3 Vectors

"My own idea of a useful course is to begin with arithmetic, and then not Euclid but algebra. Next, not Euclid, but practical geometry, solid as well as plane; not demonstration, but to make acquaintance. Then not Euclid, but elementary vectors, conjoined with algebra, and applied to geometry." - Oliver Heaviside

3.1 A Definition of Vectors

We wish to provide a very general definition of vectors that we can use here and in later chapters as needed. We consider sets of objects that follow a very specific set of rules for their computational properties. The term 'objects' is purposely vague here. The object might be a number or a matrix or a function. For clarity one might think 'real numbers' here in order to follow the definitions and rules. Note the use of boldface characters with an arrow above to denote a vector as distinct from a 'normal' number or scalar. Our purpose here is to put forward the mechanism for identifying a vector and being able to say 'yes that is a vector' or 'no that is not a vector'. It is difficult to give examples that are easy to visualize so we will defer examples until later sections. We will represent vectors here in boldface with an arrow above. Thus:

$$\vec{\boldsymbol{a}} \in V$$

represents a vector $\vec{\boldsymbol{a}}$ belonging to the set of vectors, V .

Vectors are objects that belong to a set of similar objects called a *vector space*. The words 'vector' and 'space' are inextricably linked and when someone mentions the word 'space' the mathematician thinks immediately 'vectors' and *vice versa*. Thus, the complete set of vectors, V , comprises a vector space. In order to be classified as vectors, these objects must satisfy some very basic algebraic criteria.

The basic vector algebra rules are as follows:

1. addition is commutative:

$$\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}}=\vec{\boldsymbol{b}}+\vec{\boldsymbol{a}}$$

2. addition is associative:

$$\vec{\boldsymbol{a}}+(\vec{\boldsymbol{b}}+\vec{\boldsymbol{c}})=(\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}})+\vec{\boldsymbol{c}}$$

3. there is an additive identity element, **0**:

$$\vec{\boldsymbol{a}}+\vec{\boldsymbol{0}}=\vec{\boldsymbol{a}}$$

4. every element of the set has an additive inverse, **-a**:

$$\vec{a}+(-\vec{a})=\vec{0}$$

Scalar multiplication is also included in the algebra of vectors:

1. $\alpha(\vec{a}+\vec{b})=\alpha\vec{a}+\alpha\vec{b}$ distributive
2. $(\alpha+\beta)\vec{a}=\alpha\vec{a}+\beta\vec{b}$ distributive
3. $(\alpha\beta)\vec{a}=\alpha(\beta\vec{a})$ associative law for multiplication

The new vector(s) produced by these additions and scalar multiplications *must* also be an element of the vector space, V . This being the case, we say that vectors are *closed* under addition and scalar multiplication.

A *subspace* is a set of vectors that is a subset of a full vector space. The subspace will have the properties:

1. when $\vec{a}$ and $\vec{b}$ belong to the subspace then $\vec{a}+\vec{b}$ also belongs to the subspace.
2. If $\vec{a}$ belongs to the subspace and α is a scalar then $\alpha\vec{a}$ also belongs to the subspace.

One may write a linear combination of vectors:

$$\sum_{1}^{n} c_i\vec{e}_i=c_1\vec{e}_1+c_2\vec{e}_2+c_3\vec{e}_3+\ldots \qquad [3\text{-}1]$$

If we do this and are able to find a set of c's (other than all zero's) such that:

$$c_1\vec{e}_1+c_2\vec{e}_2+c_3\vec{e}_3+\ldots=0 \qquad [3\text{-}2]$$

then the set of vectors is said to be *linearly dependent*. Otherwise, if in order to satisfy the equation we must set all c's to zero, the set is *linearly independent*. A set such as either of these is called a *basis set* of vectors. We shall be much interested in basis sets.

A set of vectors is said to *span* a vector space if every element of the vector space can be expressed as a linear combination of members of the set of vectors:

$$\vec{a}=a_1\vec{e}_1+a_2\vec{e}_2+a_3\vec{e}_3+\ldots \qquad [3\text{-}3]$$

In other words if any vector in a vector space can be written as a linear combination of a set of vectors that set of vectors spans the vector space. What is really of interest here is that if this set of spanning vectors is *also* linearly independent it is said to form a *basis* set for the vector space. The number of vectors in the basis set is the *dimension* of the vector space. Also of interest is that if we have a basis set of vectors:

$$\{\vec{\boldsymbol{e}}_{1,}\vec{\boldsymbol{e}}_{2,}\vec{\boldsymbol{e}}_{3}\ldots\}$$

any vector in the full set of vectors of which the basis set is a part can be constructed as a linear combination of these basis vectors:

$$\vec{\boldsymbol{a}}=a_1\vec{\boldsymbol{e}}_{\mathbf{1}}+a_2\vec{\boldsymbol{e}}_{\mathbf{2}}+a_3\vec{\boldsymbol{e}}_{\mathbf{3}}+\ldots \qquad [3\text{-}4]$$

The a's are called *components* of the vector, $\vec{\boldsymbol{a}}$.

In order to qualify as a vector space the algebraic rules set out at the beginning of this section must be satisfied for all objects in the vector space. It must also be possible to produce the linear combinations of objects that have been discussed.

Having been suitably general (or perhaps vague from your point of view) let's look at example of vector spaces. Remember, to qualify as a vector space the vectors in the set must satisfy the basic algebra rules. We will start with a set of 2 x 2 matrices of real numbers:

$$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\ \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\ \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\ \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\ \begin{bmatrix}12 & 0\\0 & 0\end{bmatrix}\ \begin{bmatrix}3 & 6\\9 & 12\end{bmatrix}\ \begin{bmatrix}6 & 2\\32 & 9\end{bmatrix}$$

$$\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}\quad a_{ij}\in\boldsymbol{R}$$

These are but a few of an infinite number of 2 x 2 matrices that we could cite. The first row comprises a basis set from which all other 2 x 2 matrices can be constructed from:

$$\boldsymbol{A}=\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}=a_{11}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+a_{12}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}+a_{21}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}+a_{22}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{bmatrix}=b_{11}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+b_{12}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}+b_{21}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}+b_{22}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

How do we know that it is a basis set? The test is 'is it linearly independent?'. It is:

$$x_1\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+x_2\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}+x_3\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}+x_4\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

The only way this equation can be true is if all x's are equal to zero which means the matrices are linearly independent.

Matrices **A** and **B** can be added to give matrix C:

$$\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}=\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}+\begin{bmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{bmatrix}=\begin{bmatrix}a_{11}+b_{11} & a_{12}+b_{12}\\a_{21}+b_{21} & a_{22}+b_{22}\end{bmatrix}$$

which is another matrix in the vector space. We can multiply matrix **A** (or **B** or **C**) by a scalar to produce yet another matrix in the vector space:

$$\alpha\boldsymbol{A}=\alpha\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}=\begin{bmatrix}\alpha a_{11} & \alpha a_{12}\\\alpha a_{21} & \alpha a_{22}\end{bmatrix}$$

In fact, if you look at it carefully, you will see that these 2 x 2 matrices satisfy each of the rules of algebra for vectors. Hence we can say that they constitute a set of vectors.

Another example is the set of all polynomial functions of real numbers. For example:

$$f(x)=x^4-2x$$
$$g(x)=x^{12}+3x^5-2$$

We may add these functions to produce a new function of real numbers:

$$k(x)=f(x)+g(x)=x^{12}+3x^5+x^4-2x-2$$

We may multiply any of these functions by a scalar to produce a new function of real numbers:

$$m(x)=6f(x)=6x^4-12x$$

Thus, the set of all functions of real numbers is a set of vectors by definition. We can try to find a basis set for these vectors. Let's think in terms of constructing a vector such as $f(x)$ from the basis set:

$$f(x)=(1)\times(x^4)+(-2)\times(x)$$

What we have done here is to construct f(x) from a linear combination of powers of x so a basis set must consist of the individual powers of x:

$$\sum_{i=0}^{\infty} b_i x^i$$

These are linearly independent since no one x^i can be constructed from a linear combination of the other x^i. Also, the number of members of the basis set is infinite so the dimension of the polynomial space is infinite. Infinite dimensional spaces appear in the Schroedinger formulation of quantum mechanics.

3.2 Simple Euclidean Vector Algebra and Inner Product Spaces

We have, so far, talked about addition and multiplication by scalars but *not* about multiplication of vectors by vectors. We now do so.

The most familiar vector spaces are two- and three-dimensional Euclidean spaces, which we shall denote using $\mathbf{R}^2$ and R^3[1]. We use **R** to emphasize that we are using real numbers in our vectors (for the moment). Vectors in these spaces can be conveniently represented by ordered pairs or triples of real numbers and are frequently depicted graphically as arrows beginning at the origin. These vectors may be added together using the parallelogram rule (see figure 3.2) or multiplied by real numbers (scalar multiplication) and in general, satisfy the criteria for vectors given above. We leave it to the reader to verify this (practice makes perfect!). In addition we define the concepts of length and angles between vectors.

A vector in the Euclidean sense is a quantity that has, in addition to the above properties of vectors, both a scalar value (or a magnitude) and a direction. It is thus, easy to represent two dimensional vectors as arrows on a plane. Here in figure 3-1 we have represented a vector **a** beginning at the origin as is the general convention and pointing (in this case) into the upper right quadrant. The end point of the vector can be represented by its x-y coordinates ... this is an often-used way to compactly talk about vectors. It can of course be expanded to higher dimensions .. as high as desired. The length or *norm* of the vector is $\|\vec{\mathbf{a}}\|=\sqrt{x^2+y^2}$ in 2D and $\|\vec{\mathbf{a}}\|=\sqrt{x^2+y^2+z^2}$ in 3D, easily calculated from the Pythagorean theorem.

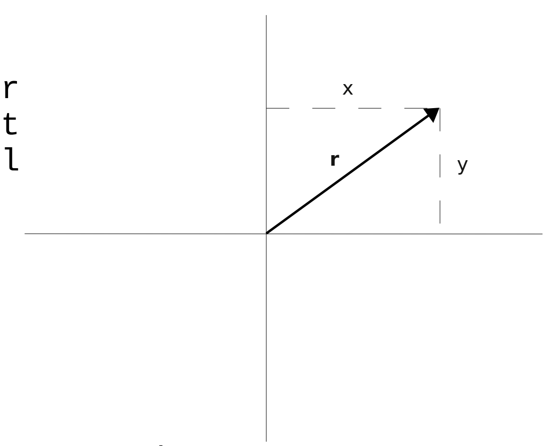


Figure 3-1:A two dimensional Euclidean vector

1 Do not confuse these superscripts with powers. They are meant instead to indicate the number of dimensions of the space.

Algebra on these types of vectors can be done in several ways. The simplest is to use the parallelogram method. Thus, addition is done by lining the vectors up head-to-tail and drawing a resultant vector from the tail of the first to the head of the last:

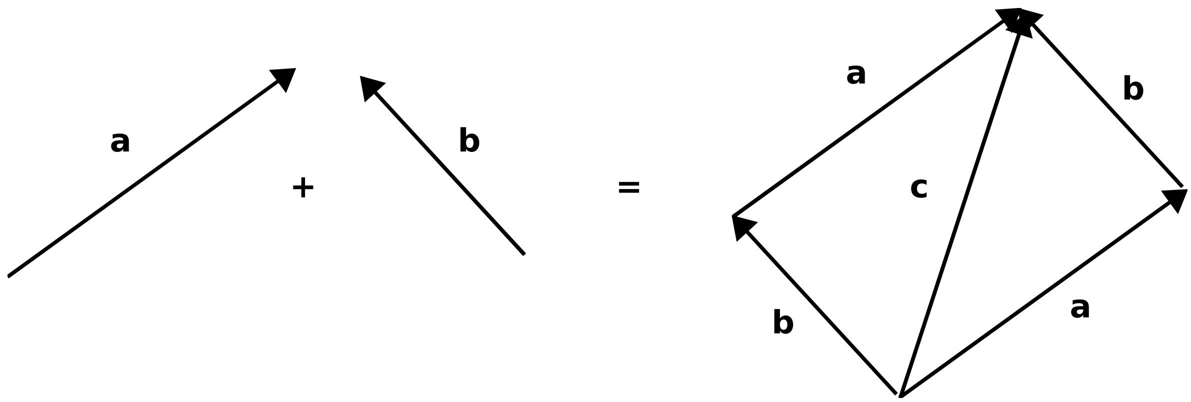


Figure 3-2:Parallelogram method of vector addition

where here, vectors **a** and **b** are placed head-to-tail to produce the resultant vector **c**. Subtraction is similar except that negative vectors are reversed. So, **a** – **b** gives:

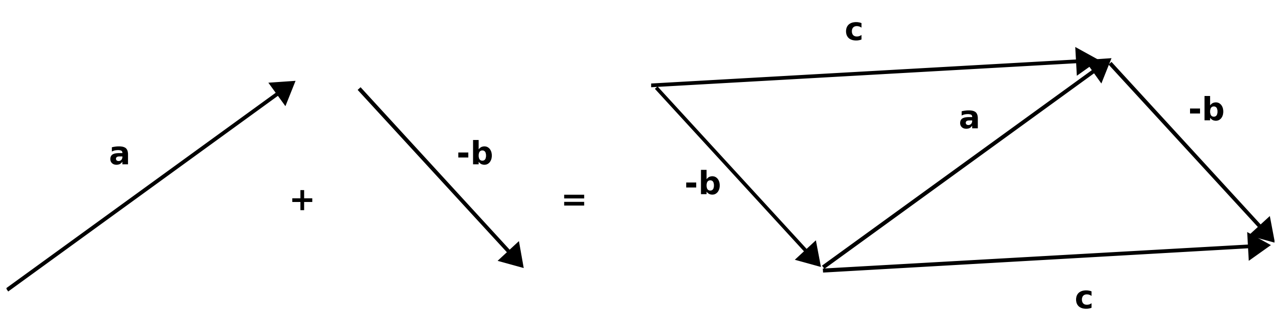


Figure 3-3: Parallelogram method of vector subtraction

A common use of vectors is to denote the spacial position of an object. The *position vector*, often give the symbol $\vec{r}$, is anchored at the origin and points to the position in space where the object resides. Position vectors can be used to calculate the vector between two points in space by simply doing a vector difference calculation.

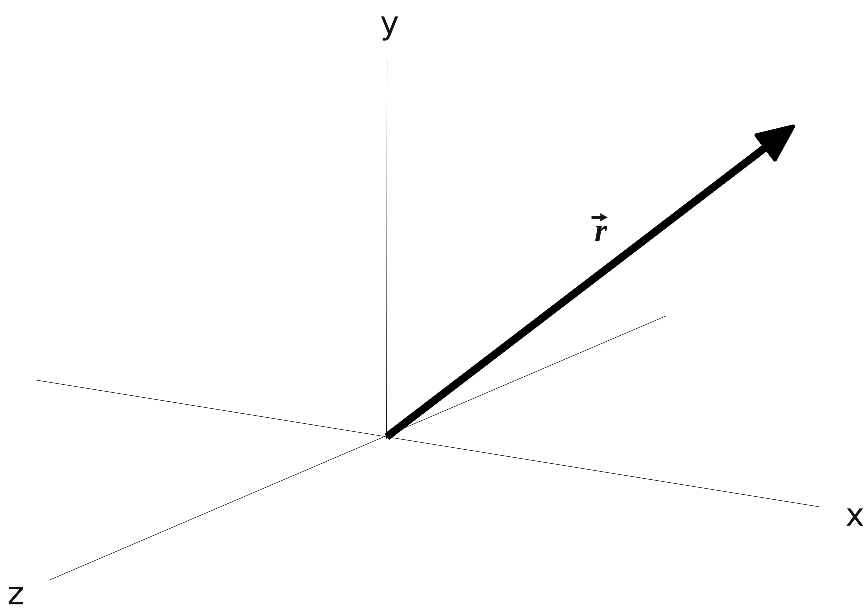


Figure 3-4: A position vector, ***r****.*

Very frequently we use *unit* vectors multiplied by a scalar number to represent vectors:

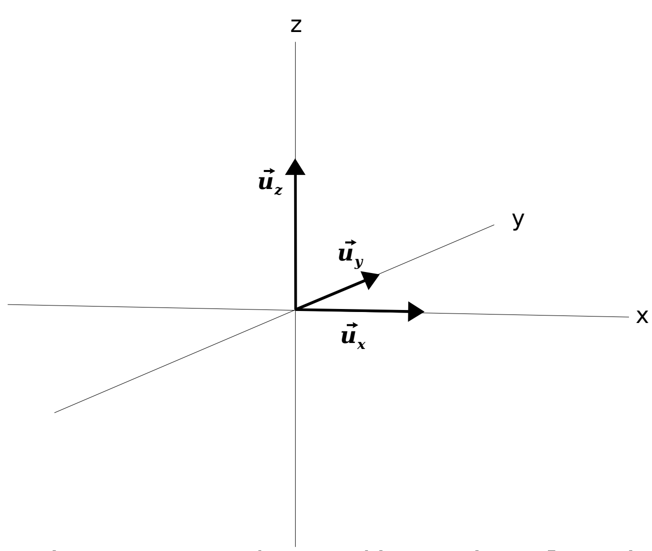


Figure 3-6:Three dimensional unit vectors

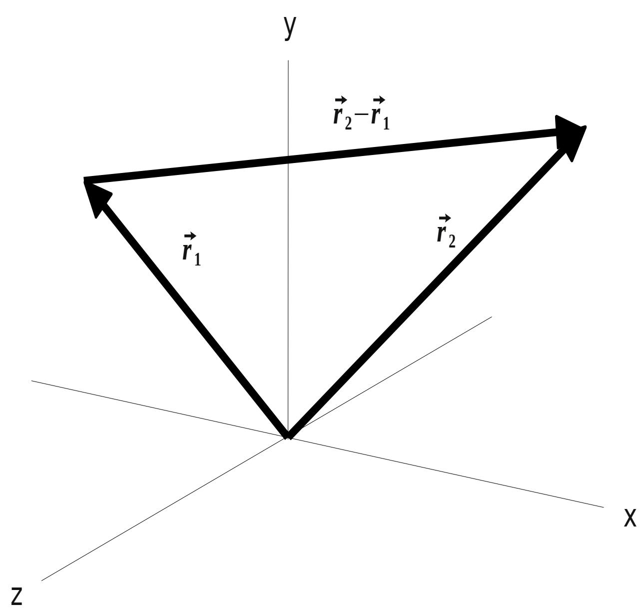


Figure 3-5: Position vector difference calculation

Thus, we might show the 3D vector **a** as the linear combination:

$$\vec{a}=3\vec{u}_x+6\vec{u}_y+\vec{u}_z$$

and vector **b** as:

$$\vec{b}=\vec{u}_x-5\vec{u}_y+2\vec{u}_z$$

$\mathbf{u_x}$, $\mathbf{u_y}$ and $\mathbf{u_z}$ are, of course, a basis set for 3D vector space, as discussed in the first section. That this is so can be seen from the definition of a basis set. First, we can construct any vector in $\mathbf{R^3}$ space from a linear combination of these three vectors, as indeed we have just done with **a** and **b**. Second, these three unit vectors are linearly independent. That is:

$$c_x \vec{\boldsymbol{u}}_x + c_y \vec{\boldsymbol{u}}_y + c_z \vec{\boldsymbol{u}}_z = 0$$

and the only way to do this is to have each c be equal to zero.

