Vector and Tensor Analysis

Ian Poon

September 2024

Selected theorems from `Gibbs Vector Analysis(Cornell University)` and `Amuthan A Ramabathiran`'s supplementary notes on tensor analysis for `AE 639: Continuum Mechanics at IIT Bombay`. Finally for advanced tensor calculus material was sources from `David Clarke` grad school notes on tensor calculus(2011, Saint Mary's University)

Contents

1 direct and skew products

Example 1

Let x be the angle $\mathbf{a}$ makes with $\mathbf{i}$ and y be the angle $\mathbf{b}$ makes with $\mathbf{i}$ then

$$\mathbf{a} = \cos x\mathbf{i} + \sin x\mathbf{j}$$

$$\mathbf{b} = \cos y\mathbf{i} + \sin y\mathbf{j}$$

$$\mathbf{a} \cdot \mathbf{b} = \cos(\mathbf{a}, \mathbf{b}) = \cos x \cos y + \sin x \sin y = \cos(y - x)$$

$$\mathbf{a} \times \mathbf{b} = \mathbf{k} \sin(\mathbf{a}, \mathbf{b}) = \mathbf{k}(\sin y \cos x + \sin x \cos y = \mathbf{k} \sin(y - x)$$

Proposition 2

If $\mathbf{A} = a\mathbf{a} + b\mathbf{b} + c\mathbf{c}$ then

$$\mathbf{A} \cdot \mathbf{A} = a^2\mathbf{a} \cdot \mathbf{a} + b^2\mathbf{b} \cdot \mathbf{b} + c^2\mathbf{c} \cdot \mathbf{c} + 2ab\mathbf{a} \cdot \mathbf{b} + 2bc\mathbf{b} \cdot \mathbf{c} + 2ca\mathbf{c} \cdot \mathbf{a}$$

and if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ then the last 3 terms cancel to zero

Definition 3 (triple scalar product)
We denote the **triple scalar product**

$$[A\ B\ C] = A \cdot B \times C = B \cdot C \times A = C \cdot A \times B$$
$$= A \times B \cdot C = B \times C \cdot A = C \times A \cdot B$$

where

$$[A\ B\ C] = -[A\ C\ B]$$

this property follows since we may write this result may be written in the form of the determinant

$$[A\ B\ C] = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix}$$

where

$$\mathbf{A} = A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k}$$
$$\mathbf{B} = B_1\mathbf{i} + B_2\mathbf{j} + B_3\mathbf{k}$$
$$\mathbf{A} = C_1\mathbf{i} + C_2\mathbf{j} + C_3\mathbf{k}$$

Proposition 4
A triple scalar product can be expressed in general by

$$[\mathbf{A}\ \mathbf{B}\ \mathbf{C}] = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} [\mathbf{a}\ \mathbf{b}\ \mathbf{c}]$$

where

$$\mathbf{A} = a_1\mathbf{a} + a_2\mathbf{b} + a_3\mathbf{c}$$
$$\mathbf{B} = b_1\mathbf{a} + b_2\mathbf{b} + b_3\mathbf{c}$$
$$\mathbf{C} = c_1\mathbf{a} + c_2\mathbf{b} + c_3\mathbf{c}$$

and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are not necessarily unit vectors

Proof. Consider that

$$[\mathbf{A}\ \mathbf{B}\ \mathbf{C}] = \det\left(\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} \mathbf{a}\ \mathbf{b}\ \mathbf{c} \end{bmatrix}^T\right)$$

where we denote the smaller brackets to be the scalar product to distinguish with

$$\begin{bmatrix} \mathbf{a}\ \mathbf{b}\ \mathbf{c} \end{bmatrix}$$

which is the matrix that contains in each column the column vector **a,b,c**. Then by the prodct rule of determinants recall `artin algebra` we have

$$[\mathbf{A}\ \mathbf{B}\ \mathbf{C}] = \det\left(\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\right) \det\left(\begin{bmatrix} \mathbf{a}\ \mathbf{b}\ \mathbf{c} \end{bmatrix}^T\right)$$

which equals the proposition as desired.

Definition 5 (Vector Product)

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = AB\sin(\mathbf{A}, \mathbf{B})\mathbf{c}$$

where **c** is the unit vector of **C**

It is clearly to see

$$A \times B = -B \times A$$

since

$$\sin(y - x) = -\sin(x - y)$$

Fact 6

The vector product may also be written in the form of a determinant as

$$B \times C = \begin{vmatrix} i & j & k \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix}$$

You can immediately swapping $C \times B = -B \times C$ fom the properties of determinant

We will attempt to prove the triple product geometrically now. Note that a more efficient method using the levi-cevita tensor will be presented as we develop our understanding of tensor algebra in later sections

Lemma 7

$$\mathbf{A} \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \cdot \mathbf{B}\mathbf{A} - \mathbf{A} \cdot \mathbf{A}\mathbf{B}$$

Hint: Exploit Geometry and use a figure

Proof. Consider

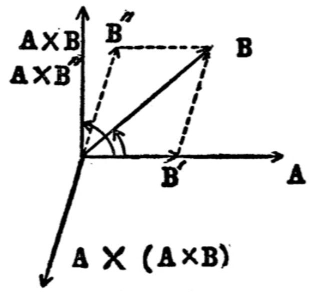

Figure 1: Triple Cross Product 1

Notice that

$$\mathbf{A} \times (\mathbf{A} \times \mathbf{B}) = -c\mathbf{B}'' = -cB\sin(\mathbf{A}, \mathbf{B})$$

at the same time it also

$$\mathbf{A} \times (\mathbf{A} \times \mathbf{B}) = -(A)(AB\sin(\mathbf{A}, \mathbf{B}))\sin\frac{\pi}{2}\mathbf{b}'' = -A^2B\sin(\mathbf{A}, \mathbf{B})\mathbf{b}''$$

Hence we have

$$c = A^2 = \mathbf{A} \cdot \mathbf{A}$$

and

$$\mathbf{B}'' = -\frac{\mathbf{A} \times (\mathbf{A} \times \mathbf{B})}{\mathbf{A} \cdot \mathbf{A}}$$

Lasty we use

$$\mathbf{B} = \mathbf{B}' + \mathbf{B}'' = \frac{\mathbf{A} \cdot \mathbf{B}}{\mathbf{A} \cdot \mathbf{A}} - \frac{\mathbf{A} \times (\mathbf{A} \times \mathbf{B})}{\mathbf{A} \cdot \mathbf{A}}$$

and the lemma follows

Theorem 8 (Triple Cross Product)

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{A} \cdot \mathbf{C}\mathbf{B} - \mathbf{A} \cdot \mathbf{B}\mathbf{C}$$

Hint: Start by expressing $\mathbf{A}$ in terms of $\mathbf{B}, \mathbf{C}$ and $\mathbf{B} \times \mathbf{C}$ first then try to reverse engineer

Proof. First express the 3D vector $\mathbf{A}$ in terms of 3 non-coplanar(independent) vectors using $\mathbf{B}, \mathbf{C}$ and $\mathbf{B} \times \mathbf{C}$ by

$$\mathbf{A} = b\mathbf{B} + c\mathbf{C} + a(\mathbf{B} \times \mathbf{C}) \tag{1}$$

where a, b, c are constants. Then computing directly

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = b\mathbf{B} \times (\mathbf{B} \times \mathbf{C}) + c\mathbf{C} \times (\mathbf{B} \times \mathbf{C}) + a(\mathbf{B} \times \mathbf{C}) \times (\mathbf{B} \times \mathbf{C})$$

we know thet last term is equal zero and by the previous lemma we althouther have

$$\begin{aligned} (\mathbf{B} \times \mathbf{C}) \times (\mathbf{B} \times \mathbf{C}) &= 0 \\ \mathbf{B} \times (\mathbf{B}) &= \mathbf{B} \cdot \mathbf{C}\mathbf{B} - \mathbf{B} \cdot \mathbf{B}\mathbf{C} \\ \mathbf{C} \times (\mathbf{B} \times \mathbf{C}) = -\mathbf{C} \times (\mathbf{C} \times \mathbf{B}) &= -\mathbf{C} \cdot \mathbf{B} + \mathbf{C} \cdot \mathbf{C}\mathbf{B} \end{aligned}$$

therefore we currently have

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = [(b\mathbf{B} \cdot \mathbf{C} + c\mathbf{C} \cdot \mathbf{C})\mathbf{B} - (b\mathbf{B} \cdot \mathbf{B} \ + c\mathbf{C} \cdot \mathbf{B})\mathbf{C}] \tag{2}$$

But from (1) we have

$$\mathbf{A} \cdot \mathbf{B} = b\mathbf{B} \cdot \mathbf{B} + c\mathbf{C} \cdot \mathbf{B}$$

and

$$\mathbf{A} \cdot \mathbf{C} = b\mathbf{B} \cdot \mathbf{C} + c\mathbf{C} \cdot \mathbf{C}$$

then upon substituting these into (2) the theorem follows

Proposition 9

Triple products are not commutative

$$A \times (B \times C) \neq (A \times B) \times C$$

Proof. Consider that

$$A \times (B \times C) = x\mathbf{B} + y\mathbf{C}$$

since it is definitely somewhere in the BC plane where it is perpendicular to both A and $B \times C$. Likewise for

$$(A \times B) \times C) = m\mathbf{A} + n\mathbf{B}$$

Fact 10

As a matter of fact the other properties of cross product such as linearity

$$A \times (C + D) + B \times (C + B) = (A + B) \times (C + D)$$

also follow by linearity in rows of a determinant even stuff like

$$i \times i = 0$$

too, repeated rows in determinant is zero

Fact 11

Stuff like $i \times k = -j$ comes from the

Proposition 12 (Dot product of cross products)

$$(A \times B) \cdot (C \times D) = (A \cdot C)(B \cdot D) - (A \cdot D)(B \cdot C)$$

which implies they may be written in determinant form as

$$(A \times B) \cdot (C \times D) = \begin{vmatrix} A \cdot C & A \cdot D \\ B \cdot C & B \cdot D \end{vmatrix}$$

Proof. See that

$$A \times B \cdot (C \times D) = A \cdot B \times (C \times D)$$

by properties scalar product. Then recalling the identity for trip cross products we have

$$B \times (C \times D) = (B \cdot D)C - (B \cdot C)D$$

then putting everything together we get

$$(A \times B) \cdot (C \times D) = (A \cdot C)(B \cdot D) - (A \cdot D)(B \cdot C)$$

as desired

Proposition 13 (Chain Rule cross and dot product)

$$(u \times v)' = u' \times v + u \times v'$$

and

$$(u \cdot v)' = u' \cdot v + u \cdot v'$$

Proof. Simply consider

$$B(t) = b_1(t)\mathbf{i} + b_2(t)\mathbf{j} + b_3(t)\mathbf{k}$$

$$C(t) = c_1(t)\mathbf{i} + c_2(t)\mathbf{j} + c_3(t)\mathbf{k}$$

and recall that vector valued differentiation is defined by for example

$$X'(t) = x_1'(t)\mathbf{i} + x_2'(t)\mathbf{j} + x_3'(t)\mathbf{k}$$

then

$$B(t) \times C(t) = \begin{vmatrix} i & j & k \\ b_1(t) & b_2(t) & b_3(t) \\ c_1(t) & c_2(t) & c_3(t) \end{vmatrix} = \mathbf{i}\begin{vmatrix} b_2(t) & b_3(t) \\ c_2(t) & c_3(t) \end{vmatrix} - \mathbf{j}\begin{vmatrix} b_1(t) & b_3(t) \\ c_1(t) & c_3(t) \end{vmatrix} + \mathbf{k}\begin{vmatrix} b_1(t) & b_2(t) \\ c_1(t) & c_2(t) \end{vmatrix}$$

