Physics 464/511 Lecture C Fall, 2016

In the last lecture we introduced the basic differential operations for vector calculus, divergence, curl, and laplacian, in terms of their expressions in cartesian coordinates. We also saw that a lot of relations could be written as vector statements without reference to particular coordinate systems. For actual calculations, especially in situations in which there is a symmetry, such as spherical symmetry, which are not well described by rectangular coordinates, it is useful to use other, non-cartesian coordinates to describe the same three dimensional Euclidean space. Thus the next step will be to discuss curvilinear coordinate systems and express the differential operators in terms of these. But before we do that, we should review both the machinery and the concepts of vector spaces.

1 More on Vector Space

In Lecture B we gave a formal definition of vector space over a field F, but considered only vectors expressed in terms of coefficients $\in F$ multiplying a fixed set of vectors $e_i, i = 1, \ldots D$. Given a set of D vectors, the space of all linear combinations of them (over the field F) is a D dimensional vector space, providing that there is not an alternate set of fewer than D basis vectors which would suffice to span the space. Thus to be a *basis* the vectors e_j must be *linearly independent*, which means $\sum_{j=1}^{D} \alpha_j e_j = 0$ only if all $\alpha_j = 0$. Otherwise we could solve for one of the e_j in terms of the others, and use just the other $D - 1$ basis vectors to express any vector.

The set of basis vectors is not unique, as for example a new set could be obtained by rotating the old set, but the number of basis vectors is unique, as will be shown below. If we have two sets of basis vectors, $\{e_j\}$ and $\{e'_j\}$, any vector in the space can be expressed as linear combinations in terms of either set, $\vec{V} = \sum_j v_j e_j = \sum_k v'_k e'_k$. In particular, each of the primed set can be expressed in terms of the unprimed, $e'_i = \sum_j M_{ij} e_j$, with $M_{ij} \in F$, and vice-versa, $e_j = \sum_k N_{jk} e'_k$, with $N_{jk} \in F$. As $e_j = \sum_k N_{jk} e'_k = \sum_{k\ell} N_{jk} M_{k\ell} e_\ell$, and as the $\{e_j\}$ are linearly independent, this requires $\sum_k N_{jk} M_{k\ell} = \delta_{j\ell}$, where the *Knonecker delta function* $\delta_{jk} = 1$ if $j = k$ and $= 0$ otherwise.

M and N are *matrices* with entries in F. If there are m e_k's and n e'_k's, M is a $n \times m$ matrix and N is an $m \times n$ matrix[1]. The multiplication of two

[1]These are not multiplications and are not equal, they represent the numbers of choices

matrices is defined only if the second dimension of the first is the same as the first dimension of the second: if A is an $m \times n$ matrix and B is a $n \times p$ matrix,

$$(A \cdot B)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}, \quad \text{for } i = 1, \ldots, m,\ j = 1, \ldots, p$$

$A \cdot B$ is an $n \times p$ matrix.

A few words about indices[2]:

[These may seem trivial, but I have seen many graduate students mess up calculations because they are careless with their use of indices.]

In an equation, there may be free indices and dummy indices. Free indices can take on any specific value within the appropriate limits, but it is the same value whereever it occurs in the equation. Dummy indices are summed over, and have no intrinsic meaning. So in

$$e'_i = \sum_j M_{ij} e_j = \sum_k M_{ik} e_k$$

the j can be replaced by any other index **which does not appear elsewhere in the equation**. (or in the same term, more accurately, in the scope of the summation). But the free index i cannot be replaced within a term, that is, in general

$$\sum_j M_{ij} e_j \neq \sum_j M_{kj} e_j.$$

This is all trivial, but note that if you want the square a primed basis vector and write $(e'_i)^2 = \sum_j M_{ij} e_j \cdot M_{ij} e_j$, you have made a mistake. The correct expression is

$$(e'_i)^2 = \sum_{j,k} M_{ij} M_{ik} e_j \cdot e_k.$$

[Actually this is a bad example, as we might well have $e_j \cdot e_k = \delta_{jk}$.]

Linear Transformations

A linear transformation T from a vector space V into another W over the same field ($T : V \to W$) must satisfy

---

of the first and second index respectively.

