Multilinear Algebra, Tensors and Wedge Products

I am going to tell you how multilinear algebra grows out of the algebra of vector spaces. This is a rich topic but I will constrain myself to that part of multilinear algebra that is useful for understanding multi-variable calculus and illuminate some issues around the vector cross product used in physics. I think of the collection of ideas around vector spaces, dual spaces, vectors and matrices as the seed of a crystal that grows into the ever more interesting structure of multilinear algebra.

I am assuming that you are familiar with vectors and matrices. Furthermore, you are aware of the distinction between row vectors and column vectors. You may be aware that row vectors are linear functionals (or covectors, or dual vectors). Row vectors can act on column vectors (via matrix multiplication) to obtain a scalar. The distinction between vectors and covectors usually becomes an issue when changing basis. Here the distinction will be important mostly for formal and structural reasons. I will use the terminology of a vector space and its dual space. I will be using Einstein index summation notation a little, but it is mostly used in redundant expressions.

For simplicity, I will focus on R^3, having basis vectors $(\hat{x}, \hat{y}, \hat{z}) = (e_1, e_2, e_3)$. Its dual R^{3*} has basis vectors $(dx, dy, dz) = (\theta^1, \theta^2, \theta^3)$. The action of a dual basis element on a basis element is 0 or 1 according to whether or not there is an index match. In index notation $\theta^i(e_j) = \delta^i_j$.

The d's in the expressions for the dual basis are suggestive of the infinitesimals of calculus. There are no infinitesimals in this part of the discussion. Our crystal vision will be directed toward straight lines and shiny facets. However, I will discuss later how the d's make sense when our linear structure is applied to calculus.

Tensor Products

Let's approach the idea of a 3x3 array with the formalism of tensor products. The tensor product $R^3 \otimes R^3$ is the vector space with basis created from the nine pairs of R^3 basis elements, represented as $e_i \otimes e_j$. An element A of $R^3 \otimes R^3$ is called a doubly contravariant tensor and can be represented as a 3x3 array A^{ij}, where the elements represent the coefficients of each $e_i \otimes e_j$. We are mostly using the convention that scalar coefficients sourced from the vector space are contravariant, and coefficients sourced from the dual space are covariant. Upper and lower indices are used to signal this difference when index notation is used. The indices for basis covectors and vectors are handled a little differently. The vector u is represented as the sum $u^i e_i$.

The tensor product symbol $\otimes$ can be viewed as a product operation on basis elements. We can extend the definition to all pairs of vectors by asserting bilinearity and the distributive law. In particular, we define the tensor product of two vectors $u \otimes v$ as the the tensor A, where $A^{ij} = u^i v^j$. Because of bilinearity, different pairs of vectors can have the same tensor product. For example, with scalars a and b: $ab(u \otimes v) = (a u) \otimes (b v) = (ab u) \otimes v = u \otimes (ab v)$. Tensor products provide a formalism for bilinear concepts.

The dual tensor product $R^{3*} \otimes R^{3*} \simeq (R^3 \otimes R^3)^*$ is the space created from the nine dual basis elements $\theta^i \otimes \theta^j$. An element ω in $R^{3*} \otimes R^{3*}$ is called a doubly covariant tensor and can be represented as a 3x3 array ω_{ij} , where the elements represent the coefficients of each $\theta^i \otimes \theta^j$. The basis element $\theta^i \otimes \theta^j$ can be interpreted as an element in the isomorphic equivalent $\left(R^3 \otimes R^3\right)^*$. The expression, $\left(\theta^i \otimes \theta^j\right)\left(e_p \otimes e_q\right)$ evaluates to 0 or 1 depending on whether or not all the corresponding indices match (i = p , j = q). In index notation: $\left(\theta^i \otimes \theta^j\right)\left(e_p \otimes e_q\right) = \delta^{ij}_{pq}$. When ω is interpreted as an element of $\left(R^3 \otimes R^3\right)^*$, $\omega_{ij} = \omega\left(e_i \otimes e_j\right)$.

As with contravariant tensors, we can extend the tensor product on dual basis elements to the tensor product of two covectors $\alpha \otimes \beta$, as the tensor ω where $\omega_{ij} = \alpha_i \beta_j$.

The dual tensor product $\left(R^3 \otimes R^3\right)^*$ is also isomorphic to the bilinear dual of the Cartesian product: $\left(R^3 \otimes R^3\right)^* \simeq \left(R^3 \times R^3\right)^{B*}$. Here, I am using non-standard terminology and notation for the space of bilinear functionals on a Cartesian product. I call it a “bilinear dual” because it is similar in flavor to a dual space. We express this isomorphism loosely as $\omega(u \otimes v) = \omega(u, v)$. This is called the “universal property”. Note that the bilinear dual is very different from the standard dual space of the Cartesian product.

When $\theta^i \otimes \theta^j$ is interpreted as an element of $\left(R^3 \times R^3\right)^{B*}$, the expression $\left(\theta^i \otimes \theta^j\right)\left(e_p, e_q\right)$ evaluates to 0 or 1 , depending on whether or not all the corresponding indices match (i = p , j = q). In index notation $\left(\theta^i \otimes \theta^j\right)\left(e_p, e_q\right) = \delta^{ij}_{pq}$. In general, $\omega(u, v) = \omega_{ij} u^i v^j$, where $\omega \in R^{3^*} \otimes R^{3^*}$.