Then apply chain rule in the elements of each component i, j, k where you will notice b and c are separately differentiated. The case of the dot product is very obvious from here

$$B(t) \cdot C(t) = a_1(t)b_1(t) + a_2(t)b_2(t) + a_3(t)b_3(t)$$

Proposition 14

Let

$$X(t) = x_1(t)\mathbf{i} + x_2(t)\mathbf{j} + x_3(t)\mathbf{k}$$

$$X'(t) = x_1'(t)\mathbf{i} + x_2'(t)\mathbf{j} + x_3'(t)\mathbf{k}$$

then

$$(AX(t))' = AX'(t)$$

where $A : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear map

Proof. after transformation you will end up with

$$AX(t) = (a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t))\,\mathbf{i} + (a_{21}x_1(t) + a_{22}x_2(t) + a_{23}x_3(t))\,\mathbf{j} + (a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t))\,\mathbf{k}$$

It is obvious to see that $(AX)' = AX'$

2 differential calculus

Definition 15

Let

$$\mathbf{V}(x,y,z) = V_1(x,y,z)\mathbf{i} + V_2(x,y,z)\mathbf{j} + V_3(x,y,z)\mathbf{k}$$

and

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

then

$$\mathbf{a}\cdot\nabla = a_1\frac{\partial}{\partial x} + a_2\frac{\partial}{\partial y} + a_3\frac{\partial}{\partial z}$$

and

$$\begin{aligned}(\mathbf{a}\cdot\nabla)\mathbf{V} &= (\mathbf{a}\cdot\nabla)V_1\mathbf{i} + (\mathbf{a}\cdot\nabla)V_2\mathbf{j} + (\mathbf{a}\cdot\nabla)V_3\mathbf{k} \\ &= \left(a_1\frac{\partial V_1}{\partial x} + a_2\frac{\partial V_1}{\partial y} + a_3\frac{\partial V_1}{\partial z}\right)\mathbf{i} \\ &+ \left(a_1\frac{\partial V_2}{\partial x} + a_2\frac{\partial V_2}{\partial y} + a_3\frac{\partial V_2}{\partial z}\right)\mathbf{j} \\ &+ \left(a_1\frac{\partial V_3}{\partial x} + a_2\frac{\partial V_3}{\partial y} + a_3\frac{\partial V_3}{\partial z}\right)\mathbf{k}\end{aligned}$$

Proposition 16

We may write the divergence as

$$\nabla\cdot\mathbf{V} = \mathbf{i}\cdot\frac{\partial\mathbf{V}}{\partial x} + \mathbf{j}\cdot\frac{\partial\mathbf{V}}{\partial y} + \mathbf{k}\cdot\frac{\partial\mathbf{V}}{\partial z}$$

and the curl as

$$\nabla\times\mathbf{V} = \mathbf{i}\times\frac{\partial\mathbf{V}}{\partial x} + \mathbf{j}\times\frac{\partial\mathbf{V}}{\partial y} + \mathbf{k}\times\frac{\partial\mathbf{V}}{\partial z}$$

Proof. We know that

$$\begin{aligned}\frac{\partial\mathbf{V}}{\partial x} &= \frac{\partial V_1}{\partial x}\mathbf{i} + \frac{\partial V_2}{\partial x}\mathbf{j} + \frac{\partial V_3}{\partial x}\mathbf{k} \\ \frac{\partial\mathbf{V}}{\partial y} &= \frac{\partial V_1}{\partial y}\mathbf{i} + \frac{\partial V_2}{\partial y}\mathbf{j} + \frac{\partial V_3}{\partial y}\mathbf{k} \\ \frac{\partial\mathbf{V}}{\partial x} &= \frac{\partial V_1}{\partial z}\mathbf{i} + \frac{\partial V_2}{\partial z}\mathbf{j} + \frac{\partial V_3}{\partial z}\mathbf{k}\end{aligned}$$

Then taking dot products like how our divergence was defined in the proposition we have

$$\nabla\cdot\mathbf{V} = \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z}$$

as desired. Similary for the curl we see that upon taking corss products we have

$$\begin{aligned}
\mathbf{i} \times \frac{\partial \mathbf{V}}{\partial x} &= \mathbf{k}\frac{\partial V_2}{\partial y} - \mathbf{j}\frac{\partial V_3}{\partial x} \\
\mathbf{j} \times \frac{\partial \mathbf{V}}{\partial y} &= \mathbf{i}\frac{\partial V_3}{\partial y} - \mathbf{k}\frac{\partial V_1}{\partial y} \\
\mathbf{k} \times \frac{\partial \mathbf{V}}{\partial z} &= \mathbf{j}\frac{\partial V_1}{\partial z} - \mathbf{i}\frac{\partial V_2}{\partial z}
\end{aligned}$$

and then grouping the same unit vectors terms together we have

$$\nabla \times \mathbf{V} = \mathbf{i}\left(\frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y}\right)$$

as desired

Proposition 17

$$\nabla \times (u\mathbf{v}) = \nabla u \times \mathbf{v} + u\nabla \times \mathbf{v}$$

Proof. Consider

$$\begin{aligned}
\nabla \times (u\mathbf{v}) &= \sum \mathbf{e}_i \times \frac{\partial (uv)}{\partial x_i} = \sum \mathbf{e}_i \times \left(\frac{\partial (u)}{\partial x_i}\mathbf{v} + u\frac{\partial \mathbf{v}}{\partial x_i}\right) \\
&= \sum \mathbf{e}_i \times \frac{\partial (u)}{\partial x_i}\mathbf{v} + \sum \mathbf{e}_i \times u\frac{\partial \mathbf{v}}{\partial x_i} \\
&= \sum \frac{\partial (u)}{\partial x_i}\mathbf{e}_i \times \mathbf{v} + \sum u\mathbf{e}_i \times \frac{\partial \mathbf{v}}{\partial x_i} \\
&= \nabla u \times \mathbf{v} + u\nabla \times \mathbf{v}
\end{aligned}$$

□

Proposition 18

$$\nabla(\mathbf{u} \cdot \mathbf{v}) = \mathbf{v} \cdot \nabla \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{v} + \mathbf{v} \times (\nabla \times \mathbf{u}) + \mathbf{u} \times (\nabla \times \mathbf{v})$$

3 linear vector functions

A case of a particular function takes the form

$$\mathbf{r}' = \mathbf{i}(c_1\mathbf{i} \cdot \mathbf{r}) + \mathbf{j}(c_2\mathbf{j} \cdot \mathbf{r}) + \mathbf{k}(c_3\mathbf{k} \cdot \mathbf{r})$$

symboliclly we could also represent this as

$$\mathbf{r}' = (\mathbf{i}c_1\mathbf{i} + \mathbf{j}c_2\mathbf{j} + \mathbf{k}c_3\mathbf{k}) \cdot \mathbf{r}$$

now in similar fashion we could also do

$$\mathbf{r}' = (\mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3) \cdot \mathbf{r}$$

Definition 19

An expression **ab** formed by the juxtaposition of two vectors without the intervention of a dot or cross is called **dyad**

Let $\Phi = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3 + \dots$ notice that $\Phi \cdot \mathbf{r}$ and $\mathbf{r} \cdot \Phi$ are generally different. Consider

$$\mathbf{r} \cdot \Phi = (\mathbf{r} \cdot \mathbf{a}_1)\mathbf{b}_1 \dots$$

while

$$\Phi \cdot \mathbf{r} = \mathbf{a}_1(\mathbf{b}_1 \cdot \mathbf{r}) \dots$$

Definition 20

The **direct product** of the **dyad** **ab** into the dyad **cd** is written

$$(\mathbf{ab}) \cdot (\mathbf{cd})$$

so we have

$$(\mathbf{ab}) \cdot (\mathbf{cd}) = \mathbf{a}(\mathbf{b} \cdot \mathbf{c})\mathbf{d} = (\mathbf{b} \cdot \mathbf{c})\mathbf{ad}$$

Example 21

$$\begin{aligned}
\Phi &= \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3 + \dots \\
\Psi &= \mathbf{b}_1\mathbf{c}_1 + \mathbf{b}_2\mathbf{c}_2 + \mathbf{b}_3\mathbf{c}_3 + \dots \\
\Phi \cdot \Psi &= (\mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3 + \dots) \cdot (\mathbf{b}_1\mathbf{c}_1 + \mathbf{b}_2\mathbf{c}_2 + \mathbf{b}_3\mathbf{c}_3 + \dots) \\
&= \mathbf{a}_1\mathbf{b}_1 \cdot \mathbf{b}_1\mathbf{c}_1 + \mathbf{a}_1\mathbf{b}_1 \cdot \mathbf{b}_2\mathbf{c}_2 + \mathbf{a}_1\mathbf{b}_1 \cdot \mathbf{b}_3\mathbf{c}_3 + \dots \\
&+ \mathbf{a}_2\mathbf{b}_2 \cdot \mathbf{b}_1\mathbf{c}_1 + \mathbf{a}_2\mathbf{b}_2 \cdot \mathbf{b}_2\mathbf{c}_2 + \mathbf{a}_2\mathbf{b}_2 \cdot \mathbf{b}_3\mathbf{c}_3 + \dots \\
&+ \mathbf{a}_3\mathbf{b}_3 \cdot \mathbf{b}_1\mathbf{c}_1 + \mathbf{a}_3\mathbf{b}_3 \cdot \mathbf{b}_2\mathbf{c}_2 + \mathbf{a}_3\mathbf{b}_3 \cdot \mathbf{b}_3\mathbf{c}_3 + \dots
\end{aligned}$$

Definition 22

The **skew** products of a dyad **ab** into a vector **r** and of a vector **r** into a dyad **ab** are define respectively by the equations

$$\begin{aligned}
(\mathbf{ab}) \times \mathbf{r} &= \mathbf{a}(\mathbf{b} \times \mathbf{r}) \\
\mathbf{r} \times (\mathbf{ab}) &= (\mathbf{r} \times \mathbf{a})\mathbf{b}
\end{aligned}$$

4 inner product spaces

Note that we have manually shown that our differentiation rules such as product rule applies cross product dot product etc. A better way to see why they apply is think of them as tensor products. For example

Example 23
It is clear to see that
$$\partial_t(\mathrm{u}\times\mathrm{v}) = \partial_t(\sum \epsilon_{ijk} u_j v_k \mathrm{e}_i) = \partial_t u \times v + u \times \partial_t v$$
Similary we learn dot product

we begin our third and final attempt at tensor analysis. This one looks promising

4.1 inner product spaces and general bases

Definition 24
inner product spaces are vector spaces with an additional structure known as an inner product.