Addition and subtraction are quite easy. One simply adds or subtracts the scalars multiplying into the unit vectors:

$$\begin{aligned} \vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} &= 3\vec{\boldsymbol{u}}_x + 6\vec{\boldsymbol{u}}_y + \vec{\boldsymbol{u}}_z + \vec{\boldsymbol{u}}_x - 5\vec{\boldsymbol{u}}_y + 2\vec{\boldsymbol{u}}_z \\ &= (3+1)\vec{\boldsymbol{u}}_x + (6-5)\vec{\boldsymbol{u}}_y + (1+2)\vec{\boldsymbol{u}}_z \\ &= 4\vec{\boldsymbol{u}}_x + \vec{\boldsymbol{u}}_y + 3\vec{\boldsymbol{u}}_z \end{aligned}$$

The scalars in front of the unit vectors are the components or coordinates of the vector .. in this case in three dimensions.

Any one of the familiar planes in $\mathbf{R^3}$ space is a subspace of $\mathbf{R^3}$ space as defined above. Thus, if we consider the x-y plane and two vectors, **a** and **b** in the plane:

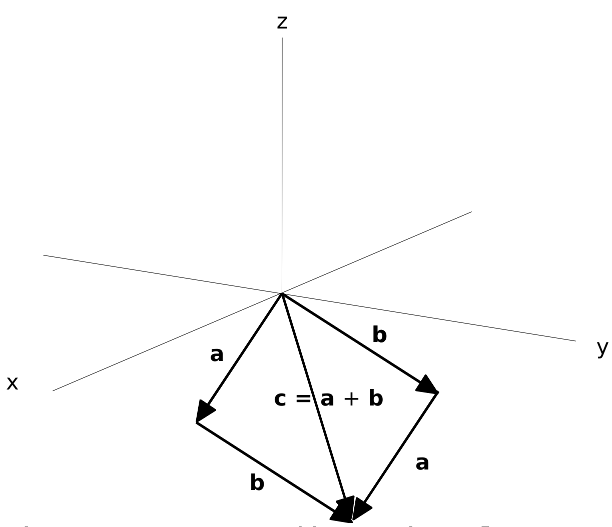


Figure 3-7:Two dimensional vector subspace of 3D space

Vector **a** plus vector **b** gives a new vector **c** that is still in the x-y plane. Also any scalar multiplied into **a** or **b** will result in a vector that is still in the x-y plane. As you will recall from the above discussion on subspaces, this fits the definition of a subspace and therefore the x-y plane is a subspace. There are two unit vectors, $\mathbf{u_x}$ and $\mathbf{u_y}$ directed along the x and y axes respectively, that form a basis for this particular subspace so this is therefore a $\mathbf{R^2}$ space.

Also from Figure 3-5, we note that vector **c** is a result of a linear combination of **a** and **b** which are themselves linear combinations of the unit vectors:

$$\begin{aligned}\vec{c}&=\vec{a}+\vec{b}\\&=a_x\vec{u}_x+a_y\vec{u}_y+a_z\vec{u}_z+b_x\vec{u}_x+b_y\vec{u}_y+b_z\vec{u}_z\\&=(a_x+b_x)\vec{u}_x+(a_y+b_y)\vec{u}_y+(a_z+b_z)\vec{u}_z\end{aligned} \qquad [3\text{-}5]$$

This being the case we can say that **a** and **b** span the subspace that is the x-y plane.

We can multiply a vector by a scalar which simply scales the vector. Thus:

$$\begin{aligned}2\vec{a}&=2(3\vec{u}_x+6\vec{u}_y+\vec{u}_z)\\&=6\vec{u}_x+12\vec{u}_y+2\vec{u}_z\end{aligned}$$

Note that $\mathbf{R}^2$ and $\mathbf{R}^3$ vectors satisfy the addition and scalar multiplication criteria set out in part 1.

It is intuitive that there is a length to a Euclidean vector. One can see it in the graphical picture of a 2-dimensional vector and calculate it using the Pythagorean theorem:

$$\vec{a}=3\vec{u}_x+6\vec{u}_y$$

and the length of **a** (represented by **|a|**) is:

$$\begin{aligned}\|\vec{a}\|&=\sqrt{3^3+6^2}\\&=6.708\end{aligned}$$

To do this calculation we have to do some multiplication. In fact, it is as though we have taken two copies of **a** and multiplied their x and y components and then taken the square root of the sum of the two. It seems a bit silly at first sight to multiply a number and then take it's square root however, since a length is a positive real quantity we do this so that any negative numbers involved result in positive lengths. So, for our vector **a**:

$$\begin{aligned}\vec{a}&=3\vec{u}_x+6\vec{u}_y\\\|\vec{a}\|&=\sqrt{\vec{a}\cdot\vec{a}}\\&=\sqrt{(3^2+6^2)}\end{aligned}$$

and generally for a $\mathbf{R}^3$ vector **a**:

$$\begin{aligned}\vec{a}&=a_x\vec{u}_x+a_y\vec{u}_y+a_z\vec{u}_z\\\|\vec{a}\|&=\sqrt{\vec{a}\cdot\vec{a}}\\&=\sqrt{(a_x\cdot a_x+a_y\cdot a_y+a_z\cdot a_z)}\end{aligned} \qquad [3\text{-}6]$$

Here we are using the convention that $\|\vec{a}\|$ represents the length or the *norm* of vector **a**.

Might we not be able to do something similar with two *different* 2-dimensional vectors? Indeed, we can do so. We define the multiplication of two different vectors, **a** and **b**, much as we did with two of copies of **a**:

$$\begin{gathered}\vec{a}=a_x\vec{u}_x+a_y\vec{u}_y+a_z\vec{u}_z\\ \vec{b}=b_x\vec{u}_x+b_y\vec{u}_y+b_z\vec{u}_z\\ \vec{a}\cdot\vec{b}=a_x b_x+a_y b_y+a_z b_z\end{gathered} \qquad [3\text{-}7]$$

Note that, just as with the length calculation, this quantity is a scalar or simply a number. Also, note that the product of **a** and **b** may be either negative or positive. This quantity is called by several different names including inner product and dot product. The inner product is the more general term while the dot product is mostly used to refer to vectors in $\mathbf{R}^2$ and $\mathbf{R}^3$. There are also several different ways that one may see this product denoted:

$$\begin{gathered}\vec{a}\cdot\vec{b}\\ \text{or}\\ (\vec{a},\vec{b})\\ \text{or}\\ \langle\vec{a}|\vec{b}\rangle\end{gathered} \qquad [3\text{-}8]$$

We will speak more of the last of these notations later. A vector space for which we can do inner product calculations is called an *inner product space*. $\mathbf{R}^2$ and $\mathbf{R}^3$ are two such spaces.

There are some properties that follow from our definition of the inner product for $\mathbf{R}^2$ and $\mathbf{R}^3$:

1. $(\vec{a},\vec{a})=\|\vec{a}\|^2\geqslant 0$ - the length of a vector is zero or greater
2. $(\vec{a},\vec{b})=(\vec{b},\vec{a})$ - the inner product is commutative (however, see below, equation 3-45)
3. $(\alpha\vec{a},\vec{b})=\alpha(\vec{a},\vec{b})$ - scalar multiplication
4. $(\vec{a},\vec{b}+\vec{c})=(\vec{a},\vec{b})+(\vec{a},\vec{c})$ - the inner product is distributive

The inner product for two different vectors in a 2 or 3D space must include the two ideas of length and angle between vectors. We have investigated the length idea but not the angle. Let's do so now using the law of cosines (appendix II, equation [A-11]) to get at it. Imagine vectors **a** and **b** in a 2-dimensional space and vector **c** = **a**-**b**:

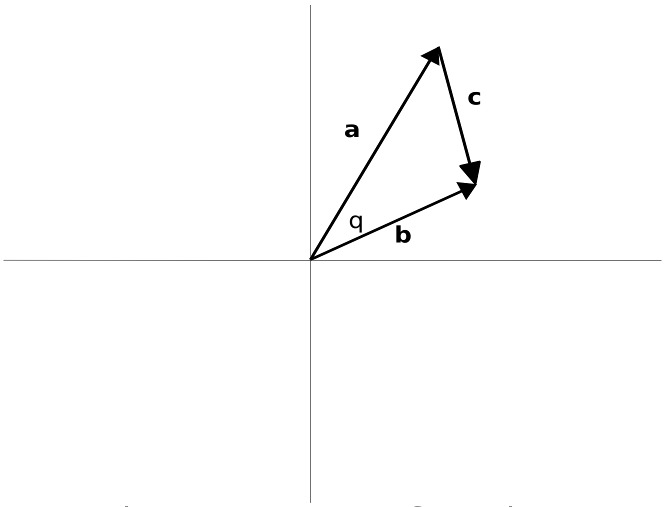


Figure 3-8:Law of Cosines

From the law of cosines:

$$\left\|\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}}\right\|^2=\|\vec{\boldsymbol{c}}\|^2=\|\vec{\boldsymbol{a}}\|^2+\left\|\vec{\boldsymbol{b}}\right\|^2-2\|\vec{\boldsymbol{a}}\|\left\|\vec{\boldsymbol{b}}\right\|\cos(\theta)$$

and from our definitions above:

$$\begin{gathered}\left\|\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}}\right\|^2=(\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}})\\ =(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}})-(\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}})\quad (\textit{distributive property})\\ =(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})-(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}})-(\vec{\boldsymbol{b}},\vec{\boldsymbol{a}})+(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})\\ =(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})+(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})-2(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}})\quad (\textit{commutative property})\\ \|\vec{\boldsymbol{a}}\|^2=(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})\quad \left\|\vec{\boldsymbol{b}}\right\|^2=(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})\end{gathered}$$

and we can then substitute into the law of cosines:

$$\begin{gathered}(\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}-\vec{\boldsymbol{b}})=(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})+(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})-2\|\vec{\boldsymbol{a}}\|\left\|\vec{\boldsymbol{b}}\right\|\cos(\theta)\\ (\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})+(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})-2(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}})=(\vec{\boldsymbol{a}},\vec{\boldsymbol{a}})+(\vec{\boldsymbol{b}},\vec{\boldsymbol{b}})-2\|\vec{\boldsymbol{a}}\|\left\|\vec{\boldsymbol{b}}\right\|\cos(\theta)\end{gathered}$$

We note that (**a**,**a**) and (**b**,**b**) and -2 cancel out on both sides so we have:

$$(\vec{\boldsymbol{a}},\vec{\boldsymbol{b}})=\|\vec{\boldsymbol{a}}\|\left\|\vec{\boldsymbol{b}}\right\|\cos(\theta) \qquad [3\text{-}9]$$

Thus, we see that the inner product is the product of the lengths of vectors **a** and **b** and the cosine of the angle between them. Note that equation [3-9] applies only to vector spaces comprised of real numbers. Complex numbers complicate things .. in particular, the commutative property of vectors is altered as we shall see.

Because there is generally an angle between two vectors **a** and **b**, the maximum value that the inner product can take occurs when that angle is zero radians. Any non-zero angle means a cosine value less than 1 and:

$$\left\|\vec{a}\cdot\vec{b}\right\| \leqslant \left\|\vec{a}\right\|\left\|\vec{b}\right\|$$
or
$$\left\|\vec{a}\cdot\vec{b}\right\|^2 \leqslant \left\|\vec{a}\right\|^2\left\|\vec{b}\right\|^2$$
or
$$\left\|\vec{a}\cdot\vec{b}\right\|^2 \leqslant (\vec{a},\vec{a})(\vec{b},\vec{b}) \qquad [3\text{-}10]$$

which is known as the Cauchy-Schwartz inequality and will be of importance to us later on in our look at quantum mechanics.

If the angle between the two vectors in question is 90 degrees then from [3-8] the cosine of the angle equals zero. Under these circumstances we say that they are *orthogonal* vectors:

$$\vec{a}\cdot\vec{b} = \left\|\vec{a}\right\|\left\|\vec{b}\right\|\cos(0) = 0 \qquad [3\text{-}11]$$

As far as a basis set in Euclidean space is concerned it is convenient under many circumstances to use one in which each element is orthogonal to the others and in which each element is normalized. Given the above definition of the dimension of a vector space it is obvious that for $\mathbf{R}^2$ and $\mathbf{R}^3$ spaces there are two and three basis vectors respectively which we represent here using $\mathbf{u}_x$, $\mathbf{u}_y$ and $\mathbf{u}_z$. Since these are both orthogonal and normalized we call this an *orthonormal* basis set.

We say that a vector is *normalized* when:

$$(\vec{a},\vec{a}) = 1 \qquad [3\text{-}12]$$

A vector may always be normalized using the definition of the inner product and the norm:

$$(\vec{a},\vec{a}) = \left\|\vec{a}\right\|^2$$
$$\frac{(\vec{a},\vec{a})}{\left\|\vec{a}\right\|^2} = 1$$
$$\text{let } \vec{a}' = \frac{\vec{a}}{\left\|\vec{a}\right\|}$$
$$\text{then } \vec{a}'\cdot\vec{a}' = 1$$

We can do an inner product for each pair of our orthogonal and normalized unit vectors:

$$\begin{array}{lll} \vec{u}_x\cdot\vec{u}_x=1 & \vec{u}_x\cdot\vec{u}_y=0 & \vec{u}_x\cdot\vec{u}_z=0 \\ \vec{u}_y\cdot\vec{u}_x=0 & \vec{u}_y\cdot\vec{u}_y=1 & \vec{u}_y\cdot\vec{u}_z=0 \\ \vec{u}_z\cdot\vec{u}_x=0 & \vec{u}_z\cdot\vec{u}_y=0 & \vec{u}_z\cdot\vec{u}_z=1 \end{array} \qquad [3\text{-}13]$$

There are two main types of vector multiplication, the inner or dot product which we have just seen and the cross product so named because of the notation used to denote it:

$$\vec{c} = \vec{a} \times \vec{b} \qquad [3\text{-}14]$$

Just to complicate things for a bit of fun, the cross product in 2D is *different* from that of 3D space. The 2D version, discovered by Argand (of Argand diagram fame) is:

$$\begin{aligned} \vec{a} \times \vec{b} &= \vec{c} \\ &= (a_x \vec{u}_x + a_y \vec{u}_y) \times (b_x \vec{u}_x + b_y \vec{u}_y) \\ &= (a_x b_x - a_y b_y) \vec{u}_x + (a_y b_x + a_x b_y) \vec{u}_y \end{aligned}$$

In diagram form this looks as in figure 3-9. This should look a bit familiar to you .. it is *exactly* the same as multiplication of two complex numbers represented in an Argand diagram. Note that by definition, the length or norm of **c** is equal to the product of the lengths of **a** and **b**.

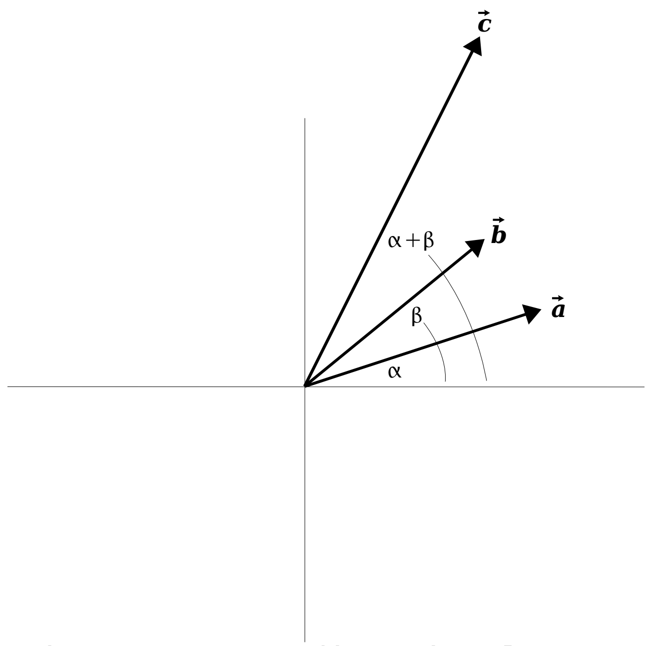


Figure 3-9:Two dimensional cross product

How do we show this? We use the fundamental geometric definitions of sine and cosine and a couple of trigonometric identities:

$$\begin{aligned} &\|\vec{c}\| = \|\vec{a}\| \|\vec{b}\| \\ &\sin(\alpha) = \frac{a_y}{\|\vec{a}\|} \quad \cos(\alpha) = \frac{a_x}{\|\vec{a}\|} \\ &\sin(\beta) = \frac{b_y}{\|\vec{b}\|} \quad \cos(\beta) = \frac{b_x}{\|\vec{b}\|} \\ &\sin(\alpha+\beta) = \frac{c_y}{\|\vec{c}\|} = \frac{c_y}{\|\vec{a}\|\|\vec{b}\|} \\ &= \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha) \\ &\cos(\alpha+\beta) = \frac{c_x}{\|\vec{c}\|} = \frac{c_x}{\|\vec{a}\|\|\vec{b}\|} \\ &= \cos(\alpha)\cos(\beta) - \sin(\beta)\sin(\alpha) \end{aligned}$$

With these in hand we write the x and y components of the new vector, **c**:

$$\frac{c_x}{\|\vec{a}\|\|\vec{b}\|}=\frac{a_x}{\|\vec{a}\|}\cdot\frac{b_x}{\|\vec{b}\|}-\frac{b_y}{\|\vec{b}\|}\cdot\frac{a_y}{\|\vec{a}\|}$$
or
$$c_x=a_x b_x-b_y a_y$$

and:

$$\frac{c_y}{\|\vec{a}\|\|\vec{b}\|}=\frac{a_y}{\|\vec{a}\|}\cdot\frac{b_x}{\|\vec{b}\|}+\frac{b_y}{\|\vec{b}\|}\cdot\frac{a_x}{\|\vec{a}\|}$$
or
$$c_y=a_y b_x+b_y a_x$$

so:

$$\vec{c}=(a_x b_x-a_y b_y)\vec{u}_x+(a_y b_x+a_x b_y)\vec{u}_y$$

This version of the cross product is little used and we mention it only for interest's sake.