2-Forms

We are interested in the subspace of anti-symmetric tensors in $R^{3^*} \otimes R^{3^*} \simeq \left(R^3 \times R^3\right)^{B*}$ acting on pairs of vectors. This is because of their relationship to the signed area spanned by a pair of vectors. Anti-symmetry means that $\omega_{ij} = -\omega_{ji}$.

The space of 2-forms $\Lambda^2\left(R^{3^*}\right)$, is the subspace of tensors in $\left(R^3 \otimes R^3\right)^*$ that are anti-symmetric. The elements $\left\{\theta^i \otimes \theta^j - \theta^j \otimes \theta^i \ : \ i < j\right\}$ form a basis for $\Lambda^2\left(R^{3^*}\right)$. So, $\Lambda^2\left(R^{3^*}\right)$ has dimension three. We can define wedge multiplication on basic covectors as: $\theta^i \wedge \theta^j = \theta^i \otimes \theta^j - \theta^j \otimes \theta^i$ which gives us a shorter notation for basis elements.

Note that $\theta^i \wedge \theta^j = -\theta^j \wedge \theta^i$ and $\omega = \sum_{ij} \omega_{ij}\left(\theta^i \otimes \theta^j\right) = \sum_{i<j} \omega_{ij}\left(\theta^i \wedge \theta^j\right)$.

We extend the definition of wedge product of basic covectors, to the wedge product of any covectors as $\alpha \wedge \beta = (\alpha \otimes \beta) - (\beta \otimes \alpha)$.

3-Forms in R^3

We can carry on with a three-fold tensor product (bilinear to multilinear):

$$R^{3^*} \otimes R^{3^*} \otimes R^{3^*} \simeq \left(R^3 \otimes R^3 \otimes R^3\right)^* \simeq \left(R^3 \times R^3 \times R^3\right)^{M*}$$

Elements of $R^{3^*} \otimes R^{3^*} \otimes R^{3^*}$ can be viewed as a 3x3x3 array (3d-matrix). Each element has three indices.

Here $\left(R^3 \times R^3 \times R^3\right)^{M*}$ means the space of multilinear functionals on the 3-fold Cartesian product (multilinear dual).

The space of 3-forms $\Lambda^3\left(R^{3*}\right)$, is the subset of tensors in $\left(R^3 \otimes R^3 \otimes R^3\right)^*$ that are anti-symmetric. Anti-symmetry here needs a little more explanation. We sometimes express the three indices as σ , which represents a sequence whose elements are from the set {1,2,3}. We will be focusing on the case when σ is a permutation of (1,2,3).

$$\text{Antisymmetry means:}\quad \omega_{ijk} = \omega_\sigma = \begin{cases} \omega_\tau & \text{if} \quad \tau \text{ is an even permutation of } \sigma \\ -\omega_\tau & \text{if} \quad \tau \text{ is an odd permutation of } \sigma \\ 0 & \text{if} \quad \tau \text{ is not a permutation of } (1,2,3) \end{cases} .$$

We determine whether a permutation is odd or even by counting the number of neighboring swaps to create the full permutation.

The basis elements for $\Lambda^3\left(R^{3*}\right)$ are a little more complicated.
For each increasing sequence of ijk (There is only one.), a basis element is created by an alternating sum of tensor products $\theta^i \wedge \theta^j \wedge \theta^k = \sum_\sigma \operatorname{sgn}(\sigma)\left(\theta^{\sigma_1} \otimes \theta^{\sigma_2} \otimes \theta^{\sigma_3}\right)$, where σ is a permutation of ijk. The sign of the permutation $\operatorname{sgn}(\sigma)$, takes the value +1 or -1 according to whether σ is even or odd. Note that $\theta^i \wedge \theta^j \wedge \theta^k$ is the determinant function and $\left(\theta^i \wedge \theta^j \wedge \theta^k\right)(u, v, w)$ is the signed volume spanned by the vectors u, v, and w.

So, $\Lambda^3\left(R^{3*}\right)$ has only one standard basis element $\theta^1 \wedge \theta^2 \wedge \theta^3$. This turns out to be the determinant function on triples of vectors. So, we see the importance of alternating sums in the calculation of areas and volumes.

About Alternating Sums and 2-Forms
Let's look at 2-forms in more detail. We will be guided by:

$$\left(\theta^i \wedge \theta^j\right)\left(e_i, e_j\right) = \left(\theta^i \wedge \theta^j\right)\left(e_i \otimes e_j\right) = \left(\theta^i \otimes \theta^j - \theta^j \otimes \theta^i\right)\left(e_i \otimes e_j\right) = 1 - 0$$

Next, let's do a detailed calculation, using bilinearity, to show how a 2-form evaluates the signed area spanned by two vectors:

$$\begin{aligned} (dx \wedge dy)(u, v) &= \left(\theta^1 \wedge \theta^2\right)(u, v) = \left(\theta^1 \otimes \theta^2 - \theta^2 \otimes \theta^1\right)(u, v) \\ &= \left(\theta^1 \otimes \theta^2 - \theta^2 \otimes \theta^1\right)\left(u^1 e_1 + u^2 e_2 + u^3 e_3 ,\ v^1 e_1 + v^2 e_2 + v^3 e_3\right) \\ &= \left(\theta^1 \otimes \theta^2\right)\left(u^1 e_1 ,\ v^2 e_2\right) - \left(\theta^2 \otimes \theta^1\right)\left(u^2 e_2 ,\ v^1 e_1\right) = u^1 v^2 - u^2 v^1 \end{aligned}$$

Notice that the final result is similar to a component in a vector cross product.