In what follows V denotes an inner product space of dimension n and (g_i) is a **general basis** of V

Definition 25
A general basis as the name suggest is "general", it is not necessarily orthonormal but it is linearly independent

Any $v \in V$ can be written in terms of the basis $(g_1, \ldots, g_n)$ of V as
$$v = \sum v_i g_i$$
where $(v_1, \ldots, v_n)$ are the components of v with respect to this basis.

To compute these compoents start with taking inner product of this equation with the basis vector g_i this yields
$$v \cdot g_i = \sum_j g_i \cdot g_j v_j$$
This equation can be written in the form of a matrix equation as follows
$$\begin{bmatrix} \mathrm{g}_1 \cdot \mathrm{g}_1 & \ldots & \mathrm{g}_1 \cdot \mathrm{g}_n \\ \vdots & \ddots & \vdots \\ \mathrm{g}_n \cdot \mathrm{g}_1 & \ldots & \mathrm{g}_n \cdot \mathrm{g}_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \mathrm{v} \cdot \mathrm{g}_1 \\ \vdots \\ \mathrm{v} \cdot \mathrm{g}_n \end{bmatrix}$$
where each $v \cdot g_i$ on the column on the RHS corresponds to that of our equation. The fact that (g_i) is a basis of V implies that the components $v_1, \ldots, v_n$ exists and are unique.

Proposition 26
This matrix (g_{ij}) is invertible

Proof. Suppose there exists a non-trivial $(v_1, \ldots, v_n)$ such that $(g_i \cdot g_1)v_1 + (g_i \cdot g_2)v_2 \ldots = 0$ for all i.(meaning there is a non-zero vector in nullspace). This implies
$$(g_1 v_1 + g_2 v_2 \ldots) \cdot g_i = v \cdot g_i, \quad \forall i$$
but because g_i are basis vector which are non-zero it must be that $v = 0$. However v cannot be zero for a non-trivial $(v_1, \ldots, v_n)$ because it is spanned by basis vectors (g_i). Hence we have a contradiction. Alternatively we know that

$g_i \cdot g_i > 0$ since the euclidean dot product is positive definite so the diagonal is non-zero and we have a full rank matrix $\square$

As preluded in the previous section it is possible express our matrix and its inverse as

$$\begin{bmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \dots & g_{nn} \end{bmatrix}^{-1} = \begin{bmatrix} g^{11} & \dots & g^{1n} \\ \vdots & \ddots & \vdots \\ g^{n1} & \dots & g^{nn} \end{bmatrix}.$$

A proper justifcation for this choice of notation will be given shortly when we study *reciprocal bases*. The fact that these two matrices are inerses of each other can be written succintly as follows

$$\sum_k g^{ik} g_{kj} = \delta_{ij} = \sum_k g_{ik} g^{kj}$$

Essentially this is so as $[(g_{ij})][(g_{ij})]^{-1} = I = [(g_{ij})]^{-1}[(g_{ij})]$ which shows the left and riht inverses.

Corollary 27

$$v = \sum_i v_i g_i \quad \Rightarrow \quad v_i = \sum_j g^{ij} g_j \cdot v$$

Proof. Consider

$$\sum_j g^{ij} g_j \cdot v = \sum_j g^{ij} g_j \cdot (\sum_k v_k g_k) = \sum_j \sum_k g^{ij} g_{jk} v_k = \sum_k \delta_{ik} v_k = v_i$$

as desired

4.2 reciprocal basis of a general basis

Definition 28

The **reciprocal basis** is defined as the basis $(g^1, \dots, g^n)$ such that

$$g^i \cdot g_j = \delta_{ij}$$

where $1 \leq i, j \leq n$.

For now assume that is exists. We will discuss its existence and more in later chapters.

Proposition 29

$$g^i = \sum_j g^{ij} g_j$$

Proof. Consider

$$g^i \cdot g_k = \delta_{ik} = \sum_j g^{ij} g_{jk} = \sum_j g^{ij} g_j \cdot g_k = (\sum_j g^{ij} g_j) \cdot g_k$$

on comparison the proposition immediatly follows

Corollary 30

$$g_i = \sum g_{ij} g^j$$

Proof. We begin by substituting the expression for g^j from the proposition into the desired formula for g_i in the corollary:

$$g_i = \sum_j g_{ij} g^j = \sum_j g_{ij} \left(\sum_k g^{jk} g_k \right).$$

We can now interchange the sums:

$$g_i = \sum_k \left(\sum_j g_{ij} g^{jk} \right) g_k.$$

Using the orthogonality property of the Gram matrix and its inverse, we simplify the expression:

$$\sum_j g_{ij} g^{jk} = \delta_i^k.$$

Thus, the equation reduces to:

$$g_i = \sum_k \delta_i^k g_k = g_i.$$

This shows that the expression in the corollary holds true, proving that:

$$g_i = \sum_j g_{ij} g^j.$$

□

Theorem 31

$$g^i \cdot g^j = g^{ij}$$

Proof. Consider

$$\begin{aligned} \mathrm{g}^i \cdot \mathrm{g}^j &= \left(\sum g^{ik} \mathrm{g}_k \right) \left(\sum g^{jl} \mathrm{g}_l \right) \\ &= \sum g^{ik} g^{jl} g_{kl} = \sum \delta_{il} g^{jl} \\ &= g^{ij} \end{aligned}$$

Theorem 32

$$\mathrm{v} = \sum \bar{v}_i \mathrm{f}_i = \sum \tilde{v}_i \mathrm{g}_i \quad \Rightarrow \quad \bar{v}_i = \sum \mathrm{f}^i \cdot \mathrm{g}_j \tilde{v}_j,$$

Proof. Taking the inner product of both sides of the equation with f^i gives:

$$\mathrm{f}^i \cdot \left(\sum_j \bar{v}_j \mathrm{f}_j \right) = \mathrm{f}^i \cdot \left(\sum_k \tilde{v}_k \mathrm{g}_k \right).$$

Using the properties of the dual basis, the left-hand side simplifies to:

$$\sum_j \bar{v}_j(\mathrm{f}^i \cdot \mathrm{f}_j) = \sum_j \bar{v}_j \delta^i_j = \bar{v}_i.$$

The right-hand side becomes:

$$\sum_k \tilde{v}_k(\mathrm{f}^i \cdot \mathrm{g}_k).$$

Equating both sides, we obtain the desired relation:

$$\bar{v}_i = \sum_k (\mathrm{f}^i \cdot \mathrm{g}_k)\tilde{v}_k.$$

□

5 linear maps

5.1 representing basic linear maps and change of basis with general bases

Definition 33

We now collect together all the constants T_{ij} as an $n \times m$ matrix $[T]$ whose $(i,j)^{\text{th}}$ entry is $[T]_{ij} = T_{ij}$. The matrix $[T]$ is called the **matrix represnetation** of $T : V \to W$ with respect to the bases (e_i) of V and (f_i) of W.

We first begin our analysis with a simplified context of **orthonormal** bases.

Proposition 34

Let V and W be finite dimensional inner product spaces. Let $T : V \to W$ be a linear map from V into W and let $(e_1, \ldots, e_m)$ and $(f_1, \ldots, f_n)$ be *orthonormal* bases of V and W respectively. Recalling, what the definition of a matrix map from basic `Linear Algebra`, recall what each column j corresponding to Te_j tells you about its transformation

$$\mathrm{Te}_i = \sum_{j=1}^{n} T_{ji}\mathrm{f}_j$$

for some constants $T_{ji} \in \mathbb{R}$ for every $1 \leq i \leq m$ an $1 \leq j \leq n$. Then

$$T_{ij} = f_i \cdot Te_j$$

Proof. To find the explicit expression for T_{ji}, take the inner product of both sides of the equation $Te_i = \sum_{j=1}^{n} T_{ji} f_j$ with f_k, where f_k is one of the orthonormal basis vectors of W. We obtain:

$$f_k \cdot Te_i = f_k \cdot \left(\sum_{j=1}^{n} T_{ji} f_j \right).$$

Using the linearity of the inner product and the orthonormality of the basis vectors f_j (i.e., $f_j \cdot f_k = \delta_{jk}$, where δ_{jk} is the Kronecker delta), we simplify the right-hand side:

$$f_k \cdot Te_i = \sum_{j=1}^{n} T_{ji}(f_k \cdot f_j) = \sum_{j=1}^{n} T_{ji}\delta_{jk} = T_{ki}.$$

Thus, we have the formula:

$$T_{ki} = f_k \cdot Te_i.$$

$\square$

Fact 35 (Component form)

Having represented the relationship between orthonormal bases previously lets see how it looks like for between elements. Simple calculations yield

$$\sum w_i \mathsf{f}_i = \mathsf{w} = \mathsf{Tv} = \sum v_j \mathsf{Te}_j = \sum T_{ij} v_j \mathsf{f}_i \quad \Rightarrow \quad w_i = \sum T_{ij} v_j.$$

Symbolically we write

$$w = Tv \quad \Rightarrow [w] = [T][v]$$

which represents

$$\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} T_{11} & \dots & T_{1m} \\ \vdots & \ddots & \vdots \\ T_{n1} & \dots & T_{nm} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}.$$

In the earlier proposition we assumed orthonormal bases so we know $f_i \cdot f_j = \delta_{ij}$. Let us do something similar again in inner product spaces but now for **general bases**. Explicitly we use the reciprocal basis relationship with a general base

$$g^i \cdot g_i = \delta_{ij}$$

Fact 36

Taking the inner product of both sides of the equation with g^i, we get:

$$\mathsf{g}^i \cdot \mathsf{Tf}_j = \mathsf{g}^i \cdot \left(\sum_{k=1}^{m} T_{kj} \mathsf{g}_k \right).$$

Using the linearity of the inner product and the orthonormality of the basis vectors g_i and their duals g^i, we can simplify the right-hand side:

$$\mathsf{g}^i \cdot \left(\sum_{k=1}^{m} T_{kj} \mathsf{g}_k \right) = \sum_{k=1}^{m} T_{kj} (\mathsf{g}^i \cdot \mathsf{g}_k) = \sum_{k=1}^{m} T_{kj} \delta^i_k = T_{ij}.$$

Thus, we have the relation:

$$T_{ij} = \mathsf{g}^i \cdot \mathsf{Tf}_j.$$

As usual having investigated the relationship between bases how about between elements?

