

Vector Spaces / Real Space

Overview and Motivation: We review the properties of a vector space. As we shall see in the next lecture, the mathematics of normal modes and Fourier series is intimately related to the mathematics of a vector space.

Key Mathematics: The concept and properties of a vector space, including addition, scalar multiplication, linear independence and basis, inner product, and orthogonality.

I. Basic Properties of a Vector Space

You are already familiar with several different vector spaces. For example, the set of all real numbers forms a vector space, as does the set of all complex numbers. The set of all position vectors (defined from some origin) is also a vector space. You may not be familiar with the concept of functions as vectors in a vector space. We will talk about that in the next lecture. Here we review the concept of a vector space and discuss the properties of a vector space that make it useful.

A. Vector Addition.

A vector space is a set (of some kind of quantity) that has the operation of **addition** (+) defined on it, whereby two elements \mathbf{v} and \mathbf{u} of the set can be added to give another element \mathbf{w} of the set,¹

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

There is also an **additive identity** included in the set; this additive identity is known as the zero vector $\mathbf{0}$, such that for any vector \mathbf{v} in the space

$$\mathbf{v} + \mathbf{0} = \mathbf{v}. \quad (2)$$

The addition rule has both **commutative**

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad (3)$$

and **associative**

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \quad (4)$$

properties.

¹ We denote vector quantities by boldface type and scalars in standard italic type. This is standard practice in most physics journals.

B. Scalar Multiplication

The vector spaces that we are interested in also have another operation defined on them known as **scalar multiplication**, in which a vector \mathbf{u} in the space can be multiplied by either a real or complex number a , producing another vector in the space $\mathbf{v} = a\mathbf{u}$. If we are interested in multiplying the elements of the space by only real numbers it is known as a **real vector space**; if we wish to multiply the elements of the space by complex numbers, then the space is known as a **complex vector space**.

Scalar multiplication must satisfy the following properties for scalars a and b and vectors \mathbf{u} and \mathbf{v} ,

$$(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}, \quad (5a)$$

$$a(b\mathbf{u}) = (ab)\mathbf{u}, \quad (5b)$$

$$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}, \quad (5c)$$

$$1\mathbf{u} = \mathbf{u}, \quad (5d)$$

$$0\mathbf{u} = \mathbf{0}. \quad (5e)$$

None of these properties should be much of a surprise (I hope!)

C. Linear Independence and Basis

The **span** of a subset of m vectors is the set of all vectors that can be written as a linear combination of the m vectors,

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \dots + a_m\mathbf{u}_m. \quad (6)$$

The subset of m vectors is **linearly independent** if none of the subset can be written as a linear combination of the other members of the subset. If the subset is **linearly dependent** then we can write at least one of the members as a linear combination of the others, for example

$$\mathbf{u}_m = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \dots + a_{m-1}\mathbf{u}_{m-1}. \quad (7)$$

For a given vector space if there is a maximum number of linearly independent vectors possible, then that number defines the **dimension** N of the vector space.² That means if we have identified N linearly independent vectors then those N vectors span the entire vector space. This means that any vector in the space can be written as a linear combination of the set of N independent vectors as

$$\mathbf{v} = v_1 \mathbf{u}_1 + v_2 \mathbf{u}_2 + \dots + v_N \mathbf{u}_N, \quad (8)$$

or in more compact notation as

$$\mathbf{v} = \sum_{n=1}^N v_n \mathbf{u}_n. \quad (9)$$

Furthermore the **coefficients**³ v_n in Eq. (9) are unique. The vectors \mathbf{u}_n in Eq. (9) are said to form a **basis** for the space. Now Eq. (9) should look strangely familiar. We have some quantity on the lhs that is written as a linear combination of quantities on the rhs. Hum... And you might even ask, assuming that I know the vectors \mathbf{u}_n in Eq. (7), how do I find the coefficients v_n ?