So, we have another example of the action of an alternating sum. We see that the 2-form expression $(dx \wedge dy)(u, v)$ gives the signed area spanned by the projections of u and v on the xy-plane.

We also see the effects of antisymmetry in degenerate expressions like $dx \wedge dx = 0$. The area spanned by the projection of two vectors onto the x-axis is zero. Any form with a repeated index is zero.

The wedge product inherits the distributive law from the tensor product. Let's calculate the wedge product of two covectors and apply wedge product rules:

$$\alpha\wedge\beta = \left(\alpha_x dx+\alpha_y dy+\alpha_z dz\right)\wedge\left(\beta_x dx+\beta_y dy+\beta_z dz\right)$$

$$\begin{aligned} = & \quad \alpha_x\beta_x dx\wedge dx && + \alpha_x\beta_y dx\wedge dy && + \alpha_x\beta_z dx\wedge dz \\ & + \alpha_y\beta_x dy\wedge dx && + \alpha_y\beta_y dy\wedge dy && + \alpha_y\beta_z dy\wedge dz \\ & + \alpha_z\beta_x dz\wedge dx && + \alpha_z\beta_y dz\wedge dy && + \alpha_z\beta_z dz\wedge dz \end{aligned}$$

$$\begin{aligned} = & \qquad 0 && + \alpha_x\beta_y dx\wedge dy && + \alpha_x\beta_z dx\wedge dz \\ & - \alpha_y\beta_x dx\wedge dy && + \qquad 0 && + \alpha_y\beta_z dy\wedge dz \\ & - \alpha_z\beta_x dx\wedge dz && - \alpha_z\beta_y dy\wedge dz && + \qquad 0 \end{aligned}$$

$$= \left(\alpha_x\beta_y-\alpha_y\beta_x\right)dx\wedge dy + \left(\alpha_x\beta_z-\alpha_z\beta_x\right)dx\wedge dz + \left(\alpha_y\beta_z-\alpha_z\beta_y\right)dy\wedge dz$$

Note the similarity between this expression and a vector cross product.

Similarly, the wedge product of a covector with a 2-form calculates as:

$$\alpha\wedge\omega = \left(\alpha_x dx + \alpha_y dy + \alpha_z dz\right)\wedge\left(\omega_{xy} dx\wedge dy + \omega_{xz} dx\wedge dz + \omega_{yz} dy\wedge dz\right)$$

$$= \left(\alpha_x\omega_{yz}-\alpha_y\omega_{xz}+\alpha_z\omega_{xy}\right)dx\wedge dy\wedge dz$$

Notice the signs of the terms and the permutations of xyz.

The expression of general wedge products based in R^n is more elaborate. This is because the distributive law introduces combinatorics into the expression.

1-Forms
Covectors are essentially in the 1-fold tensor product of the dual vector space. They are also vacuously antisymmetric. So, covectors are 1-forms. Basic 1-forms pick out the scalar value of the relevant component. For example: $dy\left(u_x\hat{x}+u_y\hat{y}+u_z\hat{z}\right) = u_y$

What about $\Lambda^4\left(R^{3*}\right)$?
There are no non-zero anti-symmetric tensors in $R^{3*}\otimes R^{3*}\otimes R^{3*}\otimes R^{3*}$. This is because all 4-tensors on R^3 will have a repeated index. So, there are essentially no 4-forms on R^3 , $\Lambda^4\left(R^{3*}\right) = \{0\}$.

What about $\Lambda^0\left(R^{3*}\right)$?
The basis elements of the 0-fold tensor product of the dual vector space are the 0-tuples of dual basis vectors. The empty-tuple is the only basis element. So, all elements of the 0-fold tensor product are given by scalar multiples of the single basis element. The 0-fold tensor product of R^3 is isomorphic to R. Each of the scalar valued elements is vacuously anti-symmetric. So, $\Lambda^0\left(R^{3*}\right) \simeq R$.

Infinitesimals and Physics Intuition

Later, I will address the application of multilinear algebra to calculus. This synthesis is related to a cultural difference in the way mathematicians and physicists approach mathematics.

This cultural difference reflects different priorities. For mathematicians, the emphasis is on establishing sound logical relationships between abstract structures. For physicists, the emphasis is on using mathematical abstractions to organize physical phenomena and perform plausibly applicable calculations.

