (ab)\mathbf{v}$
- Consistency of the multiplicative identity: $1\mathbf{v} = \mathbf{v}$
- Distributivity: $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ and $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$

Definition 4.3 (Linear Transformations)
Let $\mathbf{v} \in V$. A linear transformation $T : V \mapsto W$ maps vectors from some vector space V to a vector space W such that the following properties hold:

- Scaling: $T(c\mathbf{v}) = cT(\mathbf{v})$
- Addition: $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$

We may instead check both constraints in one go by verifying that $T(c\mathbf{v} + \mathbf{w}) = cT(\mathbf{v}) + T(\mathbf{w})$. We denote by $\mathcal{L}(V, W)$ the set of all linear transformations from V to W.

With the definition of a linear transformation, we can now formally define covariant and contravariant transformations

Definition 4.4 (Covariant and Contravariant Transformations)
Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be our original basis. Let $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be our basis after applying the transformation T. An object transforms **covariantly** if we can express it in the new basis by transforming it by T. In other words, the object transforms covariantly if we can represent it in the new basis by transforming it in the same way that we transform the basis vectors. An object transforms **contravariantly** if we can express it in the new basis by transforming it by the inverse transformation T^{-1}. Thus, objects that transform contravariantly are transformed in the *opposite* way as the basis vectors.

Now that we know what vector spaces are, let's try to prove formally that vector components transform contravariantly. Suppose that we have some arbitrary invertible linear transformation $T \in \mathcal{L}(V, V)$ from a vector space V to itself. If we imagine T as a matrix, we can let T_{ij} index the scalars of T. Suppose that we begin with some basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and we transform to $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\} \equiv \{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)\}$. Let $\mathbf{v} \in V$.

$$\begin{aligned}
\mathbf{v} &= \sum_{j=1}^{n} v^j \mathbf{e}_j \\
&= \sum_{j=1}^{n} v^j \left(\sum_{i=1}^{n} T_{ij}^{-1} \mathbf{f}_j \right) \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} T_{ij}^{-1} v^j \right) \mathbf{f}_i
\end{aligned}$$

So our components were mapped by the *inverse* transform $v^j \mapsto \left(\sum_{j=1}^{n} T_{ij}^{-1} v^j \right)$, thus proving contravariance.

5 Covector Spaces

Now that we understand vector spaces, we can define the **dual vector space** also known as the **covector space**.

Definition 5.1 (The Dual Space)
Notated V^*, the dual vector space consists of the set of *linear* maps $\mathcal{L}(V, \mathbb{F})$ from the vector space V to its field $\mathbb{F}$.

Example 5.2 (Row Vectors)
A row vector is an example of a covector. A row vector $\boldsymbol{\alpha}$ can act on a vector v through the usual row column rule for matrix multiplication:

$$\boldsymbol{\alpha}(\mathbf{v}) = [\alpha_1, \alpha_2, \alpha_3] \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix} = \alpha_1 v^1 + \alpha_2 v^2 + \alpha_3 v^3$$

To verify that $\boldsymbol{\alpha}$ is in fact a linear function, we just need to check that $\boldsymbol{\alpha}(c\mathbf{u}+\mathbf{v}) = c\boldsymbol{\alpha}(\mathbf{u})+\boldsymbol{\alpha}(\mathbf{v})$. This can be proven by expanding the above sum.

Example 5.3 (Integration)
Let $V \equiv C([0,1], \mathbb{R})$ be the set of continuous functions from the unit interval $[0,1]$ to $\mathbb{R}$. The integral $\int_0^1 f(x)dx$ is an element of the dual space V^* as

$$\int_0^1 cf(x) + g(x)dx = c\int_0^1 f(x)dx + \int_0^1 g(x)dx$$

thus making it a linear function(al).

Theorem 5.4 (The Covector Basis)
If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for V, then there exists a linearly independent set $\{\mathbf{v}^1, \dots, \mathbf{v}^n\}$ where $\mathbf{v}^i \in V^*$ and $\mathbf{v}^i(\mathbf{v}_j) = \delta^i_j$. Here, δ^i_j refers to the **Kronecker delta symbol**, which evaluates to one if $i = j$ and zero otherwise.

We omit the proof since it does not fall within the scope of this text, and it relies on methods (see Zorn's lemma) that require more mathematical background. In finite dimensional spaces, this theorem actually guarantees that $\{\mathbf{v}^1, \dots, \mathbf{v}^n\}$ spans V^*, thus demonstrating that it is a basis for the dual space. It follows that in the finite dimensional case, the dimension of V^* is equal to that of V. In the infinite dimensional case, $\{\mathbf{v}^1, \dots, \mathbf{v}^n\}$ is guaranteed to exist and be linearly independent, but it does not form a basis. From here on, we limit our discussion to finite dimensional vector spaces. We will also use this construction for defining a canonical basis. Let our initial basis be the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. We let the standard basis for the dual space be $\{\mathbf{e}^1, \dots, \mathbf{e}^n\}$ that satisfies the Kronecker delta condition.

You might have noticed that in our first example, we indexed the $\boldsymbol{\alpha}$ covector with subscript notation. This is because covector *components* transform covariantly. The covector basis, on the other hand, transforms *contravariantly.*

Proof. We first show that the covector basis transforms contravariantly. Let our initial basis for V be $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and our transformed basis be $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$. Let $\mathbf{u}_i = T(\mathbf{v}_i)$ where $T \in \mathcal{L}(V, V)$ is an arbitrary (invertible) linear transformation.

$$\begin{aligned} \mathbf{u}^i(\mathbf{u}_i) = \mathbf{v}^i(\mathbf{v}_i) = \delta^i_i = 1 \\ \mathbf{u}^i(T(\mathbf{v}_i)) = \mathbf{v}^i(\mathbf{v}_i) \\ [T(\mathbf{u}^i)](\mathbf{v}_i) = \mathbf{v}^i(\mathbf{v}_i) \\ T(\mathbf{u}^i) = \mathbf{v}^i \\ \mathbf{u}^i = T^{-1}(\mathbf{v}^i) \end{aligned}$$

In line three, we make use of the fact that $\mathbf{u}^i$ is linear, thus allowing us to pull out the linear transformation T. In line four, we make use of that fact that this must hold for *all* linear transformations T. The components of a covector $\mathbf{w}^i$ transform contravariantly with respect to a change in the *covector* basis, for the same reason that the components of a vector transform contravariantly with respect to a change in basis (think reference frames). Since covectors themselves transform contravariantly with respect to the basis, the components transform covariantly. □

As we only work with finite vector spaces, we define explicitly the isomorphisms between the vector spaces V and V^*.

Definition 5.5 (The Musical Isomophisms)
We define the isomorphisms between V and V^*. We let $\flat : V \mapsto V^*$ and $\sharp : V^* \mapsto V$ be the isomorphism and its inverse where

$$\mathbf{v}^\flat = v_1\mathbf{e}^1 + \dots v_n\mathbf{e}^n$$

for $\mathbf{v} \in V$ and

$$\mathbf{v}^\sharp = v^i\mathbf{e}_1 + \dots v^n\mathbf{e}_n$$

for $\mathbf{v} \in V^*$. The notation $\flat$ and $\sharp$ is used to denote lowering or raising indices. In music, $\sharp$ is used to denote a half-step up. Here, $\mathbf{v}^\sharp$ takes a covector with indices v_i and transforms it to v^i. The $\flat$ operator does the opposite.