D. Inner Product

This last question is most easily answered after we define one more operation on the vector space, known as the **inner product** of two vectors, which we denote (\mathbf{u}, \mathbf{v}) . The inner product returns a scalar, which is a real number for a real vector space or a complex number for a complex vector space. The inner product can be defined in any manner as long as it satisfies the following relationships

$$(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})^* \quad (10a)$$

$$(\mathbf{w}, a\mathbf{v} + b\mathbf{u}) = a(\mathbf{w}, \mathbf{v}) + b(\mathbf{w}, \mathbf{u}) \quad (10b)$$

Note the complex-conjugate symbol in Eq. (10a). If we are dealing with a real vector space, then we can just ignore the complex-conjugate symbol. Also note that Eq. (10a) implies that the inner product of a vector \mathbf{u} with itself is a real number. It can be shown that Eqs. (10a) and (10b) imply that

² If there is not a maximum number of linearly independent vectors, then the space is said to have infinite dimension.

³ The coefficients v_n are also known as the **scalar components** of \mathbf{v} in the basis $\{\mathbf{u}_1, \dots, \mathbf{u}_N\}$.

$$(\mathbf{a}\mathbf{v} + \mathbf{b}\mathbf{u}, \mathbf{w}) = a^*(\mathbf{v}, \mathbf{w}) + b^*(\mathbf{u}, \mathbf{w}) \quad (10c)$$

$$(\mathbf{a}\mathbf{v}, \mathbf{b}\mathbf{u}) = a^* b(\mathbf{v}, \mathbf{u}) \quad (10d)$$

In physics we are usually interested in vector spaces where

$$(\mathbf{u}, \mathbf{u}) \geq 0, (\mathbf{u}, \mathbf{u}) = 0 \text{ iff } \mathbf{u} = \mathbf{0}. \quad (10e)$$

Such vector spaces are said to have a **positive semi-definite norm** (the norm is defined below).

With these properties of the inner product denoted, we can define the concept of orthogonality. Two nonzero vectors \mathbf{u} and \mathbf{v} are said to be **orthogonal** if their inner product vanishes, i.e., if $(\mathbf{u}, \mathbf{v}) = 0$.

Note that if two vectors are orthogonal, then they are linearly independent. This is easy to see, as follows. Assume the converse, that they are linearly dependent. Then their (assumed) linear dependence means that $\mathbf{u} = a\mathbf{v}$, where a is some scalar [see Eq. (7)]. Then the scalar product $(\mathbf{v}, \mathbf{u}) = (\mathbf{v}, a\mathbf{v}) = a(\mathbf{v}, \mathbf{v})$ cannot be zero because \mathbf{v} is not zero [see Eq. (10e)]. Thus they must be linearly independent.

The converse is not true, two linearly independent vectors need not be orthogonal. The proof is given as one of the exercises.

One last thing regarding the inner product. The quantity $\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})}$ is generally known as the **norm** (or size) of the vector \mathbf{u} . Often we are interested in vectors whose norm is 1. We can "normalize" any vector \mathbf{u} with scalar multiplication by calculating

$$\hat{\mathbf{u}} = \frac{\mathbf{u}}{\sqrt{(\mathbf{u}, \mathbf{u})}}. \quad (11)$$

The "hat" over a vector indicates that the vector's norm is 1.

E. Orthogonal Basis

Most of the time that we deal with a basis, the vectors in that basis are orthogonal. That is, their inner products with each other vanish. In this case it is a simple matter to find the components v_n in Eq. (9). Let's say that we want to find the m th component v_m . Then we take the inner product of Eq. (9) with \mathbf{u}_m , and we get

$$(\mathbf{u}_m, \mathbf{v}) = \left(\mathbf{u}_m, \sum_{n=1}^N v_n \mathbf{u}_n \right) = \sum_{n=1}^N v_n (\mathbf{u}_m, \mathbf{u}_n) \quad (12)$$

