CSUC
Department of Physics
Mechanics: Class Notes

Introduction to ***Vector Spaces***

I. *INTRODUCTION*

Modern mathematics often constructs logical systems by merely proposing a set of elements that obey a specific set of rules. The elements needn't have any meaning whatsoever or any other reference (e.g. to the "physical world"). As we study "Geometric Vector Spaces" we are actually using one such system. Although we do, indeed, intend to model the three-dimensional *"Physical Space"* space we actually live in, the underlying structure can also be applied to a wide variety of other physical systems. The point of doing this is that we are made aware of precisely which suppositions (*axioms*) are responsible for which specific outcomes. As we add more structure, we narrow the applicability of our resultant system. It becomes ever more specific. In our development here we shall start with the very most elementary level of structure and then add on step by step.

II. *LINEAR SPACES*

1.) *Foundational Concepts*

The lowest level of underlying structure will simply be called a *"Linear Space"* . We start by defining a set of elements $\{\alpha, \beta, \gamma ...\}$ that we will call "*vectors*" (here we set aside the Greek alphabet for them) and a set of "scalars" that will be the real numbers at first and later may be chosen as the complex numbers. It's crucial that we understand that "*vectors*" are not "*numbers*" but that they interact with them. If α is a vector then 2α or $x\,\alpha$ is too. Further, vectors are given an "*addition*" relationship among themselves that looks just like addition of "numbers" but isn't. So that e.g., if α and β are vectors then $2\alpha + 3\beta$ is also a vector. We assume the following set of simple rules:

- (A)
  1. addition has "closure": if α and β are vectors , then $\alpha + \beta$ always is too.
  2. addition is "commutative": $\alpha + \beta = \beta + \alpha$
  3. addition is "associative": $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
  4. addition has an "additive identity" designated "0" such that: $\alpha + 0 = \alpha$ for every vector
  5. to every vector α there is an "additive inverse" designated $-\alpha$ so that $\alpha + (-\alpha) = 0$
- (B)
  1. multiplication by scalars is "associative": $x(y\alpha) = (xy)\alpha$
  2. $1\,\alpha = \alpha$ and $0\,\alpha = 0$ for every α .
- (C)
  1. multiplication by scalars is "distributive" : $(x + y)\alpha = x\alpha + y\alpha$ ... and
  2. $x\,(\alpha + \beta) = x\alpha + x\beta$

The key ideas underlying linear spaces are the concepts of:

1. ***linearity***
2. ***independence***

Two vectors α and β are said to be ***linearly independent*** if **no** scalars $\{x,\, y\}$ can be found *other than zero* to make the following statement true:

$$x\,\alpha + y\,\beta = 0 \tag{1}$$

Indeed, any collection of vectors $\{\alpha_i \;\; i = 1, ..., n\}$ will be said to be "linearly dependent" if we can find a set of scalars $\{x_i \;\; i = 1, ..., n\}$ that are **not all zero** such that:

$$\sum_{i=1}^{n} \alpha_i\, x_i \;=\; 0 \tag{2}$$

If, on the other hand, the statement $\sum_{i=1}^{n} \alpha_i\, x_i \;=\; 0$ **does** imply that all the scalars $\{x_i \;\; i = 1, ..., n\}$ *must* be zero, then we say that the set elements $\{\alpha_i \;\; i = 1, ..., n\}$ are linearly **independent**.

If the number of linearly independent vectors in any collection can be as large as some specific integer "**n**" but never larger, we say that the vector space is of "dimension **n**" (or in the language of physics we say that the system has **n** "degrees of freedom") . Now suppose that we have such a system of **n** linearly independent vectors $\{\alpha_i \;\; i = 1, ..., n\}$, then *any* further vector α_{n+1} can certainly be expressed as a sum over our set:

$$\alpha_{n+1} \;=\; \sum_{i=1}^{n} \alpha_i\, x_i \tag{3}$$

We say that the linearly independent set members $\{\alpha_i \;\; i = 1, ..., n\}$ form a ***basis***.