Fact 37

For any $u \in V$ letting $v = Tu$ where $T : V \to W$ and (f_i) and (g_i) be the respective bases of V and W, note that

$$\mathsf{Tu} = \mathsf{T}\left(\sum_{i=1}^{n} u_i \mathsf{f}_i\right) = \sum u_i \mathsf{T}\mathsf{f}_i.$$

Since $Tf_i \in W$ it can be expressed in terms of the basis (g_i) of W as

$$\mathsf{T}\mathsf{f}_j = \sum_{j=1}^{m} T_{ij}\mathsf{g}_i \quad \Rightarrow \quad T_{ij} = \mathsf{g}^i \cdot \mathsf{T}\mathsf{f}_j.$$

then we have

$$\mathsf{v} = \mathsf{Tu} = \sum T_{ij} u_j \mathsf{g}_i \qquad \Rightarrow \quad v_i = \sum T_{ij} u_j,$$

Remark 38. *Notice the similar result we obtained for orthonormal bases!*

Let us now explore change of basis for general bases. Recall from `Algebra I` we had

Proposition 39

Let A be the matrix of linear transformation T with respect to given bases **B** and **C**

(a) Suppose that new bases $\mathbf{B}'$ and $\mathbf{C}'$ are related by the given matrices P and Q. The matrix of T with respect to the new bases is then

$$A' = Q^{-1}AP$$

(b) The matrices A' that repsent T with respect to other bases are those of the form $A' = Q^{-1}AP$ where Q and P can be any invertible matrices of the appropriate sizes

Proof. Well go to `Algebra I` for the proof duh.

Definition 40 (Change of basis for general bases)

Now similarly using general bases notation we have Let the new bases $(\tilde{f}_i)$ of V and (g_i) of W depend on the old bases of (f_i) of V and (g_i) of W by

$$\tilde{\mathsf{f}}_i = \sum A_{ji}\mathsf{f}_j, \qquad \tilde{\mathsf{g}}_i = \sum B_{ji}\mathsf{g}_j,$$

where $A_{ij} = \mathsf{f}^i \cdot \tilde{\mathsf{f}}_j$ and $B_{ij} = \mathsf{g}^i \cdot \tilde{\mathsf{g}}_j$.

Then it follows that

$$\begin{aligned}
\tilde{T}_{ij} &= \tilde{\mathsf{g}}^i \cdot \mathsf{T}\tilde{\mathsf{f}}_j \\
&= \sum \tilde{g}^{ik}\tilde{\mathsf{g}}_k \cdot \mathsf{T}\tilde{\mathsf{f}}_j \\
&= \sum \tilde{g}^{ik} B_{ck}\, \mathsf{g}_c \cdot \mathsf{T}\mathsf{f}_b\, A_{bj} \\
&= \sum \tilde{g}^{ik} B_{ck}\, g_{ca}\, \mathsf{g}^a \cdot \mathsf{T}\mathsf{f}_b\, A_{bj} \\
&= \sum \tilde{g}^{ik} B_{ck} g_{ca} T_{ab} A_{bj},
\end{aligned}$$

Remark 41. *You should recognize the why the coloured changes follow from previous sections or assumption in this context*

where $g_{ij} = g_i \cdot g_j$ and $\tilde{g}^{ij} = \tilde{g}^i \cdot \tilde{g}^j$ Noting that

$$\begin{aligned}\sum \tilde{g}^{ik} B_{ck} g_{ca} &= \sum g_{ac} B_{ck} \tilde{g}^{ki} \\ &= \sum g_{ac} \mathrm{g}^c \cdot \tilde{\mathrm{g}}_k \tilde{g}^{ki} \\ &= \mathrm{g}_a \cdot \tilde{\mathrm{g}}^i,\end{aligned}$$

and that

$$\tilde{\mathrm{g}}_i = \sum B_{ji} \mathrm{g}_j \quad \Rightarrow \quad \mathrm{g}_i = \sum B^{-1}_{ji} \tilde{\mathrm{g}}_j \quad \Rightarrow \quad B^{-1}_{ij} = \mathrm{g}_j \cdot \tilde{\mathrm{g}}^i,$$

It follows at once that

$$\tilde{T}_{ij} = \sum B^{-1}_{ia} T_{ab} A_{bj}.$$

□

5.2 tensor product basis for L(V,W)

Now we finally move on to something relatively new to you. Suppose that V and W are inner product spaces of dimension m and n. Let us focus on the set $L(V,W)$ defined by

$$L(V,W) = \{\mathrm{T} : V \to W \,|\, \mathrm{T} \text{ is linear }\},$$

of all linear maps from V into W. Well again from `Sheldon Linear algebra` we know this is a vector space just consider

$$(\mathrm{S} + \mathrm{T})(\mathrm{u}) = \mathrm{Su} + \mathrm{Tu}, \quad (a\mathrm{T})(\mathrm{u}) = a\,\mathrm{Tu},$$

for any $S, T \in L(V,W)$. With this we define

Definition 42

The **tensor product map** $w \otimes v \in L(V,W)$ is defined as

$$(w \otimes v)(u) = (v \cdot u)w$$

for any $u \in V$

Proposition 43

Let $(f_i)_{i=1}^m$ and $(g_i)_{i=1}^n$ be *orthonormal* bases of V and W respectively. Then the mn maps

$$\mathrm{g}_i \otimes \mathrm{f}_j : V \to W,$$

where $1 \le i \le n$ and $1 \le j \le m$. This means for any $v = \sum v_i f_i \in V$ we have

$$(\mathrm{g}_i \otimes \mathrm{f}_j)(\mathrm{v}) = (\mathrm{v} \cdot \mathrm{f}_j)\mathrm{g}_i = v_j \mathrm{g}_i.$$

Now prove tha these mn maps $\mathrm{g}_i \otimes \mathrm{f}_j \in L(V,W)$ are linearly independent

Proof. Suppose

$$\sum a_{ij}\mathrm{g}_i \otimes \mathrm{f}_j = 0,$$

Then for every $1 \leq k \leq m$ we have

$$\sum a_{ij}(\mathrm{g}_i \otimes \mathrm{f}_j)(\mathrm{f}_k) = 0 \quad \Rightarrow \quad \sum a_{ik}\mathrm{g}_i = 0 \quad \Rightarrow \quad a_{ik} = 0.$$

where in the first term we put f_k into the bracket because well it can take any element in V. The second term follows when you evaluate the first term which gets since (f_i) orthonormal by assumption

$$\sum a_{ij}(f_k \cdot f_j)g_i = \sum a_{ij}(\delta_{kj})g_i$$

which clearly gets the second term. This means to get zero the only solution for $\{a_{ij}\}$ is the trivial one.So the propositino follows □

Corollary 44

Assuming orthonormal bases

$$\dim(L(V, W)) = \dim(V)\dim(W).$$

Proof. Notice that any $T \in L(V, W)$ can be rpresented with

$$\mathrm{Tv} = \sum T_{ij}v_j\mathrm{g}_i = \sum T_{ij}(\mathrm{g}_i \otimes \mathrm{f}_j)(\mathrm{v}) = \left(\sum T_{ij}\mathrm{g}_i \otimes \mathrm{f}_j\right)\mathrm{v}.$$

Therefore we have

$$\mathrm{T} = \sum T_{ij}\mathrm{g}_i \otimes \mathrm{f}_j.$$

this tells us that $\mathrm{span}(\{\mathrm{g}_i \otimes \mathrm{f}_j\}) = L(V, W)$. Therefore the theorem clearly follows □

Remark 45. *awesome! Axler mentioned but did not give a rigourous proof for this so thanks dude!!*

5.3 Transpose of a linear map

ill skip unless necessary

6 Tensor Algebra

We begin real shit proper now.

Definition 46

Recall **multilinear** functions

$$\mathrm{T}(\mathrm{v}_1, \ldots, \mathrm{u}_i + a\mathrm{v}_i, \ldots, \mathrm{v}_n) = \mathrm{T}(\mathrm{v}_1, \ldots, \mathrm{u}_i, \ldots, \mathrm{v}_n) + a\,\mathrm{T}(\mathrm{v}_1, \ldots, \mathrm{v}_i, \ldots, \mathrm{v}_n).$$

and they are elements of the set

$$\mathcal{T}(V_1 \times \ldots \times V_n, \mathbb{R})$$

Definition 47

Also recall a **tensor** of order k on V which is basically a multilinear function of the form

$$\mathrm{A} : \underbrace{V \times \ldots \times V}_{k \text{ terms}} \to \mathbb{R},$$

and it is an element of $\mathcal{T}^k(V)$

Proposition 48

Also recall that $\mathcal{T}^k(V)$ is a real linear space

Proof. Consider

$$(\mathrm{A}+\mathrm{B})(\mathrm{u}_1, \ldots, \mathrm{u}_k) = \mathrm{A}(\mathrm{u}_1, \ldots, \mathrm{u}_k) + \mathrm{B}(\mathrm{u}_1, \ldots, \mathrm{u}_k),$$
$$(a\mathrm{A})(\mathrm{u}_1, \ldots, \mathrm{u}_k) = a\,\mathrm{A}(\mathrm{u}_1, \ldots, \mathrm{u}_k),$$

Fact 49

The set $\mathcal{T}^1(V) = \{\mathrm{T} : V \to \mathbb{R} \,|\, \mathrm{T} \text{ is linear}\}$ which is the set of all linear functions on V is as you know called the **dual space** of V and is often written as V^*. It turns out that if V is a *finite dimensional inner product space* then V^* is **canonically isomporhic** to V. This means that V and V^* are identical for all practical purposes. To see this consider that the action of any $\mathrm{v} \in \mathcal{T}^1(V)$ on an $u \in V$ is defined as

$$\mathrm{v}(u) = v \cdot u$$

If you recall the existence of this is given by **riesz representation theorem** which shows that for every v a unique v exists such that this relationship holds for all $u \in V$

OMG thanks so much now you know why we made this assumption in `MIT 18.101 Analysis on Manifolds`!!

Remark 50. *Also as reminder* $\mathcal{T}^0(V) = \mathbb{R}$ *which as implies are simply treated as scalars.*

Definition 51

The set of all $\mathcal{T}^k(V)$ is said to consitute a **tensor algebra** on V

Theorem 52 (2 tensors)

Consider the second order tensor $\hat{I} \in \mathcal{T}^2(V)$ on V defined as follows: for any $u, v \in V$

$$\hat{I}(u, v) = u \cdot v$$

More generally given any map $T \in L(V, V)$ we can define a second order tensor $\hat{T} \in \mathcal{T}^2(V)$ as follows for any $u, v \in V$

$$\hat{T}(u, v) = u \cdot Tv$$

It is obvious the dot product is a unique 2nd order tensor well because it's billinear. Specifically Let $u_1, u_2 \in V$ be arbitrary vectors and let $a, b \in \mathbb{R}$ be arbitrary scalars. We check whether $\hat{T}(u, v)$ is linear in the first argument:

$$\hat{T}(au_1 + bu_2, v) = (au_1 + bu_2) \cdot Tv.$$

Using the linearity of the dot product:

$$(au_1 + bu_2) \cdot Tv = a(u_1 \cdot Tv) + b(u_2 \cdot Tv).$$

This shows:

$$\hat{T}(au_1 + bu_2, v) = a\hat{T}(u_1, v) + b\hat{T}(u_2, v).$$

Thus, $\hat{T}(u, v)$ is linear in the first argument u.