The 3D version of the cross product is however, heavily used in all areas of the physical sciences and is crucial to understanding many aspects of nmr spectroscopy. As with the 2D version, the 3D cross product (which we shall refer to simply as the cross product from now on) also produces a new vector, **c**. However, as we shall see, this new vector is at right angles to the plane defined by **a** and **b** and its magnitude depends on the angle, q, between **a** and **b** as does the inner product. We define the cross product multiplication for two **R**^{3} vectors **a** and **b** as follows:

$$\begin{gathered}\vec{a}=a_x\vec{u}_x+a_y\vec{u}_y+a_z\vec{u}_z\\ \vec{b}=b_x\vec{u}_x+b_y\vec{u}_y+b_z\vec{u}_z\\ \vec{a}\times\vec{b}=\begin{vmatrix}a_y & a_z\\ b_y & b_z\end{vmatrix}\boldsymbol{u_x}-\begin{vmatrix}a_x & a_z\\ b_x & b_z\end{vmatrix}\boldsymbol{u_y}+\begin{vmatrix}a_x & a_y\\ b_x & b_y\end{vmatrix}\boldsymbol{u_z}\\ =\begin{vmatrix}\boldsymbol{u_x} & \boldsymbol{u_y} & \boldsymbol{u_z}\\ a_x & a_y & a_z\\ b_x & b_y & b_z\end{vmatrix}\end{gathered} \qquad [3\text{-}15]$$

As with the inner product, given this definition there are properties that follow:

1. $\vec{a}\times\vec{b}=-\vec{b}\times\vec{a}$ - swapping rows (see properties of determinants in chapter 2) in the determinant [3-15] changes sign
2. $\vec{a}\times\gamma\vec{b}=\gamma\vec{a}\times\vec{b}$ - scalar multiplication (see [2-21])
3. $\vec{a}\times(\vec{b}+\vec{c})=\vec{a}\times\vec{b}+\vec{a}\times\vec{c}$ - distributive
4. $\vec{a}\times\vec{a}=\vec{0}$ - the identity vector

Property 4 follows from property 1:

$$\vec{a}\times\vec{a}=-(\vec{a}\times\vec{a})$$
therefore
$$\vec{a}\times\vec{a}=\vec{0}$$

Note the 'zero' vector in the last result. Since the result of the cross product is always a vector we define the zero vector in this fashion. *Every* vector space has a zero vector. In fact, 0 itself is a complete vector space by definition, following all of the algebraic rules set down in the beginning of this chapter. Recall that this is the identity vector that is required for a complete vector space.

We see from the definition of the cross product that the result is a new vector. What is the direction of this new vector? We can figure this out using:

$$\vec{a}\cdot(\vec{b}\times\vec{c})$$

that is, the inner product between **a** and the result of cross product of **b** and **c**. We further stipulate that **a** is in the same plane as that defined by **b** and **c**. Proceeding, we get:

$$\vec{a}\cdot(\vec{b}\times\vec{c})$$
$$=(a_x\vec{u}_x+a_y\vec{u}_y+a_z\vec{u}_z)\cdot(\begin{vmatrix}b_y & b_z\\ c_y & c_z\end{vmatrix}\vec{u}_x+\begin{vmatrix}b_x & b_z\\ c_x & c_z\end{vmatrix}\vec{u}_y+\begin{vmatrix}b_x & b_y\\ c_x & c_y\end{vmatrix}\vec{u}_z)$$
$$=a_x\begin{vmatrix}b_y & b_z\\ c_y & c_z\end{vmatrix}\vec{u}_x+a_y\begin{vmatrix}b_x & b_z\\ c_x & c_z\end{vmatrix}\vec{u}_y+a_z\begin{vmatrix}b_x & b_y\\ c_x & c_y\end{vmatrix}\vec{u}_z$$
$$=\begin{vmatrix}a_x & a_y & a_z\\ b_x & b_y & b_z\\ c_x & c_y & c_z\end{vmatrix}$$

Now, if **a** is in the same plane as that defined by **b** and **c**, then it must be possible to write **a** as a linear combination of **b** and **c** since **b** and **c** span the subspace of the plane that they define. and this being the case, the first row in the determinant is a linear combination of the second and third rows. From [2-25] this means that the result will equal zero which means in turn that **b**x**c** is orthogonal to the plane defined by vectors **b** and **c**.

As with the dot product, we ask ourselves 'what is the magnitude of **b**x**c**?':

$$
\begin{aligned}
\vec{\boldsymbol{d}} = \vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} &= \begin{vmatrix} \vec{\boldsymbol{u}}_{\boldsymbol{x}} & \vec{\boldsymbol{u}}_{\boldsymbol{y}} & \vec{\boldsymbol{u}}_{\boldsymbol{z}} \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \\
&= \begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix} \vec{\boldsymbol{u}}_{\boldsymbol{x}} + \begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix} \vec{\boldsymbol{u}}_{\boldsymbol{y}} + \begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix} \vec{\boldsymbol{u}}_{\boldsymbol{z}} \\
&= d_x \vec{\boldsymbol{u}}_{\boldsymbol{x}} + d_y \vec{\boldsymbol{u}}_{\boldsymbol{y}} + d_z \vec{\boldsymbol{u}}_{\boldsymbol{z}}
\end{aligned}
$$

From equation [3-6] the length (or the norm in vector-speak) of **d** is:

$$
\begin{aligned}
\|\vec{\boldsymbol{d}}\| &= \sqrt{d_x^2 + d_y^2 + d_z^2} \\
\|\vec{\boldsymbol{d}}\|^2 &= d_x^2 + d_y^2 + d_z^2 \\
&= \begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix}^2 + \begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix}^2 + \begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix}^2 \\
&= (b_y c_z - c_y b_z)^2 - (b_x c_z - c_x b_z)^2 + (b_x c_y - c_x b_y)^2 \\
&= b_y^2 c_z^2 + c_y^2 b_z^2 - 2\mathrm{b}_y b_z c_y c_z \\
&\quad + b_x^2 c_z^2 + c_x^2 b_z^2 - 2\mathrm{b}_x b_z c_x c_z \\
&\quad + b_x^2 c_y^2 + c_x^2 b_y^2 - 2\mathrm{b}_x b_y c_x c_y
\end{aligned}
$$

At this point we do a little 'hocus pocus' magic and add some terms into the expression:

$$
\begin{aligned}
\|\vec{\boldsymbol{d}}\|^2 = {} & b_y^2 c_z^2 + c_y^2 b_z^2 - 2\mathrm{b}_y b_z c_y c_z \\
& + b_x^2 c_z^2 + c_x^2 b_z^2 - 2\mathrm{b}_x b_z c_x c_z \\
& + b_x^2 c_y^2 + c_x^2 b_y^2 - 2\mathrm{b}_x b_y c_x c_y \\
& + b_x^2 c_x^2 - b_x^2 c_x^2 + b_y^2 c_y^2 - b_y^2 c_y^2 + b_z^2 c_z^2 - b_z^2 c_z^2
\end{aligned}
$$

Since each term in the last line has its negative counterpart we have not changed the value of the whole expression. However, it does allow us to proceed. We can now gather all of the terms and simplify things using our formula for the length of a vector, [3-6], our formulas for the inner product, [3-7] and [3-9] and a trigonometric identity (see appendix II, [AII-10]):

$$
\begin{aligned}
\|\vec{\boldsymbol{d}}\|^2 &= (b_x^2 + b_y^2 + b_z^2)(c_x^2 + c_y^2 + c_z^2) - (b_x c_x + b_y c_y + b_z c_z)^2 \\
&= \|\vec{\boldsymbol{b}}\|^2 \|\vec{\boldsymbol{c}}\|^2 - \|\vec{\boldsymbol{b}}\|^2 \|\vec{\boldsymbol{c}}\|^2 \cos^2(\theta) \\
&= \|\vec{\boldsymbol{b}}\|^2 \|\vec{\boldsymbol{c}}\|^2 (1 - \cos^2(\theta)) \\
&= \|\vec{\boldsymbol{b}}\|^2 \|\vec{\boldsymbol{c}}\|^2 \sin^2(\theta)
\end{aligned}
$$

$$\text{and}$$

$$
\|\vec{\boldsymbol{d}}\| = \|\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}\| = \|\vec{\boldsymbol{b}}\| \|\vec{\boldsymbol{c}}\| |\sin(\theta)|
$$

So, finally we reach our goal. The equation for the length of the new vector arising from the cross product of two vectors is:

$$\|\vec{c}\| = \|\vec{a} \times \vec{b}\| = \|\vec{a}\| \|\vec{b}\| |\sin(\theta)| \quad [3\text{-}16]$$

Thus, if **a** and **b** are parallel the magnitude of **c** will be zero.

We know that the new vector, **c**, will be perpendicular to the plane defined by **a** and **b** but what direction will the new vector, **c**, be in after **a** x **b**? For simplicity let's suppose that **a** lies along the +y axis and **b** along the +x axis.

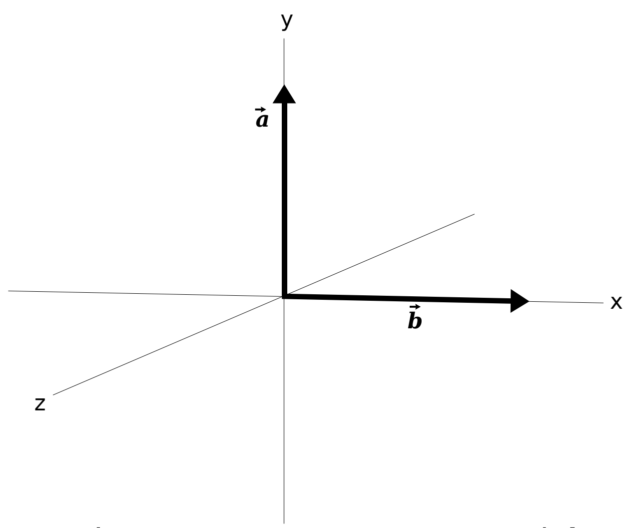


*Figure 3-10:Vectors **a** and **b***

Using our definition of the cross product, [3-15], we write:

$$\vec{a} = 0 \cdot \vec{u}_x + a_y \vec{u}_y + 0 \cdot \vec{u}_z$$
$$\vec{b} = b_x \vec{u}_x + 0 \cdot \vec{u}_y + 0 \cdot \vec{u}_z$$
$$\vec{a} \times \vec{b} = \begin{vmatrix} a_y & 0 \\ 0 & 0 \end{vmatrix} \boldsymbol{u}_x - \begin{vmatrix} 0 & 0 \\ b_x & 0 \end{vmatrix} \boldsymbol{u}_y + \begin{vmatrix} 0 & a_y \\ b_x & 0 \end{vmatrix} \boldsymbol{u}_z$$
$$= -a_y b_x \vec{u}_z$$
$$= \vec{c}$$

Evidently, **c** lies along the -z axis:

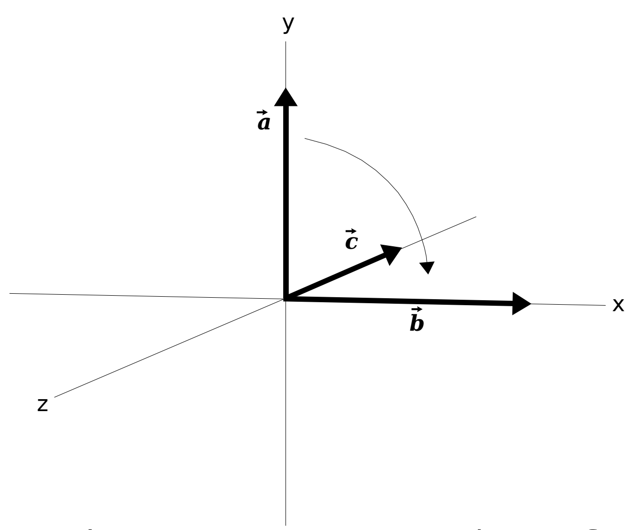


*Figure 3-11:Cross product of vectors **a** and **b***

In order to know where the resultant vector will be we use the *right-hand rule* to determine this. Imagine vector **a** rotating about the z-axis towards vector **b**. Imagine also, that this rotation corresponds to the rotational direction that a screw is being turned. The inward motion of the screw is the direction of the new vector, **c**. It should be obvious that if we write **b**x**a** that the resultant vector **c** will be along the +z-axis.

Of course, the cross product applies to our unit basis vectors as well:

$$\begin{array}{lll} \vec{u}_x \times \vec{u}_x = 0 & \vec{u}_x \times \vec{u}_y = \vec{u}_z & \vec{u}_x \times \vec{u}_z = -\vec{u}_y \\ \vec{u}_y \times \vec{u}_x = -\vec{u}_z & \vec{u}_y \times \vec{u}_y = 0 & \vec{u}_y \times \vec{u}_z = \vec{u}_x \\ \vec{u}_z \times \vec{u}_x = \vec{u}_y & \vec{u}_z \times \vec{u}_y = -\vec{u}_x & \vec{u}_z \times \vec{u}_z = 0 \end{array} \qquad [3\text{-}17]$$

It is useful to note the cyclic permutation of $\mathbf{u_x}$, $\mathbf{u_y}$ and $\mathbf{u_z}$ in these calculations. Thus if you go from left to right and cycle back to the beginning and can go from $\mathbf{u_x}$ to $\mathbf{u_y}$ to $\mathbf{u_z}$ then the result of the calculation will be positive. Otherwise, it will be negative. For example, take $\mathbf{u_x}$ x $\mathbf{u_y}$ = $\mathbf{u_z}$ then permute by cycling each unit vector one position to the right, $\mathbf{u_z}$ x $\mathbf{u_x}$ = $\mathbf{u_y}$ and once again, $\mathbf{u_y}$ x $\mathbf{u_z}$ = $\mathbf{u_x}$. Any combination that is out of order (ie. $\mathbf{u_x}$, $\mathbf{u_y}$, $\mathbf{u_z}$) will result in a negative cross product ie. $\mathbf{u_y}$ x $\mathbf{u_x}$ = -$\mathbf{u_z}$.

Using our example vectors and equations [3-13] and [3-16] we can do dot and cross products:

$$\vec{a} = 3\vec{u}_x + 6\vec{u}_y + \vec{u}_z \quad \vec{b} = \vec{u}_x - 5\vec{u}_y + 2\vec{u}_z$$

The dot (or scalar or inner) product would be:

$$\begin{gathered} \vec{a} \cdot \vec{b} = (3\vec{u}_x + 6\vec{u}_y + \vec{u}_z) \cdot (\vec{u}_x - 5\vec{u}_y + 2\vec{u}_z) \\ = 3\vec{u}_x \cdot \vec{u}_x - 15\vec{u}_x \cdot \vec{u}_y + 6\vec{u}_x \cdot \vec{u}_z + 6\vec{u}_y \cdot \vec{u}_x - 30\vec{u}_y \cdot \vec{u}_y + 12\vec{u}_y \cdot \vec{u}_z + \vec{u}_z \cdot \vec{u}_x - 5\vec{u}_z \cdot \vec{u}_y + 2\vec{u}_z \cdot \vec{u}_z \\ = 3 - 0 + 0 + 0 - 30 + 0 + 0 - 0 + 2 \\ = -25 \end{gathered}$$

and the cross product is:

$$\begin{gathered} \vec{a} \times \vec{b} = (3\vec{u}_x + 6\vec{u}_y + \vec{u}_z) \times (\vec{u}_x - 5\vec{u}_y + 2\vec{u}_z) \\ = 3\vec{u}_x \times \vec{u}_x - 15\vec{u}_x \times \vec{u}_y + 6\vec{u}_x \times \vec{u}_z + 6\vec{u}_y \times \vec{u}_x - 30\vec{u}_y \times \vec{u}_y \\ + 12\vec{u}_y \times \vec{u}_z + \vec{u}_z \times \vec{u}_x - 5\vec{u}_z \times \vec{u}_y + 2\vec{u}_z \times \vec{u}_z \\ = 0 - 15\vec{u}_z - 6\vec{u}_y - 6\vec{u}_z - 0 + 12\vec{u}_x + \vec{u}_y + 5\vec{u}_x + 0 \\ = 17\vec{u}_x - 5\vec{u}_y - 21\vec{u}_z \\ = \vec{c} \end{gathered}$$

As you can see, the dot product results in an ordinary number (or a scalar in the maths vernacular) and the cross product results in a new vector. The new vector, **c**, will be at right angles to the plane defined by **a** and **b** as we have seen. How would you show this? Well the definition of orthogonality (being at right angles) is that the dot product of two vectors equals zero.

$$\begin{gathered} \vec{a} \cdot \vec{c} = (3\vec{u}_x + 6\vec{u}_y + \vec{u}_z) \cdot (17\vec{u}_x - 5\vec{u}_y - 21\vec{u}_z) \\ = 3 \cdot 17 + 6 \cdot (-5) + 1 \cdot (-21) \\ = 51 - 30 - 21 \\ = 0 \end{gathered}$$

Therefore, **a** and **c** are orthogonal.

The cross product may also be calculated using the determinant method (equation [3-15]):

$$\vec{a}\times\vec{b}=\begin{vmatrix}\vec{u}_x & \vec{u}_y & \vec{u}_z\\ 3 & 6 & 1\\ 1 & -5 & 2\end{vmatrix}$$
$$=(6\cdot 2-(-5)\cdot 1)\vec{u}_x-(3\cdot 2-1\cdot 1)\vec{u}_y+(3\cdot(-5)-1\cdot 6)\vec{u}_z$$
$$=17\vec{u}_x-5\vec{u}_y-21\vec{u}_z$$

The norm of vector **a** is:

$$\vec{a}\cdot\vec{a}=(3\vec{u}_x+6\vec{u}_y+\vec{u}_z)\cdot(3\vec{u}_x+6\vec{u}_y+\vec{u}_z)$$
$$=(3\vec{u}_x\cdot 3\vec{u}_x)+(3\vec{u}_x\cdot 6\vec{u}_y)+(3\vec{u}_x\cdot\vec{u}_z)$$
$$+(6\vec{u}_y\cdot 3\vec{u}_x)+(6\vec{u}_y\cdot 6\vec{u}_y)+(6\vec{u}_y\cdot\vec{u}_z)$$
$$+(\vec{u}_z\cdot 3\vec{u}_x)+(\vec{u}_z\cdot 6\vec{u}_y)+(\vec{u}_z\cdot\vec{u}_z)$$
$$=9+0+0+36+0+0+0+1$$
$$=46$$

There is a third type of multiplication which produces an entity called a *dyad* from two vectors using the 'normal' distributive rules of multiplication from grade school:

$$\vec{a}\,\vec{b}=(3\vec{u}_x+6\vec{u}_y+\vec{u}_z)(\vec{u}_x-5\vec{u}_y+2\vec{u}_z)$$
$$=3\vec{u}_x\vec{u}_x-15\vec{u}_x\vec{u}_y+6\vec{u}_x\vec{u}_z+6\vec{u}_y\vec{u}_x-30\vec{u}_y\vec{u}_y+\vec{u}_z\vec{u}_x-5\vec{u}_z\vec{u}_y+2\vec{u}_z\vec{u}_z$$

We will have more to say about this in the next chapter.