2.) *Inner Product, Magnitude and Projection*

The concepts of "magnitude" and "projection" will require adding yet more structure. We introduce the idea of defining a mapping from any ordered *pair* of vectors α, β to the scalars. This will be our "***inner product***" and be designated (α, β). It shall have the following properties:

1. $(\alpha, \beta) \;=\; (\beta, \alpha)$ for real scalars or $(\alpha, \beta) \;=\; (\beta, \alpha)^*$ for complex scalars.
2. *linearity:* $(\alpha, x_1\beta_1 \,+\, x_2\beta_2) \;=\; x_1(\alpha, \beta_1) \,+\, x_2(\alpha, \beta_2)$
3. *positive definiteness:* $(\alpha, \alpha) \;\geq\; 0$ and $(\alpha, \alpha) \;=\; 0$ iff $\alpha = 0$

Property 3. allows us to define a "magnitude" for any vector. We designate the magnitude of α by $\parallel \alpha \parallel$ and let it be defined by:

$$\parallel \alpha \parallel \;\;\equiv\;\; \sqrt{(\alpha, \alpha)} \tag{4}$$

And now using this, for any vector α, we can define a "unit vector" $\hat{\alpha}$ by $\hat{\alpha} \;\equiv\; \alpha \,/ \parallel \alpha \parallel$, so that $(\hat{\alpha}, \hat{\alpha}) \;=\; 1$. We say that the vector $\hat{\alpha}$ is " normalized". As we will discover, the inner product of any vector β with any other "normalized" vector e.g. $\hat{\alpha}$ yields the special scalar $(\hat{\alpha}, \beta)$ that takes on the meaning of "projection" of β along $\hat{\alpha}$. This will be a key idea in the expression of any vector as a linear combination of normalized "*basis*" vectors.

3.) *Orthonormal Bases*

Suppose that the linearly independent set of n vectors $\{\alpha_1, ..., \alpha_n\}$ forms a basis of the set under consideration. It is not at all necessary that we find $(\alpha_i, \alpha_j) \;=\; 0$ for $i \;\neq\; j$. However, we now show that from the original set $\{\alpha_1, ..., \alpha_n\}$ we may now construct a *new* set $\{\hat{e}_1, ..., \hat{e}_n\}$ that has the orthonormality property $(\hat{e}_i, \hat{e}_j) \;=\; 0$ wherever $i \;\neq\; j$ and $(\hat{e}_i, \hat{e}_i) \;=\; 1$ for $i = 1, ..., n$.

Step 1. Set $\hat{e}_1 \equiv \hat{\alpha}_1$
Step 2. Let $\eta_2 \equiv \alpha_2 - \hat{e}_1(\hat{e}_1, \alpha_2)$ Observe that, by construction, we now have $(\hat{e}_1, \eta_2) = 0$
Step 3. Now normalize η_2 and use it to define the next new basis vector by setting $\hat{e}_2 \equiv \eta_2 / \parallel \eta_2 \parallel$
Step 4. In like manner now let $\eta_3 \equiv \alpha_3 - \hat{e}_1(\hat{e}_1, \alpha_3) - \hat{e}_2(\hat{e}_2, \alpha_3)$ and finally set $\hat{e}_3 \equiv \eta_3 / \parallel \eta_3 \parallel$ ... etc.
Step 5. Continue in this manner for each set member α_i until we exhaust the set.

In this way we construct a new set of n linearly independent vectors $\{\hat{e}_i\}$ where any two vectors obey $(\hat{e}_i, \hat{e}_j) = \delta_{i\,j}$ and the symbol $\delta_{i\,j}$ is a shorthand way of writing "1" if $i = j$ and "0" if $i \neq j$. Since we will be taking inner products very often and the inner products between basis vectors will show up ubiquitously, this compact notation (the "Kronecker delta") will simplify our algebra enormously. This set $\{\hat{e}_i\}$ is our new *orthonormal basis*. With it we can express *any* vector as $\alpha \;=\; \sum_{i=1}^{n} \hat{e}_i\, x_i$ where $x_i = (\hat{e}_i, \alpha)$. That is, for any vector α we have:

$$\alpha \;=\; \sum_{i=1}^{n} \hat{e}_i\,(\hat{e}_i, \alpha) \tag{5}$$

As a free bonus consequence, we can easily observe that we have accidentally proven one more important and useful result. Since from *any* two vectors α, β we can form $\hat{\alpha} \;\equiv\; \alpha \;/\; \parallel \alpha \parallel$ and $\eta \;\equiv\; \beta - \hat{\alpha}(\hat{\alpha}, \beta)$ and since for *any* vector whatsoever (and therefore in particular for our vector η) we must have $(\eta,\, \eta) \geq 0$, it follows:

$$(\eta,\, \eta) \;=\; (\beta - \hat{\alpha}(\hat{\alpha}, \beta),\, \beta - \hat{\alpha}(\hat{\alpha}, \beta)) \;=\; (\beta,\, \beta) - (\beta,\, \hat{\alpha})\,(\hat{\alpha},\, \beta) \;\geq\; 0 \tag{6}$$

In detail, this last inequality can be written:

$$(\beta,\, \beta) - \frac{(\beta,\, \alpha)}{\parallel \alpha \parallel}\,\frac{(\alpha,\, \beta)}{\parallel \alpha \parallel} \;\geq\; 0 \tag{7}$$

Finally, with a simple rearrangement, we find:

$$\parallel \alpha \parallel \parallel \beta \parallel \;\geq\; |(\alpha,\, \beta)| \tag{8}$$

This is the famous Schwartz inequality and we will use it rearranged in many forms such as $|(\hat{\alpha},\, \hat{\beta})| \;\leq\; 1$.

III. *3-D GEOMETRIC VECTOR SPACES*

In attempting to model the 3-D world we live in, J. Willard Gibbs constructed a three dimensional linear space with an inner product and one additional feature viz. a *vector product* (also known as a "cross product" or even sometimes as an "outer product"). He visualized the abstract elements of this space as representing *displacements* through space rather than the Cartesian starting concept of points "in" space and found thereby that he could recover all of Euclidean Geometry. To indicate that we have given our vectors "*spatial meaning*" we commonly now distinguish vectors by a superior arrow e.g. $\vec{A}$. Our three degrees of freedom in the linear space represent the three possible spatial directions that physical space seems to make available to us. We emphasize that a Linear Space, in general, need have no inner product nor any *magnitude* concept whatsoever much less a "cross product" too . These are "extra levels of structure" that we add in to our model. Gibbs found that the dot product provided a succinct and accurate rendition of the concept of "projection" (or "overlap") of one displacement on another and, in like manner, that the cross product precisely renders the concept of oriented **areas** as well as the concepts associated with **rotations**. The combination of the two products in the "vector triple product" yields the representation, accordingly, of "oriented volume" and a full completion of the Euclidean program of Geometry.

A. *The Dot-Product*

Consider first, then, the inner product (henceforth referred to as a "dot product") which, we again assert, can concisely represent physical *projection* of one displacement on another. As a first consequence, the dot product of two *normalized* vectors will now represent the *cosine* of the angle between them. Dot products are now conventionally to be written as $\vec{A} \cdot \vec{B}$ and all the time honored Cartesian formulae are then, in fact, generated by simply first defining our products among an orthonormal basis. If we let the set $\{\hat{e}_1,\, \hat{e}_2,\, \hat{e}_3\}$ represent the three axis directions of a right-handed coordinate basis where $\hat{e}_i \cdot \hat{e}_j = \delta_{i\,j}$ then, for example, our dot product becomes:

$$\vec{A} = \sum_{i=1}^{3} \hat{e}_i \, \hat{e}_i \cdot \vec{A} = \sum_{i=1}^{3} \hat{e}_i \, A_i \tag{9}$$

$$\vec{B} = \sum_{j=1}^{3} \hat{e}_j \, \hat{e}_j \cdot \vec{B} = \sum_{j=1}^{3} \hat{e}_j \, B_j \quad \textit{and therefore...}$$

$$\vec{A} \cdot \vec{B} = \left(\sum_{i=1}^{3} \hat{e}_i \, A_i \right) \cdot \left(\sum_{j=1}^{3} \hat{e}_j \, B_j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i B_j \, \hat{e}_i \cdot \hat{e}_j = \sum_{i=1}^{3} A_i B_i$$

In this last result we have used the useful notational result $\sum_{j=1}^{3} \delta_{i\,j} \, B_j = B_i$.