[This last equality follows from Eq. (10b).] So what happens? Well, there will only be one nonzero inner product on the rhs, $(\mathbf{u}_m, \mathbf{u}_m)$, and so Eq. (12) becomes

$$(\mathbf{u}_m, \mathbf{v}) = v_m (\mathbf{u}_m, \mathbf{u}_m), \quad (13)$$

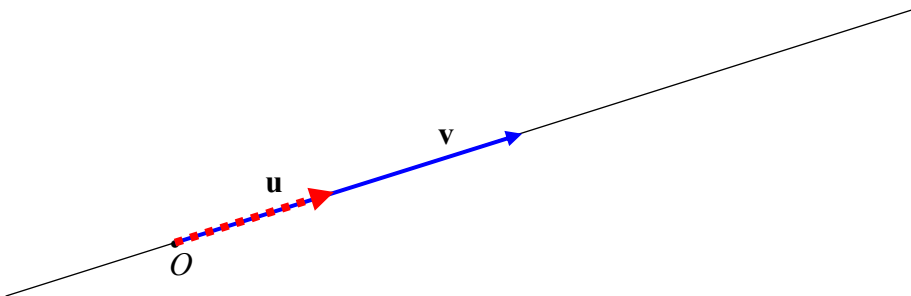
and we can now solve for v_m as

$$v_m = \frac{(\mathbf{u}_m, \mathbf{v})}{(\mathbf{u}_m, \mathbf{u}_m)}. \quad (14)$$

All of this should now look even more strangely familiar. We will get to why that is in the next lecture, but right now we will review a vector space with which you should have some familiarity.

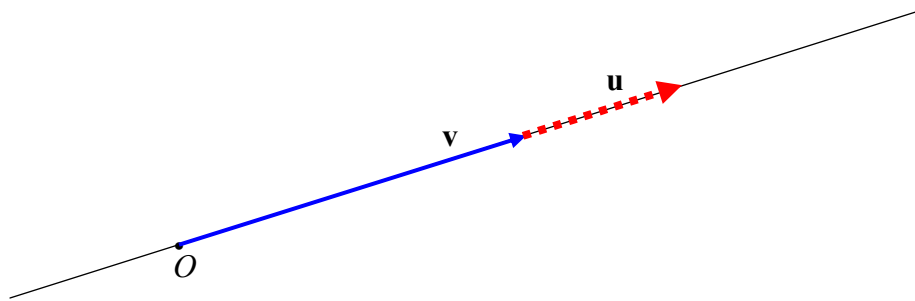
II. 1D Displacement Space

Let's look at a simple example to start. Assume that we have a line drawn somewhere, and on that line we have identified an origin O , as illustrated in the picture below. The vector space that we are interested in consists of all the arrows that start at O and end someplace on the line. The picture also illustrates two of these vectors, one denoted \mathbf{u} and one denoted \mathbf{v} .⁴



⁴ Note, this vector space is not a vector field. A vector field is the assignment of a vector to each point in space.

So let's talk about some of the math introduced above with respect to this vector space. We first have to define vector addition, which must satisfy Eqs. (1) – (4). Let's go with the standard physics definition of vector addition, whereby we add vectors by the tip-to-tail method, where the one of the arrows is translated (without any rotation) and its tail is placed at the tip of the other arrow, as illustrated in the picture below. Clearly this produces another arrow whose tails is at the origin and head is on the line (and is thus a vector in the space). Eq. (1) is thus satisfied. It should also be clear that we could have translated \mathbf{v} rather than \mathbf{u} in this example, and so this definition satisfies Eq. (3), the commutative property of vector addition. We will not illustrate it here, but you should convince yourself that Eq. (4), the associative property is satisfied by the sum of three arrows. What about the zero vector? Well, if Eq. (2) is to be satisfied, it must have no length, and so it must be the arrow that begins and ends at the origin.



What about scalar multiplication? Again, we go with the standard definition, whereby scalar multiplication by a positive number a results in an arrow that points in the same direction and is a times longer than the original arrow. Multiplication by a negative scalar b results in an arrow that points in the opposite direction and is b times longer than the original arrow. It should be clear that this definition satisfies all parts of Eq. (5).

