

YPH102 — LINEAR VECTOR SPACES

LECTURE 1 OF 3

V.H. Belvadi*

Introductory ideas—Linear dependence and independence—inner product spaces—Dirac notation—Expansion in an orthonormal basis—Adjoint operation

Yuvaraja's College, University of Mysore

September 2018

IN TRYING TO DEFINE A useful subset of quantities that come handy while studying quantum mechanics we define some requirements which have to be fulfilled by a given quantity before it can be included in this set. In the process you will end up seeing your idea of a vector being deepened—the existence of a direction and magnitude, you will soon see, is not only an incomplete requirement it is also not a compulsory one for a quantity to be called a vector. The language used in the last few sentences may seem abstract to you right now but the idea will clear itself as we progress over these lectures. Lecture 1, part 1

1 LINEAR VECTOR SPACES

What is important to know right now is that such a set is called a linear vector space. To be more precise, a **linear vector space** \mathbb{V} is a set of objects, called **vectors** (e.g. $\mathbf{U}, \mathbf{V}, \mathbf{W} \dots$), in which the following rules are compulsorily valid:

1. Closure is satisfied:

$$\mathbf{U} + \mathbf{V} \in \mathbb{V}$$

2. Scalar multiplication is *distributive* over vectors and scalars:

$$a(\mathbf{U} + \mathbf{V}) = a\mathbf{U} + a\mathbf{V}$$

$$(a + b)\mathbf{U} = a\mathbf{U} + b\mathbf{U}$$

3. Scalar multiplication with a vector is *associative*:

$$a(b\mathbf{U}) = ab\mathbf{U}$$

4. Addition is *associative*:

$$\mathbf{U} + (\mathbf{V} + \mathbf{W}) = (\mathbf{U} + \mathbf{V}) + \mathbf{W}$$

5. Addition is *commutative*:

$$\mathbf{U} + \mathbf{V} = \mathbf{V} + \mathbf{U}$$

6. A *null vector* must exist:

$$\mathbf{U} + \mathbf{0} = \mathbf{U}$$

7. An *additive inverse* must exist:

$$\mathbf{U} + (-\mathbf{U}) = \mathbf{0}$$

For more visit vhbelvadi.com/teaching.
*vh@belvadi.com

As is common in print the use of **boldface** when writing a quantity tells us that that quantity is a vector. In the next lecture we will be introducing a new representation for vectors.

Besides simply defining what vectors are we also need to define a **field** of quantities over which this vector space is defined. A field can comprise quantities a, b, c, \dots , which are scalars, and which may be real or complex. If the scalar field—over which a vector space is defined—is complex we call the vector space a *complex vector space*; we also similarly define a *real vector space*.

EXERCISE.

To see why direction and magnitude do not always feature in our definition of vectors as members of a vector space consider the matrices

$$\mathbf{M} = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 1 & 3 \\ 5 & 2 \end{bmatrix}$$

which are members of the set of 2×2 matrices. Do they satisfy all conditions (1–7) required to be called vectors? Do they have a magnitude and direction?

2 LINEAR DEPENDENCE AND INDEPENDENCE

2.1 Definitions

Consider the following linear¹ relation:

$$\sum_{i=1}^n \alpha_i \mathbf{U}_i = 0 \tag{1}$$

Clearly there is no $|0\rangle$ on the left-hand side. If there is it can be moved to the right-hand side. For example if we have the equation

$$\alpha_i \mathbf{U}_i + \alpha_j \mathbf{U}_j + \dots + \alpha_m(0) + \dots + \alpha_p \mathbf{U}_p = 0$$

we can always rewrite it as

$$\alpha_i \mathbf{U}_i + \alpha_j \mathbf{U}_j + \dots + \alpha_p \mathbf{U}_p = 0 + \alpha_m(0) = 0$$

which is similar to the general form given in eq. (1) with $\mathbf{U}_i \neq 0$. A given set of vectors is said to be **linearly independent** if the only way eq. (1) can be written is with all $\alpha_i = 0$, i.e. if $\sum_{i=1}^n \alpha_i \mathbf{U}_i = 0$ iff all $\alpha_i = 0$. If we can write eq. (1) with at least one $\alpha_i \neq 0$ then the set of vectors is **linearly dependent**. For example in the term $\alpha_j \mathbf{U}_j$ if $\alpha_j \neq 0$ we can write $\sum_{i=1}^n \alpha_i \mathbf{U}_i = 0$ as $\mathbf{U}_j = \sum_{i=1, \neq j}^n -\alpha_i \mathbf{U}_i / \alpha_j$ since $\alpha_j \neq 0$.