Notice especially that, if we make the identification of the projection of one unit vector on another with the cosine of the angle between them (i.e. the "included angle: $\theta_{included}$ ") and if we write e.g. $\vec{A} = \| \vec{A} \| \, \hat{A}$, then we have the following very useful set of equalities:

$$\vec{A} \cdot \vec{B} = \sum_{i=1}^{3} A_i B_i = \| \vec{A} \| \| \vec{B} \| \, \hat{A} \cdot \hat{B} = \| \vec{A} \| \| \vec{B} \| \, cos(\theta_{included}) \ . \tag{10}$$

These expressions link our initial "abstract" expression to the conventional Cartesian "component" expression and finally to the "polar" or "absolute" expression. Such developments will be typical from now on and we will use whatever form seems most convenient in the discussion at hand. Notice that certain results immediately become clear, e.g. that $\vec{A}$ and $\vec{B}$ are perpendicular if and only if $\vec{A} \cdot \vec{B} = 0$.

B. *The Cross-Product*

Our final development here is the vector "cross product". The key geometric match-up to be made stems from the observation that any ordered pair of vectors defines an oriented *area*. The vectors actually define a parallelogram and the information in their ordering (given by the so-called "*Right-Hand-Rule*") allows us to assign one side as "special" or "positive" (or "*up*"). In the figure we see vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ and their resultant parallelogram with the attendant normal direction $\hat{\mathbf{n}}$.

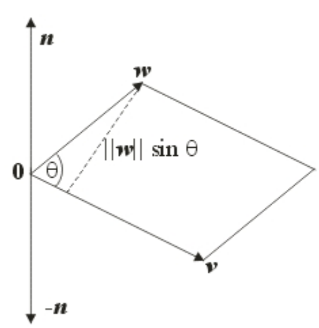

FIG. 1: The geometry of the cross-product.

Our abstract specification amounts to:

$$\vec{A} \times \vec{B} = (parallelogram \ area)\hat{n} = \| \vec{A} \| \| \vec{B} \| \, sin(\theta_{included}) \, \hat{n} \ . \tag{11}$$

The fact that any area can be oriented by a normal vector which has three components allows us to express the fact that areas "project " just like displacement vectors. This may seem surprising at first (and is, in fact, due to the three-fold dimensionality of physical space), but soon appears natural when one recognizes that there are exactly three coordinate planes (x-y, y-z, z-x) and that the projection of one area onto another is exactly proportional to the cosine of the angle between the planes (the "dihedral angle" θ) and that this angle is just the same as the angle between the two unit-normal directions (you may think of them as "opening up" together).

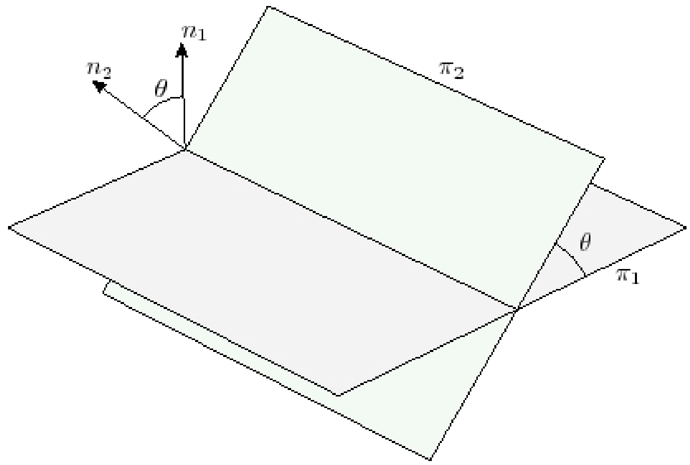

FIG. 2: The dihedral angle between planes equals the angle between their normals.