What about linear independence and dimension? Pick an arrow, any arrow. Now ask yourself the following question: can I find another arrow that is not a multiple of my first arrow. If the answer is no (which it is), then the vector space has one dimension, and you can use any arrow as the basis for the space. For example, let's say you pick the arrow \mathbf{u} in the above drawing as your basis. Then the space is one dimensional because you can write any other arrow \mathbf{v} as

$$\mathbf{v} = a\mathbf{u}, \quad (15)$$

where a is some scalar. Although we have not yet defined what the inner product is, notice that if we take the inner product of Eq. (15) with \mathbf{v} we get

$$(\mathbf{v}, \mathbf{v}) = (a\mathbf{u}, a\mathbf{u}) = a^2(\mathbf{u}, \mathbf{u}), \quad (16)$$

so that

$$a = \pm \frac{\sqrt{(\mathbf{v}, \mathbf{v})}}{\sqrt{(\mathbf{u}, \mathbf{u})}} = \pm \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|}, \quad (17)$$

with the sign depending upon the sign of a . Now scalar multiplication was defined as multiplying an arrow's length by the multiplying scalar. Thus a is also the + or - ratio of the two vector's lengths. Therefore, for this space the norm must be proportional to the length of the arrow.

So what about the inner product? Also notice the following. Because this is a one dimensional space, this basis $\{\mathbf{u}\}$ is trivially orthogonal, and we can use Eq. (14) (where here a takes the place of v_m) to express the coefficient a in Eq. (15) as

$$a = \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{u}, \mathbf{u})}. \quad (18)$$

Together Eqs. (17) and (18) imply

$$(\mathbf{u}, \mathbf{v}) = \pm \sqrt{(\mathbf{u}, \mathbf{u})} \sqrt{(\mathbf{v}, \mathbf{v})} = \pm \|\mathbf{u}\| \|\mathbf{v}\|. \quad (19)$$

So which sign do we use? As we now show, it depends upon the relative directions of the two arrows. Let's first consider the case where \mathbf{u} and \mathbf{v} are in the same direction. Then we can write $\mathbf{v} = a\mathbf{u}$, where $a > 0$. Then we have the following

$$(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, a\mathbf{u}) = a(\mathbf{u}, \mathbf{u}) \quad (20)$$

Because $(\mathbf{u}, \mathbf{u}) > 0$, $(\mathbf{u}, \mathbf{v}) > 0$, and we must use the positive sign if \mathbf{u} and \mathbf{v} are in the same direction. Similarly, if \mathbf{u} and \mathbf{v} are in opposite directions then $a < 0$, and we must use the negative sign.

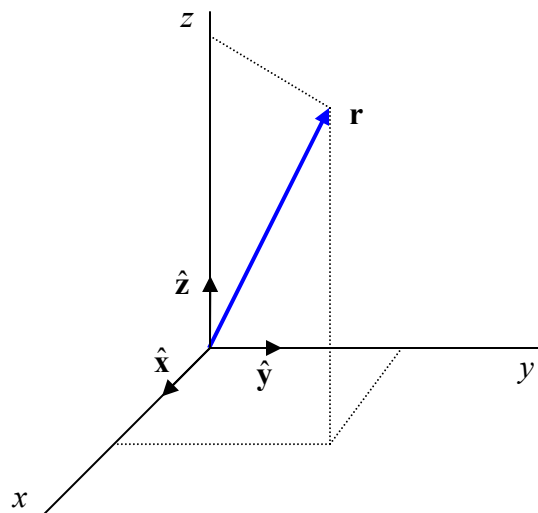
One last comment: notice that nothing we have done here makes us chose the norm to be exactly equal the length of the arrows; it must only be proportional to the length of the arrows. For this space, however, the standard definition of the vector norm is simply the arrow length.

III. Real Space (R^3)

What is real space? It is the extension of this 1D displacement space that we have been discussing to 3 dimensions. That is, it is simply the set of all displacement vectors \mathbf{r} defined with respect to some fixed origin.