When we study physics, we often think about what happens when there is a small displacement. This can result in expressions that look like dz = f(x,y)dx + g(x,y)dy. This can be especially disorienting in the study of thermodynamics. In elementary calculus it is unusual to see an expression like dx by itself. It is normally incorporated into symbols expressing derivatives or integrals and understood as limits. The limit $dx \to 0$ doesn't make much sense. A natural intellectual activity for a physicist, is to break a phenomenon into small pieces that are simpler to understand. This leads to a useful style of infinitesimal thinking.

Infinitesimal thinking requires well tuned intuition in order to reach plausible results. In the case of physics, there is an unspoken assumption that a phenomenon on small increments is almost linear. This predisposition toward linearity in physics may be related to our concept of causation. In other areas of mathematics, such as statistics, the data may be locally noisy and not linear at small scales.

Do you remember when you learned about the derivative of a function and you drew a tangent line at a point on a curve? You did this to visualize the rate of change (slope). The slope of a line is the same whether you look at a small increment or a large piece. In this context, your mind does not have to be burdened by infinitesimals.

The concept of local linearity is used throughout physics. The tangent line to a curve is its simplest application. Note that we think of the tangent line at any point on the curve. I will show how the toolbox of local linearity expands to multilinear algebra attached to any point in space. The local linearity concept adds logical rigor to the more instinctive procedure of infinitesimal thinking.

2-Forms and Surface Integration

I will use surface integration as an example of how local linearity replaces infinitesimal thinking. I will explain the relationship between the 2-form $dx \wedge dy$ and the infinitesimal rectangle dx dy. We will look at integrals over some domain in the Cartesian base plane like $\int f(x,y)\,dx\,dy$.

Previously, we considered a tangent line at any point on a curve. Similarly we will have copies of R^2 glued to every point in the base plane. They have standard basis vectors $(\hat{x},\hat{y})$. Their duals R^{2^*} have standard basis vectors (dx,dy) . All the copies of R^2 are aligned. Each of these local spaces have tensor spaces $R^2 \otimes R^2$ and 2-forms $\Lambda^2(R^{2^*})$. These glittering local multilinear structures are spread over the surface like diamond dust.,

You are familiar with sketching vector fields showing arrows all over the plane. The sketch represents a vector valued function v(x,y). We view each v(x,y) in the local copy of R^2 at the point (x,y) in the base plane.

Similarly, we can consider a 2-form field, which represents local 2-forms at each point in the base plane $\omega(x,y)=f(x,y)\, dx \wedge dy$. A differentiable 2-form field is called a differential 2-form.

Consider the small rectangular shaped increments of area involved with integration. Integration involves a limiting sum like: $\sum f(x,y)\Delta x \Delta y \to \int f(x,y)\, dx\, dy$. Here, Δx and Δy are scalars representing grid size.

Lets change the way we express the area of the small rectangles and rewrite the above sum as: $\sum f(x,y)\, \det(\Delta x\, \hat{x}\, ,\, \Delta y\, \hat{y})$.

We can view the summand as a differential form acting on the spanning vectors of the small rectangles $f(x,y)\det(\Delta x\, \hat{x}\, ,\, \Delta y\, \hat{y})=\omega(\Delta x\, \hat{x}\, ,\, \Delta y\, \hat{y})$.

Then we write: $\sum \omega(\Delta x\, \hat{x}\, ,\, \Delta y\, \hat{y}) \to \int \omega = \int f(x,y)\, dx\, dy$.

The same differential form can be applied to any dissection of the region into approximate parallelograms, each spanned by small vectors Δu and Δv and we can write:
$\int f(x,y)\, dx \wedge dy \approx \sum f(x,y)\, (dx \wedge dy)(\Delta u\, ,\, \Delta v) = \sum f(x,y)\, \det(\Delta u\, ,\, \Delta v) \to \int \omega$
where the Δu's and Δv's are expressed in terms of the basis $\hat{x}$ and $\hat{y}$.

So the expression $dx \wedge dy$ does not directly express an infinitesimal concept, but it guides the linear limiting process on arbitrarily small patches.

As a consequence of its relationship to the determinant, the order of the variables in the expression $dx \wedge dy$ expresses a 2-d orientation. In particular, $dy \wedge dx = -dx \wedge dy$. Skew symmetry of differential forms corresponds to the alternating nature of determinants. The orientation of the area expressed by $\int f(x,y)\, dx \wedge dy$ is given by the differential form in the integrand rather than limits on the domain of integration.

Line Integration and 1-Forms

Consider a differential 1-form $\alpha(x,y) = \alpha_x(x,y)\, dx + \alpha_y(x,y)\, dy$. A typical line integral along a path s(t) might look like: $\int_s (\alpha_x(x,y), \alpha_y(x,y)) \cdot ds$

An approximating sum for the integral can be found by dissecting the path with a sequence of points t_i and creating an anchored shingled path approximation of small vectors:
$\dot{s}(t_i)\Delta t_i = (\dot{s}_{ix}, \dot{s}_{iy})\Delta t_i$

A limiting sum expression for the integral looks like:

$$\int_s (\alpha_x(x,y), \alpha_y(x,y)) \cdot ds \approx \sum_s \alpha(\dot{s}_i \Delta t_i) = \sum_s (\alpha_x\, dx(\dot{s}_i \Delta t_i) + \alpha_y\, dy(\dot{s}_{iy} \Delta t_i))$$

$$= \sum_s \alpha_x(\dot{s}_{ix} \Delta t_i) + \sum_s \alpha_y(\dot{s}_{iy} \Delta t_i) \to \int_s \alpha$$