Next, let $v_1, v_2 \in V$ be arbitrary vectors and let $a, b \in \mathbb{R}$ be arbitrary scalars. We check whether $\hat{T}(u, v)$ is linear in the second argument:

$$\hat{T}(u, av_1 + bv_2) = u \cdot T(av_1 + bv_2).$$

Using the linearity of the transformation T:

$$T(av_1 + bv_2) = aTv_1 + bTv_2,$$

and then applying the linearity of the dot product:

$$u \cdot (aTv_1 + bTv_2) = a(u \cdot Tv_1) + b(u \cdot Tv_2).$$

This shows:

$$\hat{T}(u, av_1 + bv_2) = a\hat{T}(u, v_1) + b\hat{T}(u, v_2).$$

Thus, $\hat{T}(u, v)$ is linear in the second argument v.

6.1 tensor products

Definition 53

Again recall given two tensors $A \in \mathcal{T}^k(V)$ and $B \in \mathcal{T}^\ell(V)$ we have

$$\mathrm{A} \otimes \mathrm{B}(\mathrm{u}_1, \ldots, \mathrm{u}_k, \mathrm{u}_{k+1}, \ldots, \mathrm{u}_{k+l}) = \mathrm{A}(\mathrm{u}_1, \ldots, \mathrm{u}_k)\mathrm{B}(\mathrm{u}_{k+1}, \ldots, \mathrm{u}_{k+l}).$$

Example 54
As a specical case given $v, w \in V$ their tensor product yields a second order tensor $v \otimes w \in \mathcal{T}^2(V)$: for any $u_1, u_2 \in V$

$$\mathrm{v} \otimes \mathrm{w}(\mathrm{u}_1, \mathrm{u}_2) = (\mathrm{v} \cdot \mathrm{u}_1)(\mathrm{w} \cdot \mathrm{u}_2).$$

Well we can treat v, w to be each a 1 order tensor by defining

$$\mathrm{v}(u_1) = v \cdot u_1$$

and

$$\mathrm{w}(u_2) = w \cdot u_2$$

as mentioned previously. Naturally this becomes a second order tensor like so.

Extending to a finite k number of vectors we have

Fact 55 (Euclidean Tensor Product)
Consider any set of k vectors $u_1, \ldots, u_k \in V$

$$\mathrm{u}_1 \otimes \ldots \otimes \mathrm{u}_k : V^k \to \mathbb{R},$$

their tensor product yields a k order tensor like so

$$\mathrm{u}_1 \otimes \ldots \otimes \mathrm{u}_k(\mathrm{v}_1, \ldots, \mathrm{v}_k) = (\mathrm{u}_1 \cdot \mathrm{v}_1) \ldots (\mathrm{u}_k \cdot \mathrm{v}_k).$$

We now have a problem. A sharp reader would notice that $v \otimes w$ can now stand for either a linear map $V \mapsto V$ or an element of $L(V, V)$(recall previous section) or a 2nd order tensor $\mathcal{T}^2(V)$. Indeed $L(V, V)$ and $\mathcal{T}^2$ are isomorphic - meaning they can be naturally idenitified with each other. Simply recall 52 and consider:

Theorem 56
$L(V, V)$ and $\mathcal{T}^2$ are isomorphic

Proof. Suppose $\Phi(T) = 0$. This means that for all $u, v \in V$, $\hat{T}(u, v) = u \cdot T(v) = 0$. Since this holds for all u, we conclude that $T(v) = 0$ for all $v \in V$. Therefore, T must be the zero map. Hence, $\Phi(T) = 0$ implies $T = 0$, so Φ is injective.

Let $\hat{T} \in \mathcal{T}^2(V)$. We want to find a linear transformation $T \in L(V, V)$ such that $\Phi(T) = \hat{T}$. For any fixed $v \in V$, define a linear map $T : V \to V$ by the condition that:

$$u \cdot T(v) = \underbrace{\hat{T}(u, v)}_{\text{see as a linear map of some } f(u)}, \quad \text{for all } u \in V.$$

This defines a unique linear map T because for each fixed v, $u \cdot T(v)$ represents a linear functional in u, and the Riesz Representation Theorem guarantees that there exists a unique vector $T(v) \in V$ such that this holds for all u. Thus, $\Phi(T) = \hat{T}$, proving that Φ is surjective.

6.2 basis representation

We will now explore how to represent any kth order tensor in terms of a basis.

With general bases we have

Theorem 57

$$\mathsf{A} = \sum A^*_{i_1\dots i_k} \mathsf{g}^{i_1} \otimes \dots \otimes \mathsf{g}^{i_k}.$$

where $A^*_{i_1\dots i_k} = \mathsf{A}(\mathsf{g}_{i_1}, \dots, \mathsf{g}_{i_k})$ which are called the **covariant components** of $A \in \mathcal{T}^k(V)$

Proof. Consider

$$\begin{aligned}
\mathsf{A}(\mathsf{u}_1, \dots, \mathsf{u}_k) &= \mathsf{A}\left(\sum u_{1i_1}\mathsf{g}_{i_1}, \dots, \sum u_{ki_k}\mathsf{g}_{i_k}\right) \\
&= \sum \mathsf{A}(\mathsf{g}_{i_1}, \dots, \mathsf{g}_{i_k}) u_{1i_1} \dots u_{ki_k}, \\
&= \sum A^*_{i_1\dots i_k} (\mathsf{g}^{i_1} \otimes \dots \otimes \mathsf{g}^{i_k})(\mathsf{u}_1, \dots, \mathsf{u}_k)
\end{aligned}$$

Note the sum in the second line represents $\sum_{i_1, i_2, \dots, i_k}$. This follows by multilinearity. The last line follows by

$$(\mathsf{g}^{i_1} \otimes \dots \otimes \mathsf{g}^{i_k})(\mathsf{u}_1, \dots, \mathsf{u}_k) = (g^{i_1} \cdot u_1)(g^{i_2} \cdot u_2)\dots$$

But consider that

$$(g^{i_1} \cdot u_1) = g^{i_1} \cdot (\sum_{i_1} u_{1i_1} g_{i_1}) = \sum_{i_1} \delta^{i_1}_{i_1} u_{1i_1} = \sum_{i_1} u_{1i_1}$$

as desired

Corollary 58

$$\mathsf{A} = \sum A_{i_1\dots i_k} \mathsf{g}_{i_1} \otimes \dots \otimes \mathsf{g}_{i_k},$$

where

$$A_{i_1\dots i_k} = \sum g^{i_1 j_1} \dots g^{i_k j_k} A^*_{j_1 \dots j_k}.$$

which are known as th **contravariant components** of A

Proof. Essentially we have to show the RHS equals the LHS

$$\mathsf{A} = \sum_{i_1, i_2, \dots, i_k} \sum_{j_1, j_2, \dots, j_k} g^{i_1 j_1} \dots g^{i_k j_k} A^*_{j_1 \dots j_k} \mathsf{g}_{i_1} \otimes \dots \otimes \mathsf{g}_{i_k} = \sum_{i_1, i_2, \dots, i_k} A^*_{i_1\dots i_k} \mathsf{g}^{i_1} \otimes \dots \otimes \mathsf{g}^{i_k}$$

Since

$$\mathsf{g}^{i_1} = \sum_{j_1} g^{i_1 j_1} \mathsf{g}_{j_1}, \quad \mathsf{g}^{i_2} = \sum_{j_2} g^{i_2 j_2} \mathsf{g}_{j_2}, \quad \dots, \quad \mathsf{g}^{i_k} = \sum_{j_k} g^{i_k j_k} \mathsf{g}_{j_k}.$$

Thus, the contravariant tensor product:

$$\mathsf{g}^{i_1} \otimes \mathsf{g}^{i_2} \otimes \dots \otimes \mathsf{g}^{i_k}$$

can be written as:

$$\left(\sum_{j_1} g^{i_1 j_1} \mathsf{g}_{j_1}\right) \otimes \left(\sum_{j_2} g^{i_2 j_2} \mathsf{g}_{j_2}\right) \otimes \dots \otimes \left(\sum_{j_k} g^{i_k j_k} \mathsf{g}_{j_k}\right).$$

Expanding the tensor product yields:

$$\sum_{j_1, \dots, j_k} g^{i_1 j_1} g^{i_2 j_2} \dots g^{i_k j_k} (\mathsf{g}_{j_1} \otimes \mathsf{g}_{j_2} \otimes \dots \otimes \mathsf{g}_{j_k}).$$

Substituting this back into the contravariant expression for A, we have:

$$\mathsf{A} = \sum_{i_1,\ldots,i_k} A^*_{i_1\ldots i_k} \left(\sum_{j_1,\ldots,j_k} g^{i_1 j_1} \ldots g^{i_k j_k} \mathsf{g}_{j_1} \otimes \ldots \otimes \mathsf{g}_{j_k} \right).$$

Now, comparing this to the earlier expression for A (which we wrote in the covariant form), we see that the two expressions match term by term. Thus, we have shown that:

$$\sum_{i_1,i_2,\ldots,i_k} \sum_{j_1,j_2,\ldots,j_k} g^{i_1 j_1} \ldots g^{i_k j_k} A^*_{j_1\ldots j_k} \mathsf{g}_{i_1} \otimes \ldots \otimes \mathsf{g}_{i_k} = \sum_{i_1,i_2,\ldots,i_k} A^*_{i_1\ldots i_k} \mathsf{g}^{i_1} \otimes \ldots \otimes \mathsf{g}^{i_k}.$$

6.3 change of basis

Proposition 59

For general bases suppose that $A \in \mathcal{T}^k(V)$ has components $A^*_{i_1\ldots i_k}$ with respect to bases $(g^{i_1} \otimes \ldots \otimes g^{i_k})$. Similarly suppose that $\tilde{A} \in \mathcal{T}^k(V)$ has components $\tilde{A}^*_{i_1\ldots i_k}$ with respect to bases $(\tilde{g}^{i_1} \otimes \ldots \otimes \tilde{g}^{i_k})$.Then can express the change of bases as have

$$\begin{aligned}
\tilde{A}^*_{i_1\ldots i_k} &= \mathsf{A}(\tilde{\mathsf{g}}_{i_1}, \ldots, \tilde{\mathsf{g}}_{i_k}) \\
&= \mathsf{A}\left(\sum (\tilde{\mathsf{g}}_{i_1} \cdot \mathsf{g}^{j_1}) \mathsf{g}_{j_1}, \ldots, \sum (\tilde{\mathsf{g}}_{i_k} \cdot \mathsf{g}^{j_k}) \mathsf{g}_{j_k} \right) \\
&= \sum (\tilde{\mathsf{g}}_{i_1} \cdot \mathsf{g}^{j_1}) \ldots (\tilde{\mathsf{g}}_{i_k} \cdot \mathsf{g}^{j_k}) A^*_{j_1\ldots j_k}.
\end{aligned}$$