We now give geometric intuition for covectors. We can visualize vectors as physical "arrows" that have a magnitude and direction. We can likewise visualize covectors as oriented "stacks" of lines (see sub-figure 1 of Figure 1). We can geometrically visualize a covector eating a vector by placing the vector in the oriented stack and counting the number of lines that it pierces (see sub-figure 2).

If we have an orthonormal basis, we can also geometrically define the special covector basis that satisfies the Kronecker delta condition. For each basis vector, we can find its associated basis covector by considering the oriented stack where the separation between the lines is equal to the magnitude of the basis vector, and the orientation is given by the orientation of the basis vector. Given this construction, we can split a covector into its components (see sub-figure 3).

Lastly, we can see visually how covectors are *linear* operators, as we can let the covector independently eat each of the basis vectors and the summed result will be equal to the original result (see sub-figure 4).

We conclude this section on covectors by introducing the double dual vector space.

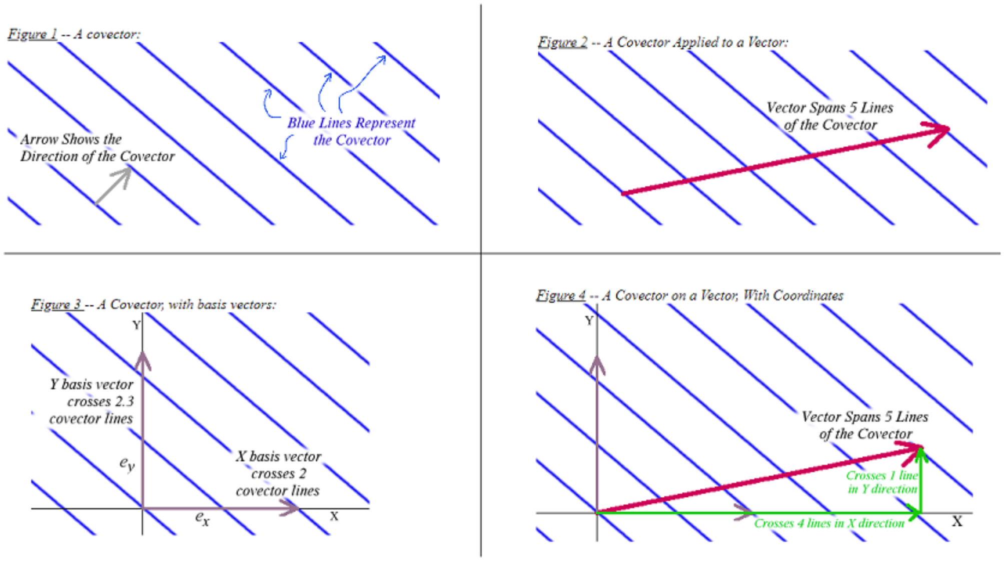

Figure 1: Covectors as Geometric Objects [4]

Definition 5.6 (Double Dual)
The **double dual** V^{**} of a vector space V is equal to $(V^*)^*$, or $\mathcal{L}(V^*, \mathbb{F}) = \mathcal{L}(\mathcal{L}(V, \mathbb{F}), \mathbb{F})$.

We now show a natural isomorphism φ between V and V^{**}. Let $\varphi(\mathbf{v}) = \hat{\mathbf{v}}$. We define $\hat{\mathbf{v}}$ by letting $\hat{\mathbf{v}}(\mathbf{v}^*) = \mathbf{v}^*(\mathbf{v})$ for $\mathbf{v}^* \in V^*$. The reason we care about the double dual is because we do not usually think of vectors as acting on some objects. However, elements of the double dual act on elements of the dual. We introduce this isomorphism as a way of noting explicitly that vectors can be thought of as objects that act on dual vectors.

Remark. The isomorphism is called a **natural** isomorphism because it does not depend on a coordinate system–it can be defined entirely with maps.

6 Understanding tensors (try 3)

We now have all (but one) of the mathematical tools necessary to complete our definition of tensors.

Definition 6.1 (Tensors)
A tensor is given by the **tensor product** of vectors and covectors.

Looks like we have to define this "tensor product" thing now. Keep in mind that the following definition is *not* the most general one (see monoidal categories), but it'll suffice for what we're going to be doing. The tensor product allows us to string together vectors and covectors together to create a larger multilinear map.

In the previous section we learned that elements of V^* "eat" vectors (elements of V) and make scalars. We also learned that elements of V^{**} (also known as V) eat covectors to make scalars. So what if we wanted to make a map that eats a vector *and* a covector to make a scalar? To do that, we'd need a covector to eat the vector, and a vector to eat the covector. We combine these two using the tensor product.

Formally, let $\mathcal{T}_q^p = \bigotimes_{i=1}^p V \otimes \bigotimes_{i=1}^q V^*$ be the set of all (p, q) tensors. What we mean by this is that any tensor $T \in \mathcal{T}_q^p$ has p contravariant components and q covariant components (remember that vector components are contravariant and covector components are covariant). Since T has p elements of the double dual, it will act on p elements of the dual. Since T has q elements of the dual, it will act on q elements of the vector space. Thus $\mathcal{T}_q^p \simeq \mathcal{L}(\bigtimes_{i=1}^p V^* \times \bigtimes_{i=1}^q V, \mathbb{F})$.

Remark. It might seem as though the spaces $V \otimes V$ and $V \times V$ should be isomorphic. This is most certainly not the case–$V \otimes V$ is much larger. This is because $V \otimes V$ is the space of *all* bilinear maps from $V^* \times V^* \mapsto \mathbb{F}$. There are some bilinear maps that cannot be expressed as the tensor product of two vectors, rather only as a linear combination of them.

Let's look at how some $\mathbf{v} \otimes \mathbf{w}^* \in V \otimes V^*$ acts on some covector and vector pair $(\mathbf{x}^*, \mathbf{y})$.

$$[\mathbf{v} \otimes \mathbf{w}^*](\mathbf{x}^*, \mathbf{y}) = \mathbf{x}^*(\mathbf{v})\mathbf{w}^*(\mathbf{y})$$

Now let's look at the dimension of the space $V \otimes V$. Any element of $V \otimes V$ can be formed from the linear combination of its basis. But what is its basis?

We know that

$$\begin{aligned} \mathbf{v} \otimes \mathbf{w} &= \left(\sum_i v^i \mathbf{e}_i\right) \otimes \left(\sum_i w^i \mathbf{e}_i\right) \\ &= \sum_i \sum_j (v^i \mathbf{e}_i) \otimes (w^j \mathbf{e}_j) \\ &= \sum_i \sum_j v^i w^j (\mathbf{e}_i \otimes \mathbf{e}_j) \end{aligned}$$

What we've just shown is that the set of all $\mathbf{e}_i \otimes \mathbf{e}_j$ can be used as a basis for $V \otimes V$. Thus, $\dim V \otimes V = (\dim V)^2$ as for each basis vector $\mathbf{e}_i$, we add all of the basis vectors corresponding to $\mathbf{e}_i \otimes \mathbf{e}_j$. It follows that if we were to take the tensor product of two arbitrary vector spaces $V \otimes W$, $\dim V \otimes W = (\dim V)(\dim W)$. Contrast this with the dimension of $V \oplus W$ (think Cartesian product, but with the algebraic structure carried over as well) which is given by $\dim V + \dim W$.