Again, the expressions dx and dy do not refer to infinitesimals. They direct the limiting process. Notice the component picking feature of dx: $dx(\dot{s}_i \Delta t_i) = \dot{s}_{ix} \Delta t_i$

Changing Basis (Polar Coordinates)

To illustrate the meaning of differential expressions, I want to take a closer look at local vector spaces and changing basis vectors. We will parameterize the base plane using polar coordinates as $x = r\cos\theta$ and $y = r\sin\theta$. You may remember analyzing this situation by dividing the base plane into $\Delta r \times \Delta\theta$ sectors. Here, we will do our analysis by focusing on the local vector spaces at each point in the base plane.

So, we parameterize the base space in r, θ variables (except at the origin). In this case the parameter space is just like R^2, except that the axes are labeled r and θ. The local coordinate systems at each point in the parameter space have basis vectors $\hat{r}$ and $\hat{\theta}$. The map from the parameter space to the base space is denoted by:

$$\phi\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \begin{pmatrix} x(r,\theta) \\ y(r,\theta) \end{pmatrix}$$

Consider a fixed point in the parameter space and its image in the base space:

$u_0 = \begin{pmatrix} r_0 \\ \theta_0 \end{pmatrix}$ and $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \phi\begin{pmatrix} r_0 \\ \theta_0 \end{pmatrix} = v_0$

For a local vector $u = \begin{pmatrix} u_r \\ u_\theta \end{pmatrix}$ at $u_0 = \begin{pmatrix} r_0 \\ \theta_0 \end{pmatrix}$, we can use a linear approximation of ϕ to get:

$$\phi(u_0 + u) - \phi(u_0) \approx J\big|_{u_0} u = \begin{pmatrix} \left.\frac{\partial \phi_x}{\partial r}\right|_{u_0} & \left.\frac{\partial \phi_x}{\partial \theta}\right|_{u_0} \\ \left.\frac{\partial \phi_y}{\partial r}\right|_{u_0} & \left.\frac{\partial \phi_y}{\partial \theta}\right|_{u_0} \end{pmatrix} \begin{pmatrix} u_r \\ u_\theta \end{pmatrix} = \begin{pmatrix} \left.\frac{\partial x}{\partial r}\right|_{u_0} & \left.\frac{\partial x}{\partial \theta}\right|_{u_0} \\ \left.\frac{\partial y}{\partial r}\right|_{u_0} & \left.\frac{\partial y}{\partial \theta}\right|_{u_0} \end{pmatrix} \begin{pmatrix} u_r \\ u_\theta \end{pmatrix} .$$

$$= \begin{pmatrix} \cos\theta_0 & -r_0\sin\theta_0 \\ \sin\theta_0 & r_0\cos\theta_0 \end{pmatrix} \begin{pmatrix} u_r \\ u_\theta \end{pmatrix} = \begin{pmatrix} x_0/r_0 & -y_0 \\ y_0/r_0 & x_0 \end{pmatrix} \begin{pmatrix} u_r \\ u_\theta \end{pmatrix} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

This tells us how to carry local vectors from the parameter space to local vectors in the base space. This induced map is designated by $\phi_*(u) = v$.

We can also define a natural map taking local covectors in the base plane back to local covectors in the parameter space. The image of a covector β is defined by:

$(\phi^*(\beta))(u) = \beta(\phi_*(u))$ where u is a local vector in the parameter space.

As an example, lets use $\beta = \beta_x dx + \beta_y dy$,

$$\text{then } \left(\phi^*(\beta)\right)(u) = \begin{pmatrix}\beta_x & \beta_y\end{pmatrix}\begin{pmatrix} x_0/r_0 & -y_0 \\ y_0/r_0 & x_0 \end{pmatrix}\begin{pmatrix} u_r \\ u_\theta \end{pmatrix} = \begin{pmatrix}\alpha_r & \alpha_\theta\end{pmatrix}\begin{pmatrix} u_r \\ u_\theta \end{pmatrix}$$

We can say

$$\phi^*(\beta) = \begin{pmatrix}\beta_x & \beta_y\end{pmatrix}\begin{pmatrix} x_0/r_0 & -y_0 \\ y_0/r_0 & x_0 \end{pmatrix} = \begin{pmatrix}\alpha_r & \alpha_\theta\end{pmatrix} = \alpha$$

We can go the other direction and say:

$$\phi^{*-1}(\alpha) = \begin{pmatrix}\beta_x & \beta_y\end{pmatrix} = \begin{pmatrix}\alpha_r & \alpha_\theta\end{pmatrix}\begin{pmatrix} \cos\theta_0 & -r_0\sin\theta_0 \\ \sin\theta_0 & r_0\cos\theta_0 \end{pmatrix}^{-1} = \begin{pmatrix}\alpha_r & \alpha_\theta\end{pmatrix}\begin{pmatrix} \cos\theta_0 & \sin\theta_0 \\ -\frac{1}{r_0}\sin\theta_0 & \frac{1}{r_0}\cos\theta_0 \end{pmatrix} = \beta$$

Images of the standard local bases in the parameter space carry over to the local base spaces as:

$$\phi_*(\hat{r}) = \begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix} = \frac{1}{r_0}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \quad \text{and} \quad \phi_*(\hat{\theta}) = r_0\begin{pmatrix} -\sin\theta_0 \\ \cos_0 \end{pmatrix} = \begin{pmatrix} -y_0 \\ x_0 \end{pmatrix}$$

I will refer to these alternative basis elements as $\hat{r}$ and $\hat{\theta}$ even though they are not in the parameter space. Also $\hat{\theta}$ is not a unit vector in the local base spaces.