Our discussion here will center on the more practical, at least from a physics point of view. The picture below illustrates the following discussion. In dealing with this space, we typically define a set of three mutually perpendicular axes that pass through the origin, which we label x , y , and z . We also denote three special vectors in this space, the three **unit-norm** vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, which are three arrows that point along the three axes, respectively. The relative orientations of these three unit vectors are defined by the right-hand-rule (i.e., cross product) through the equation $\hat{\mathbf{z}} = \hat{\mathbf{x}} \times \hat{\mathbf{y}}$. You should convince yourself that these three vectors are linearly independent (given our definitions of vector addition and scalar multiplication discussed in the last section). It is also true that these three vectors are a basis for our vector space, so this vector space is three dimensional. Thus we can write any vector in the space as a linear combination of these three vectors as

$$\mathbf{r} = r_x \hat{\mathbf{x}} + r_y \hat{\mathbf{y}} + r_z \hat{\mathbf{z}} \quad (21)$$



So now we are back to the ever occurring problem of determining the coefficients of some quantity of interest that is expressed as a linear combination of some other quantities. To do this we can again use the inner product, once it is defined. We use the standard definition of the inner product of two vectors in this space

$$(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta) \quad (22)$$

where θ is the angle between the directions of the two arrows. Notice that this definition reduces to the definition that we came up with for two vectors in our 1D space above, $(\mathbf{u}, \mathbf{v}) = \pm \|\mathbf{u}\| \|\mathbf{v}\|$, where the sign depends upon the relative directions of the two vectors.

With Eq. (22) it is easy to see that the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ satisfy the following relationships,

$$(\hat{\mathbf{x}}, \hat{\mathbf{x}}) = (\hat{\mathbf{y}}, \hat{\mathbf{y}}) = (\hat{\mathbf{z}}, \hat{\mathbf{z}}) = 1 \quad (23a)$$

$$(\hat{\mathbf{x}}, \hat{\mathbf{y}}) = (\hat{\mathbf{y}}, \hat{\mathbf{z}}) = (\hat{\mathbf{x}}, \hat{\mathbf{z}}) = 0. \quad (23b)$$

Eq. (23) defines an **orthonormal basis** for a three dimensional space. That is, the basis is made up of unit vectors [Eq. (23a)] that are all mutually orthogonal [Eq. (23b)].

Using Eq. (23) we can now express the coefficients in Eq. (21) as

$$r_x = \frac{(\hat{\mathbf{x}}, \mathbf{r})}{(\hat{\mathbf{x}}, \hat{\mathbf{x}})} = (\hat{\mathbf{x}}, \mathbf{r}), \quad r_y = \frac{(\hat{\mathbf{y}}, \mathbf{r})}{(\hat{\mathbf{y}}, \hat{\mathbf{y}})} = (\hat{\mathbf{y}}, \mathbf{r}), \quad r_z = \frac{(\hat{\mathbf{z}}, \mathbf{r})}{(\hat{\mathbf{z}}, \hat{\mathbf{z}})} = (\hat{\mathbf{z}}, \mathbf{r}). \quad (24a) - (24c)$$

which enables us to rewrite Eq. (21) as

$$\mathbf{r} = (\hat{\mathbf{x}}, \mathbf{r})\hat{\mathbf{x}} + (\hat{\mathbf{y}}, \mathbf{r})\hat{\mathbf{y}} + (\hat{\mathbf{z}}, \mathbf{r})\hat{\mathbf{z}}. \quad (25)$$

Note that Eq. (24) is the specific form of Eq. (14) for the case at hand. Notice also that because the basis vectors have unit norms, the coefficients have an especially simple form: each coefficient is simply the inner product of the respective basis vector with the particular vector of interest.

Lastly, we remark that the inner product between two vectors \mathbf{r} and \mathbf{s} can be simply written in terms of the components of those vectors in an orthonormal basis. Let's assume that \mathbf{r} is given by Eq. (21) and \mathbf{s} by an analogous equation. Then we can write

$$(\mathbf{r}, \mathbf{s}) = (r_x \hat{\mathbf{x}} + r_y \hat{\mathbf{y}} + r_z \hat{\mathbf{z}}, s_x \hat{\mathbf{x}} + s_y \hat{\mathbf{y}} + s_z \hat{\mathbf{z}}). \quad (26)$$

Now it can be shown that the definition of the inner product [Eq. (22)] satisfies Eq. (10) and so Eq. (26) simplifies to

$$(\mathbf{r}, \mathbf{s}) = r_x s_x + r_y s_y + r_z s_z. \quad (27)$$

That is, the inner product of two vectors can be simply expressed as the sum of the products of corresponding components of the two vectors.