¹A linear relation is one that is of first order i.e. without squares, cubes etc. and with only power one in all its terms, e.g. $2x + 5y = 0$.

EXERCISE.

Show that $(1, 0, 1)$, $(1, 1, 0)$ and $(3, 2, 1)$ are linearly dependent. Also show that $(1, 1, 0)$, $(1, 0, 1)$ and $(0, 1, 1)$ are linearly independent.

SOLUTION.

If $\mathbf{U}_1 \equiv (1, 0, 1)$, $\mathbf{U}_2 \equiv (1, 1, 0)$ and $\mathbf{U}_3 \equiv (3, 2, 1)$ we can write $\mathbf{U}_1 + 2\mathbf{U}_2 + (-\mathbf{U}_3) = 0$ showing that $\alpha_1 \neq 0$, $\alpha_2 \neq 0$ and $\alpha_3 \neq 0$ making the set \mathbf{U}_i linearly dependent. On the other hand if $\mathbf{V}_1 \equiv (1, 0, 1)$, $\mathbf{V}_2 \equiv (1, 1, 0)$ and $\mathbf{V}_3 \equiv (0, 1, 1)$ the only way $\sum \alpha \mathbf{V} = (0, 0, 0)$ is if all $\alpha = 0$. (Try various linear combinations to check.)

The next idea we have to define is that of a **basis**. A basis is simply a set of n linearly independent vectors in an n dimensional space. Mathematically if

$$\mathbf{U} = \sum_{i=1}^n a_i \mathbf{u}_i \quad (2)$$

then we say that the vectors \mathbf{u}_i form a basis of the set \mathbf{U} . Another term worth keeping in mind is that if every vector in a set \mathbf{U} can be written as a linear combination of all other vectors \mathbf{u}_i in the set then we say this set of vectors \mathbf{u}_i **spans** the set \mathbf{U} . Therefore the basis for a vector space is a set of vectors that span that vector space and are linearly independent.

Here is a simple example: for the space \mathbb{R}^3 —our regular three-dimensional space or, alternatively, a space of 3 dimensional column vectors with real elements—the set $\{i, j, k\}$ forms a basis. That is $i \equiv \mathbf{r}(1, 0, 0)$, $j \equiv \mathbf{r}(0, 1, 0)$ and $k \equiv \mathbf{r}(0, 0, 1)$ are linearly independent (since $a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (0, 0, 0)$ is possible only when $a_1 = a_2 = a_3 = 0$) and their linear combination gives other elements of \mathbb{R}^3 (e.g. $2i + j + 3k = (2, 1, 3) \in \mathbb{R}^3$).

2.2 Theorems

Theorem 1. Any vector \mathbf{U}_k in an n dimensional space can be expressed as a linear combination of n linearly independent vectors $\mathbf{U}_1, \dots, \mathbf{U}_n$.

Proof. We will reason in reverse to understand this statement². First let $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_{n+1}$ be a set of vectors in $n + 1$ dimensional space. Let us now assume that the first n vectors here form a linearly dependent set. By definition $\exists a_1, a_2, \dots, a_n$, all scalars, such that at least one a_i is non-zero. Consequently

$$\begin{aligned} a_1 \mathbf{U}_1 + \dots + a_n \mathbf{U}_n &= 0 \\ \implies a_1 \mathbf{U}_1 + \dots + a_n \mathbf{U}_n + 0 \cdot \mathbf{U}_{n+1} &= 0 \end{aligned}$$

which means if we assume that the first n vectors are linearly dependent then the set of $n + 1$ vectors is also linearly dependent.

²Although the recommended text for this course, R. Shankar's *Principles of quantum mechanics*, does not show any proof for this there are multiple proofs. See [1] for others besides the one given in these notes.

What if the first n are linearly independent? Then $\mathbf{U}_1, \dots, \mathbf{U}_n$ form a basis. In this case we have all scalars b_1, b_2, \dots, b_n zeroes. Additionally, since they span the vector space their linear combination should give us the element \mathbf{U}_{n+1} (see definition of spanning above) which we know is in the vector space:

$$\begin{aligned} b_1 \mathbf{U}_1 + \dots + b_n \mathbf{U}_n &= \mathbf{U}_{n+1} \\ \implies b_1 \mathbf{U}_1 + \dots + b_n \mathbf{U}_n + (-1) \mathbf{U}_{n+1} &= 0 \end{aligned}$$

with all $b_i = 0$ except for b_{n+1} . In other words even if the first n vectors are linearly independent the set of $n + 1$ is still linearly dependent.

