1 Linear Functions

1.1 of Real Numbers

The key concept behind differential geometry is the idea of homogeneous linearity. For a real-valued function $f$ with one real argument, *homogeneous linearity over the reals* is defined by the requirements

$$f(\lambda x) = \lambda f(x)$$
$$f(x + y) = f(x) + f(y)$$

for any real numbers $\lambda, x, y$. To save space, I will often abbreviate homogeneous linearity to just ‘linearity’.

1.2 of Complex Numbers

Homogeneous linearity over the complex numbers is defined in the same way as for real numbers except that the function $f$ and the numbers $\lambda, x, y$ are allowed to be complex numbers. There is, however, a new possibility because each complex number $x = a + ib$ has a complex conjugate partner, $x^* = a - ib$. For a complex-valued function $g$ of one complex argument, *homogeneous anti-linearity* is defined by the requirements

$$g(\lambda x) = \lambda^* g(x)$$
$$g(x + y) = g(x) + g(y).$$

1.3 of Vectors

As you will see from the first homework assignment, linear functions of one real or complex number are pretty trivial. What about functions whose arguments or values are not single real or complex numbers? To define linearity for such functions, it must be possible to add their arguments or values together and to multiply them by real numbers in a consistent way. Such arguments and values belong to vector spaces.

A *vector space* over the real numbers (or over the complex numbers) is defined to be a set $V$ with an addition operation and a scalar multiplication operation with the usual associative and commutative properties and a unique zero element. The addition operation $+$ takes two elements $x$ and $y$ of $V$ and forms a new object $x + y$ that also belongs to $V$. Similarly, the scalar multiplication operation $\ast$ takes a real (or complex) number $a$ and an element $x$ of $V$ and forms an object $a \ast x$ that also belongs to $V$. In the usual terse notation of mathematics we would describe these operations as mappings:

$$+ : V \times V \to V$$
$$\ast : \mathbb{R} \times V \to V$$
The usual associative and commutative properties for real (or complex) numbers $a, b$ and vectors $x, y, z$ are (omitting the * as is the usual convention)

\[
(a + b) x = ax + bx \\
a(bx) = (ab)x \\
x + (y + z) = (x + y) + z \\
x + y = y + x \\
a(x + y) = ax + ay
\]

and the zero element 0 is a member of $V$ such that

\[
x + 0 = x \\
a0 = 0
\]

for any real (or complex) number $a$ and vector $x$.

A real-valued linear function defined on a vector space $V$ is called a linear form over $V$. The set of linear forms over $V$ is itself a vector space, called the space $^*V$ dual to $V$.

### 1.4 Isomorphisms

The definition of a linear function works perfectly well if the function assigns vectors in a vector space rather than just real numbers. A vector-valued linear function on a vector space is best thought of as a mapping $f : V \rightarrow U$ of a vector space $V$ to a vector space $U$. If the mapping is one-to-one, it is called an isomorphism. Isomorphisms give us a convenient way to compare different vector spaces. If the set of possible values $f(V)$ is actually equal to $U$, then we say that $U$ and $V$ are isomorphic. In such a case, the spaces $U$ and $V$ are interchangeable copies of each other and we can use the function $f$ and its inverse $f^{-1}$ to go back and forth between them.

### 1.5 Examples of Vector Spaces

#### 1.5.1 Real n-tuples

The set $\mathbb{R}$ of real numbers is obviously a vector space with + and * just the usual operations of addition and multiplication.

The set $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ of number pairs such as $(c, d)$ is a vector space with the definitions:

\[
a \ast (c, d) = (ac, ad) \\
(a, b) + (c, d) = (a + c, b + d)
\]

The zero vector in this case is the pair $(0, 0)$. The set $\mathbb{C}$ of complex numbers is also an example of a vector space.

Similarly, the set $\mathbb{R}^n$ of real-number n-tuples with the corresponding definitions of addition and multiplication by a scalar is also a vector space.
An even larger, but still manageable vector space is the set $\mathbb{R}^\infty$ of real-number infinite sequences such as \( \{a_0, a_1, a_2, \ldots\} \), which can be multiplied by a scalar and added to one another in the obvious way. This space is an example of a vector space with a countable basis.