Similarly, images of the standard local dual basis in the parameter space carry over to the local dual spaces as:

$$\phi^{*-1}(dr) = \begin{pmatrix}\cos\theta_0 & \sin\theta_0\end{pmatrix} = \cos\theta_0\, dx + \sin\theta_0\, dy = \frac{x_0}{r_0}dx + \frac{y_0}{r_0}dy$$

$$\text{and} \quad \phi^{*-1}(d\theta) = -r_0\sin\theta_0\, dx + r_0\cos\theta_0\, dy = -y_0\,dx + x_0\,dy$$

As with basis vectors, I will be ambiguous about which space the basis covectors are in by writing : $dr = \cos\theta_0\, dx + \sin\theta_0\, dy$ and $d\theta = -r_0\sin\theta_0\, dx + r_0\cos_0\theta\, dy$

Note that $dr(\hat{r}) = 1$ and $d\theta(\hat{\theta}) = 1$.

You will be happy to know that the expressions for dr and $d\theta$ can be calculated from simple differentiation. For example $r = \sqrt{x^2 + y^2}$. The gradient of r can be calculated as:

$$\nabla r\,\big|_{x_0, y_0} = \left(\frac{\partial r}{\partial x}\bigg|_{x_0, y_0} , \frac{\partial r}{\partial y}\bigg|_{x_0, y_0} \right) = \frac{1}{2\sqrt{x_0^2 + y_0^2}}\left(2x_0 , 2y_0\right) = \left(\frac{x_0}{r_0} , \frac{y_0}{r_0}\right) = \left(\cos\theta_0 , \sin\theta_0\right)$$

As a 1-form, this is expressed as: $dr = \frac{x_0}{r_0}dx + \frac{y_0}{r_0}dy$

Interpreting dr

We can interpret the expression $dr = \frac{x_0}{r_0}dx + \frac{y_0}{r_0}dy$ using infinitesimal thinking or local linear thinking. With infinitesimal thinking the expression says how much r changes with small changes in x and y. This changes from place to place in the base plane.

With local linear thinking, the expression says that the two sides of the equation act on any local vector the same way, no matter how small it is.

Consider a local vector $u = u_x\hat{x} + u_y\hat{y}$.

$$dr(u) = \frac{x_0}{r_0}dx(u) + \frac{y_0}{r_0}dy(u) = \frac{x_0}{r_0}u_x + \frac{y_0}{r_0}u_y$$

This expression is true regardless of the size of u.

If you study differential geometry, dr can also be interpreted as the exterior derivative of the coordinate function r.

History of “d”: Leibniz , Grassmann , and Cartan

This is a story about the synthesis of smooth ideas in calculus with the rigid structures of algebra.

Gottfried Leibniz introduced the "d" symbol to express infinitesimal notions in order to develop calculus in the late 1600's. This lead to the intellectual tool of infinitesimal thinking that is widely used in physics. However, this tool lacked the same logical rigor found in Euclidian geometry, and could lead to contradictions if not informed by sound physical intuition.

From the middle 1700's, Bolzano, d’Alembert, and Cauchy developed the ideas of derivatives and continuity as limit concepts. By the middle 1800's Karl Weierstrass had a fully formal logical system for calculus.

In a seemingly unrelated effort, Hermann Grassmann examined the implications of extended structures on vector spaces in the middle 1800's. This resulted in an early version of the multilinear algebra that we use today. At the time, it went unnoticed.

In the early 1900's, Elie Cartan recognized the local linearity of calculus, made rigorous by Weierstrass. He then fused these concepts with the linear structures developed by Grassmann. The recruitment of "d" for use with differential forms happened at this time. Thanks Elie.

If you study differential geometry, you will learn more about the exterior algebra of k-forms and different types of tensors. A main goal in this direction is learning about the exterior derivative of differential k-forms and the general stokes theorem. Cartan also uses the “d” symbol here.

Vector Cross Products

Multilinear algebra has a lot to say about the vector cross products used in R^3 . Let's focus on how cross products are used in physics. Physical concepts should be independent of the choice of coordinate system or cultural artifacts like the right-hand rule. Multiple sign issues pop up when a cross product expression is expressed in a coordinate system with reversed orientation.

