ASSOCIATED VECTOR SPACES
as used in Differential Geometry

William Schulz

1. INTRODUCTION

In this module[1] we will discuss the concept of Associated Vector Spaces based on ideas from Differential Geometry, specifically the theory of Associated Bundles. This is the algebra that lies behind the construction. We will not discuss the analysis or manifolds; just the algebra, which is merely linear algebra.

The notation is interesting and efficient and forms a mathematically coherent way of incorporating the old Tensor Algebra of Riemann and Levi-Civita into modern mathematics.

We will also discuss a less flexible matrix formulation of the same ideas.

2. REPRESENTATIONS

Let V and W be vector spaces over a field k. In these notes we will take $k = \mathbb{R}$ for simplicity but our theory will work just as well over any field.

Def $\mathrm{Hom}_k(V, W)$ is the set of k-linear maps from V to W, which satisfy

$$\begin{aligned} T(v_1 + v_2) &= T(v_1) + T(v_2) \qquad \text{for } v_1, v_2 \in V \\ T(\alpha v) &= \alpha T(v) \qquad \text{for } \alpha \in k \end{aligned}$$

Purely for convenience we will write $\mathrm{Hom}(V, W)$ for $\mathrm{Hom}_k(V, W)$ since the field will be fixed and in fact be $\mathbb{R}$. We will denote by $\mathrm{Aut}(V)$ the subset of $\mathrm{Hom}(V, V)$ consisting of invertible tranformations. Everything we do will work over any field.

Def A Representation of a Group G is a map $\rho : G \to \mathrm{Hom}(V, V)$ where V is a vector space, satisfying

$$\begin{aligned} \rho(\mathrm{Id}_G) &= \mathrm{I} \qquad \mathrm{I} \text{ is the identity in } \mathrm{Hom}(V, V) \\ \rho(g_1 g_2) &= \rho(g_1)\rho(g_2) \end{aligned}$$

The axioms for a group representation actually imply that $\rho(g) \in \mathrm{Aut}(V)$. Since $\rho(g) \in \mathrm{Aut}(V)$, ρ eats group elements and excretes linear transformations. Thus

$$\rho(g)v = w, \qquad v, w \in V$$

makes sense. We sometimes write

$$g.v \text{ for } \rho(g)v$$

[1] 31 Jan 2012

and we speak of a group action on V. However, "group action" is a more general concept, and groups may act on many different kinds of mathematical objects, not just vector spaces. Group actions are one of the core ideas of modern mathematics and we should do a better job of introducing students to them.

Example Consider the group G whose table is

	1	a	b	c
1	1	a	b	c
a	a	b	c	1
b	b	c	1	a
c	c	1	a	b

Let

$$V = \mathbb{R}^2 = \left\{ \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \middle| r_i \in \mathbb{R} \right\}.$$

Let

$$\rho(a) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} -r_2 \\ r_1 \end{pmatrix}$$

Then

$$\rho(b) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \rho(a^2) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \rho(a)\rho(a) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \rho(a) \begin{pmatrix} -r_2 \\ r_1 \end{pmatrix} = \begin{pmatrix} -r_1 \\ -r_2 \end{pmatrix}$$

and

$$\rho(c) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \rho(a)\rho(b) \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \rho(a) \begin{pmatrix} -r_1 \\ -r_2 \end{pmatrix} = \begin{pmatrix} r_2 \\ -r_1 \end{pmatrix}$$

Geometrically, $\rho(a)$ rotates a vector in $\mathbb{R}^2$ positively by $\pi/2$, so that with a little more effort we could create the complex numbers $\mathbb{C}$ from this representation.

If we introduce a basis into $\mathbb{R}^2$ we can get a *matrix* representation of G. In general, if $\{e_1, \ldots, e_n\}$ is a basis of V then

$$\begin{aligned} \rho(e_i) &= \sum_j e_j \alpha_i^j \\ \rho(g) &\longleftrightarrow \begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1^n & \alpha_2^n & \ldots & \alpha_n^n \end{pmatrix} \end{aligned}$$

Clearly if

$$\begin{aligned} \rho(g_1) &\longleftrightarrow (\alpha_i^j) \\ \rho(g_2) &\longleftrightarrow (\beta_i^j) \end{aligned}$$

then

$$\rho(g_2)\rho(g_1) \longleftrightarrow (\beta_i^j)(\alpha_i^j)$$

Also it is clear that

$$\rho(\mathrm{Id}_G) \longleftrightarrow (\delta_i^j) = \text{Identity matrix}$$

Example Let in our previous example

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

then

$$\begin{aligned} \rho(a)e_1 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e_1 0 + e_2 1 \\ \rho(a)e_2 &= \begin{pmatrix} -1 \\ 0 \end{pmatrix} = e_1(-1) + e_2 0 \end{aligned}$$

so that

$$\rho(a) \longleftrightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Then we have

$$\begin{aligned} \rho(\mathrm{Id}_G) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \rho(a) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \rho(b) &= \rho(a)^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \\ \rho(c) &= \rho(a)\rho(b) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned}$$

Thus we have a matrix representation of G; that is map $G \to \mathrm{GL}(2, \mathbb{R})$ which is isomorphic to $\mathrm{Aut}(\mathbb{R}^2)$.