### 1.5.2 Function spaces

An example of a vector space that is much larger than $\mathbb{R}^\infty$ is the set $\mathbb{R}^\mathbb{R}$ of real-valued functions of real numbers. Functions $f$ and $g$ can be added to give a function $f + g$ defined by

\[
(f + g)(x) = f(x) + g(x)
\]

for all real numbers $x$ and a function $f$ can be multiplied by a real number $a$ to give a function $af$ defined in the obvious way by

\[
(af)(x) = af(x)
\]

The space $\mathbb{R}^\mathbb{R}$ is extremely large – larger than we ever need in physics. In classical physics, we pick out subspaces of $\mathbb{R}^\mathbb{R}$ that consists of functions with finite derivatives up to some order.

In quantum theory, the state of a particle that is moving in one dimension can be described by a complex-valued wave function $\psi$. The set of all such functions is the vector space $\mathbb{C}^\mathbb{R}$. That space is also much larger than we need in physics, so we pick out a subspace, $L_2(\mathbb{C}^\mathbb{R})$ of square-integrable functions to represent physical states.

### 1.5.3 Directional Derivatives as Vectors

Because differential operators can be added and multiplied by constants, they can be regarded as vectors. Consider the set $T_{(0,0)}$ of directional derivatives of functions on $\mathbb{R}^2$ at the point $(0,0)$.

The operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ act on a function $f(x, y)$ to produce its partial derivatives $\frac{\partial f}{\partial x}\bigg|_{x,y=0}$ and $\frac{\partial f}{\partial y}\bigg|_{x,y=0}$. A linear combination of these operators such
as \( v = 2 \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \) acts on a function \( f \) to produce

\[
vf = 2 \left. \frac{\partial f}{\partial x} \right|_{x=y=0} + \left. \frac{\partial f}{\partial y} \right|_{x=y=0}
\]

These operators are direction derivatives. The vector space of direction derivatives at the point \((0,0)\) is called the tangent space at \((0,0)\). There is a separate tangent space for each point in \( \mathbb{R}^2 \). Moving the evaluation point to \((1.0,1.2)\) would produce the tangent space \( T_{(1.0,1.2)} \).

A few notational conventions for operators such as the directional derivative \( v \) should be recalled. First of all, when parentheses are omitted, they are always assumed to be nested from right to left as in the expression

\[
uvgwf = u(v(g(w(f)))).
\]

where \( u, v, w \) are directional derivatives and \( f, g \) are functions. Thus, each operator acts on everything to the right of it. Second, one can use “single-term parentheses” to turn off this right-action as in the expression

\[
u(vg)wf
\]

where \( v \) acts on \( g \) to produce the function \( vg \) and does nothing else. Third, a relation between operators always assumes that the operators are acting on arbitrary functions, even if those functions are not explicitly shown as in the relation

\[
v = e_1v^1 + e_2v^2
\]

which really means

\[
vf = (e_1v^1 + e_2v^2)f = e_1v^1f + e_2v^2f
\]

for any function \( f \). In this last example, notice that the parenthesis does not turn off the right action of the operators \( e_1, e_2 \) because it includes more than one term. An example that uses both multiple and single-term parentheses would be

\[
(e_1v^1 + e_2v^2)f = e_1v^1f + e_2v^2f
\]

\[
= (e_1v^1)f + v^1e_1f + (e_2v^2)f + v^2e_2f
\]

where the product rule is applied to show \( e_1, e_2 \) acting on each factor separately.

### 1.5.4 Differential forms

The space \( \hat{T}_{(0,0)} \) of linear forms over the tangent space \( T_{(0,0)} \) is the set of real-valued linear functions of directional derivatives. One way to produce such a function is to let each vector \( v \) act on a particular function \( f \). The resulting linear form is called \( df \) and is defined by

\[
df(v) = (vf)|_{(0,0)}.
\]
Some alternative notations for the result of $df$ acting on the vector $v$ are:

$$df (v) = df \cdot (v) = df \cdot v = v \cdot df.$$  

One also sometimes sees the notation

$$df (v) = \langle df, v \rangle.$$  

Corresponding to the coordinate functions $x$ and $y$ there are linear forms $dx$ and $dy$. These objects are called differential forms and will be discussed at length later.

# 2 Basis Vectors

## 2.1 Linear Independence and Dimensionality

A set $\{v_1, v_2, v_3, ..., v_n\}$ of vectors is said to be linearly independent if the equation

$$a_1v_1 + a_2v_2 + a_3v_3 + ... + a_nv_n = 0$$

implies that all of the coefficients $a_i$ must be zero. The maximum number of linearly independent vectors in a space is called the dimensionality of the space.