[2] A more extensive appeal for correct use of indices is available at `http://www.physics.rutgers.edu/~shapiro/507/lects/indices.pdf`.

- $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$
- $T(\alpha\vec{v}_1) = \alpha T(\vec{v}_1)$

for all $\vec{v}_1 \in V$, $\vec{v}_2 \in V$ and $\alpha \in F$. V is called the **domain** of T.

These properties of linear transformations means they respect the vector space operations, or vector space structure. More generally, a mapping of one space V with an algebraic structure into another space W with the same structure, preserving the structure, is called a **homomorphism**. The structure here is vector space, but the concept of homomorphism also applies to rings, fields, and groups[3].

As any vector $\vec{v} \in V$ can be written $\vec{v} = \sum v_i e_i$, and any vector in W, including $T(\vec{v})$, can be written as a sum of coefficients times basis vectors e'_j of W, we can write $\vec{w} = T(\vec{v}) = \sum W_j e'_j$. In particular, each $T(e_i)$ can be so written, with coefficients in F we will call[4] T_{ki}, so $T(e_i) = \sum_k T_{ki} e'_k$. But

$$T(\vec{v}) = \sum_i v_i \, T(e_i) = \sum_{ik} v_i \, T_{ki} \, e'_k \quad = \sum_k W_k e'_k$$

so the coefficients of $\vec{w}$, $W_k = \sum_i T_{ki} v_i$, because the e'_k are linearly independent. So the action of a linear transformation from a vector space of dimension m into a vector space of dimension n is specified by a $n \times m$ matrix with elements in F, written as

$$T = \begin{pmatrix} T_{11} & T_{12} & \dots & T_{1m} \\ \vdots & & & \vdots \\ T_{n1} & T_{n2} & \dots & T_{nm} \end{pmatrix}.$$

If S is a subset of points in V, $T(S)$ is the set of points in W which are images of points in S. That is, $\vec{w} \in T(S)$ if and only if $\exists \vec{s} \in S \ni T(\vec{s}) = \vec{w}$. The image of the whole domain, $T(V)$, is called the **range** of T.

If T is a linear transformation from V into W ($T : V \to W$) and U is a linear transformation from W into X, then T followed by U, written $U \circ T$ or just UT, is a linear transformation from V to X. If we have bases for V,

[3] Whether or not a map is a homomorphism depends on the structure considered. For example, the doubling map D on $\mathbb{C}$, $D : z \mapsto 2z$, is a homomorphism on $\mathbb{R}^2$ or $\mathbb{C}$ considered as a vector space, but not when considered as a field, as $D(z_1) \cdot D(z_2) = 4z_1 \cdot z_2 \neq D(z_1 \cdot z_2) = 2z_1 \cdot z_2$.

[4] Note the order of the indices, reversed from that in $W_k = \sum_i T_{ki} v_i$ below.

W and X in terms of which T and U are represented by the matrices T_{jk} and U_{ij}, then the matrix which represents $U \circ T$ is

$$(UT)_{ik} = \sum_j U_{ij} T_{jk},$$

which we defined above as the multiplication of the two matrices.

A subset of elements in V which is closed under addition and multiplication by scalars is called a *subspace*, and is itself a vector space. Because it must include the vector 0 (because multiplying any vector with the scalar 0 gives the vector 0) such a space is a hyperplane within V passing through the origin.

If T is a linear transformation from V to W, the image of V is a subspace of W.

If T is a linear transformation from V to W, and S is a subspace of V, the image of S under T is a subspace of W.

If $\{e_i\}$ is a basis of a subspace S of V, any vector in $T(S)$ may be expressed in terms of $\{T(e_i)\}$. If the $T(e_i)$ are linearly independent, they form a basis of $T(S)$, and if not, some subset of these form a basis. Thus the dimension of $T(S)$ is less than or equal to the dimension of S. In the discussion of alternative bases for V above, the matrix M represents a linear transformation $T : V \to V$ with $\{e'_i\}$ a basis of $T(V)$, so $n =$ number of e'_i's is $\leq m =$ number of e_i's. But the reverse is also true, the e_i's are in the image of V under the linear transformation given by the matrix N, so $m \leq n$, and therefore $m = n$, justifying our statement that all bases of V have the same number of elements, and the dimension D of V is clearly defined. Note also that the relation $\sum_k N_{jk} M_{k\ell} = \delta_{j\ell}$, together with the relation $\sum_k M_{j\ell} N_{\ell k} = \delta_{jk}$ we get from reversing the two bases, states that M and N are inverses of each other, and the product is the unit matrix $\mathbb{1}_{jk} = \delta_{jk}$.