3. BASIS CHANGE

In this section we will introduce some matrix methods for handling basis change, and then using these as a springboard we will show how to construct the equivalence classes that exhibit V as an associated vector space to P which is the set of all bases, and also (essentially) $\mathrm{GL}(n, \mathbb{R})$. If we wish to consider some subset of bases, for example, Orthonormal bases, it is only necessary to replace $\mathrm{GL}(n, \mathbb{R})$ by $\mathrm{O}(n, \mathbb{R})$, the Orthogonal Group.

Let $\{e_1, \ldots, e_n\}$ (old basis) and $\{f_1, \ldots, f_n\}$ (new basis) be bases of V. Let us write bases of V, for purposes of matrix manipulation, as rows, $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$. In the e_i basis let $v \in V$ be $v = \sum e_i v^i$ and in the f_j basis as

$v = \sum f_j \tilde{v}^j$. Then we can write these as

$$v = (e_1, \dots, e_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

or, using the f_i basis as

$$v = (f_1, \dots, f_n) \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix}$$

If we use the abbreviations $\underline{\mathrm{e}} = (e_1, \dots, e_n)$ and $\underline{\mathrm{f}} = (f_1, \dots, f_n)$ and

$$\vec{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad \vec{\tilde{v}} = \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix}$$

Then we can write the previous equations as

$$v = \underline{\mathrm{e}}\vec{v} = \underline{\mathrm{f}}\vec{\tilde{v}}$$

Now suppose that the e_i and f_j bases are connected by the relation

$$f_j = \sum e_i g^i_j$$

which can be written in matrix form as

$$(f_1, \dots, f_n) = (e_1, \dots, e_n) \begin{pmatrix} g^1_1 & g^1_2 & \cdots & g^1_n \\ g^2_1 & g^2_2 & \cdots & g^2_n \\ \cdots & \cdots & \cdots & \cdots \\ g^n_1 & g^n_2 & \cdots & g^n_n \end{pmatrix}$$

or in more abbreviated form

$$\underline{\mathrm{f}} = \underline{\mathrm{e}}\, g$$

Now we can easily see how to change coordinates when the basis changes:

$$v = \underline{\mathrm{e}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \underline{\mathrm{e}}\, g g^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \underline{\mathrm{f}}\, g^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Since

$$v = \underline{\mathrm{f}} \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix}$$

we have

$$\boxed{\begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix} = g^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}}$$

Thus our matrix formalism has easily found the correct transformation formula for the components of vectors.

We now denote by P the set of all bases of the vector space V. The elements of P are ordered bases $e = \{e_1, \ldots, e_n\}$ The difference between e and $\underline{\mathrm{e}}$ is mystical; $\underline{\mathrm{e}}$ is e when it is written as a row vector. The equation $f = eg$ means the same thing as $\underline{\mathrm{f}} = \underline{\mathrm{e}}\, g$, namely

$$f_j = \sum e_i g^i_j$$

This equation $f = eg$ is an example of a *right action* on P. If we have $g, h \in \mathrm{GL}(n, \mathbb{R})$ it is easily checked that for $e \in P$ we have

$$(eg)h = e(gh)$$

which is the characteristic equation of right actions. Moreover, this right action has two special properties. First it is *transitive*, which means that given two bases $e, f \in P$ there exists some $g \in \mathrm{GL}(n, \mathbb{R})$ for which

$$f = eg$$

Second, the right action is *effective* which means that if $eg = e$ then g is the identity. Effectivity guarantees that the g provided by transitivity is unique

Taken together, these two properties mean that P and $\mathrm{GL}(n, \mathbb{R})$ are not very different, since given a fixed basis $e_0 \in P$ there is a one to one correspondence between P and $\mathrm{GL}(n, \mathbb{R})$ given by $e \longleftrightarrow g$ if and only if $e = e_0 g$. The bijectivity is the consequence of the transitivity and effectiveness of the group right action.

Now let ρ be a representation of $G = \mathrm{GL}(n, \mathbb{R})$ on some vector space W. We form the Cartesian product

$$P \times W$$

and then define an equivalence relation on $P \times W$ by

$$(eg, w) \sim (e, \rho(g)w)$$

We will denote the equivalence class of (e, w) by $[e, w]$ so that

$$[eg, w] = [e, \rho(g)w]$$

The equivalence classes form a vector space which is denoted by

$$P \times_\rho W$$

and the resulting vector space is called an *Associated Vector Space* to P.