## 2.2 The Basis Representation

A set of $n$ linearly independent vectors $\{e_1, e_2, e_3, ..., e_n\}$ in an $n$-dimensional vector space $V$ generates an isomorphism that connects $\mathbb{R}^n$ to $V$. Denote the array of coefficients by the row vector

$$[v] = \begin{bmatrix} v^1 & v^2 & v^3 & \cdots & v^n \end{bmatrix}$$

and define the isomorphism $e : \mathbb{R}^n \to V$ by the matrix product:

$$e ([v]) = [v] [e]$$

where the basis vectors are presented as a column array

$$[e] = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

The vector $v$ is then represented as $v = e ([v])$. 

5
To gain familiarity with this mixture of vector and matrix notation, write the expression out:

\[
v = e([v]) = [v] [e] = \begin{bmatrix} v^1 & v^2 & v^3 & \cdots & v^n \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix} = v^1 e_1 + v^2 e_2 + v^3 e_3 + \cdots + v^n e_n
\]

Because the map \( e \) is an isomorphism, it can be inverted to yield a unique map \( e^{-1} : V \rightarrow \mathbb{R}^n \). Thus, the array of expansion coefficients is given by \([v] = e^{-1}(v)\). It is useful to define the column vectors

\[
W^1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad W^2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad W^n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}
\]

and use them to extract the individual expansion coefficients.

\[
v^r = [v] W^r = e^{-1}(v) W^r
\]

For example,

\[
v^2 = \begin{bmatrix} v^1 & v^2 & v^3 & \cdots & v^n \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = v^2.
\]

### 3 Dual Basis Forms

#### 3.1 Abstract definition

For each value of the index, \( r \), the function

\[
\omega^r(v) := e^{-1}(v) W^r
\]

is a real-valued linear function on the vector space \( V \). Thus, it is a linear form over \( V \) and belongs to the dual space \( \text{\textdual V} \). The expansion coefficients for a vector \( v \) are then given by the expressions

\[
v^r = \omega^r(v) \quad (1)
\]

Using these expansion coefficients, any vector \( v \) has the expansion:

\[
v = \sum_{r=1}^n \omega^r(v) e_r \quad (2)
\]
Now suppose that $\alpha$ belongs to $\hat{V}$. If we know the value $\alpha(v)$ that $\alpha$ assigns to any vector $v$, then we know everything there is to know about $\alpha$. From the expansion above,

$$\alpha(v) = \alpha \left( \sum_{r=1}^{n} \omega^r(v) e_r \right) = \sum_{r=1}^{n} \alpha(\omega^r(v)) e_r = \sum_{r=1}^{n} \omega^r(v) \alpha(e_r)$$

so that if we know the numbers

$$\alpha_r = \alpha(e_r)$$

then we can find $\alpha(v)$ for any vector $v$.

$$\alpha(v) = \sum_{r=1}^{n} \alpha_r \omega^r(v)$$

Thus, we have found an expansion for the arbitrary linear form $\alpha$.

$$\alpha = \sum_{r=1}^{n} \alpha_r \omega^r$$

The linear forms $\omega^r$ are therefore a basis for the dual space $\hat{V}$ and are called the basis dual to $e$.

### 3.2 Dot and matrix notations

Some alternative notations for the relations between forms and their components and vectors and their components may be helpful.

$$v^T = v \cdot \omega$$

$$\alpha_r = e_r \cdot \alpha$$

In words, one finds components by dotting things into the dual basis objects. Vector components are found by dotting vectors into the form basis and form components are found by dotting forms into the vector basis. In terms of matrices, one can define the column vector of components of the form $\alpha$ as

$$[\alpha] = [e] \cdot \alpha$$

and the row vector of components of the vector $v$ as

$$[v] = v \cdot [\omega]$$

where $[\omega]$ is a row matrix whose elements are the forms $\omega^r$. In terms of these matrices, the expansions of vectors and forms becomes

$$v = [v] [e] = (v \cdot [\omega])[e]$$

and

$$\alpha = [\omega] [\alpha] = [\omega] ([e] \cdot \alpha)$$
3.3 Useful definition of dual basis forms

In general, we do not know the inverse map $e^{-1}$ even though we do know that it has to exist. Thus the formula defining the dual basis forms in terms of $e^{-1}$ is reassuring but not of much practical use. For a more useful definition, use the expansion of an arbitrary vector to expand one of the basis vectors:

$$e_s = \sum_{r=1}^{n} \omega^r (e_s) e_r$$

and note that the linear independence of the basis vectors means that

$$\omega^r (e_s) = \delta_s^r = \begin{cases} 0; r \neq s \\ 1; r = s \end{cases} \quad (5)$$