Lastly, we remark that when working with vectors in real space we often use a more notationally compact form than that in Eq. (20): we often simply express the vector \mathbf{r} as its triplet of components

$$\mathbf{r} = (r_x, r_y, r_z), \quad (28)$$

leaving the basis vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ as implied. But, when using this notation one must keep in mind that lurking in the background is an implied set of basis vectors.

Exercises

*13.1 The inner product

- (a) Show that Eq. (10a) implies that the inner product of a vector \mathbf{u} with itself is a real number.
- (b) Using Eqs. (10a) and (10b) show that Eq. (10c) follows.
- (c) Using Eqs. (10a) and (10b) show that Eq. (10d) follows

***13.2 Projection.** The projection of a vector \mathbf{v} onto the direction of another vector \mathbf{u} is defined as $\mathbf{p}(\mathbf{v}, \mathbf{u}) = \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{u}, \mathbf{u})} \mathbf{u}$. Consider an orthogonal (but not necessarily normal) basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. Using this basis any vector \mathbf{v} can be written as $\mathbf{v} = v_1 \mathbf{u}_1 + v_2 \mathbf{u}_2 + v_3 \mathbf{u}_3$. Determine expressions for v_1, v_2 , and v_3 and thus show that \mathbf{v} can be written as $\mathbf{v} = \mathbf{p}(\mathbf{v}, \mathbf{u}_1) + \mathbf{p}(\mathbf{v}, \mathbf{u}_2) + \mathbf{p}(\mathbf{v}, \mathbf{u}_3)$. That is, the vector \mathbf{v} is simply the sum of its projections onto the orthogonal basis set. In physics we often call these projections the *vector* components of \mathbf{v} in the $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ basis.

*13.3 Consider two linearly independent vectors \mathbf{u} and \mathbf{v} and the vector

$\mathbf{w} = \mathbf{v} - \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{u}, \mathbf{u})} \mathbf{u}$ made from these two vectors. Assume that the vector space is

complex. In this problem you are going to do two separate calculations, both of which show that \mathbf{w} is orthogonal to \mathbf{u} . You may find Eqs. (10a) – (10d) useful here.

(a) Easy way: Calculate the inner product (\mathbf{u}, \mathbf{w}) to show that \mathbf{w} is orthogonal to \mathbf{u} .

(b) Slightly harder way: Calculate the inner product (\mathbf{w}, \mathbf{u}) to show that \mathbf{w} is orthogonal to \mathbf{u} .

(This important result can be used to create an orthogonal basis out of any basis.)

*13.4 Show that two linearly independent vectors need not be orthogonal. (Hint: you may find the result of Exercise 13.3 to be helpful here.)

*13.5 Assuming that Eq. (10) applies, show that Eq. (27) follows from Eq. (26).

*13.6 Use Eq. (27) to find the norm of the vector $\mathbf{r} = r_x \hat{\mathbf{x}} + r_y \hat{\mathbf{y}} + r_z \hat{\mathbf{z}}$. Does your result look familiar?

13.7 **Real space. A vector \mathbf{r} in real space has components $(4, -1, 10)$ in one orthonormal basis. In this same basis a set of vectors is given by $\hat{\mathbf{u}}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$, $\hat{\mathbf{u}}_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$, $\hat{\mathbf{u}}_3 = (0, 0, -1)$.

(a) Show that this set of vectors is orthonormal (and is thus another orthonormal basis).

(b) Find the components of \mathbf{r} in this new basis.

(c) From the components given in the statement of the problem, find $\|\mathbf{r}\|$.

(d) From the components determined in part (b), find $\|\mathbf{r}\|$. Is $\|\mathbf{r}\|$ the same as calculated in part (c)?