Once again, our newly defined product may be given variously in the: "Abstract", "Cartesian Component", or "Absolute Polar" specifications that are all perfectly equivalent. We state the definitions in abstract form first and follow them with the Cartesian development. The "Cross Product" $\vec{A} \times \vec{B}$ of two vectors is specified by the following rules:

1. $\vec{A} \times \vec{B}$ is itself a vector and is **linear** in each vector of the product (we say it is *bilinear*) .
2. $\vec{A} \times \vec{B}$ is perpendicular to both $\vec{A}$ and $\vec{B}$, so that e.g. $\vec{A} \cdot (\vec{A} \times \vec{B}) \equiv 0$
3. $\vec{A} \times \vec{B} = -\vec{B} \times \vec{A}$ so that $\vec{A}$, $\vec{B}$, and $\vec{A} \times \vec{B}$ form a right-handed triple of vectors.
4. $\| \vec{A} \times \vec{B} \| = \| \vec{A} \| \| \vec{B} \| \; sin(\theta_{included})$

The Cartesian decomposition is facilitated by first stating the results for the three planes defined by the coordinate axes. Any two axes, then, define a plane and there are only three combinations: $\{\hat{e}_1, \hat{e}_2\}$, $\{\hat{e}_2, \hat{e}_3\}$, $\{\hat{e}_3, \hat{e}_1\}$. We write out our fundamental specifications in "Abstract" form first:

$$\begin{aligned} \hat{e}_1 \times \hat{e}_2 &= 1\,\hat{e}_3 \\ \hat{e}_2 \times \hat{e}_3 &= 1\,\hat{e}_1 \\ \hat{e}_3 \times \hat{e}_1 &= 1\,\hat{e}_2 \end{aligned} \tag{12}$$

Now we proceed very much as we did with our dot product in equation (9) viz. :

$$\begin{aligned} \vec{A} &= \sum_{j=1}^{3} \hat{e}_j\, \hat{e}_j \cdot \vec{A} = \sum_{j=1}^{3} \hat{e}_j\, A_j \\ \vec{B} &= \sum_{k=1}^{3} \hat{e}_k\, \hat{e}_k \cdot \vec{B} = \sum_{k=1}^{3} \hat{e}_k\, B_k \quad \textit{and therefore...} \\ \vec{A} \times \vec{B} &= \left(\sum_{j=1}^{3} \hat{e}_j\, A_i \right) \times \left(\sum_{k=1}^{3} \hat{e}_k\, B_k \right) = \sum_{j=1}^{3} \sum_{k=1}^{3} A_j B_k\, \hat{e}_j \times \hat{e}_k = \sum_{j,k=1}^{3} A_j B_k\, \hat{e}_j \times \hat{e}_k \end{aligned} \tag{13}$$

Since the Cartesian approach is committed to expressing all vector expressions in their component forms, we complete the derivation by finding a *typical component* of $\vec{A} \times \vec{B}$. We recall that, in all cases, the i^{th} component of any vector $\vec{V}$ is found by computing the projection of that vector along the i^{th} coordinate direction by means of the dot product. That is, $V_i = \hat{e}_i \cdot \vec{V}$. It then follows that:

$$(\vec{A} \times \vec{B})_i \ = \hat{e}_i \cdot (\vec{A} \times \vec{B}) \ = \ \hat{e}_i \cdot \sum_{j,k=1}^{3} A_j B_k \, \hat{e}_j \times \hat{e}_k \ = \ \sum_{j,k=1}^{3} A_j B_k \, \hat{e}_i \cdot (\hat{e}_j \times \hat{e}_k) \tag{14}$$