Proof. recall 40

6.4 contraction

First consider orthonormal basis

Definition 60

Let V be a finite dimensional inner product space and let (e_i) be an *orthonormal* basis of V. Given a tensor $A \in \mathcal{T}^k(V)$ of order k where $k \geq 2$ the (i,j) **contraction** of A is the tensor $\mathcal{C}_{i,j}A \in \mathcal{T}^{k-2}(V)$ of order $(k-2)$ defined as follows for any $v_1, \ldots, v_k \in V$

$$\begin{aligned}
&\mathcal{C}_{i,j}\mathsf{A}(\mathsf{v}_1, \ldots, \mathsf{v}_{i-1}, \mathsf{v}_{i+1}, \ldots, \mathsf{v}_{j-1}, \mathsf{v}_{j+1}, \ldots, \mathsf{v}_k) \\
&= \sum_a \mathsf{A}(\mathsf{v}_1, \ldots, \mathsf{v}_{i-1}, \mathsf{e}_a, \mathsf{v}_{i+1}, \ldots, \mathsf{v}_{j-1}, \mathsf{e}_a, \mathsf{v}_{j+1}, \ldots, \mathsf{v}_k).
\end{aligned}$$

Example 61

Suppose that $A \in \mathcal{T}^2(V)$ is a second order tensor. In this case there is only one possible contraction of the tensor A namely $\mathcal{C}_{1,2}A \in \mathbb{R}$. This is easily computed as follows: if (e_i) is an orthonormal basis of V then

$$\mathcal{C}_{1,2}\mathsf{A} = \sum_a \mathsf{A}(\mathsf{e}_a, \mathsf{e}_a) = \sum_a A_{i_1,i_2}(e_{i_1} \otimes e_{i_2})(e_a, e_a) = \sum_{i_1,i_2} A_{i_1,i_2}(e_{i_1} \cdot e_a)(e_{i_2} \cdot e_a) = \sum_{i_1,i_2} A_{i_1,i_2} \delta^{i_1}_a \delta^{i_2}_a = \sum A_{aa}$$

In this special case the contraction $\mathcal{C}_{1,2}$ is known as the **trace** of A and is often written as $\mathrm{tr}(A)$

Now general basis

Definition 62

Let (g_i) be a general basis of V. Given a tensor $A \in \mathcal{T}^k(V)$ of order k where $k \geq 2$ the (i,j) contraction of A can be defined as follows for any $v_1, \ldots, v_k \in V$

$$\begin{aligned}
&\mathcal{C}_{i,j}\mathrm{A}(\mathrm{v}_1, \ldots, \mathrm{v}_{i-1}, \mathrm{v}_{i+1}, \ldots, \mathrm{v}_{j-1}, \mathrm{v}_{j+1}, \ldots, \mathrm{v}_k) \\
&= \sum_a \mathrm{A}(\mathrm{v}_1, \ldots, \mathrm{v}_{i-1}, \mathrm{g}_a, \mathrm{v}_{i+1}, \ldots, \mathrm{v}_{j-1}, \mathrm{g}^a, \mathrm{v}_{j+1}, \ldots, \mathrm{v}_k).
\end{aligned}$$

Proposition 63

Thee order of basis g_i andd its reciprocal basis g^i inserted is not important i.e

$$\sum \mathrm{A}(\ldots, \mathrm{g}_a, \ldots, \mathrm{g}^a, \ldots) = \sum \mathrm{A}(\ldots, \mathrm{g}^a, \ldots, \mathrm{g}_a, \ldots).$$

Proof. Simply consider the example below

Proposition 64

$$\begin{aligned}
\mathcal{C}_{i,j}\mathrm{A} = \sum g_{a_i a_j} A_{a_1 \ldots a_{i-1} a_i a_{i+1} \ldots a_{j-1} a_j a_{j+1} \ldots a_k} \\
\mathrm{g}_{a_1} \otimes \mathrm{g}_{a_{i-1}} \otimes \mathrm{g}_{a_{i+1}} \otimes \ldots \otimes \mathrm{g}_{a_{j-1}} \otimes \mathrm{g}_{a_{j+1}} \otimes \ldots \otimes \mathrm{g}_{a_k}.
\end{aligned}$$

Proof. Simply consider the example below

Example 65

Upon contraction of $A \in \mathcal{T}^2$ by $\mathcal{C}_{1,2}$ we have

$$\mathrm{A} = \sum A_{ij}\, \mathrm{g}_i \otimes \mathrm{g}_j \quad \Rightarrow \quad \mathrm{tr}(\mathrm{A}) = \sum g_{ab} A_{ab}$$

Consider

$$
\begin{aligned}
\mathcal{C}_{1,2}\mathsf{A} &= \sum_a \mathsf{A}(\mathsf{g}_a, \mathsf{g}^a) \\
&= \sum_a \left(\sum_{i,j} A_{ij} (\mathsf{g}_i \otimes \mathsf{g}_j)(\mathsf{g}_a, \mathsf{g}^a) \right) \\
&= \mathcal{C}_{1,2}\mathsf{A} = \sum_a \left(\sum_{i,j} A_{ij} (\mathsf{g}_i \cdot \mathsf{g}_a)(\mathsf{g}_j \cdot \mathsf{g}^a) \right) \\
&= \mathcal{C}_{1,2}\mathsf{A} = \sum_a \left(\sum_{i,j} A_{ij} (\mathsf{g}_i \cdot \mathsf{g}_a)\delta_j^a \right) \\
&= \mathcal{C}_{1,2}\mathsf{A} = \left(\sum_{i,j} A_{ij} (\mathsf{g}_i \cdot \mathsf{g}_j) \right) \\
&= \mathcal{C}_{1,2}\mathsf{A} = \left(\sum_{i,j} A_{ij} g_{ij} \right)
\end{aligned}
$$

as desired

6.5 generalized dot product of tensors

Example 66

Note that in the simplest case given two vectors $u, v \in V$ their inner product can be expressed in terms of an contraction as follows

$$u \cdot v = \mathcal{C}_{1,2}(u \otimes v)$$

This means

$$
\begin{aligned}
\mathcal{C}_{1,2}(u \otimes v) &= \sum_a (u \otimes v)(g_a, g^a) \\
&= \sum_a \sum_{i,j} u_i v_j (g_i \otimes g_j)(g_a, g^a) \\
&= \sum_a \sum_{i,j} u_i v_j g_{ia} \delta_j^a \\
&= \sum_{i,j} u_i v_j g_{i,j}
\end{aligned}
$$

Remark 67. *You can clearly see in the standard orthonormal bases then $g_{ij} = \delta_{ij}$ which then gives your familiar euclidean dot product*

Example 68

Suppose that $u, v, w \in V$ ar any three vectors in V. THen the vector $(u \otimes v) \cdot w \in V$ is defined as follows

$$(u \otimes v) \cdot w = \mathcal{C}_{2,3}(u \otimes v \otimes w)$$

Note this makes sense by definition of dot product

To see what this means consider an arbitary vector $z \in V$: if (g_i) is a general basis of V then

$$\begin{aligned}
(u \otimes v) \cdot w &= \mathcal{C}_{2,3}(\mathrm{u} \otimes \mathrm{v} \otimes \mathrm{w})(\mathrm{z}) \\
&= \sum (\mathrm{u} \otimes \mathrm{v} \otimes \mathrm{w})(\mathrm{z}, \mathrm{g}_a, \mathrm{g}^a) \\
&= \sum u_i v_j w_k (\mathrm{g}_i \otimes \mathrm{g}_j \otimes \mathrm{g}_k) \left(\sum z_b \mathrm{g}_b, \mathrm{g}_a, \mathrm{g}^a \right) \\
&= \sum g_{ib} u_i v_j w_k z_b g_{ja} \delta_{ka} \\
&= \sum g_{jk} v_j w_k \sum g_{ia} u_a z_a \\
&= (\mathrm{v} \cdot \mathrm{w})(\mathrm{u} \cdot \mathrm{z}) \\
&= (\mathrm{v} \cdot \mathrm{w}) \mathrm{u}(\mathrm{z}).
\end{aligned}$$

where the 2nd last line follows by seeing the resemblence with the previous example while the last line follows because z is arbitary and by reisz represnentation theorem(recall previously) this is possible.

Corollary 69

Therefore we have

$$(\mathrm{u} \otimes \mathrm{v}) \cdot \mathrm{w} = (\mathrm{v} \cdot \mathrm{w}) \mathrm{u}$$

Consider a second order tensor $A \in \mathcal{T}^2(V)$. Using this corollary we have

$$\begin{aligned}
\mathrm{A} \cdot \mathrm{v} &= \left(\sum A_{ij} \mathrm{g}_i \otimes \mathrm{g}_j \right) \cdot \left(\sum v_k \mathrm{g}_k \right) \\
&= \mathrm{A} \cdot \mathrm{v} \\
&= \sum_{i,j,k} A_{ij} v_k (\mathrm{g}_i \otimes \mathrm{g}_j) \cdot \mathrm{g}_k \\
&= \sum A_{ij} v_k g_{jk} \mathrm{g}_i \in V.
\end{aligned}$$

where the the corllary is speically applied to the 2nd last line

Now let us extend our definition of dot products for tensor of second order.