Remark. You might be wondering at this point why we care so much about linear transformations. If we're eventually going to do calculus, then shouldn't we care about arbitrary transformations? The reason behind this is similar to the fundamental assumption made in calculus: that "well-behaved" functions look like lines when you zoom in really close. Thus, if we assign a linear transformation to every point, then give some initial point, we can recreate the function. This should sound familiar...

7 An Extended Example: The Metric Tensor

The metric tensor is a very useful construct in differential geometry. While its applications are not within the scope of this paper, we include its definition as an extended example of the concepts defined above. Before we introduce the metric tensor, we first introduce a standard system of notation that cleans up a lot of our math.

Definition 7.1 (Einstein Summation Convention)
As we do arithmetic with tensors, we end up with a lot of sums. Einstein summation convention aims to eliminate these sums. Suppose we have some sum $\mathbf{v} = \sum v^i \mathbf{e}_i$. In Einstein notation, we drop off the sum, giving $\mathbf{v} = v^i \mathbf{e}_i$. It is implied in Einstein notation that if one side has fewer indices than the other, the extra indices are being summed over.

Using Einstein notation also allows us to create rules that make computing these sums easier. For example, suppose we have the sum notation expression $T_{ij}v^k w^l \delta^i_k \delta^j_l$. We can instantly simplify it by

$$T_{ij}v^k w^l \delta^i_k \delta^j_l = T_{ij}v^i w^j$$

We can cancel out the upper index k and lower index k and replace the upper index with an i, since when we sum over all k, we only get 1 when $k = i$. We can do something similar when multiplying tensors together. If the same index appears as a lower and upper index, we can cancel them out. We might say, for example, that $C^i_k = A^i_j B^j_k$. From here on, we will be using Einstein notation.

As we start doing calculus with tensors and move our discussion into differential geometry, one fundamental tensor that we'll run into a lot is the **metric tensor**. The metric tensor is a type $(0, 2)$ tensor. Recall that this means that we have 0 contravariant components and two covariant components. Thus, the metric tensor is an element of $V^* \otimes V^*$, and acts on $V \times V$.

Definition 7.2 (Metric Tensor)
For a type $(0, 2)$ tensor g to be a metric, we require that the following properties hold:

1. g is bilinear: we get this for free since we've already defined it as a $(0, 2)$ tensor
2. g is symmetric: $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}, \mathbf{x})$
3. g is nondegenerate: If for all $\mathbf{y}$, $g(\mathbf{x}, \mathbf{y}) = 0$, $\mathbf{x} = \mathbf{0}$.

We can describe g in terms of its components. The g tensor is a type $(0, 2)$ tensor. We can therefore expand g as

$$g = g_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j)$$

Given this, we can interpret the conditions listed above as properties that the matrix of coefficients must have. The property of symmetry ensurest that $g_{ij} = g_{ji}$. The nondegeneracy constraint ensures that g must be invertible. Proofs of these properties are left as exercises for the reader.

We now consider how g transforms under a change of basis. Consider some element $T \in V^* \otimes V$. Note that elements of this space eat a vector and a covector to create a scalar. However, if we input only a vector, we will get back a vector, as the covector component acts on the vector while the vector component does nothing. Thus, T is a $(1, 1)$ tensor. Suppose that we have some initial basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ that we transform to $\{\mathbf{f}_1, \dots, \mathbf{f}_n\} \equiv \{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$.

$$\begin{aligned}
g &= g_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) \\
&= g_{ij}(T^i_k \mathbf{f}^k \otimes T^j_l \mathbf{f}^l) \\
&= T^i_k T^j_l g_{ij}(\mathbf{f}^k \otimes \mathbf{f}^l) \\
&= T^i_k T^j_l g_{ij}(\mathbf{f}^k \otimes \mathbf{f}^l)
\end{aligned}$$

The components transform by not one, but *two* forward transformations ($g_{ij} \mapsto T^i_k T^j_l g_{ij}$). This should make sense, since g has *two* covariant components.

The metric tensor gets its name from the fact that we can use it to compute the length of a vector (or alternatively the distance between two points) in our space.

Definition 7.3 (Length)
Given a metric g, the **length** of a vector $\mathbf{v}$ is given by $\sqrt{g(\mathbf{v}, \mathbf{v})}$.

When we're measuring the length of a vector $\mathbf{v} \in \mathbb{R}^n$ in Euclidean space, we take the square root of the sum of the squares of the components

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{n} (v^i)^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

The dot product is a motivating example of a metric. In fact, it is the simplest example of a metric: the components of its corresponding metric tensor are given by the Kronecker delta $g_{ij} = \delta^i_j$. We prove this below:

$$\begin{aligned} \mathbf{v} \cdot \mathbf{v} &= g(\mathbf{v}, \mathbf{v}) \\ &= g_{ij} v^j v^i \\ &= \delta^i_j v^j v^i \\ &= v^i v^i \end{aligned}$$

We've actually already dealt with different metric tensors in physics when we were discussing relativity. In classical physics we work in four dimensions (given by space and time). When we changed reference frames, we used the velocity and time transformation rules that preserve the Euclidean lengths of our space and time parameters. These transformations were called **Galilean transformations**. At relatavistic speeds, our velocity and time transformations rules (called **Lorentz transformations**) actually preserve a *different* metric called the **Minkowski metric** (this name should sound familiar). Discussion of these specific examples goes beyond the scope of this text, but will serve as good supplemental material.

8 The Exterior Algebra

In this section, we define a special class of tensors that make up the *exterior algebra*. The exterior algebra will form the foundations for when we start doing calculus with tensors. We'll start by defining the exterior algebra for type $(0, 2)$ tensors. Recall that these tensors are elements of $V^* \otimes V^*$ and takes as input two elements of V to return a scalar. The exterior algebra is constructed from special tensors called *alternating* tensors.

Definition 8.1 (Alternating $(0, 2)$ Tensors)
An **alternating** type $(0, 2)$ tensor is a tensor T such that $T(\mathbf{x}, \mathbf{y}) = -T(\mathbf{y}, \mathbf{x})$ for all vectors $\mathbf{x}$ and $\mathbf{y}$. It follows that for such tensors, $T_{ij} = -T_{ji}$ and $T_{ii} = 0$. We denote by $\Lambda^2(V^*)$ the set of alternating $(0, 2)$ tensors. We call it (for reasons given later) the second **exterior power** of V.

Remark. So why do we care about alternating tensors in calculus? While this concept is explored in a lot more depth later, we give an intuitive explanation here. Recall from multivariable calculus that when we do line integrals, surface integrals, and volume integrals, we require that the object that we're integrating over is *orientable*. Alternating tensors are important in calculus as they define the geometric objects that are orientable. There's a lot of geometric intuition behind the construction of alternating tensors. We cover the exterior algebra from an entirely algebraic perspective in this section. In the next section, we introduce the geometric interpretations.