There are two useful extensions of inner and cross products, the *scalar triple product* and the *vector triple product*. The scalar triple product is:

$$\vec{a}\cdot(\vec{b}\times\vec{c})$$

in which the three vectors are non-coplanar. This is the formula for the volume of a parallelepiped whose sides are the three non-coplanar vectors, $\vec{a}$, $\vec{b}$ and $\vec{c}$:

$$V=\vec{a}\cdot(\vec{b}\times\vec{c})$$

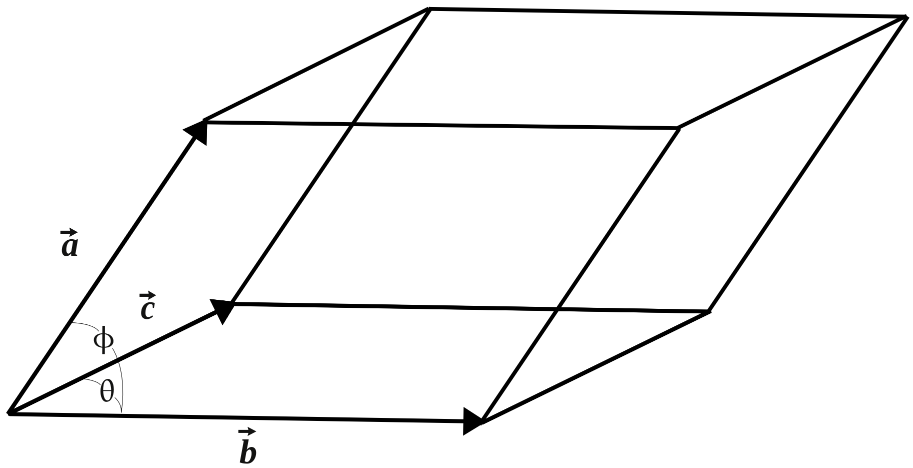


Figure 3-12: Parallelepiped defined by vectors $\vec{a}$, $\vec{b}$ *and* $\vec{c}$

That this is so is evident from the figure 3-10 and the following argument. The volume of the parallelepiped is the area of the parallelogram defined by vectors $\vec{b}$ and $\vec{c}$ times the height, h, of the parallelepiped. The area of the base of the parallelepiped is:

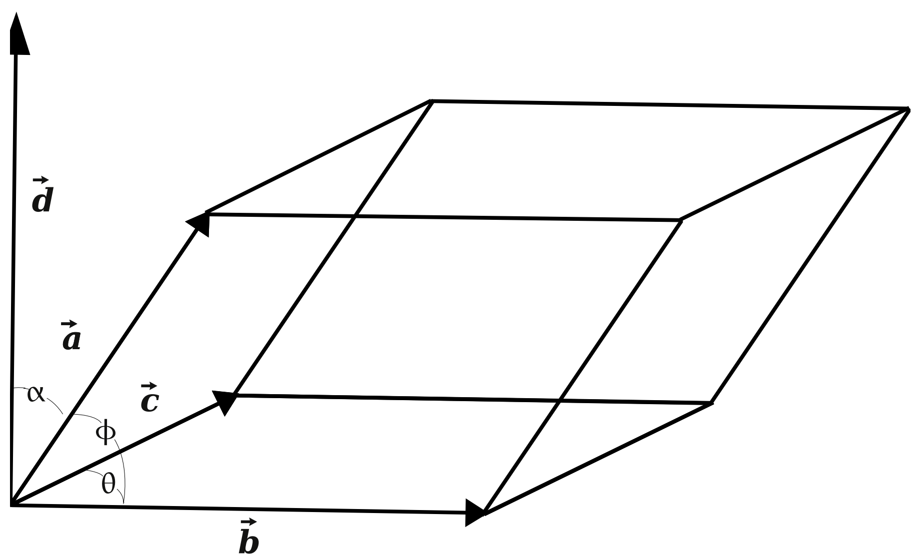
$$\|\vec{b}\times\vec{c}\|=\|\vec{b}\|\|\vec{c}\|\sin\theta$$

that is, the magnitude of the cross product (see Appendix II). The vector defined by the cross product is the *area vector* and is orthogonal to the plane described be vectors $\vec{b}$ and $\vec{c}$:

$$\vec{d}=\vec{b}\times\vec{c}$$

We now are in a position to show that our assertion concerning the volume of the parallelepiped is true:

$$\begin{aligned} V&=\vec{a}\cdot(\vec{b}\times\vec{c}) \\ &=\vec{a}\cdot\vec{d} \\ &=\|\vec{a}\|\|\vec{b}\times\vec{c}\|\cos\alpha \\ &=h\|\vec{b}\times\vec{c}\| \end{aligned}$$



In other words, the height, h, of the parallelepiped times the area defined by the base (not basis!) vectors of the parallelpiped.

We may represent this result in determinant form:

$$\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} = \begin{vmatrix} \vec{\boldsymbol{u}}_1 & \vec{\boldsymbol{u}}_2 & \vec{\boldsymbol{u}}_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) = \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{d}} = \vec{\boldsymbol{a}} \cdot \begin{vmatrix} \vec{\boldsymbol{u}}_1 & \vec{\boldsymbol{u}}_2 & \vec{\boldsymbol{u}}_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

The proof of this uses the matrix notation for the inner product operations in the next section. See also, the problems at the end of the chapter.

We may also write the scalar triple product as:

$$(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{c}}$$

which is equivalent to our first scalar triple product:

$$\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) = (\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{c}}$$

That this is so follows from our geometrical arguments. If we define the base of the parallelepiped using vectors **a** and **b** then, as above, $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$ is the area of the base and then the height of the parallelepiped is given by $\vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{d}}$.

What if the result of the scalar triple product is zero? This could be so if one or more of the vectors is the zero vector; however this is trivial and of more interest is the situation in which all three vectors are coplanar and therefore dependent vectors. Given our above geometrical arguments this would naturally be expected to give a zero volume for the parallelepiped.

$$\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0$$

Next, if in our expression $(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}) \cdot \vec{\boldsymbol{c}}$ the vector $\vec{\boldsymbol{c}}$ is a unit vector with a norm of 1, then the expression corresponds to the area of the parallelepiped defined by $\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$:

$$(\vec{a}\times\vec{b})\cdot\vec{c} = |\vec{a}\times\vec{b}||\vec{c}|\cos(\theta)$$
$$= |\vec{a}\times\vec{b}|\cos(\theta)$$
$$= |a||b|\sin(\phi)\cos(\theta)$$

As in the appendix, $\vec{a}\times\vec{b}$ is the area of a parallelogram.

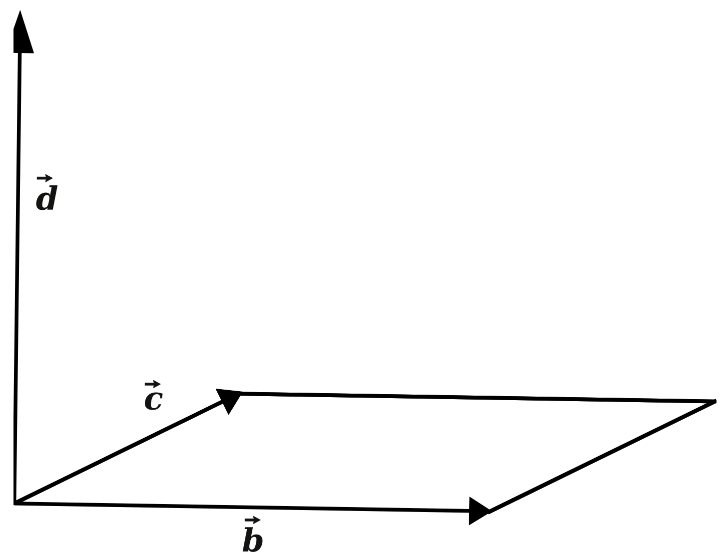
Finally, we can use the commutative property of vectors in 'normal' real space to show that:

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = (\vec{b}\times\vec{c})\cdot\vec{a}$$
$$\vec{a}\cdot\vec{d} = \vec{d}\cdot\vec{a}$$

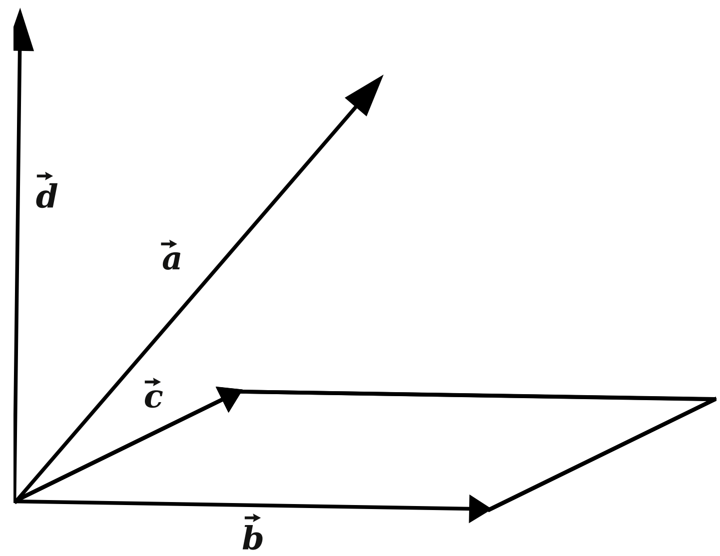
The vector triple product is:

$$\vec{a}\times(\vec{b}\times\vec{c}) = \vec{b}(\vec{a}\cdot\vec{c}) - \vec{c}(\vec{a}\cdot\vec{b})$$

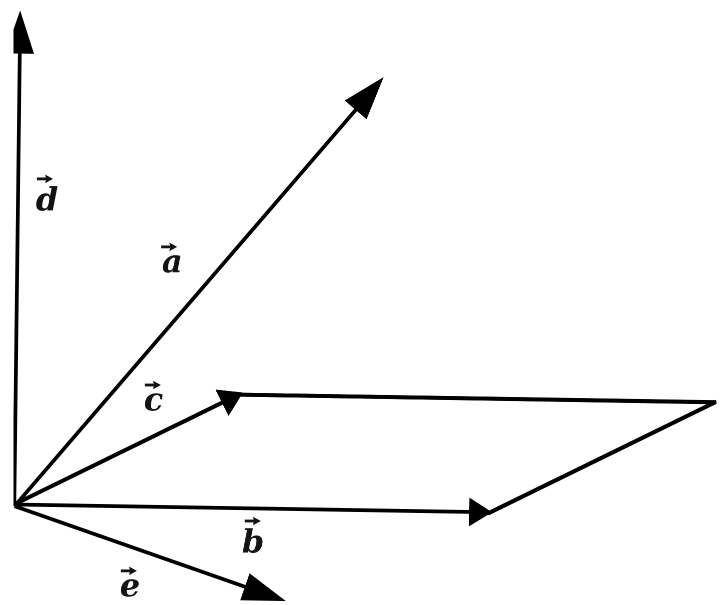
The proof for this is straightforward, if somewhat tedious. Consider first the operation $\vec{b}\times\vec{c}$. Geometrically, the resultant vector, $\vec{d}$, will be situated at right angles to the plane defined by $\vec{b}$ and $\vec{c}$:



If we add in vector $\vec{a}$ to the picture then the plane defined by vectors $\vec{a}$ and $\vec{d}$ is orthogonal to the plane defined by vectors $\vec{b}$ and $\vec{c}$ by virtue of the fact that vector $\vec{d}$ is orthogonal to the plane:



Now, the operation $\vec{a} \times \vec{d}$ will produce a new vector, $\vec{e}$, which will be orthogonal to the plane defined by vectors $\vec{a}$ and $\vec{d}$ which means that it *must* be in the same plane as vectors $\vec{b}$ and $\vec{c}$:



This being the case, it must be possible to express vector $\vec{e}$ as a linear combination of vectors $\vec{b}$ and $\vec{c}$:

$$\vec{e} = a\vec{b} + b\vec{c}$$

because all vectors in a plane belong to the same vector space and the appropriate combination of $\vec{b}$ and $\vec{c}$ will produce $\vec{e}$.

Now, again consider the operation $\vec{a} \times \vec{d} = \vec{e}$ but now from the point of view of the determinant representation:

$$\vec{e} = \begin{vmatrix} \vec{u}_x, & \vec{u}_y & \vec{u}_z \\ a_{1,} & a_2 & a_3 \\ d_{1,} & d_2 & d_3 \end{vmatrix}$$
$$= (a_2 d_3 - a_3 d_2)\vec{u}_x - (a_1 d_3 - a_3 d_1)\vec{u}_y + (a_1 d_2 - a_2 d_1)\vec{u}_z$$
$$= (a_2 d_3 - a_3 d_2)\vec{u}_x + (a_3 d_1 - a_1 d_3)\vec{u}_y + (a_1 d_2 - a_2 d_1)\vec{u}_z$$
$$= e_1 \vec{u}_x + e_2 \vec{u}_y + e_3 \vec{u}_z$$

Let's look closely at e_1:

$$e_1 = a_2 d_3 - a_3 d_2$$

The components d_2 and d_3 result from the first cross product, $\vec{b} \times \vec{c}$, and have the values:

$$d_2 = b_3 c_1 - b_1 c_3$$
$$d_3 = b_1 c_2 - b_2 c_1$$

and:

$$e_1 = a_2(b_1 c_2 - b_2 c_1) - a_3(b_3 c_1 - b_1 c_3)$$

Now, some 'hocus pocus' … add and subtract $a_1b_1c_1$:

$$e_1 = a_2 b_1 c_2 - a_2 b_2 c_1 - a_3 b_3 c_1 + a_3 b_1 c_3 + a_1 b_1 c_1 - a_1 b_1 c_1$$

and collect terms in b_1:

$$e_1 = b_1(a_1 c_1 + a_2 c_2 + a_3 c_3) - c_1(a_1 b_1 + a_2 b_2 + a_3 b_3)$$
$$= b_1(\vec{a} \cdot \vec{c}) - c_1(\vec{a} \cdot \vec{b})$$

Similarly:

$$e_2 = b_2(\vec{a} \cdot \vec{c}) - c_2(\vec{a} \cdot \vec{b})$$
$$e_3 = b_3(\vec{a} \cdot \vec{c}) - c_3(\vec{a} \cdot \vec{b})$$

and from our original linear combination of $\vec{b}$ and $\vec{c}$:

$$\vec{e} = a\vec{b} + b\vec{c}$$
$$= \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})$$

as was originally asserted.

3.3 Reciprocal Basis Vectors

Most of the time it is convenient to use basis sets that are orthogonal and normalized. Most of us are used to the standard Cartesian axes which are defined by orthogonal vectors. However, it is not strictly necessary to do so and there are situations in which one might contemplate the use of basis vectors that are not orthogonal and possibly not normalized as, for example, in crystallography. In 3D space it is necessary only that the three basis vectors be non-coplanar as are vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ in figure 3-10. Using these three vectors we may construct any other vector in 3D Cartesian space.

3.4 Matrix Notation

We can use matrix notation to compactly represent our vectors. Rather than using:

$$\vec{a}=3\vec{u}_x+6\vec{u}_y+\vec{u}_z$$

we can simply use:

$$\vec{a}=\begin{bmatrix}3\\6\\1\end{bmatrix} \qquad [3\text{-}18]$$

The inner product of two vectors:

$$\begin{gathered}\vec{a}\cdot\vec{b}=(3\vec{u}_x+6\vec{u}_y+\vec{u}_z)\cdot(\vec{u}_x-5\vec{u}_y+2\vec{u}_z)\\ =3\vec{u}_x\cdot\vec{u}_x-15\vec{u}_x\cdot\vec{u}_y+6\vec{u}_x\cdot\vec{u}_z+6\vec{u}_y\cdot\vec{u}_x-30\vec{u}_y\cdot\vec{u}_y+12\vec{u}_y\cdot\vec{u}_z+\vec{u}_z\cdot\vec{u}_x-5\vec{u}_z\cdot\vec{u}_y+2\vec{u}_z\cdot\vec{u}_z\\ =3-0+0+0-30+0+0-0+2\\ =-25\end{gathered}$$

is represented in matrix notation by:

$$\boldsymbol{a}^T \cdot \boldsymbol{b} = \begin{bmatrix} 3 & 6 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -5 \\ 2 \end{bmatrix} = (-25) = -25 \qquad [3\text{-}19]$$

Here, superscript "T" indicates the transpose of the corresponding column vector. Many texts leave this out but if we are thinking 'matrices' when looking at vectors this notation serves as a reminder that this operation is necessary in order to do the inner product calculation. Note that in order to accomplish the dot-product multiplication the first is a single-row matrix and the second a single-column matrix and that transforming from a column matrix to a row matrix is the equivalent of taking the transpose of the column matrix. We shall continue to use this notation throughout the text where it serves to remove any ambiguity however it is understood that the dot product notation implies the transpose of the first of the two vectors.

Orthogonality can be easily demonstrated from the dot product of two vectors at right angles to each other:

$$\begin{bmatrix} 0 & 6 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = (0) = 0$$

This is the matrix equivalent of writing:

$$6 \vec{\boldsymbol{u}}_y \cdot \vec{\boldsymbol{u}}_x = 0$$

What happens if, instead of multiplying a row vector by a column vector, we reverse the order of multiplication? Let's give it a try:

$$\boldsymbol{a}\,\boldsymbol{b}^T = \begin{bmatrix} 3 \\ 6 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -5 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -15 & 6 \\ 6 & -30 & 12 \\ 1 & -5 & 2 \end{bmatrix} \qquad [3\text{-}20]$$

This is obviously *not* the scalar that we are used to by now but rather a new matrix. This operation is called the *outer product*. And is considered to be a special case of the Kronecker product (see [2-3]). If we do the outer product of a m x 1 vector and a 1 x n vector we will obtain an m x n matrix. In the situation in which neither of the vectors is labeled as being transposed we use the " $\otimes$ " operator here to clearly distinguish between the inner product and the outer product when which one is the transposed vector is not obvious:

$$\boldsymbol{a}\,\boldsymbol{b}^T = \boldsymbol{a} \otimes \boldsymbol{b}$$

As with our dot product, it is now understood from this notation that the *second* of the two vectors is transposed.

For our familiar **R**^{2} and **R**^{3} spaces the outer product is the matrix product of a column matrix with 2 or three rows and a row matrix with 2 or 3 columns: The resulting matrix is a 2x2 or 3x3 matrix which is not a normal 2D or 3D vector.