In short, if there existed some vector \mathbf{U}_k which we could not express as a linear combination of n linearly independent vectors then that vector would form a set of $n + 1$ linearly independent vectors in n dimensional space which, as we proved above, is impossible. \square

Theorem 2. *Given a basis set its components are unique. Mathematically, in*

$$\mathbf{U} = \sum_{i=1}^n a_i \mathbf{u}_i \quad (2)$$

the components \mathbf{u}_i are unique.

Proof. Once again we provide a negative proof: say we have an alternate definition like

$$\mathbf{U} = \sum_{i=1}^n b_i \mathbf{u}_i$$

we can simply subtract the two definitions to get

$$0 = \sum_{i=1}^n (a_i - b_i) \mathbf{u}_i \implies a_i = b_i$$

Therefore, an alternate definition is identical to the original one; in other words, given a basis to a vector space, its components are unique. \square

It is worth keeping in mind that as soon as we change the basis set e.g. $\mathbf{u}_i \rightarrow \mathbf{v}_i$ the components need not be unique i.e. $a_i = b_i$ is not necessarily true.

2.3 Vector arithmetic

Vectors $\mathbf{U} = \sum_i a_i \mathbf{m}_i$ and $\mathbf{V} = \sum_i b_i \mathbf{m}_i$ can be added or scaled straightforwardly:

$$\mathbf{U} + \mathbf{V} = \sum_i (a_i + b_i) \mathbf{m}_i \quad (3)$$

$$p\mathbf{U} = p \sum_i a_i \mathbf{u}_i = \sum_i p a_i \mathbf{u}_i \quad (4)$$

3 INNER PRODUCT SPACES

Although scaling vectors is simple multiplying them is not, particularly because, in the past, when we multiplied vectors we relied heavily upon the idea of direction. Recall, for example, the dot product $\mathbf{A} \cdot \mathbf{B} = AB \cos \theta$ which required knowledge of the angle θ between the vectors thereby indirectly requiring knowledge of the directions of said vectors. Earlier in this lecture, however, we said that the idea of associating directions and magnitudes to vectors is unnecessary. As a result we will have to somewhat better define the dot product.

Recall also the component definition of the dot product which said $\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z$ which requires knowledge of the exact components of these vectors and therefore their orientations even if not their mutual angle itself. The dot product, regardless of how you define it, obeys three properties:

- | | |
|--|--|
| <p>1. <i>Symmetry</i> exists:</p> $\mathbf{U} \cdot \mathbf{V} = \mathbf{V} \cdot \mathbf{U}$ | <p>3. Shows <i>linearity</i>:</p> $\mathbf{U} \cdot (a\mathbf{V} + b\mathbf{W}) = a\mathbf{U} \cdot \mathbf{V} + b\mathbf{U} \cdot \mathbf{W}$ |
| <p>2. It is <i>positive semidefinite</i>:</p> $\mathbf{U} \cdot \mathbf{U} \geq 0 \text{ in general;}$ $\text{But } \mathbf{U} \cdot \mathbf{U} = 0 \iff \mathbf{U} = 0$ | |

3.1 Dirac notation

We will quickly deviate from inner products to set up a simple notation which will help us write down the inner product of two vectors. We showed earlier how matrices can be vectors too since they satisfy all requirements necessary for them to be elements inside a linear vector space. We find that any ‘traditional’ vector with n components can be represented as a column (or row) matrix of rank n . For example³

$$\mathbf{V} \equiv (1, 0, 0) \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

as is convenient for us in a given scenario. Consequently this allows us to represent the matrix notation of a vector in two forms: the column vector \mathbf{V} is represented as $|V\rangle$ and read as **ket** \mathbf{V} while its row vector is represented as $\langle V|$ and read as **bra** \mathbf{V} . This is called **Dirac notation** in honour of its creator the English theoretical physicist P.A.M. Dirac.

³Keep in mind that in this example we say these representations are equivalent ‘ \equiv ’ and not equal ‘ $=$ ’. They are certainly *not* equal: we get to choose between either representation so long as we remember to keep things consistent. We will see their real relationship soon.

