

Appendix B. Vector Spaces

Throughout this text we have noted that various objects of interest form a *vector space*. Here we outline the basic structure of a vector space. You may find it useful to refer to this Appendix when you encounter this concept in the text.

B.1. What are vector spaces?

In physics and engineering one of the first mathematical concepts that gets introduced is the concept of a vector. Usually, a vector is defined as a quantity that has a direction and a magnitude, such as a position vector, velocity vector, acceleration vector, *etc.* However, the notion of a vector has a considerably wider realm of applicability than these examples might suggest. The set of all real numbers forms a vector space, as does the set of all complex numbers. The set of functions on a set (*e.g.*, functions of one variable, $f(x)$) form a vector space. Solutions of linear homogeneous equations form a vector space. There are many more examples, some of these are highlighted in the text. We begin by giving the abstract rules for forming a space of vectors, also known as a vector space.

A *vector space* V is a set equipped with an operation of “addition” and an additive identity. The elements of the set are called vectors, which we shall denote as \vec{u} , \vec{v} , \vec{w} , *etc.* For now, you can think of them as position vectors in order to keep yourself sane. *Addition*, is an operation in which two vectors, say \vec{u} and \vec{v} , can be combined to make another vector, say, \vec{w} . We denote this operation by the symbol “+”:

$$\vec{u} + \vec{v} = \vec{w}. \quad (\text{B.1})$$

Do not be fooled by this simple notation. The “addition” of vectors may be quite a different operation than ordinary arithmetic addition. For example, if we view position vectors in the x - y plane as “arrows” drawn from the origin, the addition of vectors is defined by the parallelogram rule. Clearly this rule is quite different than ordinary “addition”. In general, any operation can be used to define addition if it has the commutative and associative properties:

$$\vec{v} + \vec{w} = \vec{w} + \vec{v}, \quad (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}), \quad (\text{B.2})$$

The requirement of an additive identity means that there exist an element of V , called the zero vector and denoted by $\vec{0}$, such that for any element $\vec{v} \in V$,

$$\vec{v} + \vec{0} = \vec{v}. \quad (\text{B.3})$$

As an *exercise* you can check that the set of position vectors relative to the origin in the x - y plane forms a vector space with (i) the vectors being viewed as arrows with the parallelogram rule for addition, and (ii) the position of the origin being the zero vector.

In applications to physics and engineering one is normally interested in a vector space with just a little more structure than what we defined above. This type of vector space has an additional operation, called *scalar multiplication*, which is defined using either real or complex numbers, called *scalars*. Scalars will be denoted by $a, b, c, \text{ etc.}$. When scalar multiplication is defined using real (complex) numbers for scalars, the resulting gadget is called a real (complex) vector space.* Scalar multiplication is an operation in which a scalar a and vector \vec{v} are combined to make a new vector, denoted by $a\vec{v}$. Returning to our example of position vectors in the plane, the scalar multiplication operation is defined by saying that the vector $a\vec{v}$ has the same direction as \vec{v} , *provided* $a \geq 0$, but the length of \vec{v} is scaled by the amount a . So, if $a = 2$ the vector is doubled in length, and so forth. If the scalar is negative, then the vector is reversed in direction, and its length is scaled by $|a|$. In general, any rule for scalar multiplication is allowed provided it satisfies the properties:

$$(a + b)\vec{v} = a\vec{v} + b\vec{v}, \quad a(b\vec{v}) = (ab)\vec{v}, \quad a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}, \quad 1\vec{v} = \vec{v}, \quad 0\vec{v} = \vec{0}. \quad (\text{B.4})$$

Again, you can check that the scalar multiplication we use for position vectors satisfies all these properties.

As an exercise, prove that the vector $-\vec{v}$, defined by

$$-\vec{v} = (-1)\vec{v} \quad (\text{B.5})$$

is an *additive inverse* of \vec{v} , that is,

$$\vec{v} + (-\vec{v}) = \vec{0}. \quad (\text{B.6})$$

We often use the notation

$$\vec{v} + (-\vec{w}) \equiv \vec{v} - \vec{w}, \quad (\text{B.7})$$

so that

$$\vec{w} - \vec{w} = \vec{0}. \quad (\text{B.8})$$

One of the most important features of a (real or complex) vector space is the existence of a *basis*. To define it, we first introduce the notion of linear independence. A subset of vectors $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_k)$ is *linearly independent* if no non-trivial *linear combination* of them vanishes, *i.e.*, a relation

$$a_1\vec{e}_1 + a_2\vec{e}_2 + \dots + a_k\vec{e}_k = \vec{0} \quad (\text{B.9})$$

exists between the elements of the set only if $a_1 = a_2 = \dots = a_k = 0$. If such a relation (B.9) exists, the subset is called *linearly dependent*. For example, if \vec{v} and \vec{w} are position