Rank of a Matrix

If T is a linear transformation represented by the matrix M, we have seen that the range of T, $T(V)$, is a vector space, with a dimension less than or equal to the dimension n_V of V. We call this dimension r the *rank* of the matrix. If the rank of the matrix is less than that of V ($r < n_V$), the images of the basis of V satisfy a linear relation $\sum_j \alpha_k T(e_k) = 0$, and the vector $\vec{a} = \sum \alpha_k e_k$ is *annihilated* by T, that is, $T(\vec{a}) = 0$. The set of all vectors which are annihilated by T is called the *kernel* of T. It is a vector space of dimension $n_V - r$.

$\mathbb{R}^n$ and $\mathbb{C}^n$

If we have an n dimensional vector space V over the reals, and we take a fixed basis set, the vectors $\vec{v}$ are in 1-1 correspondence to ordered n-tuples of real numbers, $(v_1, v_2, \ldots, v_n)$, the coefficients v_i of each of those basis vectors required to make $\vec{v} = \sum v_i e_i$. The set of such ordered n-tuples of real numbers is called $\mathbb{R}^n$. Then we may identify the points $(v_1, v_2, \ldots, v_n)$ of $\mathbb{R}^n$ with the vector $\vec{v}$. The same could be said for an n dimensional vector space over the complexes, and n-tuples of complex numbers. Thus vectors in an n-dimensional vector space are equivalent to n-tuples of numbers (in F), given a fixed basis for the space. But for many of the vector spaces we wish to consider, there is no preferred basis. There is often, however, a concept of distance or norm. Formally this is a function[5] $f : V \to \mathbb{R}$ with

- $f(v) > 0$ for $v \neq 0$
- $f(v_1 + v_2) \leq f(v_1) + f(v_2)$
- $f(\lambda v) = |\lambda| f(v)$

for all $v, v_1, v_2 \in V$, $\lambda \in F$. We generally write the norm as $\|v\|$. Note the norm is always real, even for complex vector spaces.

There are a number of norms one might consider[6], but for our purposes we generally want to also impose that the space be an inner product space[7], one with a scalar product $\langle \cdot, \cdot \rangle : V \times V \to F$ with[8]

- $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle^*$
- $\langle v_1, \lambda v_2 \rangle = \lambda \langle v_1, v_2 \rangle$
- $\langle v_1, v_2 + v_3 \rangle = \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle$
- $\langle v_1, v_1 \rangle > 0$ unless $v_1 = 0$.

---

[5] I think this applies only if the field is $\mathbb{R}$ or $\mathbb{C}$.

[6] Common examples of normed spaces which are not inner product spaces are placing a norm $\|x\|_p = \left(\sum |x_j|^p\right)^{1/p}$ for $p \geq 1$, which is called the L^p norm. For $p = 2$ this is the Euclidean norm and is that of an inner product space, but that is not true for $p \neq 2$. The case $p = 1$ is called the Manhattan norm, the distance you would have to walk if you could only walk parallel to one or another of the axes. L^∞, which says $\|x\| = \max |x_j|$ is another used norm.

[7] Vaughn calls this a *unitary* vector space

[8] Some mathematicians make it linear in the first argument rather than the second, which matters if $F = \mathbb{C}$. Physicists *never* do.

for all $v_1, v_2, v_3 \in V$, $\lambda \in F$. Note that $f(v) = \sqrt{\langle v, v\rangle}$ (the L^2 norm) satisfies the conditions for a norm if $\langle \cdot, \cdot\rangle$ is an inner product. We usually call this inner product the scalar product.