Definition 70

Given $u, v, w, z \in V$ the dot product

$$(u \otimes v) \cdot (w \otimes z) \in \mathbb{R}$$

between the second order tensors $u \otimes v \in \mathcal{T}^2(V)$ and $w \otimes z \in \mathcal{T}^2(V)$ we have

$$(\mathrm{u} \otimes \mathrm{v}) \cdot (\mathrm{w} \otimes \mathrm{z}) = \mathcal{C}_{1,2} \left(\mathcal{C}_{1,3} \left(\mathrm{u} \otimes \mathrm{v} \otimes \mathrm{w} \otimes \mathrm{z} \right) \right) = (\mathrm{u} \cdot \mathrm{w})(\mathrm{v} \cdot \mathrm{z}).$$

To understand this consider $\mathsf{A} = \sum A_{ij}\mathbf{g}_i \otimes \mathbf{g}_j$ and $\mathsf{B} = \sum B_{ij}\mathbf{g}_i \otimes \mathbf{g}_j$ then note that

$$\mathsf{A} \cdot \mathsf{B} = \sum A_{ij}B_{ij}.$$

6.6 volume forms

Definition 71

Let V be an inner product space of dimension n. A tensor $A \in \mathcal{T}^k(V)$ is said to be **symmetric** if for any $u_1, \ldots, u_k \in V$

$$\mathsf{A}(\mathsf{u}_1, \ldots, \mathsf{u}_i, \ldots, \mathsf{u}_j, \ldots, \mathsf{u}_n) = \mathsf{A}(\mathsf{u}_1, \ldots, \mathsf{u}_j, \ldots, \mathsf{u}_i, \ldots, \mathsf{u}_n),$$

where $1 \leq i < j \leq n$. If the sign reverses when any two arguments are interchanged then the tensor is said to be **skew-symmetric**

Definition 72

a **volume form** on V is a skew symmetric tensor $\varepsilon : \times^N V \to \mathbb{R}$ of order n. It is customary to denote the set of all volume forms on V using the notation $\Omega^n(V)$

$$\Omega^n(V) = \{\epsilon \in \mathcal{T}^n(V) \,|\, \epsilon \text{ is skew-symmetric}\}.$$

Theorem 73

Given two volume forms $\varepsilon, \omega \in \Omega^n(V)$ it can be shown that there exists a scalar $a \in \mathbb{R}$ such that $\omega = a\varepsilon$. In other words the set of all antisymmetric tensors of order n over an n dimesnional vector space is a vector space of dimension 1. That is

$$\dim(\Omega^n(V)) = 1$$

Proof. Recall from your `MIT 18.101 Analysis on Manifold` notes the dimension of alternating k-tensors(which is exactly what we have here) is $\binom{n}{k}$ but since in our case $k = n$ therefore the dimension of the standard volume form is 1. Because we have assumed on trivial volume forms(the zero map is not in Ω^n) it follows that every volume form are obviously scalar multiples of each other since dimension 1

Definition 74

Picking a particular volume form $\varepsilon \in \Omega^n(V)$ the set of all volume forms $w \in \Omega^n(V)$ such that $w = a\varepsilon$ for some $a > 0$ are said t consitute an **orientation** of V. Given this orientation on V a basis (g_i) of V is said to be **right-handed** if $\varepsilon(g_1, \ldots, g_n) > 0$ and **left-handed** otherwise

Definition 75
the **standard volume form** on $\mathbb{R}^3$ written $\varepsilon \in \Omega^3(\mathbb{R}^3)$ is sometimes called the **levi-civita tensor** is defined as follows: if (e_i) denotes the standard basis of $\mathbb{R}^3$ then

$$\epsilon = \sum \epsilon_{ijk} \mathrm{e}_i \otimes \mathrm{e}_j \otimes \mathrm{e}_k,$$

where ε_{ijk} is the levi-citvita symbol defined as follows

$$\epsilon_{ijk} = \begin{cases} 1, & \{i, j, k\} \text{ is an even permutation of } \{1, 2, 3\}, \\ -1, & \{i, j, k\} \text{ is an odd permutation of } \{1, 2, 3\}, \\ 0, & \text{otherwise}, \end{cases}$$

Note that $\varepsilon(e_1, e_2, e_3) = \varepsilon_{123} = 1$ the standard basis of $\mathbb{R}^3$ is thus right handed with respect to the standard volume form of $\mathbb{R}^3$

Definition 76
In three dimensional euclidean space $\mathbb{R}^3$ there exists a special algebraic operation called the **cross product** defined as follows: given two vectors $u = \sum u_i e_i \in \mathbb{R}^3$ and $v = \sum v_i e_i \in \mathbb{R}^3$ where (e_i) is the standard basis of $\mathbb{R}^3$ their cross product is defined as the vector $u \times v \in \mathbb{R}^3$ given by

$$\mathrm{u} \times \mathrm{v} = (u_2 v_3 - u_3 v_2)\mathrm{e}_1 + (u_3 v_1 - u_1 v_3)\mathrm{e}_2 + (u_1 v_2 - u_2 v_1)\mathrm{e}_3.$$

which can be written with levi-civita symbols ε_{ijk} as follows

$$\mathrm{u} \times \mathrm{v} = \sum \epsilon_{ijk} u_j v_k \mathrm{e}_i,$$

alternatively

$$\mathrm{u} \cdot (\mathrm{v} \times \mathrm{w}) = \epsilon(\mathrm{u}, \mathrm{v}, \mathrm{w}).$$

the equivalence between the the two defintiions are obvious. Just let u be the *standard basis* e_i which then immediately yields

$$(\mathrm{v} \times \mathrm{w})_i = \sum \epsilon_{ijk} v_j w_k$$

recall standard basis allows us to assume $e_i \cdot e_j = \delta_{ij}$ since orthonormal. Also see that

$$\epsilon(u, v, w) = \sum \epsilon_{ijk} \mathrm{e}_i \otimes \mathrm{e}_j \otimes \mathrm{e}_k(u, v, w) = \sum \epsilon_{ijk}(u \cdot e_i)(v \cdot e_j)(w \cdot e_k) = \sum \epsilon_{ijk} u_j v_k \mathrm{e}_i$$

Example 77

Consider the following what we call $\boldsymbol{\varepsilon} - \boldsymbol{\delta}$ identities

$$
\begin{aligned}
\epsilon_{ijk}\epsilon_{lmn} &= \det \begin{bmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{bmatrix}, \\
\epsilon_{ijk}\epsilon_{imn} &= \det \begin{bmatrix} \delta_{jm} & \delta_{jn} \\ \delta_{km} & \delta_{kn} \end{bmatrix}, \\
\epsilon_{ijk}\epsilon_{ijn} &= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}, \\
\epsilon_{ijk}\epsilon_{ijk} &= 6.
\end{aligned}
$$

To see why this is true just consider if ijk and lmn have repeated that obvioulsy we will have either repeating columns or rows in this square matrix in which case determinant is zero. Now consider what happens when you swap one of the orders in either $\{i, j, k\}$ or $\{l, m, n\}$. Recall that a swap in rows or columns multiplies the determinant by -1 from `Algebra I`. So clearly the first identity holds. Having proven the 1st identity just expand out the derminant

$$
\begin{aligned}
\epsilon_{ijk}\epsilon_{lmn} =& \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\
& - \delta_{in}\delta_{jm}\delta_{kl} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{im}\delta_{jl}\delta_{kn}
\end{aligned}
$$

and let $l = i$ to prove the second case like so

$$
\begin{aligned}
\epsilon_{ijk}\epsilon_{imn} =& \delta_{ii}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{ki} + \delta_{in}\delta_{ji}\delta_{km} \\
& - \delta_{in}\delta_{jm}\delta_{ki} - \delta_{ii}\delta_{jn}\delta_{km} - \delta_{im}\delta_{ji}\delta_{kn} \\
=& 3\delta_{jm}\delta_{kn} + \delta_{km}\delta_{jn} + \delta_{jn}\delta_{km} \\
& - \delta_{kn}\delta_{jm} - 3\delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn} \\
=& \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}
\end{aligned}
$$

Do the same to get the 3rd and 4th idenity by letting $i = l, m = j$ and $i = l, m = j, n = k$ respectively. For the finally relation we have

$$\epsilon_{ijk}\epsilon_{ijk} = 2\varepsilon_{ii} = 6$$

this is because there are 3 possible values of i

Remark 78. *Alternatively let K_{ijk} be the 3×3 matrix whose columns are the standard basis vectors e_i, e_j, e_k. So* $\det K_{ijk} = \varepsilon_{ijk}$ *Then we have*

$$\begin{pmatrix} \delta_{ip} & \delta_{iq} & \delta_{ir} \\ \delta_{jp} & \delta_{jq} & \delta_{jr} \\ \delta_{kp} & \delta_{kq} & \delta_{kr} \end{pmatrix} = K_{ijk}^T K_{pqr},$$

In which case taking the derminant of the above yields the first identity

Example 79

$$\det[A] = \frac{1}{6}\varepsilon_{ijk}\varepsilon_{pqr}A_{ip}A_{jq}A_{kr}$$

and

$$\varepsilon_{ijk}\det A = e_{ijk}A_{ip}A_{jq}A_{kr}$$

Proof. Recall that

$$\det[A] = \varepsilon_{ijk}A_{1i}A_{2j}A_{3k}$$

and that

$$\det[A] = -\varepsilon_{ijk}A_{2i}A_{1j}A_{3k} = \varepsilon_{jik}A_{1j}A_{2i}A_{3k}$$

Therefore for any fixed p, q, k in the right order we will get the determinant. The correct sign that will get the determinant is assigned by $\varepsilon_{p,q,k}$. Given that p, q, k permutes a total of $3! = 6$ and any one can get $\det A$ the below follows

$$\det[A] = \frac{1}{6}\varepsilon_{ijk}\varepsilon_{pqr}A_{ip}A_{jq}A_{kr}$$

. The other identity follows similarly.

$$\varepsilon_{ijk}\det A = e_{ijk}A_{ip}A_{jq}A_{kr}$$

See that ε_{ijk} assigns a sign to $\det A$ following the order of p, q, r

Example 80 (Triple Cross Product Levi Cevita)

Triple cross product proof with levi cevita

$$\begin{aligned}
\mathrm{u} \times (\mathrm{v} \times \mathrm{w}) &= \sum \epsilon_{ijk}u_j(\mathrm{v} \times \mathrm{w})_k \mathrm{e}_i \\
&= \sum \epsilon_{ijk}\epsilon_{klm}u_j v_l w_m \mathrm{e}_i \\
&= \sum (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})u_j v_l w_m \mathrm{e}_i \\
&= \left(\sum u_m w_m\right)\sum v_l \mathrm{e}_l - \left(\sum u_j v_j\right)\sum w_m \mathrm{e}_m \\
&= (\mathrm{u}\cdot\mathrm{w})\mathrm{v} - (\mathrm{u}\cdot\mathrm{v})\mathrm{w}.
\end{aligned}$$

Example 81

prove

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}\nabla \cdot \mathbf{b} + \mathbf{b}\cdot\nabla\mathbf{a} - \mathbf{b}\nabla\cdot\mathbf{a} - \mathbf{a}\cdot\nabla\mathbf{b}$$

where recall by definition of cross product with levita cevita we have

$$\mathbf{a} \times \mathbf{b} = \mathbf{e}_i\epsilon_{ijk}a_i b_j$$

Proof.

$$\begin{aligned}
\nabla \times (\vec{A} \times \vec{B}) &= \partial_l \hat{e}_l \times (a_i b_j \hat{e}_k \epsilon_{ijk}) &(3)\\
&= \partial_l a_i b_j \epsilon_{ijk} \underbrace{(\hat{e}_l \times \hat{e}_k)}_{(\hat{e}_l \times \hat{e}_k) = \hat{e}_m \epsilon_{lkm}} &(4)\\
&= \partial_l a_i b_j \hat{e}_m \underbrace{\epsilon_{ijk}\epsilon_{mlk}}_{\text{contracted epsilon identity}} &(5)\\
&= \partial_l a_i b_j \hat{e}_m \underbrace{(\delta_{im}\delta_{jl} - \delta_{il}\delta_{jm})}_{\text{They sift other subscripts}} &(6)\\
&= \partial_j (a_i b_j \hat{e}_i) - \partial_i (a_i b_j \hat{e}_j) &(7)\\
&= a_i \partial_j b_j \hat{e}_i + b_j \partial_j a_i \hat{e}_i - (a_i \partial_i b_j \hat{e}_j + b_j \partial_i a_i \hat{e}_j) &(8)\\
&= \vec{A}(\nabla \cdot \vec{B}) + (\vec{B} \cdot \nabla)\vec{A} - (\vec{A} \cdot \nabla)\vec{B} - \vec{B}(\nabla \cdot \vec{A}) &(9)\\
& &(10)
\end{aligned}$$

6.7 special groups of linear maps

In the above we have considered skew symmetric tensors.