With no context whatsoever, we're going to independently define the *wedge product* for type $(0,2)$ tensors (we give the general definition later).

Definition 8.2 (The Wedge Product)
The **wedge product** (denoted $\wedge$) is a binary operation on covectors given by

$$\mathbf{v}^1 \wedge \mathbf{v}^2 = \mathbf{v}^1 \otimes \mathbf{v}^2 - \mathbf{v}^2 \otimes \mathbf{v}^1$$

which returns a type $(0,2)$ tensor. We get immediately from this definition that

$$\mathbf{v}^1 \wedge \mathbf{v}^2 = -\left(\mathbf{v}^2 \wedge \mathbf{v}^1\right)$$

We first show that whenever we take the wedge product of two vectors, we end up with an *alternating* type $(0,2)$ tensor. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be vectors in V. Let's see how the tensor $\mathbf{v}^1 \wedge \mathbf{v}^2$ acts on these vectors:

$$\begin{aligned}
\left[\mathbf{v}^1 \wedge \mathbf{v}^2\right](\mathbf{u}_1, \mathbf{u}_2) &= \left[\mathbf{v}^1 \otimes \mathbf{v}^2 - \mathbf{v}^2 \otimes \mathbf{v}^1\right](\mathbf{u}_1, \mathbf{u}_2) \\
&= \left[\mathbf{v}^1 \otimes \mathbf{v}^2\right](\mathbf{u}_1, \mathbf{u}_2) - \left[\mathbf{v}^2 \otimes \mathbf{v}^1\right](\mathbf{u}_1, \mathbf{u}_2) \\
&= \mathbf{v}^1(\mathbf{u}_1)\mathbf{v}^2(\mathbf{u}_2) - \mathbf{v}^2(\mathbf{u}_1)\mathbf{v}^1(\mathbf{u}_2) \\
&= \mathbf{v}^2(\mathbf{u}_2)\mathbf{v}^1(\mathbf{u}_1) - \mathbf{v}^1(\mathbf{u}_2)\mathbf{v}^2(\mathbf{u}_1) \\
&= -\left(\mathbf{v}^1(\mathbf{u}_2)\mathbf{v}^2(\mathbf{u}_1) - \mathbf{v}^2(\mathbf{u}_2)\mathbf{v}^1(\mathbf{u}_1)\right) \\
&= -\left(\left[\mathbf{v}^1 \otimes \mathbf{v}^2\right](\mathbf{u}_2, \mathbf{u}_1) - \left[\mathbf{v}^2 \otimes \mathbf{v}^1\right](\mathbf{u}_2, \mathbf{u}_1)\right) \\
&= -\left[\mathbf{v}^1 \otimes \mathbf{v}^2 - \mathbf{v}^2 \otimes \mathbf{v}^1\right](\mathbf{u}_2, \mathbf{u}_1) \\
&= -\left[\mathbf{v}^1 \wedge \mathbf{v}^2\right](\mathbf{u}_2, \mathbf{u}_1)
\end{aligned}$$

Thus, by switching the order of the vector argument $\mathbf{u}_1$ and $\mathbf{u}_2$, we showed that the result was negated. Thus the wedge product returns an alternating tensor. Just as we showed earlier that the $\mathbf{e}_i \otimes \mathbf{e}_j$ formed the basis of $V \otimes V$, we show now that the $\mathbf{e}^i \wedge \mathbf{e}^j$ form the basis of $\Lambda^2(V^*)$. Suppose we have some tensor $T \in \Lambda^2(V^*)$

$$\begin{aligned}
T &= \sum_{i=1}^{n}\sum_{j=1}^{n} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) \\
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) + \sum_{j=i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) + \sum_{j>i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) + 0 + \sum_{j>i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) \\
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) + \sum_{i>j} T_{ji}(\mathbf{e}^j \otimes \mathbf{e}^i) \\
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j) - \sum_{i>j} T_{ij}(\mathbf{e}^j \otimes \mathbf{e}^i) \\
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \otimes \mathbf{e}^j - \mathbf{e}^j \otimes \mathbf{e}^i) \\
&= \sum_{j<i} T_{ij}(\mathbf{e}^i \wedge \mathbf{e}^j)
\end{aligned}
$$

In the third line, we use the fact that $T_{ii} = 0$ in any alternating tensor, thus allowing us to eliminate the middle sum. Since the names of our index variables i and j don't matter, we swap them in the fourth line. In the fifth line, we make use of the fact that $T_{ji} = -T_{ij}$. Since summing over $i > j$ is the same as summing over $j < i$, we can collect like terms. What we've shown here is that *any* alternating $(0, 2)$ tensor can be written as a linear combination of the $\mathbf{e}^i \wedge \mathbf{e}^j$. Thus, these tensors form a basis for $\Lambda^2(V^*)$. Note that while there are n^2 linearly independent $\mathbf{e}^i \otimes \mathbf{e}^j$, there are only $\binom{n}{2}$ linearly independent $\mathbf{e}^i \wedge \mathbf{e}^j$ as $\mathbf{e}^i \wedge \mathbf{e}^j = -(\mathbf{e}^i \wedge \mathbf{e}^j)$.

We now extend these definitions to the general case of type $(0, q)$ tensors. Before we can do that, we have to define permutations.

Definition 8.3 (Permutations)
A **permutation** $\sigma : S \mapsto S$ of a set S is a rearrangement of its elements. More formally, it is a bijection from the set to itself. For example, consider the set $\{1, 2, 3\}$. We might rearrange its elements to be $1, 3, 2$. To accomplish this, we define the function σ such that $\sigma(1) = 1$, $\sigma(2) = 3$, and $\sigma(3) = 2$.

Associated with any permutation is its *sign*, which we denote by $\text{sgn}(\sigma)$. Suppose that we wanted construct the permutation in successive steps by swapping two elements of the set at a time. In the above example, we can construct the permutation in one step by simply swapping 3 and 2. We let $\text{sgn}(\sigma) = 1$ if the number of steps necessary to construct the permutation is even, and we let $\text{sgn}(\sigma) = -1$ if the number of steps necessary to construct the permutation is odd.

We use permutations to define alternating $(0, q)$ tensors.

Definition 8.4 (Alternating $(0, q)$ Tensors)
We denote by $\Lambda^q(V^*)$ the set of alternating $(0, q)$ tensors. We similarly call $\Lambda^q(V^*)$ the qth exterior power of V. An alternating $(0, q)$ tensor T satisfies the property that for all permutations Let $\sigma(1), \ldots, \sigma(q)$ of the indices $1, \ldots, q$.

$$T(\mathbf{v}_1, \ldots, \mathbf{v}_q) = \text{sgn}(\sigma) T(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(q)})$$

Thus, *every time* we swap two vectors in the input of T, the sign of the result flips. If we were to just swap two vectors as we did with type $(0, 2)$ tensors, the sign would flip once. If we were to swap another two vectors, the sign would flip again. It is this behavior that gives these tensors the name "alternating", as the sign alternates upon any swap of two vectors.