Here is the definition of addition; given $[e, w_1]$ and $[f, w_2]$, determine the unique g for which $f = eg$. Then

$$[e, w_1] + [f, w_2] = [e, w_1] + [eg, w_2] = [e, w_1] + [e, \rho(g)w_2] = [e, w_1 + \rho(g)w_2]$$

Next we present the critical example in which we show that V itself is an associated vector space of P, it's set of bases. To do this we must have two things; a vector space W and a representation $\rho : G \rightarrow \text{Aut}(W)$. For W we take $\mathbb{R}^n$ (which we write as column vectors) and for $\rho(g)$ we take left multiplication by g itself so that

$$\rho(g) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = g \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

Now if $v = \sum e_i v^i$ we set up an isomorphism

$$v \longleftrightarrow [e, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}]$$

Now the critical thing in this game is how well the correspondence survives change of basis. So let $f = eg$. Then the rules give us

$$\begin{aligned} [e, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}] &= [egg^{-1}, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}] = [eg, \rho(g^{-1}) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}] \\ &= [eg, g^{-1} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}] = [f, \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix}] \end{aligned}$$

exactly as we would wish; that is, the isomorphism does not depend on whether we use basis e or f. Why does this work? Because we selected the correct W and the correct group representation on W of G.

In ancient times, mathematicians often defined (contravariant) vectors as quantities which had n ordered components and transformed under basis change as indicated above. Historically the basis changes were caused by coordinate changes on the manifold, which we don't see here in the algebra. The ancients new all about the group actions we have considered, but had no easy way to make these actions clear. Our methodology succeeds in doing this in a very convenient way.

The equivalence class $[e, \vec{v}] = [e, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}]$ actually captures the vector better than other methods, because it contains the representation of the vector in all possible bases, thus eliminating the emotional tendency to pick some "canonical" basis and treat the coordinates in that basis better than others. This method is totally fair.

For simple purposes this method may be counterfeited as we did above, by setting

$$\underline{e}\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \longleftrightarrow [e, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}]$$

4. DUAL SPACE AND ITS ρ

A linear functional is a linear function from a vector space V to its field k. Here $k = \mathbb{R}$. The set of linear functionals is denoted by V^*. It is a vector space. For the value of a linear functional $\lambda \in V^*$ on $v \in V$ we write

$$\lambda(v) \text{ or } \langle \lambda, v \rangle$$

The second form better brings out the duality between V^* and V. It is easy to show (but it requires using a basis) that V^{**} is naturally isomorphic to V for finite dimensional spaces. It is also important to know that this fails in general for infinite dimensional spaces, and should it be true the infinite dimensional space is said to be *reflexive*.

Each basis $\{e_1, \ldots, e_n\}$ of V has a corresponding basis $\{e^1, \ldots, e^n\}$ of V^*, called the dual basis. The e^i are defined by

$$e^i(e_j) = \langle e^i, e_j \rangle = \delta^i_j$$

Notice that if $v = e_1 v^1 + \ldots + e_n v^n$ then

$$\begin{aligned} e^i(v) &= \langle e^i, \sum_j e_j v^j \rangle \\ &= \sum_j v^j \langle e^i, e_j \rangle \\ &= \sum_j v^j \delta^i_j \;=\; v^i \end{aligned}$$

so e^i picks out the e_i component of v.

If $\{f_1, \ldots, f_n\}$ is another basis of V with dual basis $\{f^1, \ldots, f^n\} \subseteq V^*$, then we must have

$$f^i = \sum h^i_j e^j$$

for some (h^i_j), and this (h^i_j) must relate to (g^i_j) where $f_i = \sum_j e_i g^i_j$ in V. We will determine this relationship with an amusing matrix technique, although it is easy enough to do it using standard techniques. This will serve to introduce some notation.

Let $\lambda = \sum_j \lambda_j e^j$ be a typical element of V^* and write

$$e^* = \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix}$$

and the coefficients as a row vector $(\lambda_1, \ldots, \lambda_n)$ so

$$\lambda = \vec{\lambda} e^* = (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = \sum_j \lambda_j e^j$$

so that we have a matrix representation of λ.

Next we write

$$\langle e^*, e \rangle = \langle \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix}, (e_1, \ldots e_n) \rangle = \begin{pmatrix} \langle e^1, e_1 \rangle & \cdots & \langle e^1, e_n \rangle \\ \ldots & \ldots & \ldots \\ \langle e^n, e_1 \rangle & \cdots & \langle e^n, e_n \rangle \end{pmatrix} = (\delta^i_j) = \mathrm{I}$$

Similarly

$$\langle f^*, f \rangle = \mathrm{I}$$

Now $f^i = \sum_j h^i_j e^j$ translates into

$$\begin{aligned} \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix} &= (h^i_j) \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} \\ f^* &= h e^* \end{aligned}$$

Then we have

$$\mathrm{I} = \langle f^*, f \rangle = \langle h e^*, e g \rangle = h g \langle e^*, e \rangle = h g (\delta^i_j) = h g \mathrm{I}$$

so we have shown that

$$h = g^{-1}$$

Next we have

$$\lambda = (\lambda_1, \ldots \lambda_n) \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = (\lambda_1, \ldots \lambda_n) g g^{-1} \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} = (\lambda_1, \ldots \lambda_n) g \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix}$$

and since

$$\lambda = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n) \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix}$$

we have

$$\boxed{(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n) = (\lambda_1, \ldots \lambda_n)g}$$

Let us now verify that everything is consistent:

$$\begin{aligned}
\langle \lambda, v \rangle &= \sum_i \lambda_i v^i = (\lambda_1, \ldots \lambda_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\
&= (\lambda_1, \ldots \lambda_n) g g^{-1} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \\
&= (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n) \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix} = \sum_j \tilde{\lambda}_j \tilde{v}^j
\end{aligned}$$

so that we can compute $\langle \lambda, v \rangle$ in either coordinate system in the same way.