An alternative notation for this expression would be

$$e_s \cdot \omega^r = \delta_s^r$$

and, in terms of the column matrix $[e]$ and the row matrix $[\omega]$

$$[e] \cdot [\omega] = [1]$$

where $[1]$ is the identity $n \times n$ matrix. Written out, the matrix operation looks like this:

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \cdot \begin{bmatrix} \omega^1 \\ \omega^2 \\ \vdots \\ \omega^n \end{bmatrix} = \begin{bmatrix} \omega^1 \cdot e_1 & \omega^1 \cdot e_2 & \cdots & \omega^1 \cdot e_n \\ \omega^2 \cdot e_1 & \omega^2 \cdot e_2 & \cdots & \omega^2 \cdot e_n \\ \vdots & \vdots & \ddots & \vdots \\ \omega^n \cdot e_1 & \omega^n \cdot e_2 & \cdots & \omega^n \cdot e_n \end{bmatrix}$$

so that the relation is

$$\begin{bmatrix} \omega^1 \cdot e_1 & \omega^1 \cdot e_2 & \cdots & \omega^1 \cdot e_n \\ \omega^2 \cdot e_1 & \omega^2 \cdot e_2 & \cdots & \omega^2 \cdot e_n \\ \vdots & \vdots & \ddots & \vdots \\ \omega^n \cdot e_1 & \omega^n \cdot e_2 & \cdots & \omega^n \cdot e_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

4 Names and Notations

4.1 Covariant and contravariant vectors

Notice that there are automatically the same number of dual basis forms as there are dual basis vectors. As a result, so long as $V$ is finite dimensional, the dual space $\tilde{V}$ is isomorphic to $V$. It is therefore possible to think of the forms as ‘other kinds of vectors’ and older treatments of tensor analysis do exactly that. In the ancient texts of geometrical physics, the elements of $V$ are called ‘contravariant vectors’ and the elements of $\tilde{V}$ are called ‘covariant vectors’.
4.2 The dual of the dual space and ‘dot products’

The elements of \( \hat{V} \) are functions that assign real (or complex) numbers to forms in the dual space \( \hat{V} \). Some of these functions are easy to specify. Given a vector \( v \) in \( V \) we can assign a number to each form \( \alpha \), namely the number \( \alpha(v) \). Thus, we define the ‘form on forms’ \( \hat{v} \) that is associated with the vector \( v \) by

\[
\hat{v}(\alpha) := \alpha(v)
\]

Instead of keeping track of two different objects, \( \hat{v} \) and \( v \) it is customary to think of one object that happens to have two different actions and define the action of \( v \) on a form by

\[
v(\alpha) := \alpha(v).
\] (6)

The symmetrical relationship between forms and vectors is often emphasized by using a different notation.

\[
\alpha \cdot v = v \cdot \alpha := \alpha(v)
\]

When we wish to emphasize that forms and vectors are equally fundamental but wish to keep them separate, we use yet another notation

\[
(\alpha, v) = \alpha(v)
\]

where the form is always placed to the left of the comma and the vector to the right.

4.3 Components and sum conventions

From Equation 1 for the components of a vector, and Equation 6 which defines the action of a vector on a form, the components of a vector can be found by evaluating the vector acting on the basis forms

\[
v^r = v(\omega^r)
\]

in just the same way that Equation 3 indicates that the components of a form can be found by evaluating the form acting on the basis vectors. As you discovered when you did the homework, each expansion of a vector or a form introduces a new index and a new summation sign. Not only do you have to take the time to write each summation sign, but you need to explicitly change the order of the summation signs even though that order never really matters. The Einstein summation convention takes advantage of the fact that every sum that arises from forms, basis vectors, and related objects can be summarized by a simple rule: Sum over any index that is repeated once as a subscript and once as a superscript. The rule works only if you are careful to never use the same index for different sums. For example, instead of writing

\[
v = \sum_r v^r e_r, \quad \text{and} \quad \alpha = \sum_r \alpha_r \omega^r
\]
we just write
\[ v = v^r e_r \quad \text{and} \quad \alpha = \alpha_s \omega^s. \]

So long as you are careful to avoid coincident names for the summed indexes, the summation convention lets you ignore the summations entirely and just do algebra on a typical term of the sum. Be sure to try this technique in Homework 04. We will return to this idea when we get to tensors, where the number of indexes and summations can become quite large.