The $\mathbf{e}^i \wedge \mathbf{e}^j$ gave the basis of $\Lambda^2(V^*)$. We now provide the generalization for this without justification (the proofs are tedious and resemble the one we did for $\Lambda^2(V^*)$). We claim that the basis for $\Lambda^q(V^*)$ consists of the $\binom{n}{q}$ linearly independent $\mathbf{e}^{i_1} \wedge \mathbf{e}^{i_2} \wedge \cdots \wedge \mathbf{e}^{i_q}$. We note that the highest exterior power we can raise V to is $q = n$ as for $q > n$, $\binom{n}{q} = 0$. By convention, we let $\Lambda^1(V^*) = V^*$ and $\Lambda^0(V^*) = \mathbb{F}$ where $\mathbb{F}$ is the field of the vector space. We do this so that we obey the rule that $\dim(\Lambda^k(V^*)) = \binom{n}{k}$.

So far, we have only used the wedge product as a binary operation on covectors. However, it can in fact act as a binary operation on any two alternating tensors. For example, suppose that we have some tensor $P \in \Lambda^2(V^*)$ and a tensor $Q \in \Lambda^3(V^*)$. Let's try and compute $P \wedge Q$.

$$\begin{aligned} P \wedge Q &= \left(P_{ij}(\mathbf{e}^i \wedge \mathbf{e}^j)\right) \wedge \left(Q_{klm}(\mathbf{e}^k \wedge \mathbf{e}^l \wedge \mathbf{e}^m)\right) \\ &= P_{ij}Q_{klm}\left((\mathbf{e}^i \wedge \mathbf{e}^j) \wedge (\mathbf{e}^k \wedge \mathbf{e}^l \wedge \mathbf{e}^m)\right) \\ &= P_{ij}Q_{klm}\left(\mathbf{e}^i \wedge \mathbf{e}^j \wedge \mathbf{e}^k \wedge \mathbf{e}^l \wedge \mathbf{e}^m\right) \end{aligned}$$

In the second line we used the bilinearity of the wedge product. In the third line we used the fact that the wedge product is associative. While we have justified that the wedge product is bilinear, as it is generated from the tensor product, we have not justified why the wedge product is associative. While this justification isn't too difficult, it is beyond the scope of this text. Using these properties, we were able to wedge together an element of $\Lambda^2(V^*)$ and an element of $\Lambda^3(V^*)$ to get an element of $\Lambda^5(V^*)$.

In general, wedging together elements of $\Lambda^p(V^*)$ and $\Lambda^q(V^*)$ gives us an element of $\Lambda^{p+q}(V^*)$. Note that since field elements belong to $\Lambda^0(V^*)$, multiplying by a field element is the same as wedging by a field element. More formally, for $\boldsymbol{\alpha} \in \Lambda^k(V^*)$

$$c\boldsymbol{\alpha} = c \wedge \boldsymbol{\alpha}$$

as $c \in \Lambda^0(V^*)$ since the tensor $(c \wedge \boldsymbol{\alpha}) \in \Lambda^{k+0}(V^*) = \Lambda^k(V^*)$.

We now have the necessary machinery to define the *exterior algebra*. The exterior algebra gets its name because we can wedge together elements of some exterior power $\Lambda^k(V^*)$ we get an element of a *different* space $\Lambda^l(V^*)$. The exterior algebra consists of all of the elements that we could possibly wedge together.

Definition 8.5 (Exterior Algebra)
Let V be a vector space of dimension n. The **exterior algebra**

$$\Lambda(V^*) = \Lambda^0(V^*) \oplus \Lambda^1(V^*) \oplus \Lambda^2(V^*) \oplus \Lambda^3(V^*) \oplus \cdots \oplus \Lambda^n(V^*)$$

is the aggregation of all of the exterior powers.

We reiterate that in this section we provide no geometric intuition for the constructs we define – we will provide this intuition in the next section. We now define algebraically an important property of exterior powers called *Hodge duality*.

Definition 8.6 (Hodge Duality)
Recall that the dimension of $\Lambda^k(V^*) = \binom{n}{k}$. Note that $\binom{n}{k} = \binom{n}{n-k}$. Thus, $\dim(\Lambda^k(V^*)) = \dim(\Lambda^{n-k}(V^*))$. It follows that $\Lambda^k(V^*)$ is isomorphic to $\Lambda^{n-k}(V^*)$. We call these spaces **Hodge duals**. We define the **Hodge star** as an operator that acts on elements of the exterior algebra and maps it to the corresponding element of its Hodge dual. Formally, if we let T be an element of $\Lambda^k(V^*)$, we denote by $\star T$ its corresponding element in $\Lambda^{n-k}(V^*)$.

9 Tangent and Cotangent Spaces

We now transition to doing calculus with tensors. Consider this section as a large extended example of the concepts above in much the same way that differential equations are an application of linear algebra (in fact, in *exactly* the same way). In this section, we will work in $\mathbb{R}^3$, but all of the concepts can be easily extended to n dimensions.

We motivate this section by considering the fundamental assumption of calculus: that smooth objects "look linear" when you zoom in really close. We formally call these smooth objects **manifolds**. We call the linear approximation of the manifold when we zoom in to a point the **tangent space** of the manifold.

Example 9.1 (The Tangent Space of a Function)
The simplest type of smooth manifold that we've dealt with extensively are real-valued differentiable functions $f : \mathbb{R} \mapsto \mathbb{R}$. When we zoom in really close to the function at some point $f(x_0)$, we get the tangent line of the function (given by the derivative of f). The tangent space of the function f *is* just the tangent line. Thus, it is a one-dimensional vector space.

Example 9.2 (The Tangent Space of a Surface)
In multivariable calculus, we found the tangent plane of a surface at a given point. This plane gives a linear approximation of the surface. Thus, the tangent space of a surface is a two-dimensional vector space.

Now that we've established what tangent spaces are, let's show more formally that tangent spaces are in fact vector spaces.

To show that the tangent space is a vector space, we need to establish its basis. We first do this for the simplest type of manifolds that we deal with in multivariable calculus: 2-dimensional surfaces. Suppose we have a surface $z = f(x, y)$. We can get its tangent plane as

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

As z is a function of x and y, a vector in the tangent plane can be associated with a choice of $c_1 = (x - x_0)$ and $c_2 = (y - y_0)$ where offset our x and y position by the vector

$$\mathbf{v} = c_1 \frac{\partial f}{\partial x} + c_2 \frac{\partial f}{\partial y}$$

Thus, the tangent space is a 2-dimensional vector space with a *basis* given by $\left\{\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right\}$. In the general case where $z = f(x^1, \ldots, x^n)$, this argument would extend to show that the tangent space of this surface is the n-dimensional space with the basis $\left\{\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right\}$.