Now we wish to put V^* into the form of an associated vector space to P. We need a vector space W and a representation ρ of G on W which mimics the way V^* works. We will take

$$W = \mathbb{R}^{n*} = \{(\lambda_1, \ldots, \lambda_n) \big| \lambda_i \in \mathbb{R}\}$$

and let ρ be defined by

$$\rho(g)(\lambda_1, \ldots, \lambda_n) = (\lambda_1, \ldots, \lambda_n)g^{-1}$$

This is indeed a group representation (left action) as we now show.

$$\begin{aligned}
\rho(gh)(\lambda_1, \ldots, \lambda_n) &= (\lambda_1, \ldots, \lambda_n)(gh)^{-1} \\
&= (\lambda_1, \ldots, \lambda_n)h^{-1}g^{-1} \\
&= \Big(\rho(h)(\lambda_1, \ldots, \lambda_n)\Big)g^{-1} \\
&= \rho(g)\Big(\rho(h)(\lambda_1, \ldots, \lambda_n)\Big)
\end{aligned}$$

so indeed $\rho(gh) = \rho(g)\rho(h)$ as required.

Now we can show the isomorphism

$$V^* \longleftrightarrow P \times_\rho W = P \times_\rho \mathbb{R}^{n*}$$

as

$$\lambda = (\lambda_1, \ldots, \lambda_n) \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix} \longleftrightarrow [e, (\lambda_1, \ldots, \lambda_n)]$$

Note particularly that the first element in $[e, (\lambda_1, \ldots, \lambda_n)]$ is $e = (e_1, \ldots, e_n)$, the basis of V, NOT the basis of V^*. This is critical to the system.

We must of course verify that this isomorphism does not depend on the choice of coordinate system. With all our equipment this is very easy.

$$\begin{aligned}
[e,(\lambda_1,\ldots,\lambda_n)] &= [egg^{-1},(\lambda_1,\ldots,\lambda_n)] \\
&= [eg,\rho(g^{-1})(\lambda_1,\ldots,\lambda_n)] \\
&= [eg,(\lambda_1,\ldots,\lambda_n)g] \\
&= [f,(\tilde{\lambda_1},\ldots,\tilde{\lambda_n})]
\end{aligned}$$

5. TENSOR PRODUCT REPRESENTATIONS

We will briefly recall how tensor products work. We will work with a product of two vector spaces but the definitions will work just the same for any finite number of vector spaces.

Tensor products in $V \otimes W$ are subject to two laws, namely for all $v, v_1, v_2 \in V$ and $w, w_1, w_2 \in W$

$$\begin{aligned}(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2 \\ \alpha(v \otimes w) &= (\alpha v) \otimes w \;=\; v \otimes (\alpha w)\end{aligned}$$

A proper mathematical way to to define the tensor product would be to form the algebra R over k generated by the Cartesian product $V \times W$ and then form the ideal I generated by all elements of the form $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$ and similarly for the other equations, and then

$$V \otimes W = R/I$$

This puts the tensor product $V \otimes W$ into the standard mathematical mold and the result is a vector space over K It is important to take a very Zen view of this situation; there are the basic equations above. The only other thing one has to learn about tensor products (in this context) is that there is *nothing* more to learn.

Now on to representations. Let

$$\rho_i : G \to V_i, \quad i = 1, 2, \ldots, r$$

be a set of r representations on various vector spaces V_i. We can squish them all together into a single tensor product representation on $V_1 \otimes V_2 \otimes \ldots \otimes V_r$ by

$$\rho(g)(v_1 \otimes v_2 \otimes \ldots \otimes v_r) = \rho_1(g)(v_1) \otimes \rho_2(g)(v_2) \otimes \ldots \otimes \rho_r(g)(v_r)$$

This ρ is called the tensor product of the ρ_i; $\rho = \rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_r$. Not surprisingly such tensor products of representations are the key to tensor algebra.

We are going to illustrate these concepts by showing that $\mathrm{Hom}(V, V)$ is an associated vector space of G. First we must show that $\mathrm{Hom}(V, V)$ is indeed a tensor product, which is not particularly obvious. To save us from a notational snowstorm we will do this with a 2 by 2 numerical example. Let T be a linear operator on an 2-dimensional vector space, and after introduction of a basis e suppose the matrix comes out

$$T \overset{e}{\longleftrightarrow} \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}$$

Now we will decompose T as follows:

$$\begin{aligned}\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix} &= \begin{pmatrix} 2 & 3 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -1 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} (2, 3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (-1, 4)\end{aligned}$$

Then the isomorphism between $\mathrm{Hom}(V,V)$ and $\mathbb{R}^2 \otimes \mathbb{R}^{2*}$ is given by

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} (2,3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} (-1,4) \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (2,3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (-1,4)$$