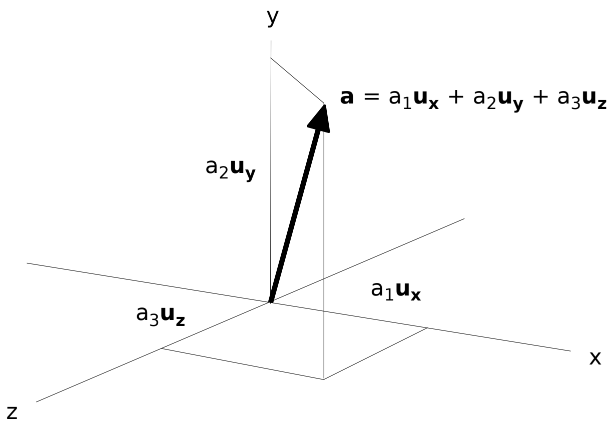


Figure 3-13:Components of vector ***a***

The outer product is of some interest later on so we will investigate it a bit further. Consider a vector **a** in **R**^{3} where **a** is, as usual, composed of a linear combination of the basis vectors **u**$_x$, **u**$_y$ and **u**$_z$ (figure 3-13). Now, consider what we get when we do an inner product calculation using one of our basis vectors, **u**$_x$ say, and **a**:

$$\vec{u}_x^T \cdot \vec{a} = \vec{u}_x^T \cdot (a_1 \vec{u}_x + a_2 \vec{u}_y + a_3 \vec{u}_z)$$
$$= a_1 \vec{u}_x^T \cdot \vec{u}_x + a_2 \vec{u}_x^T \cdot \vec{u}_y + a_3 \vec{u}_x^T \cdot \vec{u}_z$$
$$= a_1$$

Similarly, for **u**$_y$ and **u**$_z$ we have:

$$\vec{u}_y^T \cdot \vec{a} = a_2$$
$$\vec{u}_z^T \cdot \vec{a} = a_3$$

Now, consider:

$$\vec{u}_x (\vec{u}_x^T \vec{a}) = a_1 \vec{u}_x \qquad [3\text{-}21]$$

With reference to figure 3-13 you can see that what we have done is generate the *projection* of **a** onto the x-axis and have generated a vector that is proportional to the unit vector along the x-axis.

If we define the outer product:

$$\hat{P}_x = \vec{u}_x \vec{u}_x^T$$
$$\hat{P}_y = \vec{u}_y \vec{u}_y^T \qquad [3\text{-}22]$$
$$\hat{P}_z = \vec{u}_z \vec{u}_z^T$$

then [3-21] becomes:

$$\hat{P}_x \vec{a} = a_1 \vec{u}_x \qquad [3\text{-}23a]$$

Similarly:

$$\hat{\boldsymbol{P}}_y \vec{\boldsymbol{a}} = a_2 \vec{\boldsymbol{u}}_y \qquad [3\text{-}23b]$$
$$\hat{\boldsymbol{P}}_z \vec{\boldsymbol{a}} = a_3 \vec{\boldsymbol{u}}_z$$

$\mathbf{P_x}$, $\mathbf{P_y}$ and $\mathbf{P_z}$ generate the projection of a vector onto the corresponding unit basis vector's axis and are called *projection operators*. Note that they are *not* written here as vectors because they are not vectors but rather operators. For our unit vectors in $\mathbf{R^3}$:

$$\hat{\boldsymbol{P}}_x = \vec{\boldsymbol{u}}_x \cdot \vec{\boldsymbol{u}}_x^T = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\hat{\boldsymbol{P}}_y = \vec{\boldsymbol{u}}_y \cdot \vec{\boldsymbol{u}}_y^T = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\hat{\boldsymbol{P}}_z = \vec{\boldsymbol{u}}_z \cdot \vec{\boldsymbol{u}}_z^T = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Notice that:

$$\hat{\boldsymbol{P}}_x + \hat{\boldsymbol{P}}_y + \hat{\boldsymbol{P}}_z = \mathbf{1} \qquad [3\text{-}24]$$
$$\sum_i \hat{\boldsymbol{P}}_i = \mathbf{1}$$

This is generally the case with projection operators constructed from orthonormal basis sets and is referred to as the completeness *theorem* (which allows us to determine whether or not we have a complete set of basis vectors .. hence the name). We shall encounter this property later. In fact, you could think of [3-24] as saying that all of the possible projection operators for a particular space put together project a vector onto itself:

$$(\hat{\boldsymbol{P}}_x + \hat{\boldsymbol{P}}_y + \hat{\boldsymbol{P}}_z)\vec{\boldsymbol{a}} = \hat{\mathbf{1}}\vec{\boldsymbol{a}} = \vec{\boldsymbol{a}}$$

From our visual picture of the projection operator and what it does it should be no surprise that the projection of a unit basis vector onto itself is:

$$\hat{\boldsymbol{P}}_i \vec{\boldsymbol{u}}_i = \vec{\boldsymbol{u}}_i$$

With this in mind and using [3-23]:

$$\begin{aligned} \hat{\boldsymbol{P}}_i^2 \vec{\boldsymbol{a}} &= \hat{\boldsymbol{P}}_i (\hat{\boldsymbol{P}}_i \vec{\boldsymbol{a}}) \\ &= \hat{\boldsymbol{P}}_i (a_i \vec{\boldsymbol{u}}_i) \\ &= a_i \hat{\boldsymbol{P}}_i \vec{\boldsymbol{u}}_i \\ &= a_i \vec{\boldsymbol{u}}_i \end{aligned}$$

from which we infer that:

$$\hat{\boldsymbol{P}}_i^2 = \hat{\boldsymbol{P}}_i$$

and we say that $\hat{\boldsymbol{P}}_i$ is *idempotent*.

Again, from our visual picture of what the projection operator does, we can predict that $\hat{\boldsymbol{P}}_i \hat{\boldsymbol{P}}_j \vec{\boldsymbol{a}}$ (where i and j refer to orthogonal basis vectors) should be zero:

$$\begin{aligned} \hat{\boldsymbol{P}}_i \hat{\boldsymbol{P}}_j \vec{\boldsymbol{a}} &= \vec{\boldsymbol{u}}_i \vec{\boldsymbol{u}}_i^T \vec{\boldsymbol{u}}_j \vec{\boldsymbol{u}}_j^T \vec{\boldsymbol{a}} \\ &= \vec{\boldsymbol{u}}_i (\vec{\boldsymbol{u}}_i^T \vec{\boldsymbol{u}}_j) \vec{\boldsymbol{u}}_j^T \vec{\boldsymbol{a}} \\ &= \vec{\boldsymbol{u}}_i (0) \vec{\boldsymbol{u}}_j^T \vec{\boldsymbol{a}} \\ &= 0 \end{aligned}$$

We need not restrict ourselves to using basis set vectors to define projection operators. We can be completely general and say that:

$$\hat{\boldsymbol{P}}_b = \vec{\boldsymbol{b}} \vec{\boldsymbol{b}}^T$$

and this, when operating on a vector **a** will constitute a projection of vector **a** onto vector **b**:

$$\begin{aligned} \vec{\boldsymbol{b}}^T \vec{\boldsymbol{a}} &= k \\ \hat{\boldsymbol{P}}_b \vec{\boldsymbol{a}} &= \vec{\boldsymbol{b}} (\vec{\boldsymbol{b}}^T \vec{\boldsymbol{a}}) \\ &= k \vec{\boldsymbol{b}} \end{aligned}$$

and the projection of **a** onto **b** is a scalar times **b** and is obviously parallel to **b**. What is the value of k and how can we use it?

Consider vectors **u** and **v**:

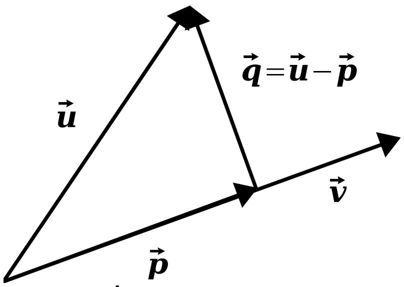


*Figure 3-14: Projection of **u** on **v***

Vector **p** is the projection of **u** on **v** and **q** = **u**-**p** is at right angles to **v** and **p** (orthogonal). Vector **p** is also proportional to **v**:

$$\vec{p}=k\vec{v}$$

Since **q** is orthogonal to **v** we can write:

$$\begin{aligned}\vec{v}\cdot\vec{q}&=0\\ &=\vec{v}\cdot(\vec{u}-\vec{p})\\ &=\vec{v}\cdot\vec{u}-\vec{v}\cdot\vec{p}\\ &=\vec{v}\cdot\vec{u}-k\vec{v}\cdot\vec{v}\end{aligned}$$

Adding $k\vec{v}\cdot\vec{v}$ to each side:

$$k\vec{v}\cdot\vec{v}=\vec{v}\cdot\vec{u}$$

and:

$$k=\frac{\vec{v}\cdot\vec{u}}{\vec{v}\cdot\vec{v}}$$

Therefore:

$$\boxed{\vec{p}=k\vec{v}=\frac{\vec{v}\cdot\vec{u}}{\vec{v}\cdot\vec{v}}\vec{v}} \qquad [3\text{-}25]$$

Equation [3-25] gives us a tool for finding an orthogonal basis set via the Gram-Schmidt orthogonalization procedure. For example, a two dimensional vector set in $\mathbf{R}^2$ is, in matrix notation:

$$\vec{e}_1=\begin{bmatrix}1\\1\end{bmatrix},\ \vec{e}_2=\begin{bmatrix}1\\0\end{bmatrix}$$

To test whether or not this is a basis set it is necessary to demonstrate that they span $\mathbf{R}^2$ and that they are linearly independent. To demonstrate the first property we imagine a vector, **a**, in $\mathbf{R}^2$ that is a linear combination of the two:

$$\vec{a}=\alpha_1\vec{e}_1+\alpha_2\vec{e}_2$$

or:

$$\begin{bmatrix}a_1\\a_2\end{bmatrix}=\begin{bmatrix}\alpha_1+\alpha_2\\\alpha_1+0\end{bmatrix}$$

This system of equations is solvable since the determinant of the coefficient matrix of the right side is non-singular (see equation [2-32]). In other words:

$$\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

is non-zero. Thus, when we solve for the a's we get:

$$\alpha_1 = a_2$$
$$\alpha_2 = a_1 - a_2$$

That these vectors are also linearly independent is shown by:

$$\alpha_1 \vec{\boldsymbol{e}}_{\mathbf{1}} + \alpha_2 \vec{\boldsymbol{e}}_{\mathbf{2}} = 0$$

or:

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

which can only be so if a_1 and a_2 are both equal to zero. Hence $\mathbf{e_1}$ and $\mathbf{e_2}$ are linearly independent and form a spanning set and therefore form a basis set. They are not however, orthogonal:

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \cdot 1 + 0 = 1$$

We can construct an orthogonal basis set using the Gram-Schmidt orthogonalization method. We first normalize $\mathbf{e_1}$:

$$\vec{\boldsymbol{e}}_{\mathbf{1}} \cdot \vec{\boldsymbol{e}}_{\mathbf{1}} = |e_1|^2 = 2$$
$$\vec{\boldsymbol{e}}_{\mathbf{1}}{}' = \frac{\vec{\boldsymbol{e}}_{\mathbf{1}}}{|e_1|}$$
$$= \frac{\vec{\boldsymbol{e}}_{\mathbf{1}}}{\sqrt{2}}$$

so that:

$$\vec{\boldsymbol{e}}_{\mathbf{1}}{}' \cdot \vec{\boldsymbol{e}}_{\mathbf{1}}{}' = 1$$

Now, since we are looking for an orthogonal basis set we project $\mathbf{e_2}$ onto $\mathbf{e_1}'$ using [3-24]:

$$\vec{p}=\frac{\vec{e}_1{}'\cdot\vec{e}_2}{\vec{e}_1{}'\cdot\vec{e}_1{}'}\vec{e}_1{}'$$

$$\vec{e}_1{}'=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$$

$$\vec{e}_1{}'\cdot\vec{e}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}$$

$$=\frac{1}{\sqrt{2}}$$

and:

$$\vec{p}=\frac{1}{\sqrt{2}}\vec{e}_1{}'$$

Obviously, **p** is not orthogonal to $\mathbf{e_1'}$ but rather, is parallel. $\mathbf{e_2}$ – **p** however *is* orthogonal to $\mathbf{e_1'}$:

$$\vec{e}_1{}'\cdot(\vec{e}_2-\vec{p})$$

$$=\vec{e}_1{}'\cdot\vec{e}_2-\vec{e}_1{}'\cdot\vec{p}$$

$$=\vec{e}_1{}'\cdot\vec{e}_2-\frac{1}{\sqrt{2}}\vec{e}_1{}'\cdot\vec{e}_1{}'$$

$$=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}-\frac{1}{\sqrt{2}}$$

$$=0$$

So, our second orthogonal basis vector is:

$$\vec{e}_2^{\,O}=\vec{e}_2-\vec{p}$$

$$=\vec{e}_2-\frac{\vec{e}_1{}'\cdot\vec{e}_2}{\vec{e}_1{}'\cdot\vec{e}_1{}'}\vec{e}_1{}'$$

$$=\vec{e}_2-\frac{1}{\sqrt{2}}\vec{e}_1{}'$$

$$=\vec{e}_2-\frac{1}{2}\vec{e}_1$$

$$=\begin{bmatrix}1\\0\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\\frac{-1}{2}\end{bmatrix}$$

where $\vec{e}_2^{\,O}$ is orthogonal to $\vec{e}_1$ but not normalized. We can easily normalize $\vec{e}_2^{\,O}$:

$$\vec{e}_2{}' = \frac{\vec{e}_2^{O}}{|e_2^{O}|}$$
$$= \sqrt{2}\,\vec{e}_2^{O}$$
$$= \sqrt{2}\begin{bmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{bmatrix}$$

and our (now) orthonormal basis set is:

$$\vec{e}_1{}' = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \vec{e}_2{}' = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{bmatrix}$$

You can verify for yourself that they are indeed now both orthogonal and normalized.

Thus we can always construct an orthogonal basis set from a non-orthogonal one via projection operations.

3.5 Vector Transformations

A given vector **a** can be transformed to a new vector **b** with the application of an *operator*:

$$\vec{b} = \hat{O}\,\vec{a}$$

We have already seen this in the case of the projection operator. Generally, the operator **O** can be represented by an n x n matrix, “n” being the dimension of the vector space. Again, we have seen this in the case of the projection operator above.

There are two things an operation on a vector can do. The length of the vector can be changed, accomplished by simple multiplication by a scalar and the direction of the vector can be changed by rotating it. These two fundamental operations can be done simultaneously. The projection operator does this by creating a new vector that is parallel to another vector (a rotation) and which has a different length from the starting vector.

An example of an operation that can be performed on a vector is a rotation. We apply the operator to vector **a** to produce vector **b** which is rotated in the vector space by the specified angle. This, of course requires a Euclidean space since angles and vector lengths are not defined in the more basic definition of vectors. Let's look at a simple example in two dimensional **R²** space. We start with a vector a that is for simplicity, along the y-axis:

$$\vec{a}=a_y\vec{u}_y$$

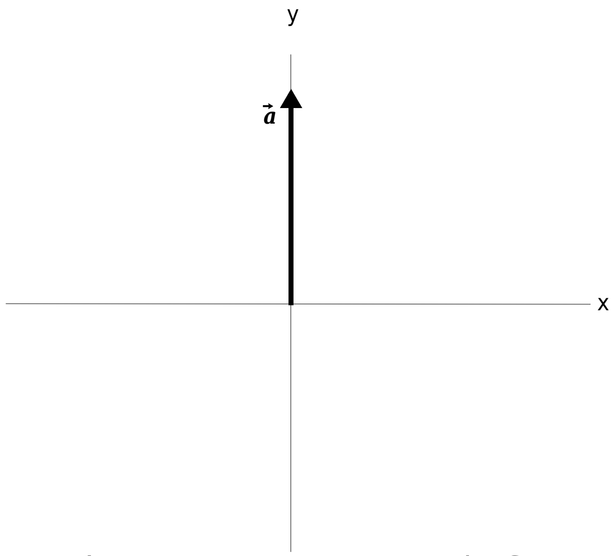


*Figure 3-15:Vector **a** before transformation*

We now apply a rotation in the form of an operator, **R**(90), so that the new vector **b** lies along the x-axis:

$$\vec{b}=\hat{R}(90)\vec{a} \quad [3\text{-}25]$$

The new vector **b** is of the form:

$$\vec{b}=a_y\vec{u}_x$$

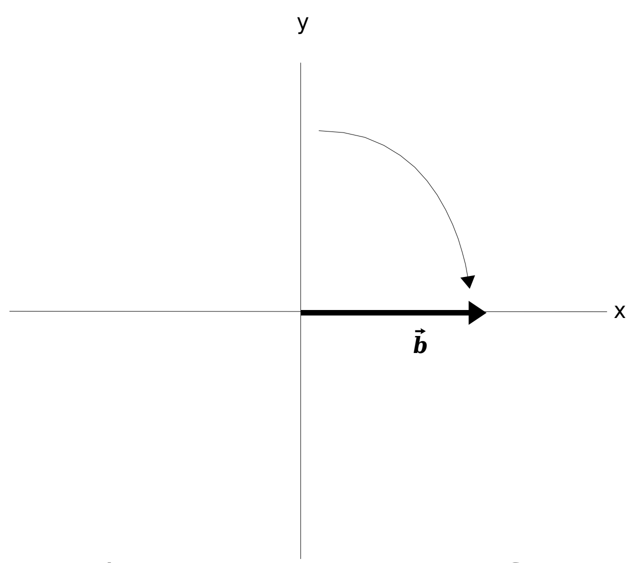


*Figure 3-16:Vector **a** after transformation to **b***

Note that the new vector, **b**, has the same length (or norm) as the old vector, a_y, as we would expect from a simple rotation of a vector. This is referred to as *invariance of the norm* and we will explore this in more detail later when we discuss tensors.

What will the matrix form of the rotation operator look like? The operator must be such that the matrix representation of **a** becomes the matrix representation of **b**. A matrix that does this is:

$$\hat{\boldsymbol{R}}(90)=\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix} \qquad [3\text{-}26a]$$

We can now write [3-25] in matrix form:

$$\begin{bmatrix}a_y\\0\end{bmatrix}=\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}\begin{bmatrix}0\\a_y\end{bmatrix}$$

R(-90) will, of course, be a rotation by 90 degrees in the opposite direction and its matrix will be:

$$\hat{\boldsymbol{R}}(-90)=\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix} \qquad [3\text{-}26b]$$

the transpose of **R(90)**. Thus, the application of **R(90)** followed by **R(-90)** will bring a vector back to its original place. In other words, sequential application of these two operators is equivalent to using the identity operator:

$$\hat{\boldsymbol{R}}(90)\hat{\boldsymbol{R}}(-90)=\hat{\boldsymbol{R}}(-90)\hat{\boldsymbol{R}}(90)=\hat{\mathbf{1}}$$

R(90) and **R(-90)** are *inverses* of each other.

A more general clockwise rotation operator in $\mathbf{R}^2$ space is:

$$\hat{\boldsymbol{R}}(\phi)=\begin{bmatrix}\cos(\phi) & \sin(\phi)\\-\sin(\phi) & \cos(\phi)\end{bmatrix}$$

It should be noted that one can view these operations as either rotations of vectors *or* as rotations of the coordinate system (in the opposite direction). It is often the case that the latter is the point of view taken and may be referred to as a *change of basis*.