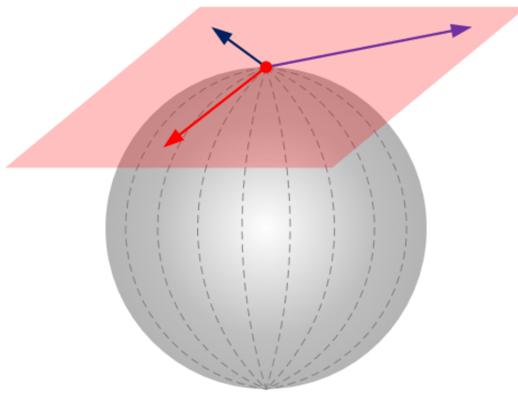

Figure 2: A sphere's tangent space [3]

But what if we don't know the function f? If we have only a *bug-eye* view of the manifold, how can we determine the tangent space? A way to visualize this is to imagine that we are walking along the surface of the earth, and we point in some direction $\mathbf{v}$. If we were to continue going straight in the direction of $\mathbf{v}$, we would end up in outer space since $\mathbf{v}$ is an element of the tangent space of the Earth at the point where we are. We don't know the equation for the surface of the Earth, but we still wish to describe the components of $\mathbf{v}$. Another way to consider this problem is through *stereographic projection*. Suppose that we have flat map (like a Mercator map) of the Earth. If all we have is the map, and we don't have an explicit description of the surface, what are our basis directions at any given point?

The way we solve this problem is by letting the basis vectors be the partial differentiation *operators* $\left\{\frac{\partial}{\partial x^1} \ldots \frac{\partial}{\partial x^n}\right\}$ themselves. Since partial differentiation is a linear operator, we can apply the $c_1, \ldots, c_n$ constants *before* we act on some function f. Think of this in much the same way that we define the linear L operator in a differential equations class. These operators define what direction means at a local scale.

If this discussion didn't make sense, feel free to think of the basis vectors as $\left\{\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right\}$ instead of $\left\{\frac{\partial}{\partial x^1} \ldots \frac{\partial}{\partial x^n}\right\}$. The only purpose of using the latter basis is to extend these concepts to *intrinsic geometries*, or manifolds that we do not have an explicit description for. In physics (specifically when we're talking about concepts like general relativity), we use intrinsic geometries extensively, as we only have a bug-eye view of spacetime. We have no way to get an explicit description of the shape of spacetime, but we can use local directions to determine its curvature.

We now define the covector basis. Recall that we want to find a covector basis $\{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$ such that $\mathbf{e}^j\mathbf{e}_i = \delta_i^j$. As our basis vectors are orthonormal, geometrically (recall from Figure 1), the basis covector associated with $\frac{\partial}{\partial x^i}$ should be an oriented stack where the separation distance is equal to the magnitude of $\frac{\partial}{\partial x^i}$ (which is infinitesimally small) and its orientation should the same as $\frac{\partial}{\partial x^i}$.

We let our covector basis be given by $\{dx^1, \ldots, dx^i\}$. We show intuitively that this basis satisfies the conditions stated above. When we write an integral in two dimensions, for example $\int f(x)dx$, we think of dx as a "little change" in the x direction (which is the same as the $\frac{\partial}{\partial x}$ direction since flat 2-d space is not curved). Thus, we may think of dx as a stack oriented *in the $\frac{\partial}{\partial x}$ direction* where the separation distance is in an infinitesimal. The dx covector is constant, as in two dimensions, for example, the density and orientation of the lines remain constant across all x and y.

To summarize, on the tangent space of a manifold, the vector basis is given by the set of partial differentiation operators $\left\{\frac{\partial}{\partial x^1}\dots\frac{\partial}{\partial x^n}\right\}$ and the covector basis is given by the set of *differentials* $\{dx^1,\dots,dx^n\}$. We denote by T_pM the tangent space at some point p on a manifold M, and we denote by T_p^*M the cotangent space at some point p on a manifold M.

10 The Exterior Calculus

We now put together everything discussed in the previous sections to define the *exterior calculus*, which is a generalization of what we do in multivariable calculus. The exterior calculus is built up from the exterior algebra of the cotangent space $\Lambda(T_p^*M)$. But what does this mean?

We first give intuitive explanations for each of the exterior powers of T_p^*M. We assume that the dimension of T_p^*M is 3 to simplify explanations, but all concepts generalize.

Definition 10.1 (0-forms)
Let's start with the zeroth exterior power $\Lambda^0(T_p^*M)$. These are just elements of the field. We would like the field elements of these vector spaces to be $\mathbb{R}$, as usual. However, since this space is specific to point p, we can also allow *scalar functions* $f : M \mapsto \mathbb{R}$ as by plugging in p, we still get scalar values. This is similar to how in differential equations with solutions y, we treat functions $\rho(x)$ sort of like constants. As scalar functions belong to $\Lambda^0(T_p^*M)$, we call them 0-forms.

Definition 10.2 (1-forms)
Now we move to the first exterior power $\Lambda^1(T_p^*M) = T_p^*M$. It's dimension is given by $\binom{3}{1} = 3$, and the basis vectors are simply the basis vectors of T_p^*M, given by $\{dx^1, dx^2, dx^3\}$. Geometrically, as described in the previous section, these represent differentials in each of the 3 directions. We call elements of $\Lambda^1(T_p^*M)$ 1-forms.

Definition 10.3 (2-forms)
The dimension of the next exterior power $\Lambda^2(T_p^*M)$ has dimension $\binom{3}{2} = 3$. As defined in Section 8, the basis of $\Lambda^2(T_p^*M)$ is given by $\{dx^1 \wedge dx^2, dx^2 \wedge dx^3, dx^3 \wedge dx^1\}$. But what do the elements of this basis mean geometrically? As hinted at earlier, since the $dx^i \wedge dx^j$ are *alternating* tensors, they have an orientation. Geometrically, these tensors are simply *oriented plane segments*. We call elements of $\Lambda^2(T_p^*M)$ 2-forms

Definition 10.4 (3-forms)
The dimension of the last exterior power $\Lambda^3(T_p^*M) = \binom{3}{3} = 1$. The basis of this space is simply $\{dx^1 \wedge dx^2 \wedge dx^3\}$. Elements of this space represent *oriented volume segments*. We call elements of $\Lambda^3(T_p^*M)$ 3-forms.

We now give geometric intuition for Hodge duality and the Hodge star operator. In our setup, $\Lambda^1(T_p^*M)$ is dual to $\Lambda^2(T_p^*M)$ and $\Lambda^0(T_p^*M)$ is dual to $\Lambda^3(T_p^*M)$ since $n = 3$. This holds geometrically