We now want the *action* of $\mathbb{R}^2 \otimes \mathbb{R}^{2*}$ on $\mathbb{R}^2$ that will counterfeit the action of $M(2,\mathbb{R})$ on $\mathbb{R}^2$. In general, there is an action of $V \otimes V^*$ on V given by

$$(v \otimes \lambda)(w) = v\lambda(w)$$

Applying this here we see, for $w = (1,2)^\top$

$$\begin{aligned} &\Big(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (2,3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (-1,4)\Big) \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Big((2,3) \begin{pmatrix} 1 \\ 2 \end{pmatrix}\Big) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Big((-1,4) \begin{pmatrix} 1 \\ 2 \end{pmatrix}\Big) \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} 8 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} 7 = \begin{pmatrix} 8 \\ 7 \end{pmatrix} \end{aligned}$$

Comparing this with the ordinary matrix multiplication

$$\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

we can see why it all works. Thus we have shown by example that

$$T \longleftrightarrow [e, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (2,3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (-1,4)]$$

All this is perfectly general so that we have an isomorphism

$$\mathrm{Hom}(V,V) \longleftrightarrow P \times_\rho (\mathbb{R}^n \otimes \mathbb{R}^{n*})$$

Now we verify that change of basis does not disturb the isomorphism. Thus let $\rho_1 : G \to \mathrm{Aut}(\mathbb{R}^n)$ and $\rho_2 : G \to \mathrm{Aut}(\mathbb{R}^{n*})$ be the isomorphisms we previously studied, namely

$$\rho_1(g)(\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}) = g \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad \text{and} \qquad \rho_2(g)\big((\lambda_1, \dots, \lambda_n)\big) = (\lambda_1, \dots, \lambda_n) g^{-1}$$

We then set $\rho = \rho_1 \otimes \rho_2$ on $\mathbb{R}^n \otimes \mathbb{R}^{n*}$ and we see that

$$\begin{aligned} \rho(g)\Big(\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \otimes (\lambda_1, \dots, \lambda_n)\Big) &= \rho_1(g)\Big(\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}\Big) \otimes \rho_2(g)\big((\lambda_1, \dots, \lambda_n)\big) \\ &= g \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \otimes (\lambda_1, \dots, \lambda_n) g^{-1} \end{aligned}$$

We can perhaps see what is happening a little more clearly if we apply this to our example. We have computed things in the e basis. Let us now shift over to the f basis and apply our rules. To compress the notation set

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes (2,3) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes (-1,4)$$

and we have

$$\begin{aligned} T &\longleftrightarrow [e, A] \\ &= [egg^{-1}, A] \\ &= [eg, \rho(g^{-1})(A)] \\ &= [f, g^{-1}Ag] \end{aligned}$$

This is the well known rule for the change in the matrix of a linear transformation when the basis is changed. So we see our system gives the correct results.

We will now show how to counterfeit the previous system by using matrices. We will use some extensions of matrix calculus in a formal manner without going into extensive justification, which could be provided for the obsessive.

Let T be a linear operator and e be a basis. We form the matrix of T by

$$Te_j = \sum_i e_i t^i_j$$

and write

$$T \overset{e}{\longleftrightarrow} (t^i_j)$$

This can also be written as follows:

$$\begin{aligned} T(e_1, \ldots, e_n) &\overset{def}{=} (Te_1, \ldots, Te_n) \\ &= (\sum_i e_i t^i_1, \ldots, \sum_i e_i t^i_n) \\ &= (e_1, \ldots, e_n)(t^i_j) \end{aligned}$$

If S is a second linear transformation with matrix (s^i_j) so that

$$S(e_1, \ldots, e_n) = (e_1, \ldots, e_n)(s^i_j)$$

Then we can find the matrix of ST as follows

$$\begin{aligned} ST(e_1, \ldots, e_n) &= S\Big(T(e_1, \ldots, e_n)\Big) \\ &= S\Big((e_1, \ldots, e_n)(t^j_k)\Big) \\ &= \Big(S(e_1, \ldots, e_n)\Big)(t^j_k) \\ &= \Big((e_1, \ldots, e_n)(s^i_j)\Big)(t^j_k) \\ &= (e_1, \ldots, e_n)\Big((s^i_j)(t^j_k)\Big) \end{aligned}$$

showing that

$$ST \overset{e}{\longleftrightarrow} (s^i_j)(t^j_k)$$

If v is a vector and $v = \sum e_j v^j$ then

$$v = (e_1, \ldots, e_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

so

$$Tv = T(e_1, \ldots, e_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = (e_1, \ldots, e_n)(t^i_j) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

so

$$Tv \overset{e}{\longleftrightarrow} (t^i_j) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

Note that if we had set up the system more flexibly, which is possible, then the second calculation would be a consequence of the first.