If V is a finite dimensional[9] inner product space, we can choose an orthonormal basis $\hat{e}_j$ with $\langle \hat{e}_j, \hat{e}_k\rangle = \delta_{jk}$. Then the norm of a vector $\vec{V} = \sum_j v_j \hat{e}_j$ is $\|\vec{V}\| = \sqrt{\sum_j |v_j|^2}$, just as Pythagoras told us. And the inner product $\langle \vec{V}, \vec{W}\rangle = \sum_j v_j^* w_j$. Note that the scalar product does not give us a vector in V, but rather a number in F. If F is the complexes, the scalar product is not symmetric, $\langle \vec{v}, \vec{w}\,\rangle \neq \langle \vec{w}, \vec{v}\,\rangle$, but rather the two complex numbers are complex conjugates of each other, $\langle \vec{v}, \vec{w}\,\rangle = (\langle \vec{w}, \vec{v}\,\rangle)^*$. But the scalar product of a vector with itself is still a positive real number (except for 0), and therefore so is the length.

Dual Space

If we have a vector space V over F, we may define another vector space, V^* to be the space of linear functions from V to F. As the action of the linear function Λ on an arbitrary vector $\sum v_i e_i$ is determined by its action on the basis vectors, $\Lambda(\sum v_i e_i) = \sum v_i \Lambda(e_i)$, we may take as a basis of V^* the functions $u_j : e_i \mapsto \delta_{ij}$. If V is of finite dimension n, clearly so is V^*. And in this case V is the dual of V^*. Things can get more complicated if the space is infinite dimensional, because of issues of convergence of infinite sums, but we will consider that later.

Note we are so accustomed to dealing with finite-dimensional vector spaces with a measure or norm, and always using orthonormal basis vectors to describe the space, that the distinction between the space and its dual seems to be nit-picking. But note that the dual space is defined even if there is no norm on V, and $u(v)$ is defined for $u \in V^*$, $v \in V$, with $u_i(v_j) = \delta_{ij}$, even when we have no meaning assigned to $\langle v, v\rangle$, and even, when we do have some scalar product on V, if we are using non-orthonormal basis vectors. This will be the case when we consider differential geometry.

An example of a vector space for which there is not a single natural definition of the inner product is the space of continuous real-valued functions $f(x)$ on $[-1, 1]$, called $C([-1, 1])$. Given any positive real function $w(x)$ on $[-1, 1]$, we may define an inner product $\langle f_1, f_2\rangle_w = \int_{-1}^{1} w(x) f_1^*(x) f_2(x) dx$. Despite the norm being w-dependent, the elements of the dual space are not.

[9] We will also be considering infinite dimensional vector spaces, but there are complications there, which we will postpone discussing until later.

For example, if we define vectors $\Lambda \in C([-1,1])^*$ by $\Lambda(f) := \int_{-1}^1 \lambda(x) f(x) dx$ for functions $\lambda \in C([-1,1])$, we have defined a map from $C([-1,1])$ into $C([-1,1])^*$, but $\Lambda(f) \neq \langle \lambda, f \rangle_w$. Notice that the Dirac $\delta(x)$ is well defined in $C([-1,1])^*$ but does not correspond to any function in $C([-1,1])$. This is an indication that things can get involved for infinite dimensional spaces.

A vector space with an inner product which is also *complete* is called a *Hilbert space*, especially if it is infinite-dimensional[10]. Completeness means that any sequence of vectors with $\|\vec{V}_n - \vec{V}_m\| \to 0$ as $m, n \to \infty$ has a limit vector $\vec{V}$ such that $\|\vec{V}_m - \vec{V}\| \to 0$ as $m \to \infty$. The space of real polynomials on $[-1,1]$ with inner product $\langle f_1, f_2 \rangle_w$ with $w = 1$ is not complete, because if we define $f_n = \sum_{j=1}^{n} x^j/j$, then for $m < n$,

$$\begin{aligned}(\|f_n - f_m\|)^2 &= \sum_{j=m+1}^{n} \sum_{k=m+1}^{n} \frac{1}{jk} \int_{-1}^{1} x^{j+k}\, dx \\ &= \sum_{j=m+1}^{n} \sum_{k=m+1}^{n} \frac{1-(-1)^{j+k+1}}{jk(j+k+1)} < 2\frac{(n-m)^2}{m^3} \underset{m,n\to\infty}{\longrightarrow} 0,\end{aligned}$$

but there is no polynomial p such that $\|f_n - p\| \to 0$, because $f_n \to -\ln(1-x)$ which is not a polynomial, and cannot be approximated perfectly by any finite polynomial. But if we take the vector space $V = L^2([-1,1])$ the set of quadratically integrable functions on $[-1,1]$, we do get a complete space, a Hilbert space. Hilbert spaces are very important in quantum mechanics.