A slightly different way to write the rotation is to use the cosines of the new coordinate axes with the old axes. Thus, we would write for for the transformation from vector $\vec{\boldsymbol{a}}$ to $\vec{\boldsymbol{b}}$:

$$\vec{\boldsymbol{a}}=a_1\vec{\boldsymbol{u}}_x+a_2\vec{\boldsymbol{u}}_y+a_3\vec{\boldsymbol{u}}_z$$
$$\vec{\boldsymbol{b}}=b_1\vec{\boldsymbol{u}}_x{}'+b_2\vec{\boldsymbol{u}}_y{}'+b_3\vec{\boldsymbol{u}}_z{}'$$

$$b_1=\cos(\vec{\boldsymbol{u}}_x{}',\vec{\boldsymbol{u}}_x)a_1+\cos(\vec{\boldsymbol{u}}_x{}',\vec{\boldsymbol{u}}_y)a_2+\cos(\vec{\boldsymbol{u}}_x{}',\vec{\boldsymbol{u}}_z)a_3$$

where $\cos(\vec{\boldsymbol{u}}_i{}',\vec{\boldsymbol{u}}_j)$ is the angle between the new axis unit vectors $\vec{\boldsymbol{u}}_i{}'$ and the old axis unit vectors $\vec{\boldsymbol{u}}_j$ and is referred to as the *direction cosine* (see Appendix II). Similarly:

$$b_2 = \cos(\vec{u}_y{}', \vec{u}_x) a_1 + \cos(\vec{u}_y{}', \vec{u}_y) a_2 + \cos(\vec{u}_y{}', \vec{u}_z) a_3$$
$$b_3 = \cos(\vec{u}_z{}', \vec{u}_x) a_1 + \cos(\vec{u}_z{}', \vec{u}_y) a_2 + \cos(\vec{u}_z{}', \vec{u}_z) a_3$$

or more succinctly:

$$b_i = \sum_j^3 \cos(\vec{u}_i{}', \vec{u}_j) a_j \qquad [3\text{-}27]$$

This can be understood easily by using a 2D example. Let's suppose that we have a vector $\vec{a}$ in coordinate system A and that we transform the basis set such that the axes rotate by 90 degrees in the clockwise direction. We then have, as initial and final state:

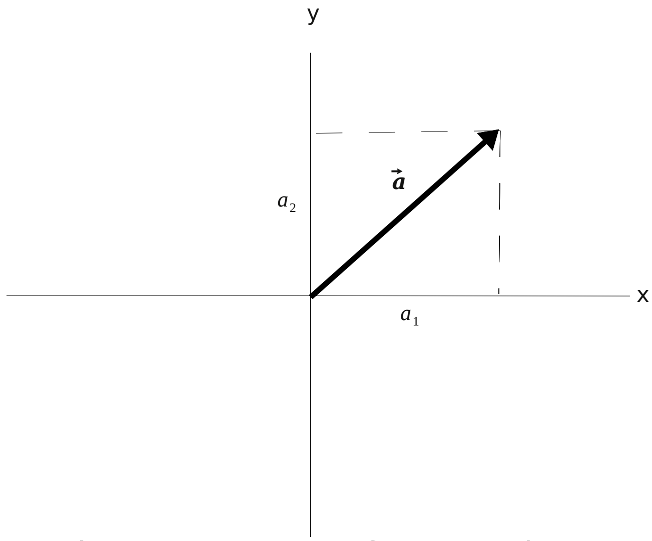


Figure 3-17a: Before 90 degree clockwise axis rotation

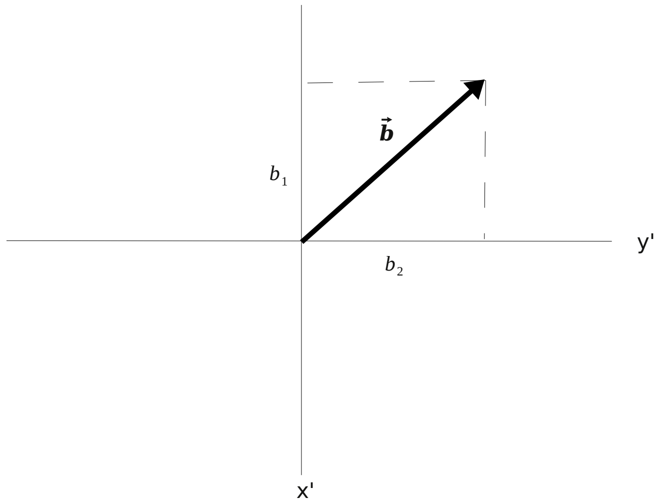


Figure 3-17b: After 90 degree clockwise axis rotation

The vector of course remains the same vector; only the coordinate axes have changed. We will have more to say about this when we discuss tensors. Our transformation equation indicates that:

$$b_1 = \cos(\vec{u}_x{}', \vec{u}_x) a_1 + \cos(\vec{u}_x{}', \vec{u}_y) a_2$$

If the new coordinate axes are superimposed on top of the old ones the angle between the new x-axis and the old x-axis is 90 degrees and its cosine value is zero. The angle between the new x-axis and the old y-axis is 180 degrees with a cosine value of -1:

$$b_1 = 0 \cdot a_1 - 1 \cdot a_2$$

Similarly:

$$b_2 = 1 \cdot a_1 + 0 \cdot a_2$$

Note that we can express this as a matrix equation:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

which is identical to our matrix equation 3-26 in which we rotated the vector by -90 degrees. Of course, a counterclockwise rotation of a vector is the equivalent of a clockwise rotation of the axes. Equation 3-27 is of convenience in the discussion of tensors.

From the discussion of matrices in chapter 2 there is another type of operator that we can consider; the eigenvalue operator. That is, the operator which when applied to a vector produces a new version of the same vector scaled by some value called the eigenvalue. The vector is called the eigenvector.

$$\hat{O}\vec{a}=\lambda\vec{a} \qquad [3\text{-}28]$$

Everything that was discussed about eigenvalue equations with respect to matrices applies here to vectors. The astute reader will be, by now, beginning to suspect that vectors and matrices are connected in many subtle and useful ways. This is so ... it is very useful to think of vectors in their matrix forms when studying their properties.

In our use of operators with vectors we assume linearity. That is:

$$\begin{gathered}\hat{O}(\vec{a}+\vec{b})=\hat{O}\vec{a}+\hat{O}\vec{b}\\(O+P)\vec{a}=O\vec{a}+P\vec{a}\\OP\vec{a}=O(P\vec{a})\\\hat{O}(\hat{P}\hat{Q})=(\hat{O}\hat{P})\hat{Q}\end{gathered}$$

We shall see that a *very* important aspect of operators is their commutative properties:

$$[\hat{O},\hat{P}]=\hat{O}\hat{P}-\hat{P}\hat{O} \qquad [3\text{-}29]$$

If **OP** = **PO** and [**O**,**P**] = 0 then we say that the operators *commute*. Since we have learned that matrices in general don't commute we expect to find the same behaviour in our vector operators.

3.6 Dirac Notation and Operator Matrices

We continue our exploration of operators but with a new notation in hand; the Dirac notation. This is a shorthand way of representing vectors introduced by Paul Dirac in 1947 and is very frequently used today.

We represent a vector **a** as:

$$\vec{a}\equiv|a> \qquad [3\text{-}30]$$

This is called a “ket” and in this particular case we would say “ket a”. Note that we are not using boldface inside the ket since this would seem to be redundant. If it is inside the ket then, by definition, it is a vector and we need not explicitly indicate it as we have been doing so far. Thus, our eigenvalue equation, for example, becomes:

$$\hat{O}|a>=\lambda|a>$$

What about the inner product? Recall that the inner product of vector **a** with itself is:

$$\vec{\boldsymbol{a}}^T \cdot \vec{\boldsymbol{a}} = \|\vec{\boldsymbol{a}}\|^2$$

where we multiply the transpose (thinking of the matrix version of vectors) of **a** by **a** to get a scalar quantity. It is defined this way since we want to get a real, non-negative value for **a**. In Dirac notation the transpose is:

$$\vec{\boldsymbol{a}}^T \equiv < a \,|$$

which is called the “bra” and we would say “bra a”. The inner product is written as:

$$< a \,|\, a > = \|a\|^2$$

which is, of course, the square of the length or norm of the vector. Can you guess how we say this? “bracket a” of course. Professor Dirac had a bit of a sense of humour[2].

From the above definitions:

$$|\, a > = (< a \,|\,)^T$$

An operator can also operate on the bra:

$$< a' \,| = < a \,|\, \boldsymbol{O}$$

Note that in the case of the ket the operator is placed to the left of the vector and operates left-to-right. For the bra, the operator is placed to the right and operates right-to-left (or at least, appears to operate right-to-left .. see below).

The Dirac notation is useful because it is very general and applies to *any* vectors for which an inner product is defined, not just our familiar two and three dimensional ones. Also useful is that in order to understand a concept in an abstract vector space we need only ask ourselves what would the case be in simple 3D space.

There is another generalization that needs to be introduced. Up until now we have been using numbers from the set of real numbers for our coefficients of the unit vectors:

2 To some this must have been somewhat surprising. Professor Dirac had a reputation for being a very quiet, non-verbose person who simply worked away at his work. However, he did love to read comics and watch Bugs Bunny cartoons and must therefore have enjoyed humour.

$$|a> = a_x|u_x> + a_y|u_y> + a_z|u_z>$$
$$a_i \in \mathbb{R}$$

We wish to now generalize this so that the coefficients are in the set of complex numbers:

$$|a> = a_x|u_x> + a_y|u_y> + a_z|u_z>$$
$$a_i \in \mathbb{C}$$

From here on we will assume complex numbers in order to be completely general and that the vector spaces are defined as **C^n** indicating n-dimensional vector spaces using complex numbers.

This change requires some minor alterations in how we do our inner product calculations. Since we are still looking for a real, non-negative number the bra portion of the product now represents the transpose of the complex conjugate:

$$|a> = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$
$$<a| = \begin{bmatrix} a_1^* & a_2^* & \cdots & a_n^* \end{bmatrix}$$

and:

$$<a|a> = \begin{bmatrix} a_1^* & a_2^* & \cdots & a_n^* \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$
$$= \sum_i^n a_i^* a_i = |a|^2 \quad (eq.\ [1-6])$$

We can multiply our bra and ket by a scalar however we must remember the complex conjugate nature of the bra:

$$\begin{gathered} \alpha|a> = |\alpha a> \\ \alpha<a| = <a\alpha^*| \\ <\alpha a| = \alpha^*<a| \end{gathered} \qquad [3\text{-}31]$$

Also:

$$|a> = (<a|)^* \qquad [3\text{-}32a]$$

and:

$$<a\,|\,b>=(<b\,|\,a>)^* \qquad [3\text{-}32b]$$

Note that "*" in the context of the Dirac notation refers to the transpose of the complex conjugate. We can best illustrate this with a 2 x 2 matrix example:

$$|\,a>=\begin{bmatrix}\alpha_1+\beta_1 i\\ \alpha_2+\beta_2 i\end{bmatrix} \quad |\,b>=\begin{bmatrix}\alpha_3+\beta_3 i\\ \alpha_4+\beta_4 i\end{bmatrix}$$

$$\begin{aligned}<a\,|\,b>&=\begin{bmatrix}\alpha_1-\beta_1 i & \alpha_2-\beta_2 i\end{bmatrix}\begin{bmatrix}\alpha_3+\beta_3 i\\ \alpha_4+\beta_4 i\end{bmatrix}\\ &=(\alpha_1-\beta_1 i)(\alpha_3+\beta_3 i)+(\alpha_2-\beta_2 i)(\alpha_4+\beta_4 i)\\ &=\alpha_1\alpha_3-\beta_1\beta_3 i^2-\alpha_3\beta_1 i+\alpha_1\beta_3 i+\alpha_2\alpha_4-\beta_2\beta_4 i^2-\alpha_4\beta_2 i+\alpha_2\beta_4 i\\ &=(\alpha_1\alpha_3+\beta_1\beta_3+\alpha_2\alpha_4+\beta_2\beta_4)+(\alpha_1\beta_3-\alpha_3\beta_1+\alpha_2\beta_4-\alpha_4\beta_2)i\end{aligned}$$

$$\begin{aligned}<b\,|\,a>&=\begin{bmatrix}\alpha_3-\beta_3 i & \alpha_4-\beta_4 i\end{bmatrix}\begin{bmatrix}\alpha_1+\beta_1 i\\ \alpha_2+\beta_2 i\end{bmatrix}\\ &=(\alpha_3-\beta_3 i)(\alpha_1+\beta_1 i)+(\alpha_4-\beta_4 i)(\alpha_2+\beta_2 i)\\ &=\alpha_3\alpha_1-\beta_3\beta_1 i^2-\alpha_1\beta_3 i+\alpha_3\beta_1 i+\alpha_4\alpha_2-\beta_4\beta_2 i^2-\alpha_2\beta_4 i+\alpha_4\beta_2 i\\ &=(\alpha_3\alpha_1+\beta_3\beta_1+\alpha_4\alpha_2+\beta_4\beta_2)-(\alpha_1\beta_3-\alpha_3\beta_1+\alpha_2\beta_4-\alpha_4\beta_2)i\\ &=<a\,|\,b>^*\end{aligned}$$

This is different from the normal two and three dimensional vector dot product and is referred to as *skew-symmetry* and results in some unusual properties of inner products involving complex numbers. For example if the ket is a linear combination of vectors and we take an inner product, all is as we suspect it should be:

$$\begin{aligned}&<a\,|\,\alpha b+\beta c>\\ &=<a\,|\,\alpha b>+<a\,|\,\beta c>\\ &=\alpha<a\,|\,b>+\beta<a\,|\,c>\end{aligned}$$

However, if the linear combination of vectors is in the bra then the result is not quite as expected:

$$\begin{aligned}<\alpha b+\beta c\,|\,a>&=<a\,|\,\alpha b+\beta c>^*\\ &=(\alpha<a\,|\,b>+\beta<a\,|\,c>)^*\\ &=\alpha^*<b\,|\,a>+\beta^*<c\,|\,a>\end{aligned}$$

using [3-32] in the first and last lines and [3-31] in the last line. So, the skew-symmetry property results in an *anti-symmetry* property with respect to the bra. Also, from [3-32]:

$$<a\,|\,a>=(<a\,|\,a>)^*$$

which means that the inner product of **a** with itself is real as we know it should be.

We now turn our attention back to operators; specifically we want to produce a general way of writing the matrix form of an operator. To do so we work with our unit vectors since any vector in our vector space can be built up from a linear combination of them and they are thus the 'lowest common denominator' so to speak. We assume that the operator produces a new vector that is in the same vector space as the old vector. Consider then, the effect of an operator on a unit vector:

$$|u_i'> = \hat{O}|u_i>$$

Now let's form an inner product of this result with another unit vector:

$$<u_j|u_i'> = <u_j|\hat{O}|u_i> = \hat{O}_{ji} \qquad [3\text{-}33]$$

We use this to find our matrix representation of operator **O**:

$$\hat{O} = \begin{bmatrix} \hat{O}_{11} & \hat{O}_{12} & \cdots \\ \hat{O}_{21} & \hat{O}_{22} & \cdots \\ \vdots & \vdots & \ddots \\ \hat{O}_{n1} & \hat{O}_{n2} & \cdots \end{bmatrix} \qquad [3\text{-}34]$$

Note that we could just as well have done this 'backwards' with the operator acting on a bra unit vector instead:

$$<u_i|\hat{O} = <u_i'|$$
$$<u_i'|u_j> = <u_i|\hat{O}|u_j>$$

Thus, the expression:

$$<u_i|\hat{O}|u_j>$$

can refer to the operator acting in either direction on the bra or the ket.

We can use equation [3-33] to construct our **R(90)** operator from above. Recall that **R(90)** means rotate by 90 degrees clockwise; thus unit vector $\mathbf{u_x}$, for example, becomes -$\mathbf{u_y}$ etc.:

$$\hat{\boldsymbol{R}}(90)=\begin{bmatrix} \hat{\boldsymbol{R}}_{11} & \hat{\boldsymbol{R}}_{12} \\ \hat{\boldsymbol{R}}_{21} & \hat{\boldsymbol{R}}_{22} \end{bmatrix}$$

$$=\begin{bmatrix} <u_x \mid \hat{\boldsymbol{R}} \mid u_x> & <u_x \mid \hat{\boldsymbol{R}} \mid u_y> \\ <u_y \mid \hat{\boldsymbol{R}} \mid u_x> & <u_y \mid \hat{\boldsymbol{R}} \mid u_y> \end{bmatrix}$$

$$=\begin{bmatrix} -<u_x \mid u_y> & <u_x \mid u_x> \\ -<u_y \mid u_y> & <u_y \mid u_x> \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Now that we know how to construct the matrix form of the operator it must be said that there is another way to look at the rotation operator (and by extension, operators in general). We have been looking at it as though it causes a rotation in the vector however it is also possible to look at it as causing a rotation of the basis vectors, the vector itself remaining motionless.

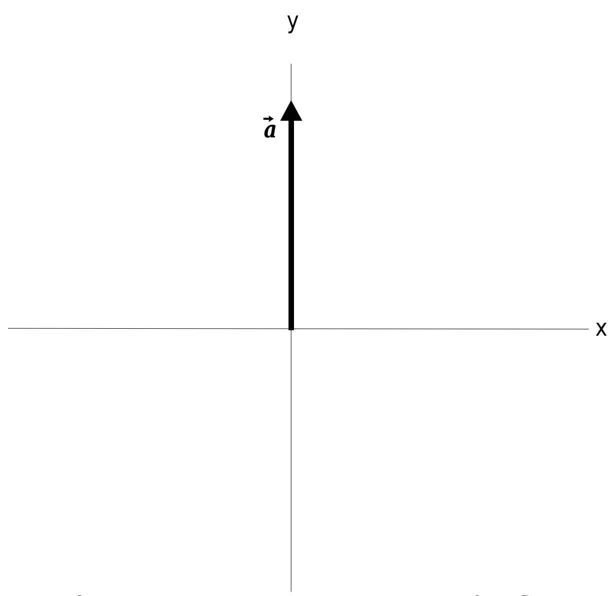


*Figure 3-18: Vector **a** before basis set transformation*

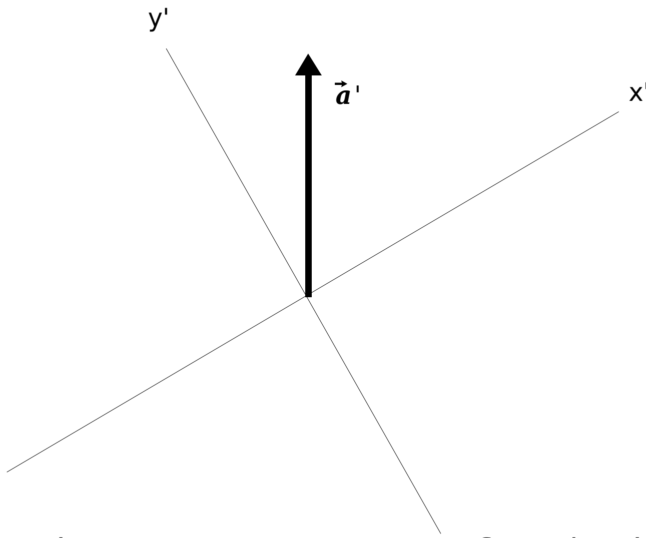


*Figure 3-19: Vector **a** after basis set transformation*

Thus, vector **a** remains motionless but attains a new set of coordinates in the new basis set:

$$\vec{\boldsymbol{a}}' = \hat{\boldsymbol{R}}(\theta)\vec{\boldsymbol{a}}$$
$$\begin{bmatrix} a_x' \\ a_y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix}$$

So, what we viewed earlier as a clockwise motion of the vector can also be viewed as a counterclockwise motion of the basis vectors. This is referred to as a *change of basis* and is generalized:

$$\begin{gathered} e_1' = a_{11} e_1 + a_{12} e_2 \cdots a_{1\mathrm{n}} e_n \\ e_2' = a_{21} e_1 + a_{22} e_2 \cdots a_{2\mathrm{n}} e_n \\ e_3' = a_{31} e_1 + a_{32} e_2 \cdots a_{3\mathrm{n}} e_n \\ \vdots \\ e_n' = a_{n1} e_1 + a_{n2} e_2 \cdots a_{nn} e_n \end{gathered}$$

or in matrix notation:

$$\begin{bmatrix} e_1' \\ e_2' \\ e_3' \\ \vdots \\ e_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2\mathrm{n}} \\ a_{11} & a_{32} & a_{33} & \cdots & a_{3\mathrm{n}} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

where the e_i are the old basis set and the e_i' are the new basis set. One can see that for any given vector there can be an infinite number of basis sets corresponding to all possible rotations.