Now we change the basis; the new basis will be $\{f_1, \ldots, f_n\}$ and the change will be

$$f_j = \sum_i e_i g^i_j$$

This change may be written as

$$(f_1, \ldots, f_n) = (e_1, \ldots, e_n)(g^i_j) \qquad f = eg$$

Note fortuitous resemblance to the linear transformation formula; do not confuse them. Also note that

$$(e_1, \ldots, e_n) = (f_1, \ldots, f_n)(g^i_j)^{-1} \qquad e = fg^{-1}$$

Now we have

$$\begin{aligned} T(e_1, \ldots, e_n) &= (e_1, \ldots, e_n)(t^i_j) & Te = et \\ T(f_1, \ldots, f_n) &= (f_1, \ldots, f_n)(\tilde{t}^k_l) & Tf = f\tilde{t} \end{aligned}$$

and also

$$\begin{aligned} T(f_1, \ldots, f_n) &= (f_1, \ldots, f_n)(\tilde{t}^l_k) & Tf = f\tilde{t} \\ T(e_1, \ldots, e_n)(g^i_j) &= (e_1, \ldots, e_n)(g^k_l)(\tilde{t}^l_j) & Teg = eg\tilde{t} \\ (e_1, \ldots, e_n)(t^k_i)(g^i_j) &= (e_1, \ldots, e_n)(g^k_l)(\tilde{t}^l_j) & etg = eg\tilde{t} \end{aligned}$$

and thus

$$\begin{aligned} (g^k_l)(\tilde{t}^l_j) &= (t^k_i)(g^i_j) & g\tilde{t} = tg \\ (\tilde{t}^l_j) &= (g^k_l)^{-1}(t^k_i)(g^i_j) & \tilde{t} = g^{-1}tg \end{aligned}$$

Thus we have derived the matrix change formula with our matrix technique.

6. BILINEAR FORMS

The material in this section is a little trickier than in the previous section for mostly psychological reasons.

Def A bilinear form is a map $b : V \times V \to k$ satisfying the following

1a. $b[v_1 + v_2, w] = b[v_1, w] + b[v_2, w]$

1b. $b[w, v_1 + v_2] = b[w, v_1] + b[w, v_2]$

2a. $b[\alpha\, v, w] = \alpha\, b[v, w]$

2b. $b[v, \alpha\, w] = \alpha\, b[v, w]$

where $v, v_1, v_2, w \in V$ and $\alpha \in k$.

There is a similar definition where b is a function with r slots, which is called a multilinear form. The extension is obvious. All of the development below (except for the matrix material at the end) can be extended to the multilinear case in an obvious way.

The tensor product allows us to replace the bilinear form by a linear form B on the vector space $V \otimes V$. B is defined by $B(v \otimes w) = b[v, w]$. We will not be emphasizing this. One can prove this by showing that $b[u, v] = 0$ when either v or w is in the ideal used in the definition of $V \otimes V$.

There are special kinds of bilinear forms; for example

Def A bilinear form is symmetric if and only if $b[v, w] = b[w, v]$

Def A bilinear form is positive definite if and only if $b[v, v] > 0$ for all $v \neq 0$.

An inner product on V is just a symmetric positive definite bilinear form. The standard inner product on $\mathbb{R}^3$ is an example. This is one of the most useful uses of bilinear forms but there are many others. They are especially useful in functional analysis and number theory.

Given a basis, $\{e_1, \ldots, e_n\}$ each multilinear form produces a system of coefficients as follows

$$\begin{array}{lrcl} \text{bilinear form} & b_{ij} & = & b[e_i, e_j] \\ \text{trilinear form} & b_{ijk} & = & b[e_i, e_j, e_k] \end{array}$$

In the case of a bilinear form it is tempting to put the b_{ij} into a matrix (b_{ij}). This has good and bad features which we will discuss later, but obviously it will not easily generalize to multilinear forms, so it is suspect right from the outset.

These systems of coefficients can be used to compute the value of $b[v, w]$ when v and w are given in a specific basis: $v = \sum e_i v^i$ and $w = \sum e_j w^j$. Then

$$b[v, w] = b[\sum_i e_i v^i, \sum_j e_j w^j] = \sum_{ij} b[e_i, e_j] v^i w^j = \sum_{ij} b_{ij} v^i w^j$$

Now we wish to put bilinear forms into the Associated vector space milieu. To do this, we will have to represent bilinear forms in a tensor product context.

This requires a choice of W and ρ. The natural choice of W is $W = \mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$. We will need also an action of W on $\mathbb{R}^n \times \mathbb{R}^n$ which we will take to be

$$\begin{aligned}
(\lambda_1^1, \dots \lambda_n^1) \otimes (\lambda_1^2, \dots \lambda_n^2)[\begin{pmatrix} v_1^1 \\ \vdots \\ v_1^n \end{pmatrix}, \begin{pmatrix} v_2^1 \\ \vdots \\ v_2^n \end{pmatrix}] \\
&= (\lambda_1^1, \dots \lambda_n^1) \begin{pmatrix} v_1^1 \\ \vdots \\ v_1^n \end{pmatrix} (\lambda_1^2, \dots \lambda_n^2) \begin{pmatrix} v_2^1 \\ \vdots \\ v_2^n \end{pmatrix} \\
&= (\sum_i \lambda_i^1 v_1^i)(\sum_i \lambda_i^2 v_2^i)
\end{aligned}$$