To view cross products through the lens of multilinear algebra, recall our previous discussion about the 2-forms $\Lambda^2\left(R^{3^*}\right)$, as a subspace of tensors in $R^{3^*} \otimes R^{3^*}$. There is an analogous theory about 2-vectors (bivectors) as a subspace of tensors in $R^3 \otimes R^3$. In order to talk about the vector cross product, I will say more about bivectors and introduce the Hodge star (Hodge dual) operation. In the end, you will understand the vector cross product as: $u \times v = *\left(u \wedge v\right)$. It will take a little work to get there. But, this helps with understanding the organization of sign issues with reversed orientation. This is not just a change in notation.

Simple Torque Example

Let's look at torque expressed as a cross product $T = r \times F$, with standard coordinates. If we are given r and T, we can find an equivalent perpendicular force f on the lever arm so that :

$T = r \times F = r \times f$

We can find f with this cross product expression: $f = \left(T \times r\right) \frac{1}{\|r\|^2}$

This is a little like saying $f = \frac{T}{r}$ (you can't do this with vectors).

Note that there are many f's that satisfy $r \times F = r \times f$. For example: $r \times F = r \times \left(F + r\right)$

I will be referring to a simple case where $r = \hat{x}$ and $F = \hat{y}$ to get $T = \hat{z}$. Then, one more step gives us $f = \hat{y}$.

Physical Space and Coordinates

Before discussing left handed coordinates, I want to talk about how I am looking at physical space and coordinates.

Physical space has points, angles and distances. But, points don't have coordinates and there are no named directions. Think Euclidean geometry where we use concepts of superposition rather than numbers. This provides a framework where we can talk about all the concepts in Newtonian physics. We can enhance our framework to support numerical calculations, when we embed a copy of R^3. The embedding must be compatible with distance and angles. This makes the coordinate basis orthonormal.

The physics that happens in space does not depend on the particular embedding of R^3. When describing physics, we try not to refer to the coordinates.

For example, we can embed R^3 in space so that the coordinate axes point east, north and up. Or they could point east, north and down. We see a difference between these embeddings by relating them to the patterns of fingers on our hands.

A less anatomical way to express orientation is to specify the signed volume spanned by the basis vectors (volume element). We usually express this as a 3-form expressed in terms of the det() function. For a “reversed” coordinate system we choose $\mathrm{vol}() = -\det()$.

Choosing the embedding of R^3 and the volume element is part of the process of using mathematical abstractions to organize physical phenomena.

Torque Example with Reversed Orientation

Let's use our previous torque example with coordinate axes having reversed orientation. The idea of reversed orientation makes sense if we have another coordinate axes for comparison. The new orthonormal basis is $(\hat{a}, \hat{b}, \hat{c}) = (\hat{x}, \hat{y}, -\hat{z})$. Designating which one is reversed is a matter of choice. Here, the relationship with standard coordinates gives us
$\mathrm{vol}(\hat{a}, \hat{b}, \hat{c}) = -1$.

Our vector calculations must result in vectors pointing in the right physical direction.
We must make a rule that all cross products follow a left-hand rule, so that: $\hat{a} \times \hat{b} = -\hat{c}$

We have, $r = \hat{a}$ and $F = \hat{b}$. Then $T = r \times F = \hat{a} \times \hat{b} = -\hat{c}$.

If we want to calculate f from T and r, We have to recognize that we are in a reversed coordinate system and calculate:

$$f = (T \times r)\frac{1}{\|r\|^2} = ((-\hat{c}) \times \hat{a})\frac{1}{\|r\|^2} = \hat{b}$$

The sign issues that we see here are handled better when we recognize $T = *(r \wedge F)$.

Bivectors

Bivectors $\Lambda^2(R^3)$ are the antisymmetric tensors in $R^3 \otimes R^3$. We model the definition of wedge product of two vectors u and v, on the definition of wedge product on covectors as:
$u \wedge v = (u \otimes v) - (v \otimes u)$
It is bilinear and antisymmetric.

The space of bivectors has three standard basis elements like $e_i \wedge e_j$ where i <j .

Bivectors are naturally isomorphic to the dual space of 2-forms. Evaluation follows this rule:
$(u \wedge v)(\omega) = \omega(u, v)$

Another way to describe bivectors is the space generated by equivalence classes of vector pairs $[(u, v)]$. The equivalence relation is given by $(u, v) \approx (s, t)$ whenever $\omega(u, v) = \omega(s, t)$ for all 2-forms ω . More intuitively, pairs of vectors are equivalent if they span an equal area in the same plane.

We have three views of bivectors:

1. defined as a subspace of a tensor space
2. isomorphic to the dual of 2-forms
3. isomorphic to a space of equivalence classes

It takes a little work to prove the isomorphisms between these spaces, but it all follows from bilinearity.

Bivectors are also called 2-vectors, and 1-vectors are ordinary vectors. The space of 1-vectors is isomorphic to the dual space of 1-forms.

The Hodge Star on k-Forms

Most people are using the Hodge Star (or Hodge dual) when doing flux integration through a surface in 3D Cartesian space. A 2-form acting on surface elements becomes a 1-form acting on surface normals. The 1-form is represented as a vector field projected onto surface normals.