You would be quite correct to ask 'are the new basis vector orthogonal if the old ones were?' for our rotation example. Visually, from figure [3-14] you can see that they are but this really isn't proof. We can prove it though. We must show using equation [3-11] that:

$$\vec{\boldsymbol{u}}_x{}'\cdot\vec{\boldsymbol{u}}_y{}'=0$$

The transformation from the old basis to the new basis is:

$$\begin{bmatrix} u_x{}' \\ u_y{}' \end{bmatrix}=\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} u_x \\ u_y \end{bmatrix}$$

or:

$$\begin{gathered} \vec{\boldsymbol{u}}_x{}'=\cos(\theta)\vec{\boldsymbol{u}}_x+\sin(\theta)\vec{\boldsymbol{u}}_y \\ \vec{\boldsymbol{u}}_y{}'=-\sin(\theta)\vec{\boldsymbol{u}}_x+\cos(\theta)\vec{\boldsymbol{u}}_y \end{gathered}$$

each new basis vector being a linear combination of the old basis vector set. Writing the inner product:

$$\begin{gathered} \vec{\boldsymbol{u}}_x{}'\cdot\vec{\boldsymbol{u}}_y{}'=(\cos(\theta)\vec{\boldsymbol{u}}_x+\sin(\theta)\vec{\boldsymbol{u}}_y)\cdot(-\sin(\theta)\vec{\boldsymbol{u}}_x+\cos(\theta)\vec{\boldsymbol{u}}_y) \\ =-\cos(\theta)\sin(\theta)\vec{\boldsymbol{u}}_x\cdot\vec{\boldsymbol{u}}_x+\cos(\theta)\cos(\theta)\vec{\boldsymbol{u}}_x\cdot\vec{\boldsymbol{u}}_y \\ -\sin(\theta)\sin(\theta)\vec{\boldsymbol{u}}_y\cdot\vec{\boldsymbol{u}}_x+\sin(\theta)\cos(\theta)\vec{\boldsymbol{u}}_y\cdot\vec{\boldsymbol{u}}_y \\ =-\sin(\theta)\cos(\theta)+0+0+\sin(\theta)\cos(\theta) \\ =0 \end{gathered}$$

Thus the new basis set is orthogonal as was the old one which means that the angle between basis vectors is invariant. We can similarly show that the new basis vectors are normalized if the old ones were:

$$\begin{gathered} \vec{\boldsymbol{u}}_x{}'\cdot\vec{\boldsymbol{u}}_x{}'=(\cos(\theta)\vec{\boldsymbol{u}}_x+\sin(\theta)\vec{\boldsymbol{u}}_y)\cdot(\cos(\theta)\vec{\boldsymbol{u}}_x+\sin(\theta)\vec{\boldsymbol{u}}_y) \\ =\cos^2(\theta)\vec{\boldsymbol{u}}_x\cdot\vec{\boldsymbol{u}}_x+2\sin(\theta)\cos(\theta)\vec{\boldsymbol{u}}_x\cdot\vec{\boldsymbol{u}}_y+\sin^2(\theta)\vec{\boldsymbol{u}}_y\cdot\vec{\boldsymbol{u}}_y \\ =1 \end{gathered}$$

using the orthonormal properties of the original basis set.

With reference to equation [[3-31] we draw the analogy between scalar multipliers and operators:

$$\begin{gathered} \hat{\boldsymbol{O}}\,|\,a>=|\,\hat{\boldsymbol{O}}a> \\ <a\,|\,\hat{\boldsymbol{O}}=<(\hat{\boldsymbol{O}}^T)^*a\,| \end{gathered} \qquad \text{[3-35a]}$$

or the equivalent:

$$|\hat{\boldsymbol{O}}a> = \hat{\boldsymbol{O}}|a>$$
$$<\hat{\boldsymbol{O}}a| = <a|(\hat{\boldsymbol{O}}^T)^* \qquad [3\text{-}35b]$$

There is a non-intuitive subtlety of notation in equation [3-35] that should be mentioned. As far as the ket goes, whether the operator is inside or outside doesn't affect anything much. The bra is a different story. To understand this think 'matrices' again. The expression:

$$<\hat{\boldsymbol{O}}a|$$

has **Oa** inside the bra. In terms of matrices this means multiply the column vector **a** by the square matrix **O** *inside the bra*. We could not write it as **aO** outside of the bra and have it make sense since the corresponding matrix multiplication doesn't work. However, because they are enclosed in the bra we take the complex transpose and use it for any further calculations. If we take **O** and **a** out of the bra altogether and use our original vector notation we must take the complex transpose (and reverse the order, equation [2-10]):

$$<\hat{\boldsymbol{O}}a| \rightarrow \vec{\boldsymbol{a}}^T(\hat{\boldsymbol{O}}^T)^*$$

or, if we remove only the operator:

$$<\hat{\boldsymbol{O}}a| = <a|(\hat{\boldsymbol{O}}^T)^*$$

In a sense one could view the bra as itself being an operator meaning 'take the complex transpose of whatever is inside me'. To convince yourself that this is so you need only set up a simple two dimensional vector/operator pair and see how it works.

There are special types of operators linked very closely to the special matrices mentioned in chapter 2. When dealing with real numbers only, an *orthogonal* operator is one whose transpose is equal to its inverse:

$$\hat{\boldsymbol{O}}^T = \hat{\boldsymbol{O}}^{-1} \qquad [3\text{-}36]$$

In this case the product of the operator and its transpose equals the unit operator:

$$\hat{\boldsymbol{O}}\hat{\boldsymbol{O}}^T = \hat{\boldsymbol{O}}^T\hat{\boldsymbol{O}} = \hat{\boldsymbol{1}}$$

We have already seen an example of this in our two dimensional rotation operator.

The next type of operator is the *unitary* operator and is associated with complex numbers in the vector. This is one in which the complex conjugate of the transpose of the operator equals the inverse of the operator:

$$(\hat{\boldsymbol{O}}^T)^* = \hat{\boldsymbol{O}}^{-1} \qquad [3\text{-}37]$$

and:

$$\hat{\boldsymbol{O}}(\hat{\boldsymbol{O}}^T)^* = (\hat{\boldsymbol{O}}^T)^* \hat{\boldsymbol{O}} = \hat{\mathbf{1}}$$

A bit of thought will reveal that if unitary operator matrix consists of all real numbers then it is also an orthogonal operator. These two operators are ones that leave the vector norms (lengths) and angles between vectors invariant. In equation form, using the unitary operator, this is:

$$<\hat{\boldsymbol{O}} a \mid \hat{\boldsymbol{O}} b> = <a \mid b>$$

This is easily shown to be so:

$$\begin{aligned} <\hat{\boldsymbol{O}} a \mid \hat{\boldsymbol{O}} b> &= <a \mid (\hat{\boldsymbol{O}}^T)^* \hat{\boldsymbol{O}} \mid b> \\ &= <a \mid \hat{\mathbf{1}} \mid b> \\ &= <a \mid b> \end{aligned}$$

The unitary operator has an interesting feature in that it's eigenvalues are all equal to 1. As usual, we can show this. We start with:

$$\lambda^* \lambda <a \mid a>$$

where |a> is a non-zero vector and l is the eigenvalue of operator **O**. Since the l's are scalars we can move them into the bra and ket:

$$<\lambda a | \lambda a>$$

using equation [3-31] in the bra. Now we substitute our unitary operator:

$$\begin{aligned} &<\hat{\boldsymbol{O}} a \mid \hat{\boldsymbol{O}} a> \\ &= <a \mid (\hat{\boldsymbol{O}}^T)^* \hat{\boldsymbol{O}} \mid a> \\ &= <a \mid \hat{\mathbf{1}} \mid a> \\ &= <a \mid a> \end{aligned}$$

So:

$$\lambda^* \lambda <a \mid a> = <a \mid a>$$

and since <a|a> is not equal to zero $\lambda^* \lambda$ must be equal to one.

Finally, the *hermitian* operator:

$$(\hat{\boldsymbol{O}}^T)^* = \hat{\boldsymbol{O}} \qquad [3\text{-}38]$$

and

$$\hat{\boldsymbol{O}}^{-1}(\hat{\boldsymbol{O}}^T)^* = (\hat{\boldsymbol{O}}^T)^* \hat{\boldsymbol{O}}^{-1} = \hat{\mathbf{1}}$$

We can use the hermitian operator in an inner product calculation:

$$<a\,|\,\hat{\boldsymbol{O}}\,|\,b> = <a\,|\,(\hat{\boldsymbol{O}}^T)^*\,|\,b>$$

Let's take the complex conjugate of this:

$$\begin{aligned}(<a\,|\,\hat{\boldsymbol{O}}\,|\,b>)^* &= (<a\,|\,(\hat{\boldsymbol{O}}^T)^*\,|\,b>)^* \\ &= (<\hat{\boldsymbol{O}}a\,|\,b>)^* \\ &= (<\hat{\boldsymbol{O}}a\,|\,)^*(\,|\,b>)^* \\ &= <b\,|\,\hat{\boldsymbol{O}}a> \\ &= <b\,|\,\hat{\boldsymbol{O}}\,|\,a>\end{aligned} \qquad [3\text{-}39]$$

Products of operators are worth a mention here. We treat a product of two or more operators acting on a vector **a** as follows:

$$\hat{\boldsymbol{O}}_1\hat{\boldsymbol{O}}_2\,|\,a> = \hat{\boldsymbol{O}}_1(\hat{\boldsymbol{O}}_2\,|\,a>) = |\,\hat{\boldsymbol{O}}_1(\hat{\boldsymbol{O}}_2 a)>$$

That is, operator $\mathbf{O_2}$ acts first on ket a followed by $\mathbf{O_1}$. Recall that the order of operations is generally important since we do not expect that operators will commute (equation [3-28]). The situation with the bra is not quite the same, as a result of anti-linearity:

$$<a\,|\,\hat{\boldsymbol{O}}_1\hat{\boldsymbol{O}}_2 = <(\hat{\boldsymbol{O}}_1\hat{\boldsymbol{O}}_2)^{\mathrm{T}*} a\,| = <(\hat{\boldsymbol{O}}_2^T)^*[(\hat{\boldsymbol{O}}_1^T)^* a]\,|$$

From our look at matrices we saw that taking the transpose of a product of matrices is the same as the product of the transposes of each matrix in reverse order (equation [2-10]). The tight relationship between operators and matrices suggests the same behaviour for operators.

There are some interesting and, as we shall see later, useful properties of these types of operators. First we show that commuting operators have simultaneously defined eigenvectors. This may seem a bit backwards but … we begin by letting $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ be two linear operators that have a common set of eigenvectors:

$$\hat{\boldsymbol{A}}\,|\,f> = a\,|\,f> \quad \hat{\boldsymbol{B}}\,|\,f> = b\,|\,f>$$

and then show that $[\hat{\boldsymbol{A}},\hat{\boldsymbol{B}}]$ is always equal to 0 for any arbitrary vector, $\vec{f}$, or, what is equivalent, that $(\hat{\boldsymbol{A}}\hat{\boldsymbol{B}}-\hat{\boldsymbol{B}}\hat{\boldsymbol{A}})\vec{f}$ = 0. We start by defining $\vec{f}$ in terms of the complete set of basis vectors:

$$\vec{f}=\sum_i c_i \vec{e}_i$$

then:

$$\begin{aligned}(\hat{A}\hat{B}-\hat{B}\hat{A})|f>&=(\hat{A}\hat{B}-\hat{B}\hat{A})|\sum_i c_i e_i>\\ =\sum_i c_i[\hat{A}\hat{B}-\hat{B}\hat{A}]|e_i>&=\sum_i c_i[\hat{A}(\hat{B}|e_i>)-\hat{B}(\hat{A}|e_i>)]\\ =\sum_i c_i[\boldsymbol{A}b_i|e_i>-\boldsymbol{B}a_i|e_i>]&=\sum_i c_i[a_i b_i|e_i>-b_i a_i|e_i>]=0\end{aligned} \qquad [3\text{-}40]$$

Therefore, $\hat{A}$ and $\hat{B}$ are commuting operators and have the common eigenvector, $\vec{f}$. Furthermore, [3-40] shows that two operators with a common eigenvector will *always* commute.

Next we look at the hermitian operator and assert that the eigenvalues of hermitian operators are real. That is, for:

$$\hat{O}|a>=\lambda|a>$$

where it is understood that $\hat{O}$ is hermitian as defined above and that $|a>$ is normalized, the eigenvalue, λ , is a real number. We construct an inner product and do the eigenvalue calculation:

$$<a|\hat{O}|a>=<a|\lambda|a>=\lambda<a|a>=\lambda$$

Now, we do that same for the complex conjugate:

$$(<a|\hat{O}|a>)^*=(<a|\lambda|a>)^*=\lambda^*(<a|a>)^*=\lambda^*$$

Looking at equation [3-39] we realize that, if $\hat{O}$ is hermitian:

$$<a|\hat{O}|a>=(<a|\hat{O}|a>)^*$$

and that therefore:

$$\lambda=\lambda^*$$

Assuming that λ is not equal to zero then since $\lambda = \lambda^*$, λ *must* be real.

We also assert that the eigenvectors of a hermitian operator are orthogonal. Consider the eigenvalue equations on two different eigenvectors:

$$\hat{O}|a>=\lambda_a|a> \qquad \hat{O}|b>=\lambda_b|b>$$

where, as before, $\hat{\boldsymbol{O}}$ is hermitian. We form the inner products and do the eigenvalue calculation:

$$<b|\hat{\boldsymbol{O}}|a>=\lambda_a<b|a> \quad <a|\hat{\boldsymbol{O}}|b>=\lambda_b<a|b>$$

Now we take the complex conjugate of one of them (it doesn't matter which for our purposes) and, again, do the eigenvalue calculation:

$$<a|\hat{\boldsymbol{O}}|b>^*=\lambda_b^*<a|b>^*=\lambda_b^*<b|a>$$

using equation [3-32b]. Now we subtract the complex conjugate calculation from our first eigenvalue calculation:

$$<b|\hat{\boldsymbol{O}}|a>-<a|\hat{\boldsymbol{O}}|b>^*=(\lambda_a-\lambda_b^*)<b|a> \qquad [3\text{-}41]$$

and, using equation [3-39] we can say:

$$<b|\hat{\boldsymbol{O}}|a>=<a|\hat{\boldsymbol{O}}|b>^*$$

and since they are equal because **O** is hermitian, their difference is zero:

$$<b|\hat{\boldsymbol{O}}|a>-<a|\hat{\boldsymbol{O}}|b>^*=0$$

which means that in equation [3-41]:

$$(\lambda_a-\lambda_b^*)<b|a>=0$$

If $|a>$ and $<b|$ are not degenerate so that λ_a and λ_a^* are not equal and their difference is non-zero, then:

$$<b|a>=0$$

which means that $|a>$ and $<b|$ are orthogonal.

We can write our operators in terms of outer products. Specifically, if we use an orthonormal basis set for our outer products we write:

$$\hat{\boldsymbol{O}}=\sum_{i,j}^{n} O_{ij}|u_i><u_j| \qquad [3\text{-}42]$$

As an example let's consider two dimensional space R2, and the standard orthonormal basis set:

$$|u_x>=\begin{bmatrix}1\\0\end{bmatrix} \quad u_y>=\begin{bmatrix}0\\1\end{bmatrix}$$

The four possible outer products of these two vectors are:

$$|u_x\rangle\langle u_x| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$|u_x\rangle\langle u_y| = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
$$|u_y\rangle\langle u_x| = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$
$$|u_y\rangle\langle u_y| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

That the first and last of these are also projection operators follows from our definition of the projection operator in equations [3-22] and [3-23]. Recall that our 90 degree rotation operator in two dimensional space is:

$$\hat{\boldsymbol{R}}(90) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

In terms of the outer product matrices this is:

$$\hat{\boldsymbol{R}}(90) = R_{11}|u_x\rangle\langle u_x| + R_{12}|u_x\rangle\langle u_y| + R_{21}|u_y\rangle\langle u_x| + R_{22}|u_y\rangle\langle u_y|$$
$$\hat{\boldsymbol{R}}(90) = |u_x\rangle\langle u_y| - |u_y\rangle\langle u_x|$$

Note that in this special case the projection operators, $|u_x\rangle\langle u_x|$ and $|u_x\rangle\langle u_y|$, are not part of the operator. Had we been more general and specified a rotation angle of θ then the projection operators would have been part of the rotation operator:

$$\hat{\boldsymbol{R}}(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

and:

$$\hat{\boldsymbol{R}}(\theta) = R_{11}|u_x\rangle\langle u_x| + R_{12}|u_x\rangle\langle u_y| + R_{21}|u_y\rangle\langle u_x| + R_{22}|u_y\rangle\langle u_y|$$
$$= \cos(\theta)(|u_x\rangle\langle u_x|) + \sin(\theta)(|u_x\rangle\langle u_y|) - \sin(\theta)(|u_y\rangle\langle u_x|) + \cos(\theta)(|u_y\rangle\langle u_y|)$$

Another place where outer products appear is in *spectral decomposition*. The complete set of eigenvalues of an operator is called the *spectrum*.

