

Vectors

By now you have a pretty good experience with “vectors”. 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. 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 “+”:

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

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:

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

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 $\vec{v} \in V$,

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

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 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 , *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

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

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:

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

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

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

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

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

We often use the notation

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

so that

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

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

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

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

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

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} :

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

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 :

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

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

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

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.

Dual vectors

Given a vector space V of dimension n , there is also defined a closely related vector space, also n -dimensional, called the *dual vector space*, denoted by V^* whose elements are called *dual vectors*. The dual vector space is the set of all *linear functions on V* . The elements of the space will (at least for now) be denoted with underlined Greek letters; the value of the linear function $\underline{\alpha}$ on the vector \vec{v} is a scalar; it will be denoted by $\underline{\alpha}(\vec{v})$. The linearity requirement is very important; it means that for any scalars a, b and vectors \vec{v}, \vec{w}

$$\underline{\alpha}(a\vec{v} + b\vec{w}) = a\underline{\alpha}(\vec{v}) + b\underline{\alpha}(\vec{w}).$$

^{*} 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*.

As a simple example, let us just consider a two-dimensional vector space with a basis $\{\vec{e}_1, \vec{e}_2\}$. Let $\underline{\omega}^1$ be the function which takes any vector $\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2$ and returns its first component:

$$\underline{\omega}^1(\vec{v}) = v^1.$$

You should check (*Exercise*) that the linear function $\underline{\omega}^1$ so-defined is indeed linear.

The general form of a linear function $\underline{\alpha}$ of a vector \vec{v} is a linear combination of its components:

$$\underline{\alpha}(\vec{v}) = \alpha_a v^a.$$

Evidently, the set of linear functions — the set of dual vectors — is the same as the set of all possible n -tuples $\alpha_a = (\alpha_1, \alpha_2, \dots, \alpha_n)$. This is, of course, an n -dimensional vector space in the same way as the n -tuples characterizing vectors, $v^a = (v^1, v^2, \dots, v^n)$ form a vector space.

To formally view the set of dual vectors as a vector space, we need a notion of addition and scalar multiplication. For any \vec{v} and scalar c we define

$$(\underline{\alpha} + \underline{\beta})(\vec{v}) = \underline{\alpha}(\vec{v}) + \underline{\beta}(\vec{v}),$$

$$(c\underline{\alpha})(\vec{v}) = c\underline{\alpha}(\vec{v}).$$

This definition corresponds to the usual rule for adding and scalar multiplying the corresponding n -tuples, as you can easily check.

We can define a basis for V^* as follows. Let $\underline{\omega}^a$ be the linear function

$$\underline{\omega}^a(\vec{v}) = v^a.$$

It is now not hard to check that the general linear function can be written

$$\underline{\alpha} = \alpha_a \underline{\omega}^a.$$

Indeed, we have

$$\underline{\alpha}(\vec{v}) = \alpha_a \underline{\omega}^a(\vec{v}) = \alpha_a v^a.$$

The n dual vectors $\underline{\omega}^a = (\underline{\omega}^1, \underline{\omega}^2, \dots, \underline{\omega}^n)$ are called the *dual basis* to the basis \vec{e}_a .

As you know, we represent the n -tuple of components of a vector (in a given basis) as a column.

$$\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2 + \dots + v^n\vec{e}_n \quad \longleftrightarrow \quad \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}.$$

The n -tuple for the components of a dual vector is represented as a row:

$$\underline{\alpha} = \alpha_1 \underline{\omega}^1 + \alpha_2 \underline{\omega}^2 + \dots + \alpha_n \underline{\omega}^n \quad \longleftrightarrow \quad (\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n).$$

Dual vectors are important in general relativity, of course. But they also feature elsewhere in physics. For example, in quantum mechanics: if $|\psi\rangle$ are vectors – the “kets” – representing the state of a quantum system, then the dual vectors are the “bras”, $\langle\psi|$. Another example comes from solid state physics. Recall that one can define a crystalline lattice via certain linear combinations of a fundamental basis of vectors. The so-called “reciprocal lattice” is just the corresponding linear combinations of the dual basis.