Some examples:

$$*(dx \wedge dy) = dz \quad , \quad *dx = (dy \wedge dz) \quad , \quad *(dx \wedge dy \wedge dz) = 1 \quad , \quad *(dx \wedge dz) = -dy$$

The defining expression for a k-form ω in R^3 (where ω is a wedge product of dual basis elements) is $\omega \wedge {*\omega} = dx \wedge dy \wedge dz = \mathrm{vol}() = +\det()$. This definition is extended linearly from the k-form basis to all k-forms. For R^3 and any k-form ω $\; **\omega = \omega$

Changing Basis and the Hodge Star

When we embed R^3 into physical space, we can encode orientation by choosing the definition of our vol() function and how it acts on the basis vectors. In our case $\mathrm{vol}_{abc}(\hat{a}, \hat{b}, \hat{c}) = -1$. Then we can use the following expression to calculate $*(da \wedge db)$:

$$((da \wedge db) \wedge *(da \wedge db))(\hat{a}, \hat{b}, \hat{c}) = \mathrm{vol}_{abc}(\hat{a}, \hat{b}, \hat{c}) = -1 = ((da \wedge db) \wedge (-dc))(\hat{a}, \hat{b}, \hat{c})$$

So, $*(da \wedge db) = -dc$ Note: $(da \wedge db \wedge dc)(\hat{a}, \hat{b}, \hat{c}) = 1$ always.

Similarly $*(da \wedge dc) = db$ and in R^3 , the Hodge star introduces a sign flip in a reversed coordinate system. For ω a k-form basis element, the defining expression $\omega \wedge *\omega = \mathrm{vol}()$ is universal in R^3 , as long as vol() correctly calculates the signed volume spanned by the basis elements.

The Hodge Star on k-Vectors

There is a notion of Hodge star for k-vectors that is just like the Hodge star for k-forms. It is difficult to directly define the Hodge star in terms of a wedge product and the vol() function. However, we can view k-vectors as elements of the dual of k-forms. Then the defining expression for the Hodge star for a k-vector A is given by $(*A)(*\omega) = A(\omega)$, for any k-form ω .

Some examples: $*(\hat{x} \wedge \hat{y}) = \hat{z} \quad , \quad *\hat{x} = (\hat{y} \wedge \hat{z}) \quad , \quad *(\hat{x} \wedge \hat{y} \wedge \hat{z}) = 1 \quad , \quad *(\hat{x} \wedge \hat{z}) = -\hat{y}$

Verify with a sample calculation:

$$1 = (\hat{x} \wedge \hat{z})(dx \wedge dz) = (*(\hat{x} \wedge \hat{z}))(*(dx \wedge dz)) = (-\hat{y})(-dy) = 1$$

Let's do a sample calculation of the Hodge star of a more general bivector $u \wedge v$.

Expanding, we get:

$$u \wedge v = \left(u_x \hat{x} + u_y \hat{y} + u_z \hat{z}\right) \wedge \left(v_x \hat{x} + v_y \hat{y} + v_z \hat{z}\right)$$
$$= \left(u_x v_y - u_y v_x\right) \hat{x} \wedge \hat{y} + \left(u_x v_z - u_z v_x\right) \hat{x} \wedge \hat{z} + \left(u_y v_z - u_z v_y\right) \hat{y} \wedge \hat{z}$$

Applying the Hodge star gives:

$$*\left(u \wedge v\right) = \left(u_x v_y - u_y v_x\right) \hat{z} - \left(u_x v_z - u_z v_x\right) \hat{y} + \left(u_y v_z - u_z v_y\right) \hat{x}$$
$$= \left(u_y v_z - u_z v_y\right) \hat{x} + \left(u_z v_x - u_x v_z\right) \hat{y} + \left(u_x v_y - u_y v_x\right) \hat{z} = u \times v$$

Torque as the Hodge Star of a Bivector
Previously we saw torque as a cross product $T = r \times F$, and then tinkered with the signs if we were in a coordinate system with left handed orientation.

Let's see what happens if we define torque as $T = *\left(r \wedge F\right)$. The result is the right vector, but we do not have to directly tinker with the signs.

With the standard coordinate system, $T = *\left(\hat{x} \wedge \hat{y}\right) = \hat{z}$.
With the alternate coordinate system, $T = *\left(\hat{a} \wedge \hat{b}\right) = -\hat{c} = \hat{z}$.

The sign issues are handled by the Hodge star, which is based on vol(), which is based on the handling of signed volume.

Similarly, to find f, we use the bivector analog of $f = \left(T \times r\right) \frac{1}{\|r\|^2}$:

$$f = *\left(T \wedge r\right) \frac{1}{\|r\|^2}$$

With the standard coordinate system, $f = *\left(\hat{z} \wedge \hat{x}\right) = \hat{y}$.
With the alternate coordinate system, $f = *\left(-\hat{c} \wedge \hat{a}\right) = *\left(\hat{a} \wedge \hat{c}\right) = \hat{b} = \hat{y}$.

Concluding
We followed three lines of thought:

- There is a rich tensor structure generated by vector spaces.
- Tensors provide a variety of local linear concepts which are applied to differential geometry.
- Philosophy of physics issues involving the cross product are better addressed in the context of bivectors.

I have used R^3 to introduce several multilinear concepts. If you study differential geometry, you will encounter the same concepts in greater generality.

Questions and comments are welcome: jed@islandnet.com