### 4.4 Matrices and automatic symbol manipulation

The use of matrix multiplication to replace sums and indexes has a long tradition in differential geometry, going back at least to the work of the mathematician Élie Cartan in 1928. It yields simple and elegant expressions that can usually be written in just one way. A practical advantage is that it avoids the need to decorate symbols with subscripts and superscripts and avoids the arbitrariness of naming dummy indexes. A practical disadvantage is that computations with specific examples require one to write out all of the terms of each matrix in each expression. Another practical disadvantage is that the manual procedure of matching up the row and column entries that are to be multiplied together is both tedious and error-prone. These disadvantages have usually led physicists to revert to index notation to organize their calculations.

A key feature of matrix expressions is that they obey the associative rule even when they are mixed with dot products. Parentheses are never needed in these expressions. For example, there is no ambiguity about the meaning of the expressions \( v \cdot [\omega] [e] \) and \([\omega] [e] \cdot \alpha \) once one knows that \( v \) is a vector and \( \alpha \) is a form. Thus, one can regroup these expressions and define an object
\[ Id = [\omega] [e] \]
that acts as an identity operation on both the space \( V \) of vectors and the space \( \hat{V} \) of forms according to
\[ v = v \cdot Id \]
\[ \alpha = Id \cdot \alpha \]

Some of the practical disadvantages of using matrix notation in actual calculations are now less important than they once were. Consider the dot product of a vector
\[ v = y \partial_x + 2z \partial_y + 3x \partial_z \]
and a form
\[ \alpha = y dx + z dy + x dz. \]
A direct calculation of \( v \cdot \alpha \) is straightforward but tedious since one must consider nine possible terms. The index expression \( v^r \alpha_r \) is better but requires the additional step of writing out six definitions for \( v^1, v^2, v^3, \alpha_1, \alpha_2, \alpha_3 \) and then
plugging those definitions into the resulting sum of three terms. The matrix version of the calculation

\[
v \cdot \alpha = [v] \alpha = \begin{bmatrix} y & 2z & 3x \end{bmatrix} \begin{bmatrix} y \\ z \\ x \end{bmatrix}
\]

can be completed in an instant (if you are using Scientific Notebook or Scientific Workplace) by marking the matrix product with your mouse and hitting CTRL-E.

You will find that symbol manipulation programs are confused by superscripts (which get interpreted as powers by default) and by multiple subscripts. There are special packages for most symbol manipulation programs that let them handle such notation but those packages are usually a step or two removed from the process of entering and editing expressions. Thus the avoidance of indexes is a strong advantage of matrix notation in a symbol manipulating environment such as we are using here.

### 4.5 The interior trace operation

Symbol manipulation routines are also confused by operator symbols such as the dot that we have been using to indicate the action of forms on vectors or *vice versa*. For example, try evaluating the expression

\[
\begin{bmatrix} \partial_x \\ \partial_y \\ \partial_z \end{bmatrix} \cdot \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}
\]

within the SN or SW environment and you will find that the software performs a hermitean transpose on the second array and then dots the two arrays together which is not what we want here. A work-around is to introduce the *interior trace* operation \( \mathcal{T} \) which instructs one to let each vector act on the form (if any) that follows it. Evaluating the expression

\[
\mathcal{T} \begin{bmatrix} \partial_x \\ \partial_y \\ \partial_z \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}
\]

yields the matrix

\[
\begin{bmatrix} \mathcal{T}\partial_x dx & \mathcal{T}\partial_y dy & \mathcal{T}\partial_z dz \\ \mathcal{T}\partial_y dx & \mathcal{T}\partial_y dy & \mathcal{T}\partial_y dz \\ \mathcal{T}\partial_z dx & \mathcal{T}\partial_z dy & \mathcal{T}\partial_z dz \end{bmatrix}
\]

which is interpreted as exactly what we want:

\[
\begin{bmatrix} \partial_x \cdot dx & \partial_x \cdot dy & \partial_x \cdot dz \\ \partial_y \cdot dx & \partial_y \cdot dy & \partial_y \cdot dz \\ \partial_z \cdot dx & \partial_z \cdot dy & \partial_z \cdot dz \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]
Using this notation, we can write the identities

\[ v = \mathcal{F}v Id = \mathcal{F}v [\omega] [e] \]
\[ \alpha = \mathcal{F}Id \alpha = \mathcal{F}v [\omega] [e] \alpha \]

The usual matrix trace operation may also be used in these expressions. For example, it is easy to see that, for any vector \( v \) and form \( \alpha \) we have all of the following notations for the dot product:

\[ v \cdot \alpha = \mathcal{F}v \alpha = \text{Tr} (v \alpha) = [v] [\alpha]. \]