* One often gets lazy and calls a real/complex vector space just a “vector space”.

vectors, then they are linearly dependent if they have parallel or anti-parallel directions, *i.e.*, they are colinear. If they are not colinear, then they are linearly independent (exercise). Note that in a linearly dependent subset of vectors it will be possible to express some of the vectors as *linear combinations* of the others. In general, there will be a unique maximal size for sets of linearly independent vectors. If all sets with more than n vectors are linearly dependent, then we say that the vector space is n -dimensional, or has n dimensions. In this case, any set of n linearly independent vectors, say $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$, is said to form a *basis*.^{*} The utility of a basis is that every element of V can be uniquely expressed as a linear combination of the basis vectors:

$$\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2 + \dots + v^n\vec{e}_n. \quad (\text{B.10})$$

The scalars v^i , $i = 1, 2, \dots, n$ are called the *components* of \vec{v} in the basis \vec{e}_i . (Note that in expressions like v^1, v^2, v^3, \dots the superscripts are simply numerical labels — not exponents!) Thus a vector can be characterized by its components in a basis. As a nice exercise, you can check that, in a given basis, the components of the sum of two vectors \vec{v} and \vec{w} are the ordinary sums of the components of \vec{v} and \vec{w} :

$$(\vec{v} + \vec{w})^i = v^i + w^i. \quad (\text{B.11})$$

Likewise, you can check that the components of the scalar multiple $a\vec{v}$ are obtained by ordinary multiplication of each component of \vec{v} by the scalar a :

$$(a\vec{v})^i = av^i. \quad (\text{B.12})$$

Let us take a deep, relaxing breath and return to our running example, position vectors in the plane. As you know, in the x - y plane we can introduce a basis consisting of a (unit) vector \vec{e}_1 along the x direction and a (unit) vector \vec{e}_2 along the y direction. Every position vector can then be expressed as

$$\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2, \quad (\text{B.13})$$

where v^1 is the “ x -component” of \vec{v} (sometimes denoted by v^x) and v^2 is the “ y -component” of \vec{v} (sometimes denoted by v^y). Evidently, the set of position vectors in the plane is a 2-dimensional, real vector space.

B.2. Scalar Products

Often times we augment the properties of a vector space with an extra bit of structure called a *scalar product* (also known as an “inner product”). The scalar product is a way of

^{*} It can be shown that a vector space of dimension n admits infinitely many sets of basis vectors, but each basis will always consist of precisely n (linearly independent) vectors. Speaking a little loosely, if $n = \infty$ we say that the vector space is *infinite dimensional*.

making a scalar from a pair of vectors. We shall denote this scalar by (\vec{v}, \vec{w}) . The scalar product generalizes the familiar notion of a “dot product” of position vectors, velocity vectors, *etc.* Any rule for forming a scalar from a pair of vectors will be allowed as a scalar product provided it satisfies†

$$(\vec{v}, \vec{w}) = (\vec{w}, \vec{v}), \quad (a\vec{v} + b\vec{w}, \vec{u}) = a(\vec{v}, \vec{u}) + b(\vec{w}, \vec{u}), \quad (\vec{v}, \vec{v}) \geq 0, \quad (\text{B.14})$$

and $(\vec{v}, \vec{v}) = 0$ if and only if $\vec{v} = \vec{0}$. As a good exercise you can check that the dot product of position vectors,

$$(\vec{v}, \vec{w}) \equiv \vec{v} \cdot \vec{w}, \quad (\text{B.15})$$

which you certainly should have some experience with by now, provides an example of a scalar product.

Borrowing terminology from, say, position vectors in the plane, a pair of vectors \vec{v} and \vec{w} are called *orthogonal* if

$$(\vec{v}, \vec{w}) = 0. \quad (\text{B.16})$$

If two vectors are orthogonal, then they are linearly independent. The converse is not true, however (*exercise*). Likewise we define the *length* or *norm* of a vector, $\|\vec{v}\|$ by

$$\|\vec{v}\| = \sqrt{(\vec{v}, \vec{v})}. \quad (\text{B.17})$$

We say that a vector is *normalized* if it has unit length. Any vector can be normalized by scalar multiplication (*exercise*). A basis \vec{e}_i , $i = 1, 2, \dots, n$ is called *orthonormal* if it has the following scalar products among its elements:

$$(\vec{e}_i, \vec{e}_j) = \delta_{ij}, \quad i, j = 1, 2, \dots, n. \quad (\text{B.18})$$

Here we have used the very convenient symbol δ_{ij} , known as the *Kronecker delta*, by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \quad (\text{B.19})$$

If \vec{e}_i , $i = 1, 2, \dots, n$ is an orthonormal basis, you can verify as an exercise that the components v^i of a vector \vec{v} ,

$$\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2 + \dots + v^n\vec{e}_n, \quad (\text{B.20})$$

can be computed by

$$v^i = (\vec{e}_i, \vec{v}). \quad (\text{B.21})$$

† Here, for simplicity, we restrict ourselves to a real vector space, where the scalars are always real numbers.