With this equipment we will replicate the technique we used for linear transformations. Given a basis e and forming $b_{ij} = B[e_i, e_j]$ we calculate

$$\begin{aligned}
\begin{pmatrix} b_1^1 & b_2^1 & \dots & b_n^1 \\ b_1^2 & b_2^2 & \dots & b_n^2 \\ \dots & \dots & \dots & \dots \\ b_1^i & b_2^i & \dots & b_n^i \\ \dots & \dots & \dots & \dots \\ b_1^n & b_2^n & \dots & b_n^n \end{pmatrix} &= \sum_i \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_1^i & b_2^i & \dots & b_n^i \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix} \\
&= \sum_i \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} (b_1^i, b_2^i \dots, b_n^i)
\end{aligned}$$

Thus a reasonable isomorphism to $\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$ would be

$$(b_{ij}) \longleftrightarrow b_e = \sum_i (0, \dots, \overset{i\text{–th slot}}{0, 1, 0}, \dots, 0) \otimes (b_1^i, \dots, b_n^i)$$

Now let's check it works correctly on $\mathbb{R}^n \times \mathbb{R}^n$.

$$\begin{aligned}
b_e[\begin{pmatrix} v_1^1 \\ \vdots \\ v_1^r \end{pmatrix}, \begin{pmatrix} v_2^1 \\ \vdots \\ v_2^r \end{pmatrix}] &= \sum_i (0, \dots, \overset{i\text{–th slot}}{0, 1, 0}, \dots, 0) \begin{pmatrix} v_1^1 \\ \vdots \\ v_1^r \end{pmatrix} (b_1^i, \dots, b_n^i) \begin{pmatrix} v_2^1 \\ \vdots \\ v_2^r \end{pmatrix} \\
&= \sum_{ij} v_1^i b_{ij} v_2^j = \sum_{ij} b_{ij} v_1^i v_2^j
\end{aligned}$$

which is correct. Thus we can use $b_e \in \mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$ to represent the activity of the bilinear form b, and it will be no surprise that we now set up the isomorphism

(with ρ to be defined soon)

$$\begin{array}{rcl} \text{Bilinear Forms} & \longleftrightarrow & P \times_\rho (\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}) \\ b & \longleftrightarrow & [e, b_e] \end{array}$$

which shows that the bilinear forms are an associated vector space of P, provided that the change of basis works correctly. For this we need the representation of G on $\mathbb{R}^{n*} \otimes \mathbb{R}^{n*}$. We already have a representation $\rho_2 : G \to \mathrm{Hom}(\mathbb{R}^{n*}, \mathbb{R}^{n*})$ given by

$$\rho_2(g)(\lambda_1, \ldots, \lambda_n) = (\lambda_1, \ldots, \lambda_n) g^{-1}$$

so the ρ we want will of course by $\rho = \rho_2 \otimes \rho_2$.

Of course to check if the answer is right we must know the answer. If $f_j = \sum e_i g^i_j$ is the change of basis from e to new basis f, then we have

$$\begin{array}{rcl} \tilde{b}_{kl} & = & b[f_k, f_l] \\ & = & b[\sum_i e_i g^i_k, \sum_j e_j g^j_l] = \sum_{ij} g^i_k g^j_l b[e_i, e_j] \\ & = & \sum_{ij} g^i_k g^j_l b_{ij} \end{array}$$

so we know how it should come out. Now the next part is a little tricky. First we have according to the isomorphism we have set up

$$b \overset{f}{\longleftrightarrow} b_f = \sum_i (0, \ldots, 0, 1, 0 \ldots 0) \otimes (\tilde{b}_{i1}, \ldots, \tilde{b}_{in})$$

and then

$$b \longleftrightarrow [f, b_f]$$

We must show that this is consistent with the rules. So we have

$$\begin{array}{rcl} b \longleftrightarrow [e, b_e] & = & [egg^{-1}, b_e] = [eg, \rho(g^{-1}) b_e] \\ & = & [f, \rho(g^{-1}) b_e] \end{array}$$

Hence it comes down to showing that $\rho(g^{-1}) b_e = b_f$ Here we go

$$\begin{array}{rcl} \rho(g^{-1}) b_e & = & \rho(g^{-1}) \Big(\sum_i (0, \ldots, 0, 1, 0, \ldots, 0) \otimes (b_{i1}, \ldots, b_{in}) \Big) \\ & = & \big(\rho_2(g^{-1}) \otimes \rho_2(g^{-1})\big) \Big(\sum_i (0, \ldots, 0, 1, 0, \ldots, 0) \otimes (b_{i1}, \ldots, b_{in}) \Big) \\ & = & \sum_i \Big(\rho_2(g^{-1})(0, \ldots, 0, 1, 0, \ldots, 0) \otimes \rho_2(g^{-1})(b_{i1}, \ldots, b_{in}) \Big) \\ & = & \sum_i \Big((0, \ldots, 0, 1, 0, \ldots, 0) g \otimes (b_{i1}, \ldots, b_{in}) g \Big) \\ & = & \sum_i \Big((g^i_1, \ldots, g^i_n) \otimes \big(\sum_\ell b_{i\ell} g^\ell_1, \ldots, \sum_\ell b_{i\ell} g^\ell_n \big) \Big) \end{array}$$