This time only the outer products corresponding to projection operators are used. Suppose we have the eigenvalue equation:

$$\hat{\boldsymbol{O}}\,|\,a_i\!>=\lambda_i\,|\,a_i\!>$$

where there are several eigenvectors and associated eigenvalues. We can also write the projection operator on vector $|a_i\!>$ as:

$$\hat{\boldsymbol{P}}_{\boldsymbol{i}}=|\,a_i\!><\!a_i\,|$$

and a typical matrix representation of a projection operator is:

$$\hat{\boldsymbol{P}}_{\boldsymbol{i}}=\begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 \\ 0 & \cdots & 1 & \cdots \\ \vdots & \vdots & \vdots & 0 \end{bmatrix}$$

and the closure theorem indicates that:

$$\sum_i \hat{\boldsymbol{P}}_{\boldsymbol{i}}=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & 1 \end{bmatrix}=\hat{\mathbf{1}}$$

The eigenvalue matrix looks like:

$$\boldsymbol{\Lambda}=\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots \\ \vdots & \vdots & \vdots & \lambda_n \end{bmatrix}$$

Looking carefully at the last two matrices one can write:

$$\begin{aligned} \boldsymbol{\Lambda} &= \sum_i \lambda_i \mathbf{1} \\ &= \sum_i \lambda_i \hat{\boldsymbol{P}}_{\boldsymbol{i}} \\ &= \sum_i \lambda_i\,|\,a_i\!><\!a_i\,| \end{aligned}$$

Where is this of use? Well, one can use it for determining the reciprocal of a matrix.

What about the square of an operator? What does:

$$\boldsymbol{O}^2\,|\,a\!>$$

mean? What we do here is successively apply the operator to the vector:

$$\begin{aligned}&\boldsymbol{O}^2|a>\\&=\boldsymbol{O}(\boldsymbol{O}|a>)\end{aligned}$$

If $|a>$ is an eigenfunction of $\hat{O}$ then we simply get the square of the eigenvalue:

$$\begin{aligned}&\boldsymbol{O}^2|a>\\&=\boldsymbol{O}(\boldsymbol{O}|a>)\\&=\boldsymbol{O}(\lambda|a>)\\&=(\lambda\boldsymbol{O}|a>)\\&=\lambda^2|a>\end{aligned}$$

Alternately, if $\hat{O}$ is a rotation operator we simply get two successive rotations by an angle, θ .

3.7 Hilbert Space

The mathematical concept of a Hilbert space, named after the mathematician David Hilbert, generalizes the notion of a vector space and is referred to frequently in quantum mechanics. It extends the methods of vector algebra from the two-dimensional plane, **R2**, and three-dimensional space **R3** to include infinite-dimensional spaces in $\boldsymbol{R}^\infty$ and $\boldsymbol{C}^\infty$, where **R**, as always, refers to real numbers and **C** refers to the complex numbers. In fact, a Hilbert space can be comprised of vectors that are *functions* rather than numbers!

The characteristics of a Hilbert space that are important to us are:

1. It has a scalar or inner product as have most of the vector spaces that we have looked at. This scalar product is not necessarily commutative but:

$$<a|b>=<b|a>^* \qquad [3-47]$$

where, of course, the asterisk indicates the complex conjugate. Thus, the scalar product of a vector with itself is:

$$<a|a>=<a|a>^*$$

which we have seen is necessarily real.

2. The real value of the scalar product, $<a|a>$, is positive and its square root is called the norm of the vector $|a>$.

For a Hilbert space there will always be a set of basis vectors of unit norm such that:

$$<e_i|e_i>=1$$
$$<e_i|e_j>=0$$

or:

$$<e_i|e_j>=\delta_{ij} \qquad [3\text{-}48]$$

where:

$$\delta_{ij}=0 \quad if\, i\neq j$$
$$\delta_{ij}=1 \quad if\, i=j$$

(δ is the Kronecker delta). These vectors constitute a orthonormal basis set from which we may construct any other arbitrary vector:

$$\vec{\boldsymbol{A}}=\sum_i a_i\vec{\boldsymbol{e}}_{\boldsymbol{i}}$$

in exactly the same way we did above for a vector in Euclidean 3D-space using $\vec{\boldsymbol{u}}_x$, $\vec{\boldsymbol{u}}_y$ and $\vec{\boldsymbol{u}}_z$.

With this and orthonormality and the scalar product, we can write:

$$<e_i|A>=a_i$$
$$<A|e_i>=a_i^*$$

The complication that arises here is the inclusion of infinite dimensional spaces. In order that we have something that can be managed mathematically in terms of our inner products we must add to the definition of a Hilbert space that the sum of the squares of the components of a vector be a finite number. Otherwise we could possibly end up with a vector of infinite length! Recall that the definition of the inner product is:

$$\begin{aligned}<a|a>&=<e_1|a_1^*a_1|e_1>+<e_2|a_2^*a_2|e_2>+<e_3|a_3^*a_3|e_3>\cdots<e_n|a_n^*a_n|e_n>\\&=a_1^*a_1<e_1|e_1>+a_2^*a_2<e_2|e_2>+a_3^*a_3<e_3|e_3>\cdots a_n^*a_n<e_n|e_n>\\&=a_1^*a_1+a_2^*a_2+a_3^*a_3\cdots a_n^*a_n\\&=\|a_1\|^2+\|a_2\|^2+\|a_3\|^2\cdots\|a_n\|^2\\&=\sum_i^n\|a_i\|^2\end{aligned}$$

It is this sum (a Cauchy sequence) that we cannot allow to be infinite. In other words it must converge to a finite number in order to be of use to us, even if n goes to infinity. That this sequence converges is the third requirement of a Hilbert space, the first two being the normal requirements of an inner product vector space that we have discussed so far.

This is not really so very different from the Euclidean vector spaces that we have already seen and, in fact, Euclidean vector spaces *are* Hilbert spaces. They satisfy the regular requirements of a vector space and the sums of the squares of their vector components are finite. We will deal almost exclusively with finite dimensional spaces in our exploration of nmr spectroscopy.

3.8 Merging Vector Spaces

There will come a point in our deliberations when we will be considering two spaces which we will want to combine into one. The way that we will do this is by using the direct product of the two spaces to produce the new, larger space. Let's consider two spaces, each of dimension two and for simplicity let us assume that each uses the same simple orthonormal basis set:

$$\begin{array}{ccc} \text{space A} & & \text{space B} \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{and} \begin{bmatrix} 0 \\ 1 \end{bmatrix} & & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{and} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \text{or} & \text{and} & \text{or} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array}$$

The direct product (equation [2-3]) of these basis sets gives:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which is the orthonormal basis set in the new, combined space. We started with two spaces of dimension 2 each and ended up with a new space of dimension 4. Generally, if space A is of dimension n and space B is of dimension m then the direct product will be of dimension n x m:

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \otimes \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} = \begin{bmatrix} A_1 B_1 \\ A_1 B_2 \\ \vdots \\ A_1 B_m \\ \vdots \\ A_n B_m \end{bmatrix} \qquad [3\text{-}49]$$

In equation [3-49] we begin with a vector of dimesion n and a vector of dimension m and end up with a vector of dimension m x n. The situation is similar with respect to operators. If the matrix representation of operator $\hat{\boldsymbol{M}}$ is of dimension n x n and operator $\hat{\boldsymbol{N}}$ is of dimension m X m then the direct product matrix will be of dimension n x m.

In the literature one sees the following notation used to indicate the combining of two spaces:

$$|\phi>\otimes|\eta>$$

or more simply (if somewhat confusingly to the uninitiated):

$$|\phi>|\eta>$$
or
$$|\phi\eta>$$

We shall encounter this concept of merging vector spaces in our exploration of the coupling of angular momenta and in product operators.

There are a few things to consider with respect to the direct product. First, for $|\psi_A>$ and $|\psi_B>$ in spaces A and B respectively:

$$a(|\psi_A>\otimes|\psi_B>)=(a|\psi_A>)\otimes|\psi_B>=|\psi_A>\otimes(a|\psi_B>) \quad (a\in\mathbb{C}) \qquad [3\text{-}50]$$

This is easily shown to be so using the matrix notation for the direct product of vectors $|\psi_A>$ and $|\psi_B>$:

$$a\times\begin{bmatrix}A_1\\A_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}B_1\\B_2\\\vdots\end{bmatrix}=\begin{bmatrix}aA_1\\aA_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}B_1\\B_2\\\vdots\end{bmatrix}=\begin{bmatrix}aA_1\begin{bmatrix}B_1\\B_2\\B_3\\\vdots\end{bmatrix}\\aA_2\begin{bmatrix}B_1\\B_2\\B_3\\\vdots\end{bmatrix}\\\vdots\end{bmatrix}=\begin{bmatrix}aA_1B_1\\aA_1B_2\\\vdots\\aA_2B_1\\\vdots\end{bmatrix}=\begin{bmatrix}A_1aB_1\\A_1aB_2\\\vdots\\A_2aB_1\\\vdots\end{bmatrix}=\begin{bmatrix}A_1\begin{bmatrix}aB_1\\aB_2\\aB_3\\\vdots\end{bmatrix}\\A_2\begin{bmatrix}aB_1\\aB_2\\aB_3\\\vdots\end{bmatrix}\\\vdots\end{bmatrix}=\begin{bmatrix}A_1\\A_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}aB_1\\aB_2\\\vdots\end{bmatrix}=\begin{bmatrix}A_1\\A_2\\\vdots\end{bmatrix}\otimes a\times\begin{bmatrix}B_1\\B_2\\\vdots\end{bmatrix}$$

Second, for $|\psi_1>$ and $|\psi_2>$ in space A and $|\psi_3>$ in space B:

$$(|\psi_1>+|\psi_2>)\otimes|\psi_3>=|\psi_1>\otimes|\psi_3>+|\psi_2>\otimes|\psi_3> \qquad [3\text{-}51]$$

which, again, is easily shown using matrix notation for vectors $|\psi_1>$, $|\psi_2>$ and $|\psi_3>$ with basis coefficients A, B and C respectively:

$$\left(\begin{bmatrix}A_1\\A_2\\\vdots\end{bmatrix}+\begin{bmatrix}B_1\\B_2\\\vdots\end{bmatrix}\right)\otimes\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}=\begin{bmatrix}A_1+B_1\\A_2+B_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}=\begin{bmatrix}(A_1+B_1)\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}\\(A_2+B_2)\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}\\\vdots\end{bmatrix}$$

$$=\begin{bmatrix}A_1C_1+B_1C_1\\A_1C_2+B_1C_2\\\vdots\\A_2C_1+B_2C_1\\\vdots\end{bmatrix}=\begin{bmatrix}A_1C_1\\A_1C_2\\\vdots\\A_2C_1\\\vdots\end{bmatrix}+\begin{bmatrix}B_1C_1\\B_1C_2\\\vdots\\B_2C_1\\\vdots\end{bmatrix}=\begin{bmatrix}A_1\\A_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}+\begin{bmatrix}B_1\\B_2\\\vdots\end{bmatrix}\otimes\begin{bmatrix}C_1\\C_2\\\vdots\end{bmatrix}$$

Third, for $|\psi_A>$ and $|\psi_B>$ in spaces A and B respectively and operators $\hat{\boldsymbol{M}}$ and $\hat{\boldsymbol{N}}$, also of spaces A and B respectively we define:

$$(\hat{\boldsymbol{M}}\otimes\hat{\boldsymbol{N}})(|\psi_A>\otimes|\psi_B>)\equiv\hat{\boldsymbol{M}}|\psi_A>\otimes\hat{\boldsymbol{N}}|\psi_B> \qquad [3\text{-}52]$$

We can, of course, show that this is a reasonable definition .. again by using matrices (it is a bit tedious so you are allowed to take our word for it if you wish):

$$\hat{\boldsymbol{M}} \otimes \hat{\boldsymbol{N}} = \begin{bmatrix} M_{11} & M_{12} & \cdots \\ M_{21} & M_{22} & \cdots \\ \vdots & \cdots & \cdots \\ M_{n1} & \cdots & M_{nn} \end{bmatrix} \otimes \begin{bmatrix} N_{11} & N_{12} & \cdots \\ N_{21} & N_{22} & \cdots \\ \vdots & \cdots & \cdots \\ N_{m1} & \cdots & N_{mm} \end{bmatrix}$$

$$= \begin{bmatrix} M_{11} N_{11} & M_{11} N_{12} & \cdots & M_{12} N_{11} & M_{12} N_{12} & \cdots \\ M_{11} N_{21} & M_{11} N_{22} & \cdots & M_{12} N_{21} & M_{12} N_{22} & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ M_{n1} N_{m1} & M_{n1} N_{m2} & \cdots & \cdots & \cdots & M_{nn} N_{mm} \end{bmatrix}$$

$$|\psi_A> \otimes |\psi_B> = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \otimes \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} A_1 B_1 \\ A_1 B_2 \\ \vdots \\ A_2 B_1 \\ \vdots \\ A_n B_m \end{bmatrix}$$

$$(\hat{\boldsymbol{M}} \otimes \hat{\boldsymbol{N}})(|\psi_A> \otimes |\psi_B>) = \begin{bmatrix} M_{11} N_{11} & M_{11} N_{12} & \cdots & M_{12} N_{11} & M_{12} N_{12} & \cdots \\ M_{11} N_{21} & M_{11} N_{22} & \cdots & M_{12} N_{21} & M_{12} N_{22} & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ M_{n1} N_{m1} & M_{n1} N_{m2} & \cdots & \cdots & \cdots & M_{nn} N_{mm} \end{bmatrix} \begin{bmatrix} A_1 B_1 \\ A_1 B_2 \\ \vdots \\ A_2 B_1 \\ \vdots \\ A_n B_m \end{bmatrix}$$

$$= \begin{bmatrix} \sum_j^n \sum_i^m M_{1\mathrm{j}} N_{1\mathrm{i}} A_j B_i \\ \sum_j^n \sum_i^m M_{1\mathrm{j}} N_{2\mathrm{i}} A_j B_i \\ \vdots \\ \sum_j^n \sum_i^m M_{2\mathrm{j}} N_{1\mathrm{i}} A_j B_i \\ \vdots \\ \sum_j^n \sum_i^m M_{nj} N_{mi} A_j B_i \end{bmatrix}$$

That's the left hand side of the equation. Now the right:

$$\hat{M} \mid \psi_A> = \begin{bmatrix} M_{11} & M_{12} & \cdots \\ M_{21} & M_{22} & \cdots \\ \vdots & \cdots & \cdots \\ M_{n1} & \cdots & M_{nn} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} \sum_j^n M_{1j} A_j \\ \sum_j^n M_{2j} A_j \\ \vdots \\ \sum_j^n M_{nj} A_j \end{bmatrix}$$

$$\hat{N} \mid \psi_B> = \begin{bmatrix} N_{11} & N_{12} & \cdots \\ N_{21} & N_{22} & \cdots \\ \vdots & \cdots & \cdots \\ N_{m1} & \cdots & N_{mm} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} = \begin{bmatrix} \sum_i^m N_{1i} B_i \\ \sum_i^m N_{2i} B_i \\ \vdots \\ \sum_i^m N_{mi} B_i \end{bmatrix}$$

$$\hat{M} \mid \psi_A> \otimes \hat{N} \mid \psi_B> = \begin{bmatrix} \sum_j^n M_{1j} A_j \\ \sum_j^n M_{2j} A_i \\ \vdots \\ \sum_j^n M_{nj} A_j \end{bmatrix} \otimes \begin{bmatrix} \sum_i^m N_{1i} B_i \\ \sum_i^m N_{2i} B_i \\ \vdots \\ \sum_i^m N_{mi} B_i \end{bmatrix} = \begin{bmatrix} \sum_j^n M_{1j} A_j \sum_i^m N_{1i} B_i \\ \sum_j^n M_{1j} A_j \sum_i^m N_{2i} B_i \\ \vdots \\ \sum_j^n M_{2j} A_j \sum_i^m N_{1i} B_i \\ \vdots \\ \sum_j^n M_{nj} A_j \sum_i^m N_{mi} B_i \end{bmatrix} = \begin{bmatrix} \sum_j^n \sum_i^m M_{1j} N_{1i} A_j B_i \\ \sum_j^n \sum_i^m M_{1j} N_{2i} A_j B_i \\ \vdots \\ \sum_j^n \sum_i^m M_{2j} N_{1i} A_j B_i \\ \vdots \\ \sum_j^n \sum_i^m M_{nj} N_{mi} A_j B_i \end{bmatrix}$$

As you can see, this is the same as the left side of the equation.

Finally, if the basis set of each vector is orthonormal then the basis set of the direct product will be orthonormal as well.

$$<\psi_A \psi_B \mid \psi_A \psi_B> = 1 \quad if <\psi_A \mid \psi_A> = <\psi_B \mid \psi_B> = 1 \qquad [3\text{-}53]$$

We assume here that $<\psi_A|\psi_A>$ and $<\psi_B|\psi_B>$ are normalized since they can always be made to be so. As usual, we invoke matrices to show this:

$$|\psi_A> = \sum_i^n A_i|\phi_i>$$
$$<\psi_A|\psi_A> = \sum_i^n A_i^* A_i <\phi_i|\phi_i> = \sum_i^n A_i^* A_i = 1$$
$$similarly$$
$$<\psi_B|\psi_B> = \sum_j^m B_j^* B_j <\chi_j|\chi_j> = \sum_j^m B_j^* B_j = 1$$

This is so, of course, only if the basis sets are orthonormal. Proceeding:

$$|\psi_A> \otimes |\psi_B> = |\psi_A \psi_B>$$
$$= \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \otimes \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} = \begin{bmatrix} A_1 B_1 \\ A_1 B_2 \\ \vdots \\ A_2 B_1 \\ \vdots \\ A_n B_m \end{bmatrix}$$

$$<\psi_A \psi_B|\psi_A \psi_B> =$$
$$\begin{bmatrix} A_1^* B_1^* & A_1^* B_2^* & \cdots & A_2^* B_1^* & \cdots & A_n^* B_m^* \end{bmatrix} \begin{bmatrix} A_1 B_1 \\ A_1 B_2 \\ \vdots \\ A_2 B_1 \\ \vdots \\ A_n B_m \end{bmatrix}$$
$$= \sum_i^n A_i^* A_i \sum_j^m B_j^* B_j$$
$$= <\psi_A|\psi_A><\psi_B|\psi_B> = 1$$

Evidently, if the basis sets of the separate vectors are orthonormal then so is the basis set of the new vector.

3.9 Problems

3.10 References

1. J.S. Golan, *The Linear Algebra a Beginning Graduate Student Ought to Know*, Kluwer Academic Publishers, 2004.
2. M. O'nan, *Linear Algebra,* Harcourt, Brace Jovanovich, inc. 1976.
3. R. Shankar, *Principles of Quantum Mechanics*, Plenum Press, 1994.
4. K.A. Stroud and D. Booth, *Linear Algebra*, Industrial Press Inc., 2008.
5. G. Strang, *Linear Algebra and its Applications*, Harcourt Brace Jovanovich, 1988.
6. R.I.G. Hughes, *The Structure and Interpretation of Quantum Mechanics*, Harvard University Press (1989).
7. J.C. Kolecki, *Foundations of Tensor Analysis for Students of Physics and Engineering With an Introduction to the Theory of Relativity*, NASA Science and Technical Information, TP-2005-213115.
8. J.D. Faires and B.T. Faires, *Calculus*, 2nd Edition, Random House (1988).