Notice especially that with two separate indices each sweeping through three values, that there are *nine* terms in this double sum ... yet all but two of them turn out to be zero! This is because of the curious vector triple product term $\hat{e}_i \cdot (\hat{e}_j \times \hat{e}_k)$ in the argument. You can easily convince yourself that to be anything but zero, the indices $\{i, j, k\}$ must represent the numbers $\{1, 2, 3\}$ in some order. That is, if any two of the indices represent the *same* number, then the triple product comes out to be zero. Finally, if $\{i, j, k\}$ represent $\{1, 2, 3\}$ in any *cyclic* (sequential) order i.e. $1 \to 2 \to 3 \to 1 \ldots$ etc. , then the value is 1 and in any *anti*-cyclic order it comes out -1. Because we will be doing so many component calculations we conventionally use the concise notation $\epsilon_{ijk} \equiv \hat{e}_i \cdot (\hat{e}_j \times \hat{e}_k)$. This symbol (frequently called the *Levi-Civita totally antisymmetric tensor*) embodies the component expression of the cross product, and we simply remember:

$$(\vec{A} \times \vec{B})_i \ = \ \sum_{j,k=1}^{3} \epsilon_{ijk} A_j B_k \tag{15}$$

We summarize the properties of this odd but extremely useful device in the following list:

1. $\epsilon_{ijk} = -\epsilon_{jik}$ i.e. the value changes by a *minus sign* under interchange of *any two* of its indices.
2. $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$ i.e. the value is *unchanged* under any cyclic rotation *of all the indices.*
3. $\sum_{k=1}^{3} \epsilon_{ijk} \, \epsilon_{kst} \ = \ \sum_{k=1}^{3} \epsilon_{kij} \, \epsilon_{kst} \ = \ \sum_{k=1}^{3} \ \hat{e}_k \cdot (\hat{e}_i \times \hat{e}_j) \ \hat{e}_k \cdot (\hat{e}_s \times \hat{e}_t) \ = \ (\hat{e}_i \times \hat{e}_j) \cdot (\hat{e}_s \times \hat{e}_t) \ = \ \delta_{is}\delta_{jt} - \delta_{it}\delta_{js}$

This third property may take some staring but actually follows straight forwardly enough upon consideration of what a complete set of orthonormal unit vectors is like. We will need this result to evaluate expressions involving two cross products. Next we list the *same* properties but as they appear in *abstract* notation:

1. $\vec{A} \times \vec{B} \ = \ -\vec{B} \times \vec{A}$
2. $\vec{A} \cdot \vec{B} \times \vec{C} \ = \ \vec{B} \cdot \vec{C} \times \vec{A} \ = \ \vec{C} \cdot \vec{A} \times \vec{B}$ and one often relates the first and last as $\vec{A} \cdot \vec{B} \times \vec{C} \ = \ \vec{A} \times \vec{B} \cdot \vec{C}$
3. $\vec{A} \times (\vec{B} \times \vec{C}) \ = \ \vec{B}(\vec{A} \cdot \vec{C}) \ - \ \vec{C}(\vec{A} \cdot \vec{B})$ *and also* $(\vec{A} \times \vec{B}) \times \vec{C} \ = \ \vec{B}(\vec{A} \cdot \vec{C}) \ - \ \vec{A}(\vec{B} \cdot \vec{C})$

Finally, we may combine these results in many and various utilitarian forms, of which a few are listed below:

1. $(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) \ = \ (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) \ - \ (\vec{A} \cdot \vec{D})(\vec{B} \cdot \vec{C})$ ... so, in particular ...
2. $(\vec{A} \times \vec{B}) \cdot (\vec{A} \times \vec{B}) \ = \ (\vec{A} \cdot \vec{A})(\vec{B} \cdot \vec{B}) \ - \ (\vec{A} \cdot \vec{B})(\vec{B} \cdot \vec{A}) \ = \parallel \vec{A} \parallel^2 \ \parallel \vec{B} \parallel^2 \ (1 \ - \ cos^2(\theta_{included}))$ ... so ...
3. $\parallel \vec{A} \times \vec{B} \parallel = \parallel \vec{A} \parallel \parallel \vec{B} \parallel \ sin(\theta_{included})$