$$
\begin{aligned}
&= \sum_i \Big(\big(\sum_j g^i_j (0,\ldots,0,1,0,\ldots,0) \big) \otimes \big(\sum_\ell b_{i\ell} g^\ell_1, \ldots, \sum_\ell b_{i\ell} g^\ell_n \big) \Big) \\
&= \sum_i \Big((0,\ldots,0,1,0,\ldots,0) \otimes \big(\sum_{j\ell} g^i_j b_{i\ell} g^\ell_1, \ldots, \sum_{j\ell} g^i_j b_{i\ell} g^\ell_n \big) \Big) \\
&= \sum_i \Big((0,\ldots,0,1,0,\ldots,0) \otimes (\tilde{b}_{j1}, \ldots, \sum_\ell \tilde{b}_{jn}) \Big) \\
&= b_f
\end{aligned}
$$

as required.

Next we will show how to counterfeit the activity of bilinear forms using matrices. One can tell this is a questionable activity because it can't be done for trilinear forms or higher. Hence one should be suspicious. However, there is worse in store; this will turn out to require something downright evil. However, as is often the case with evil, it confers some benefits too.

We have, in the e basis, $b[v,w] = \sum_{ij} b_{ij} v^i w^j$. The j-sum looks just like matrix multiplication. The i-sum does not, but perhaps we can cheat:

$$
b[v,w] = \sum_{ij} b_{ij} v^i w^j = (v^1, \ldots, v^n) \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \ldots & \ldots & \ldots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}
$$

Yes, that works, but we have written a contravariant vector v as a *row*. The formalism thinks it is in the dual space. When you identify the space V and it's dual space V^* you have introduced an inner product into a system which did not previously have one. This is the evil act. It will not be possible to perform complex manipulations after having done this. Things will start to go wrong soon.

However, there are also benefits. The most important is the change of basis computation. Suppose we have $f = eg$. Then, as we saw, if

$$
v = \sum_i e_i v^i \qquad v \overset{e}{\longleftrightarrow} \vec{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}
$$

then

$$
v = \sum_j f_j \tilde{v}^j \qquad v \overset{f}{\longleftrightarrow} \vec{\tilde{v}} = \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix} = g^{-1} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}
$$

Thus

$$
\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = g \begin{pmatrix} \tilde{v}^1 \\ \vdots \\ \tilde{v}^n \end{pmatrix} \qquad \text{and} \qquad (v^1, \ldots, v^n) = (\tilde{v}^1, \ldots \tilde{v}^n) g^\top
$$

and similarly for w. Then

$$b[v,w] = (v^1, \dots v^n) \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = (\tilde{v}^1, \dots \tilde{v}^n) g^\top \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} g \begin{pmatrix} \tilde{w}_1 \\ \vdots \\ \tilde{w}_n \end{pmatrix}$$

so that if

$$b \overset{e}{\longleftrightarrow} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

then

$$b \overset{f}{\longleftrightarrow} \begin{pmatrix} \tilde{b}_{11} & \dots & \tilde{b}_{1n} \\ \dots & \dots & \dots \\ \tilde{b}_{n1} & \dots & \tilde{b}_{nn} \end{pmatrix} = g^\top \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} g$$

This simple transformation law is of great utility in, among others, number theory, specifically in the theory of quadratic forms.

7. EPILOG

This completes our goal for this module. I will mention a few other things here.

First, algebraically a tensor is an inhabitant of a vector space of the form

$$V \otimes \dots \otimes V \otimes V^* \otimes \dots \otimes V^*$$

(In fact the vector spaces in the tensor product might be different but this is the commonest form). Elements of the space can be written in a basis and dual basis, for example if we have

$$V^* \otimes V \otimes V^* \otimes V^*$$

Then a typical element would look like

$$R_i{}^j{}_{k\ell}\, e^i \otimes e_j \otimes e^k \otimes e^l$$

using the Einstein summation convention, that is omitting $\sum_{ijkl}$. This would be called a once contravariant thrice covariant tensor.

Second, in differential geometry there are vector spaces, for example, the tangent space, correlated to each point of the manifold. Thus for each point on the r-dimensional Manifold M we have a series of coordinates $(u^1, \dots, u^r)$ valid over an open subset of M called a coordinate patch. All the tangent spaces of M then fit together into the *tangent bundle.* This corresponds to the V in our development, and there is a cotangent bundle corresponding to V^*. Then the coefficients and basis vectors all become functions of the u^i and these then constitute the tensor algebra on the manifold. (The standard meaning of "tensor" in differential geometry is objects formed from $V = T(M)$, the tangent bundle.) The mechanisms for base change that we have developed deal with the transitions between coordinate patches. This is all one has on a raw differentiable manifold. If one wishes to do calculus on the manifold one needs additional structures. A vector field on the manifold lets you define the Lie Derivative, but for a more complete form of Calculus one needs an Affine Connection which essentially controls the calculus of the basis vectors. The affine connection may be defined if the manifold has an inner product structure on the tangent bundle (a Riemannian Metric).

But that's a different module.