Linear Spaces

Motivation

A lot of what we do in physics deals with linear approximations and linear equations. Linear approximations are often what we get from the Taylor series, and linear equations have the property of superposition (the sum of two solutions is a solution) so when our physical system has a small oscillation or satisfies superposition, we often describe it using linear equations. Deciding what degrees of freedom to quantify and how to put them together into an appropriate mathematical structure requires that we both know how the physical system behaves and what are the properties of the mathematical structures associated with linear equations. In this handout, we will discuss the abstractions of linear spaces and inner product spaces — two structures that play a powerful and useful role in many areas of physics.

We've considered two examples of linear spaces: the description of one particle moving in three dimensions, and the description of two particles moving in one dimension. In the first case, we specify our position by identifying three objects in physical space — directions having no length (because the dimensions go with the coordinate), $\hat{i}$, $\hat{j}$, and $\hat{k}$. A vector then looks like

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}.$$

Our "vector space" consists of all vectors of this form where *x, y,* and *z* are real numbers times a dimension of length.

In the second case, we chose two different coordinate systems to describe the motion of the two particles (they had different origins), y_1 and y_2. We wrote an arbitrary vector in this space as

$$|y\rangle \leftrightarrow \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

We use a "corresponding to" arrow rather than an "=" since the expression on the left is meant to be independent of choice of basis, while the expression on the right explicitly depends on the choice of basis. We will see what this means more explicitly once we have defined our general definitions of spaces and representations below. Note that our first example could be expressed in one on three ways:

$$\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$$

$$|r\rangle \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

All of these stand for the same thing, but the first (with the $\hat{i}$, $\hat{j}$, and $\hat{k}$) gives the expression in terms of a specific set of basis vectors. The second, |*r*>, does not explicitly show the basis and does not depend on it, while the column vector implicitly assumes a(n unspecified) basis.

We are going to be using these examples as analogies or metaphors for other systems that look (physically) nothing like them. As usual, a careful mathematical definition can tell us when an analogy works and can be used to calculate results and when it can't.

Definition: Linear Space[*]

Let's begin by abstracting what we have done in the two cases above. The first is the easiest to think about. In the case of describing an object moving in three space, we have started with three objects for which we know some physical properties or have some intuitions about: the three basis directions. We then multiplied each direction be a length and added the resulting three vectors formally. We can abstract what we got as follows.

A *linear space* (or a *vector space*) is a set of elements (vectors), V, and a set of numbers (scalars), S, (where for us, S will be either the real or complex numbers) satisfying the following properties:

- The set forms a group under addition. This means: it is closed, an identity exists, and each element has an inverse. Written out mathematically:

  If $\vec{a},\vec{b} \in V$ then $\vec{a}+\vec{b} \in V$ (Note this means we have some definition of +).

  There exists a vector $\vec{0}$ such that for any $\vec{a} \in V$, $\vec{a}+\vec{0}=\vec{0}+\vec{a}=\vec{a}$.

  For every $\vec{a} \in V$, $\exists\, (-\vec{a}) \in V \quad \ni \vec{a}+(-\vec{a})=(-\vec{a})+\vec{a}=0$.

- The set forms a group under multiplication by a scalar. This means: it is closed, there is an identity, and every scalar has an inverse (except 0). Written out mathematically:

  If $\vec{a} \in V$ and $\alpha \in S$ then $\alpha\vec{a} \in V$ (Note this means we have a definition of what it means to multiply a vector by a scalar).

  There exists a scalar 1 such that for any $\vec{a} \in V$, $1 \cdot \vec{a} = \vec{a} \cdot 1 = \vec{a}$.

  For every $\alpha \in S \quad \exists\, \alpha^{-1} \in S \quad \ni \alpha^{-1}(\alpha\vec{a}) = \alpha(\alpha^{-1}\vec{a}) = \vec{a}$ except for $\alpha = 0$.

- The two operations of addition and scalar multiplication are distributive. Mathematically, this means that addition and multiplication work like for regular numbers.

  $$\begin{aligned} &\forall\, \alpha \in \mathrm{S}, \vec{a},\vec{b} \in V, \quad \alpha\left(\vec{a}+\vec{b}\right) = (\alpha\vec{a}) + \left(\alpha\vec{b}\right) \\ &\forall\, \alpha,\beta \in \mathrm{S}, \vec{a} \in V, \quad (\alpha+\beta)\vec{a} = (\alpha\vec{a}) + (\beta\vec{a}) \end{aligned}.$$

These properties specify a linear (or vector) space. It is easy to show that the two examples we started with both satisfy all these properties.