B.3. Linear Operators

An extremely useful notion that is employed in the context of vector spaces is that of a *linear operator* (sometimes just called an “operator”). Given a (real or complex) vector space V , a linear operator denoted, say, by A , is

- (i) an *operator*: it is a rule which assigns to every vector $\vec{v} \in V$ another (possibly the same) vector, denoted $A\vec{v}$.
- (ii) *linear*: for any two vectors \vec{v} and \vec{w} and any two scalars a and b ,

$$A(a\vec{v} + b\vec{w}) = aA\vec{v} + bA\vec{w}. \quad (\text{B.22})$$

This latter requirement can be viewed as insisting that the operator is compatible with the structure of the set V as a vector space.

We encounter several examples of linear operators in the text. All the differential operators appearing in the linear differential equations are linear operators. The matrices featuring in the coupled oscillator discussion define linear operators. As a simple example of a linear operator, consider the real vector space of position vectors in the x - y plane. We can define a linear operator by the rule that takes any vector and rotates it 10° clockwise about the origin. To see that this rule is linear is a very nice exercise — try it! As another exercise: Which vector(s) are left unchanged by this rule?

You may have encountered linear operators only in the context of matrices, matrix multiplication, *etc.* While the idea of a linear operator is somewhat more general than that of a matrix, for many applications one need only use matrices. This is because every linear operator on a finite dimensional vector space can be viewed as a matrix acting upon column vectors. The matrix and the column vectors are defined relative to a choice of basis.

B.4. Eigenvalue problems

Given a linear operator, A , one is often interested in its *eigenvalues* and *eigenvectors*. The eigenvalue is a scalar and the eigenvector is a vector; if we denote them by λ and \vec{e}_λ , respectively, they are solutions to

$$A\vec{e}_\lambda = \lambda\vec{e}_\lambda. \quad (\text{B.23})$$

It is a fundamental result of linear algebra (for finite dimensional vector spaces) that the eigenvectors and eigenvalues (over the complex numbers) completely characterize the linear operator A . Indeed, the word “eigen” in German means “inherent” or “characteristic”.

Assuming A is a linear operator on a finite dimensional vector space, one solves the *eigenvalue problem* (B.23) as follows. Let the vector space have dimension n . In a given

basis, the operator A is represented by a square $n \times n$ matrix, which we shall also denote by A , while the eigenvector e_λ will be represented by a column vector with n rows. It is a fundamental theorem of linear algebra that (B.23) has a solution if and only if there exists a solution λ to

$$\det(A - \lambda I) = 0, \tag{B.24}$$

where I is the identity matrix. This is the *characteristic equation* defined by A ; it says that λ is the root of an n^{th} -order polynomial. If we are working with complex numbers, then (B.24) always has a solution. But, as is often the case, if we are working with real numbers there may be no solution to (B.24), and hence there will be no eigenvalues/eigenvectors. Given a solution λ to (B.24), one can substitute it into (B.23) and solve for the eigenvec-tor(s) \vec{e}_λ corresponding to λ . Note that if \vec{e}_λ is a solution to (B.23) for some λ , then so is any scalar multiple of \vec{e}_λ (*exercise*). This means that (given a scalar product) one can always normalize the eigenvectors to have unit length.

Still assuming the real vector space of interest is finite-dimensional (so that, for exam-ple, our discussion need not apply to linear differential operators acting on vector spaces of functions) the solution to the eigenvalue problem for *symmetric* operators has very special properties. Recall that a symmetric operator has a matrix which satisfies $A^T = A$, where the superscript “T” means “transpose” (interchange rows and columns). For example, a 2×2 matrix M is symmetric if and only if it takes the form

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Symmetric operators always have real eigenvalues, *i.e.*, (B.24) always has solutions even when we are restricting to real numbers only. More significantly, the eigenvectors always form a basis. Moreover, eigenvectors belonging to two distinct eigenvalues are always orthogonal. With a little work, it can be shown that the eigenvectors of a symmetric operator can be chosen to form an orthonormal basis.