3.2 Inner products and other definitions

The definition of inner products is not a complex one. When we have matrices we know that we can only multiply a matrix with another if the column count of the first matches the row count of the second. So in multiplying two vectors represented as matrices—for their inner product—since they can only be column or row matrices, we will have to make sure we multiply a row matrix with a column matrix. Therefore we represent the inner product of two vectors as $\langle U|V \rangle$. Analogous to the properties of the dot product we have

- | | |
|---|---|
| <p>1. <i>Skew-symmetry</i>:</p> $\langle U V \rangle = \langle V U \rangle^*$ <p>2. <i>Positive semidefiniteness</i>:</p> $\langle U U \rangle \geq 0 \text{ in general;}$ <p>But $\langle U U \rangle = 0 \iff U \rangle = 0 \rangle$</p> | <p>3. <i>Linearity</i> (in ket):</p> $\langle U (a V \rangle + b W \rangle)$ $\equiv \langle U aV + bW \rangle$ $= a\langle U V \rangle + b\langle U W \rangle$ |
|---|---|

Note how the use of Dirac notation means we no longer use boldface to represent vectors. Also note how, as in property 2 above, we represent triviality (i.e. 0) with a ket as $|0 \rangle$ whereas when we used boldface for vectors we simply wrote it as the number zero. This is because $|0 \rangle$ stands for a null column matrix. Henceforth and throughout this course we will use Dirac notation exclusively and give up emboldening vectors altogether.

An interesting result pops up on combining the first and third properties:

$$\begin{aligned} \langle aV + bW|U \rangle &= \langle U|aV + bW \rangle^* \\ &= a^* \langle U|V \rangle^* + b^* \langle U|W \rangle^* \\ &= a^* \langle V|U \rangle + b^* \langle W|U \rangle \end{aligned}$$

This is known as *antilinearity* in bra. It is thus important to keep in mind how we are multiplying our bra and ket vectors.

Inner products (also called scalar or dot products) give rise to a few important definitions that we will be using quite frequently henceforth:

Orthogonality. A pair of vectors is said to be orthogonal if $\langle U|V \rangle = 0$. Geometrically this means they form a mutual angle of 90° .

Norm. A vector $|V \rangle$ has the norm or length $\sqrt{\langle V|V \rangle} = |V|$. And a normalised vector is one with $\sqrt{\langle V|V \rangle} = 1$.

Orthonormal basis. A set of vectors that are orthogonal in pairs and that are all normalised are said to form an orthonormal basis.

EXERCISE.

The importance of orthonormal bases can be illustrated with a simple example. We know that the inner product ought not yield a result different from the regular dot product. Try then to multiply two ket vectors in terms of their components e.g. $|U\rangle = \sum_i a_i |i\rangle$ and $|V\rangle = \sum_j b_j |j\rangle$.

SOLUTION.

This is quite simple initially. We need to multiply (find the inner product of) these vectors as $\langle U|V\rangle = \sum_i \sum_j a_i^* b_j \langle i|j\rangle$ which is based on the definition of an inner product. If this has to end up looking like the dot product $u_x v_x + \dots$ in component form then in multiplying $\langle i|j\rangle$ only the diagonal elements must remain i.e. all $\langle i|j\rangle$ terms must vanish (orthogonality) and only $\langle i|i\rangle$ diagonal terms must remain and be equal to 1 (normality). In other words $\langle i|j\rangle$ must be an orthonormal basis⁴.

4 DUAL SPACES

The inner product, being no different in spirit than the dot product, yields a number. Although we have already discussed the bra vector earlier it is worth revisiting it to understand what motivates us to write such a vector. Lecture 1, part 2

Conventionally we begin by assuming all vectors to be kets. This is not required and looking at them as columns will work too but convention simplifies things. Say there exists a column matrix in some basis which is associated with an abstract entity called a ‘ket vector’ $|U\rangle$. To multiply two such vectors would require us to multiply two column matrices which is impossible. We must therefore associate with every ket vector a row matrix, or a bra vector, and find a way to convert between the two. This is what we referred to as an equivalence earlier. There is, however, also an equality:

$$|U\rangle \leftrightarrow \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \leftrightarrow \begin{bmatrix} u_1^* & u_2^* & \dots & u_n^* \end{bmatrix} \leftrightarrow \langle U| \quad (5)$$

In other words bra and ket are related by their adjoints or transpose conjugates. From a broader perspective one can look at this as two separate vector spaces: one with kets (where multiplying two kets gives you their scalar product) and another, called the **dual space**, with bra vectors and a ket vector associated with each bra vector (and vice versa)—the space in which we will be working primarily as evinced by our definition of inner products as products between bras and kets. Bras and kets have their own basis vectors $\langle i|$ and $|i\rangle$ respectively.

⁴To see this orthonormality more clearly try actually multiplying $\langle i|j\rangle$ as matrices.