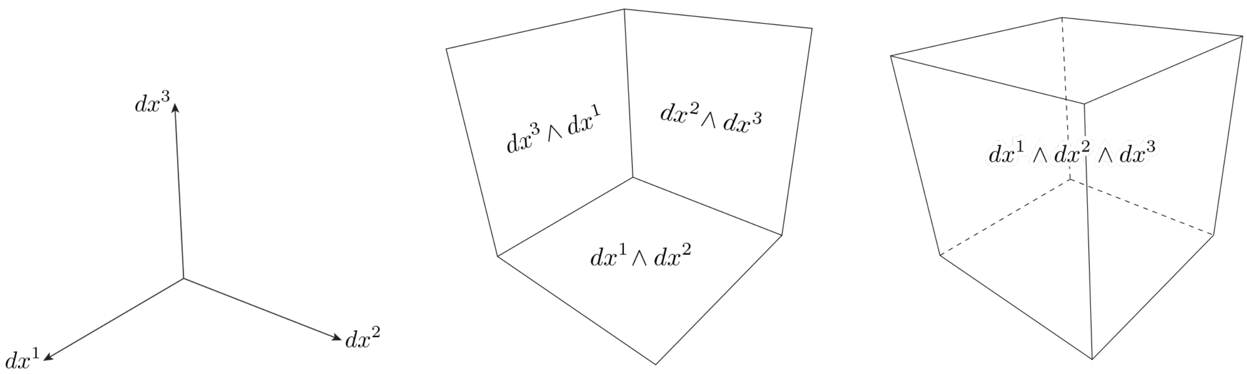

Figure 3: The one-form, two-form, and three-form bases [1]

as for any plane we can associate a unique normal vector, and for any oriented volume segment we can associate a unique scalar corresponding to the segment's volume. For any 0-form f, we let $\star f$ be the volume segment with volume f. For any 2-form X, we let $\star X$ denote its normal vector.

Formally,

$$\star f = f(dx^1 \wedge dx^2 \wedge dx^3)$$

and vice versa. Similarly,

$$\star(\alpha_1 dx^1 + \alpha_2 dx^2 + \alpha_3 dx^3) = \alpha_1(dx^2 \wedge dx^3) + \alpha_2(dx^3 \wedge dx^1) + \alpha_3(dx^1 \wedge dx^2)$$

and vice versa.

11 Exterior Differentiation

With this, we can now start doing calculus with elements of $\Lambda(T_p^* M)$. We define the *exterior derivative* which extends the concept of a differential to higher order differential forms.

Definition 11.1 (Exterior Derivative)
The **exterior derivative** is an operator d that acts on elements of $\Lambda(T_p^* M)$ such that the following hold

- $d^2 = 0$
- For 0-forms f,
$$df := \frac{\partial f}{\partial x^1} dx^1 + \ldots \frac{\partial f}{\partial x^n} dx^n$$
gives the *differential* of f
- For any p-form $\boldsymbol{\alpha}$ and q-form $\boldsymbol{\beta}$,
$$d(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) = d(\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + d(\boldsymbol{\beta}) \wedge \boldsymbol{\alpha}$$

The first property comes from the fact that the differential of a differential goes to 0. The second property establishes that taking the exterior derivative of 0-forms is the same as finding the 0-form's differential (which is something we did multivariable calculus). The last property establishes the analogue of the product rule for wedge products. It is this property that allows us to extend the differential to higher order forms. Since wedge products give alternating tensors, we can rewrite the third property as

$$d(\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) = d(\boldsymbol{\alpha}) \wedge \boldsymbol{\beta} + (-1)^p(\boldsymbol{\alpha} \wedge d(\boldsymbol{\beta}))$$

We get the $(-1)^p$ factor by making p "swaps" to get the form $\boldsymbol{\alpha}$ in front of the form $d(\boldsymbol{\beta})$.

Let's take the exterior derivative of a 0-form f. By our definition,

$$df = \frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + \frac{\partial f}{\partial x^3}dx^3$$

The exterior derivative of a 0-form resembles the *gradient*. Recall that the gradient

$$\nabla f = \left\langle \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \frac{\partial f}{\partial x^3} \right\rangle$$

Thus, the components of df are the same as the components ∇f. However, ∇f is a vector and df is a covector. We can relate the two by the musical isomorphisms:

$$\nabla f = (df)^\sharp$$

Since df is a covector, let's consider what happens when we let it act on some vector $\mathbf{u}$.

$$df(\mathbf{u}) = \left[\frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + \frac{\partial f}{\partial x^3}dx^3\right]\left(u^1\frac{\partial}{\partial x^1} + u^2\frac{\partial}{\partial x^n} + u^3\frac{\partial}{\partial x^3}\right) = \frac{\partial f}{\partial x^1}u^1 + \frac{\partial f}{\partial x^2}u^2 + \frac{\partial f}{\partial x^3}u^3$$

Note that the *directional derivative*

$$\nabla f \cdot \mathbf{u} = \frac{\partial f}{\partial x^1}u^1 + \frac{\partial f}{\partial x^2}u^2 + \frac{\partial f}{\partial x^3}u^3$$

Thus,

$$\nabla f \cdot \mathbf{u} = df(\mathbf{u})$$

It follows that the exterior derivative of a 0-form, when applied on a vector $\mathbf{u}$ gives its directional derivative.

Now let's take the exterior derivative of a 1-form $\boldsymbol{\alpha}$. We use Einstein sum notation.

$$\begin{aligned}
d(\boldsymbol{\alpha}) &= d(\alpha_i dx^i) \\
&= d(\alpha_i \wedge dx^i) \\
&= d(\alpha_i) \wedge dx^i - \alpha_i \wedge d(dx^i) \\
&= d(\alpha_i) \wedge dx^i - 0 \\
&= \frac{\partial \alpha_i}{\partial x^j} dx^j \wedge dx^i
\end{aligned}$$

Let's rewrite this final line without using Einstein notation

$$\frac{\partial \alpha_1}{\partial x^1}dx^1 \wedge dx^1 + \frac{\partial \alpha_1}{\partial x^2}dx^2 \wedge dx^1 + \frac{\partial \alpha_1}{\partial x^3}dx^3 \wedge dx^1 +$$

$$\frac{\partial\alpha_2}{\partial x^1}dx^1\wedge dx^2+\frac{\partial\alpha_2}{\partial x^2}dx^2\wedge dx^2+\frac{\partial\alpha_2}{\partial x^3}dx^3\wedge dx^2+$$
$$\frac{\partial\alpha_3}{\partial x^1}dx^1\wedge dx^3+\frac{\partial\alpha_3}{\partial x^2}dx^2\wedge dx^3+\frac{\partial\alpha_3}{\partial x^3}dx^3\wedge dx^3$$

By the definition of the wedge product, $dx^i \wedge dx^i = 0$. Furthermore, we also know that $dx^i \wedge dx^j = -(dx^j \wedge dx^i)$. We can thus simplify the above to the following

$$\begin{aligned}&0-\frac{\partial\alpha_1}{\partial x^2}dx^1\wedge dx^2+\frac{\partial\alpha_1}{\partial x^3}dx^3\wedge dx^1+\\&\frac{\partial\alpha_2}{\partial x^1}dx^1\wedge dx^2+0-\frac{\partial\alpha_2}{\partial x^3}dx^2\wedge dx^3+\\&-\frac{\partial\alpha_3}{\partial x^1}dx^3\wedge dx^1+\frac{\partial\alpha_3}{\partial x^2}dx^2\wedge dx^3+0\end{aligned}$$

which further reduces to

$$d(\boldsymbol{\alpha})=\left(\frac{\partial\alpha_3}{\partial x^2}-\frac{\partial\alpha_2}{\partial x^3}\right)dx^2\wedge dx^3+\left(\frac{\partial\alpha_1}{\partial x^3}-\frac{\partial\alpha_3}{\partial x^1}\right)dx^3\wedge dx^1+\left(\frac{\partial\alpha_2}{\partial x^1}-\frac{\partial\alpha_1}{\partial x^2}\right)dx^1\wedge dx^2$$