---

[*] Note the mathematical symbols that will be used in this handout: $\in$ (contained in — as an object in a set), $\exists$ (there exists), $\forall$ (for all), $\ni$ (such that), and $\therefore$ (therefore).

Inner Product Spaces

In our vector space in which a single particle is moving in three dimensions, our standard vectors (our unit directions) had a useful property — they were each perpendicular to each other and had a unit length. In our more general space, we don't necessarily know what it means to be "perpendicular" or even to say whether it makes sense to say the vector has a "length". In some cases it is useful to do this but in others it is not.

[For an example of a vector in which it is not useful to define a length, consider the example of geometrical optics. It is useful (approximations of Gaussian optics) to define a vector in which one component is the distance, h, away from the optical axis (a line running through the center of the optical system) and a second component is the angle, θ, the ray makes with the axis. This two component vector (h, θ) can be acted on by matrices corresponding to propagation of the ray or to bending by a lens. The "length" of this vector is meaningless since you cannot add a length and an angle.]

The tool we use to define lengths and angles in our 3D space is the same: the dot product. We can see this be the equation that gives the length of a vector:

$$a = |\vec{a}| = \sqrt{\vec{a} \cdot \vec{a}}$$

and the cosine of the angle between two vectors:

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}$$

If we can define something analogous to the dot product in our linear space that has analogous properties, we can define lengths and angles. The generalization of the dot product to an arbitrary linear space is called an *inner product* and a linear space in which an inner product can be defined is called an *inner product space*.

Just like the dot product, the inner product is a mapping — an assignment of a scalar to any pair of vectors in the space. We write this as a mapping rule:

$$I : V \times V \to S$$

that is, the inner product is a function rule, $I(\vec{a},\vec{b})$ that takes two elements of the vector and returns a single scalar. [In our base examples, we write $I(\vec{a},\vec{b}) \equiv \vec{a} \cdot \vec{b}$ but we are going to something a bit more general soon.]

In order for the inner product to be usable in the way we expect, it has to have certain properties. Check for yourself that the standard dot product satisfies these conditions.

- It has to be linear in both vector variables, that is:

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

- Scalars have to factor out. (We choose to make it factor out as a complex conjugate from the first variable in case we are working with the complex numbers as our scalars.)

$$I(\alpha\vec{a},\vec{b}) = \alpha^* I(\vec{a},\vec{b})$$
$$I(\vec{a},\alpha\vec{b}) = \alpha I(\vec{a},\vec{b})$$

- The product is (complex) symmetric, that is:

$$I(\vec{a},\vec{b}) = I(\vec{b},\vec{a})^*$$

- The lengths have to be positive, that is,

$$\forall\ \vec{a} \in V\ (\vec{a} \neq \vec{0})\ \ I(\vec{a},\vec{a}) > 0.$$
$$I(\vec{0},\vec{0}) = 0$$

  (Note that the complex symmetry condition requires that $I(\vec{a},\vec{a})$ must be real.)

- The cosines have to be between 1 and –1, that is

$$\forall\ \vec{a},\vec{b} \in V,$$
$$\left| \frac{I(\vec{a},\vec{b})}{\sqrt{I(\vec{a},\vec{a})I(\vec{b},\vec{b})}} \right| \leq 1$$

Check for yourself that these conditions hold for the standard dot product in three space.

If we can define such an inner product in whatever space we have — even if it doesn't "look" geometrical, we can define lengths of vectors and angles between vectors. This lets us define orthonormal bases and do lots of calculations very much more easily that otherwise. This is particularly true in the theory of mechanical waves and in quantum mechanics.

In general, we will use Dirac's bra-ket notation for vectors, especially for complex vector spaces since changing from a ket to a bra keeps track of whether we have taken a complex conjugate or not. In such a space, we use a simplified (and satisfying) notation for the inner product: a bra-ket, the angle brackets closing to produce a scalar — collapsing the two vectors into a number:

$$I\big(|a\rangle,|b\rangle\big) = \langle a\|b\rangle = \langle a|b\rangle$$

where we commonly replace the double bar that comes from putting the bra and ket together by a single bar.