5 OPERATIONS

5.1 Expansion in an orthonormal basis

Having discussed what vectors and bases are and what an orthonormal basis is we will now try to understand how we can find the components (of an orthonormal basis) into which we can expand a given vector.

To help you recall, $|U\rangle$ is a vector, u_i are scalars of field over which our vector space is defined, and $|i\rangle$ are vectors using which we can expand the vector $|U\rangle$.

Given $|U\rangle = \sum_i u_i |i\rangle$ we can write its inner product with the required component. For instance to find the k^{th} component we take the inner product of $|U\rangle$ with $\langle k|$ as follows:

$$\langle k|U\rangle = \sum_i u_i \langle k|i\rangle = u_k$$

This works because if i and k in fact form an orthonormal bases then, by definition, $\langle k|i\rangle = \delta_{ik}$, i.e. all u_i terms vanish except u_k .

This idea is not new. You would have done the same in case of the more familiar dot product too: given $\mathbf{U} = U_x \hat{i} + U_y \hat{j} + U_z \hat{k}$ we can find, say, U_y by taking the dotproduct of \mathbf{U} with \hat{j} as $\hat{j} \cdot \mathbf{U} = U_y$ since all other terms on the right-hand side go to zero.

The congruence between dot product and inner product is no surprise to us but what will the basis vectors in the above equations look like when we write them as matrices? Consider $|U\rangle = \sum_i u_i |i\rangle$ which may also be written as $|U\rangle = \sum_i |i\rangle u_i$ which is nothing but $|U\rangle = \sum_i |i\rangle \langle i|U\rangle$ based on the expansion discussed above. If, in this final expression, we now substitute $|U\rangle = |j\rangle$ we get

$$|j\rangle = \sum_i |i\rangle \langle i|j\rangle = \sum_i |i\rangle \delta_{ij}$$

In other words the column vector on the left hand side has all elements as zeroes except for the 1 in its j^{th} row. Again, this is not new: we already know the \mathbb{R}^3 basis vectors $|i\rangle$, $|j\rangle$ and $|k\rangle$ have their first, second and third elements as 1 respectively with all others as zeroes. These two equations are then the same:

$$|U\rangle = \sum u_i |i\rangle \Leftrightarrow \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = u_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + u_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ n \end{bmatrix} \quad (6)$$

5.2 Adjoint operation

The last idea we discuss is that of an adjoint. This word too is not new to you—we used earlier in this lecture too and you would have learnt this in your previous years as well. In defining the relationship between bra and ket

vectors we used the adjoint—see eq. (5)—so we know, for instance, that

$$a|U\rangle = |aU\rangle \equiv \begin{bmatrix} au_1 \\ au_2 \\ \vdots \\ au_n \end{bmatrix} \equiv \begin{bmatrix} a^*u_1^* & a^*u_2^* & \dots & a^*u_n^* \end{bmatrix} \equiv \langle aU| = \langle U| a^*$$

Note especially the final relation in comparison with its ket analogue:

$$\langle aU| = \langle U| a^* \quad (7a)$$

$$|aU\rangle = a|U\rangle \quad (7b)$$

These two equations are called **adjoints** of each other. We can also extend it to n terms: $a|U\rangle = b|V\rangle + c|W\rangle + \dots$ is the adjoint of $\langle U| a^* = \langle V| b^* + \langle W| c^* + \dots$. In a typical physicist's fashion let us write down a general expansion for this:

$$|U\rangle = \sum_i u_i |i\rangle \quad \text{has the adjoint} \quad \langle U| = \sum_i \langle i| u_i^* \quad (8)$$

Alternately, since $u_i = \langle i|U\rangle$ as we saw earlier and, similarly, since $u_i^* = \langle U|i\rangle$ we can quickly see that

$$|U\rangle = \sum_i |i\rangle \langle i|U\rangle \quad \text{has the adjoint} \quad \langle U| = \sum_i \langle U|i\rangle \langle i| \quad (8)$$

To see why consider $\langle U| = \sum_i u_i^ \langle i|$ which means the u_k^{*th} component is given by $\langle U|k\rangle = \sum_i u_i^* \delta_{ik}$ which is simply u_k^* .*

is a general expression for the adjoint operation on bra and ket vectors. The steps to take an adjoint is quite simple: reverse the order by switching bras and kets and take complex conjugates of all coefficients.

References

- [1] Math 23A solution set 3. Harvard: <https://bit.ly/2zpbhD6>.
- [2] R. Shankar. Principles of Quantum Mechanics, 2ed. Springer.