This should look quite similar to something else that we worked with extensively in physics and multivariable calculus. Recall that the *curl* of a vector valued function $\mathbf{F}$ is given by

$$\nabla\times\mathbf{F}=\left\langle\left(\frac{\partial F^3}{\partial x^2}-\frac{\partial F^2}{\partial x^3}\right),\left(\frac{\partial F^1}{\partial x^3}-\frac{\partial F^3}{\partial x^1}\right),\left(\frac{\partial F^2}{\partial x^1}-\frac{\partial F^1}{\partial x^2}\right)\right\rangle$$

Thus,

$$\nabla\times\mathbf{F}=\left(\star d\mathbf{F}^\flat\right)^\sharp$$

Since we applied several transformations, let's break it down. We applied the $\flat$ operator to $\mathbf{F}$ to make it a covector. We then took the exterior derivative. However, at this point, our result is an element of $\Lambda^2(T_p^*M)$. We apply the Hodge star to convert this oriented plane segment to its corresponding covector. We then apply the $\sharp$ operator to make our output a vector.

The exterior derivative of a 0-form gives us the gradient. The exterior derivative of a 1-form gives us the curl. We now take the exterior derivative of a 2-form $\boldsymbol{\beta}$.

$$\begin{aligned}d(\boldsymbol{\beta})&=d(\beta_1dx^2\wedge dx^3+\beta_2dx^3\wedge dx^1+\beta_3dx^1\wedge dx^2)\\&=d(\beta_1dx^2\wedge dx^3)+d(\beta_2dx^3\wedge dx^1)+d(\beta_3dx^1\wedge dx^2)\\&=d(\beta_1)\wedge dx^2\wedge dx^3+d(\beta_2)\wedge dx^3\wedge dx^1+d(\beta_3)\wedge dx^1\wedge dx^2\\&=\frac{\partial\beta_1}{\partial x^1}dx^1\wedge dx^2\wedge dx^3+\frac{\partial\beta_2}{\partial x^2}dx^1\wedge dx^2\wedge dx^3+\frac{\partial\beta_3}{\partial d}x^1\wedge dx^2\wedge dx^3\\&=\left(\frac{\partial\beta_1}{\partial x^1}+\frac{\partial\beta_2}{\partial x^2}+\frac{\partial\beta_3}{\partial x^3}\right)dx^1\wedge dx^2\wedge dx^3\end{aligned}$$

We go from the second line to the third line by using the fact that $d^2 = 0$. We go from the third line to the fourth line by using the fact that $dx^i \wedge dx^i = 0$. Again, this should look familiar. Recall that the *divergence* of a vector field $\mathbf{F}$ is given by

$$\nabla\cdot\mathbf{F}=\frac{\partial F^1}{\partial x^1}+\frac{\partial F^2}{\partial x^2}+\frac{\partial F^3}{\partial x^3}$$

Thus,

$$\nabla \cdot \mathbf{F} = \star d \star \mathbf{F}^\flat$$

Let's break this one down. First, we use the $\flat$ transformation to change $\mathbf{F}$ into a covector. We then use the Hodge star to turn our result into a 2-form. Next, we take the exterior derivative. To remove the $dx^1 \wedge dx^2 \wedge dx^3$ term, we take the Hodge star again, giving us a scalar.

12 Generalized Stokes' Theorem

We tie everything together in this last section on generalized Stokes' theorem. Let M be a manifold and ω be an element of the exterior algebra $\Lambda(M)$. Stokes' theorem states that

$$\int_{\partial M} \omega = \int_M d\omega$$

where ∂M denotes the *boundary* of M, and $d\omega$ refers to the exterior derivative of the differential form ω. While we won't prove the theorem, we will apply the results we obtained in the previous section to show that this theorem gives us the fundamental theorem of calculus, classical Stokes' theorem, and divergence theorem.

Let's first consider the simple case where M is a one dimensional interval and ω is a 0-form. Note that this is the *only* possible scenario as in this case

$$\Lambda(T_p^*M) = \Lambda^0(T_p^*M) \oplus \Lambda^1(T_p^*M)$$

since $n = 1$, as M is a one dimensional interval. Thus, the only forms that can be differentiated are 0 forms. In this case, the boundary ∂M consists only of the endpoints of the interval M. However, the two boundary points are *oriented* in opposite directions. Thus, integrating ω across the boundary is equivalent to computing the difference of ω at the end points. Since ω is a 0-form, its exterior derivative $d\omega$ is its differential. Thus, this is exactly the fundamental theorem of calculus.

Now let's consider the case where M is a two dimensional manifold. We now consider differential forms ω that belong to $\Lambda^1(T_p^*M)$. Recall that classical Stokes' theorem states that

$$\oint_{\partial M} \mathbf{F} \cdot d\mathbf{r} = \iint_M \nabla \times \mathbf{F} \cdot d\mathbf{S}$$

Let $\omega = \mathbf{F} \cdot d\mathbf{r}$. We note that ω is simply a 1-form

$$\omega = \mathbf{F} \cdot d\mathbf{r} = F_1 dx^1 + F_2 dx^2 + F_3 dx^3$$

In the previous section we showed that

$$d\omega = (\nabla \times \mathbf{F})_1 dx^2 \wedge dx^3 + (\nabla \times \mathbf{F})_2 dx^3 \wedge dx^1 + (\nabla \times \mathbf{F})_3 dx^1 \wedge dx^2 = \nabla \times \mathbf{F} \cdot d\mathbf{S}$$

Thus, we get that this is a special case of Stokes' theorem.

We repeat the same process to get divergence theorem. Let M be a 3-dimensional manifold and ω belong to $\Lambda^2(T_p^*M)$. Divergence theorem states that

$$\oiint_{\partial M} \mathbf{F} \cdot d\mathbf{S} = \iiint_M \nabla \cdot \mathbf{F} \;\; dV$$

Just as before, $\omega = \mathbf{F} \cdot d\mathbf{S}$ is an arbitrary 2-form. In the previous section, we showed that

$$d\omega = (\nabla \cdot \mathbf{F}) dx^1 \wedge dx^2 \wedge dx^3 = (\nabla \cdot \mathbf{F}) dV$$

Thus, we again see that divergence theorem is a special case of Stokes' theorem.

References

[1] California Institute of Technology Course 314 Professors. Discrete differential geometry, 2011. http://brickisland.net/cs177/?p=200.

[2] Jonathan H Manton. On the role of differential geometry in signal processing. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 5, pages v–1021. IEEE, 2005.

[3] Peter James Thomas. Making connections, 2016. https://peterjamesthomas.com/glimpses-of-symmetry/chapter-19-making-connections/.

[4] RInterested. 1-forms, covectors and gradient. http://rinterested.github.io/statistics/one_forms_as_gradients.html.

[5] Wikipedia contributors. Differential geometry — Wikipedia, the free encyclopedia, 2019.

[6] Stephen Wolfram. *A new kind of science*, volume 5. Wolfram media Champaign, IL, 2002.