If we have an orthonormal basis $\hat{e}_j$ of a Hilbert space, any vector $\vec{v}$ is a limit of a sequence of vectors $\vec{v}_n = \sum_{j=1}^n \langle \hat{e}_j | \vec{v} \rangle \hat{e}_j$. As the norm squared is the sum of $|\langle \hat{e}_j | \vec{v} \rangle|^2$, all of which are ≥ 0, the finite sum is $\leq ||\vec{v}||^2$, so we have the Bessel inequality:

$$\sum_{j=1}^{n} |\langle \hat{e}_j | \vec{v} \rangle|^2 \leq ||\vec{v}||^2.$$

For any two vectors $\vec{u}$ and $\vec{v}$, and for any complex λ, the norm of $\vec{u} + \lambda \vec{v}$ is ≥ 0. Thus

$$0 \leq \langle \vec{u} + \lambda \vec{v} | \vec{u} + \lambda \vec{v} \rangle = ||\vec{u}||^2 + \lambda \langle \vec{u} | \vec{v} \rangle + \lambda^* \langle \vec{v} | \vec{u} \rangle + |\lambda|^2 ||\vec{v}||^2.$$

[10]A finite-dimensional inner product space is automatically complete.

Let $\lambda = -\langle \vec{v}|\vec{u}\rangle/||\vec{v}||^2$, and this gives

$$0 \leq ||\vec{u}||^2 - 2|\langle \vec{u}|\vec{v}\rangle|^2/||\vec{v}||^2 + |\langle \vec{v}|\vec{u}\rangle|^2/||\vec{v}||^2.$$

Multiply by $||\vec{v}||^2$ and take square roots to get

$$|\langle \vec{u}|\vec{v}\rangle| \leq ||\vec{u}||\,||\vec{v}||,$$

which is the Schwarz inequality.

Of course in finite dimensional Euclidean real space these are just statements that if you leave off components in the Pythagorean evaluation for a hypotenuse you will get an underestimate, and that $|\cos\theta| \leq 1$.

1.1 Morphisms

If T is a linear transformation from V to W and U is a linear transformation $W \to V$, the two compositions, $T \circ U : W \to W$ and $U \circ T : V \to V$ are **endomorphisms** on W and on V respectively. An endomorphism is a homomorphism from a space into itself. If V and W have dimensions D_V and D_W respectively, the rank of T cannot be bigger than $\min(D_V, D_W)$. If it is an 1–1 map, ($i.e.$ $T(v_1) = T(v_2) \Longrightarrow v_1 = v_2$) it must have kernal $\{0\}$ and rank D_V, and if it is onto[11], it has rank D_W. If T is both 1–1 and onto, it gives an **isomorphism** between V and W.

We might possibly have U be the left-inverse of T, ($U \circ T = 1\!\!1 : V \to V, v \mapsto v$), but only if T is 1–1, or it might be a right-inverse of T, ($T \circ U = 1\!\!1 : W \to W, w \mapsto w$), but only if T is onto. To be both, we must have the same dimension for V and W ($D_V = D_W$), and T and U are both represented by square matrices, T_{ij} and U_{ij} respectively, and the matrix product is the $D_V \times D_V$ identity matrix, $\sum_j T_{ij}U_{jk} = \delta_{ik} = \sum_j U_{ij}T_{jk}$, and we just say U is the inverse of T.

We have strayed rather far from the topic of differential operators for solving for physical fields, just so we could use basis vectors other than cartesian ones. But the rest of this formal discussion of vector spaces will come back later, when we consider that solutions of partial differential equations give us vector spaces, infinite dimensional ones at that!

[11] That is, all points in W are the image of some point in V.