Definition 82

We now consider the **skew-symmetric linear group** on V defined by

$$\text{skw}(V) = \{\mathsf{T} \in L(V, V) \mid \mathsf{T}^T = -\mathsf{T}\}.$$

Theorem 83

Given any $v \in \mathbb{R}^3$ we may associate it with the skew symmetric linear map $\check{v} \in \text{skw}(\mathbb{R}^3)$ defined as follows: for any $u \in \mathbb{R}^3$

$$\check{v}u = v \times u$$

In other words any skew symmetric tensor in a 3D euclidean space can be realized by a cross product of some $v \in \mathbb{R}^3$

Remark 84. *Notice this is like the "Reicsz representation theorem but for cross products instead of inner products". However note that they are different in a sense that one works with functions, that is to produce scalars from vectors while the other works with linear transformations which involves vectors producing other vectors. This explains the big difference in complexity of the proof as the former requires functional analysis while the other are just linear maps where we are working with the same mathematical object - vectors.*

Proof. Let W be a skew-symmetric linear transformation on $\mathbb{R}^3$. This means that for any vectors $x, y \in \mathbb{R}^3$,

$$\langle Wx, y \rangle = -\langle x, Wy \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^3$. In terms of matrices, a skew-symmetric transformation W can be written as a 3×3 matrix $[W]$ that satisfies $W^\top = -W$. The general form of a skew-symmetric matrix in $\mathbb{R}^3$ is:

$$W = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix},$$

where $w = (w_1, w_2, w_3)^\top$ is some vector in $\mathbb{R}^3$. Notice that the entries of this matrix involve the components of the vector w. Now, let's look at the cross product $w \times x$ for some vector $x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$ with $w = (w_1, w_2, w_3)^\top$. The cross product can be expressed as:

$$w \times x = \begin{pmatrix} w_2x_3 - w_3x_2 \\ w_3x_1 - w_1x_3 \\ w_1x_2 - w_2x_1 \end{pmatrix}.$$

This can be written as a matrix-vector multiplication:

$$w \times x = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Notice that this matrix is exactly the same as the matrix W from Step 1! Therefore, we conclude that:

$$Wx = w \times x$$

for all $x \in \mathbb{R}^3$, where W is the skew-symmetric matrix corresponding to the vector w.

Remark 85. *Alternatively you could use levi cevita notation to prove it too but it doesn't make a difference in efficiency here so we don't show it*

7 euclidean tensor analysis

7.1 coordinate systems

Definition 86

A **coordinate system** on $\mathbb{R}^3$ is an open subset $U \subseteq \mathbb{R}^3$ with a smooth injective map $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ such that $\phi^{-1} : \phi(U) \to U$ is also smooth. Such a map is called a **diffeomorphism**. The map ϕ is called a **coordinate chart** and the open set U is called a **coordinate patch** on $\mathbb{R}^3$. Given any $x \in U$ the triple of real numbers $\phi(x) \in \mathbb{R}^3$ are called the **curvilinear coordinates** of x

Remark 87. *In this context it is clear that U a subset of the manifold $\mathbb{R}^3$ while $\phi(x)$, the curvilinear coordinates describe a point on an open subset in $\mathbb{R}^n$(not in the manifold) but diffeomorphic to $U \subseteq M$*

Definition 88

Also recall that the **local parameterization** ϕ^{-1} is essentially the inverse of the coordinate patch. The latter is coordinate patch on manifold $U \subseteq M$ maps to open subset $V \in \mathbb{R}^n$ while the former is open subset maps to coordinate patch

Example 89
The special feature about $\mathbb{R}^3$(and also $\mathbb{R}^n$) in general is that as you recall it admits a **global coordinate system**$(\mathbb{R}^3, \iota)$ where $\iota : \mathbb{R}^3 \to \mathbb{R}^3$ is thte identity map defined as follows: for any $x \in \mathbb{R}^3, \iota(x) = x$. In fact this is precisely what we call the **cartesian coordinate system** and that $\iota(x) \in \mathbb{R}^3$ with respect to this global coordinate system are called the **cartesian coordinates** of x

Remark 90. *Recall from* `Geometry of Manifolds MIT 18.965` *a global coordinate system is basically a single coordinate patch that covers the whole manifold.*

7.2 tensor fields

Recall that

Definition 91
A **vector field** on $U \subseteq \mathbb{R}^3$ is amap of the form

$$v : U \to TU$$

such that $v(x) \in T_xU$ for every $x \in U$

The vector field v can be expressed with respect to the cartesian coordinate system as follows

$$\mathrm{v}(\mathrm{x}) = \sum v_i(\mathrm{x})\mathrm{e}_i$$

Now we define a tensor field analagously

Definition 92
A **tensor field** of order k on $U \subseteq \mathbb{R}^3$ is defined bya map

$$A : U \to \otimes^k TU$$

such that $A(x) \in \otimes^k T_xU$ for every $x \in U$. In writing this we have introduced the following notation

$$\otimes^k T_\mathrm{x}U = \{\mathrm{T} : \times^k T_\mathrm{x}U \to \mathbb{R} \,|\, \mathrm{T} \text{ is multilinear } \} = \mathcal{T}^k(T_\mathrm{x}U),$$
$$\otimes^k TU = \cup_{\mathrm{x}\in U}\left(\otimes^k T_\mathrm{x}U\right) = \cup_{\mathrm{x}\in U}\mathcal{T}^k(T_\mathrm{x}U).$$

Definition 93
The **gradient** of f is defined as the vector field $\nabla f : U \to TU$ such that at every $x \in U$

$$\nabla f(\mathrm{x}) \cdot \mathrm{v} = \nabla_\mathrm{v} f(\mathrm{x})$$

for any $v \in T_xU$. Note that $\nabla f(x) \in T_xU$ by definition. Since $v \in T_xU$ the dot product $\nabla f(x) \cdot v$ is indeed well-defined

Example 94
Important examples include

- 0 order tensor: scalar field $\mathbb{R}^n \mapsto \mathbb{R}$ - result can be represented as a "constant"
- 1 order tensor: vector field $\mathbb{R}^n \mapsto T(\mathbb{R}^n)$ - result can be represented as a "column vector"
- 2 order tensor: tensor field $\mathbb{R}^n \mapsto \otimes^2 T(\mathbb{R}^n)$ - result can be represented by "matrix"

7.3 Covariant derivative and gradient

Definition 95 (Scalar field version)
Let $f : U \subseteq \mathbb{R}^3 \to \mathbb{R}$ where U is an open subset of $\mathbb{R}^3$ be a smooth scalar field on U. The **directional derivative** also called the **covariant derivative** of f at $x \in U$ along any $v \in T_x U$ is defined as the scalar $\nabla_v f(x) \in \mathbb{R}$ such that

$$\nabla_\mathrm{v} f(\mathrm{x}) = \frac{d}{dt}\bigg|_{t=0} f(\mathrm{x} + t\mathrm{v})$$

Note that in the term $f(x+tv)$ in the definition above $x \in U$ and $v \in T_x U$. Since U and $T_x U$ are not the same spaces the quantity $x + tv$ is ill-defined strictly speaking.

The directional derivative of f at $x \in U$ along $v \in T_x U$ can be written in term os the cartesian coordinate sysetm on $\mathbb{R}^3$ as follows

$$\begin{aligned}\nabla_\mathrm{v} f(\mathrm{x}) &= \frac{d}{dt}\bigg|_{t=0} f(x_1 + tv_1, x_2 + tv_2, x_3 + tv_3)\\ &= \sum \partial_i f(\mathrm{x}) v_i.\end{aligned}$$

Remark 96. *You should see how this makes sense. To differentiate with respect to t chain rule of partial derivatives get that. Then see $d/dt(x_i + tv_i) = v_i$ hence the v_i term*

Noting that

$$\sum \partial_i f(\mathrm{x}) v_i = \left(\sum \partial_i f(\mathrm{x})\, \mathrm{e}_i\right) \cdot \left(\sum v_j \mathrm{e}_j\right)$$

it follows that the cartesian coordinate representation o the **gradient** of f is given by

$$\boxed{\nabla f(x) = \sum \partial_i f(x) e_i}$$

The covariant derivative and gradient of a *smooth vector field $v : U \to TU$* are defined similarly.

Definition 97 (Vector field version)
The **covariant derivative** of v at $x \in U$ along $w \in T_x U$ is defined as the vector $\nabla_w v(x) \in T_x U$ such that

$$\nabla_\mathrm{w} \mathrm{v}(\mathrm{x}) = \frac{d}{dt}\bigg|_{t=0} \mathrm{v}(\mathrm{x} + t\mathrm{w})$$

note that $v(x) \in T_x U$ whereas $v(x + tw) \in T_{x+tw} U$

Remark 98. *Using the characterization of $v(x) \in T_xU$ as the pair $(x, v(x))$ the derivative in the definition of the covariant derivative of v can be understood in the following sense:*

$$\nabla_{\mathrm{w}}\mathrm{v}(\mathrm{x}) = \lim_{t\to 0} \frac{(\mathrm{x}, \mathrm{v}(\mathrm{x}+t\mathrm{w})) - (\mathrm{x}, \mathrm{v}(\mathrm{x}))}{t} = \left(\mathrm{x}, \lim_{t\to 0} \frac{\mathrm{v}(\mathrm{x}+t\mathrm{w}) - \mathrm{v}(\mathrm{x})}{t}\right) \in T_{\mathrm{x}}U.$$

Notice how the vector $v(x + tw)$ has been parallely shifted from the tangent space $T_{x+tw}U$ to T_xU. Such a shift is possible because of the fact that the euclidean space $\mathbb{R}^3$ is flat.

Definition 99 (Vector field version)

The **gradient** ∇v of the vector field v is defined as the second order tensor field $\nabla v : U \to \otimes^2 TU$ such that at every $x \in U$

$$\nabla v(x) \cdot w = \nabla_w v(x)$$

for every $w \in T_xU$

Remark 100. *Later in advanced tensor analysis section when they talk about "contracting with" some other tensor vs with itself they refer to*

$$\nabla v(x) \cdot w \quad \textit{and} \quad \mathcal{C}_{2,3}(\nabla v(x) \otimes w)$$

respectively

Recall that the term $\nabla v(x) \cdot w$ in the definition of the gradient above stands for $\mathcal{C}_{2,3}(\nabla v(x) \otimes w)$. The cartesian coordinate representation of the covariant derivative of the vector field $v : U \to TU$ at $x \in U$ along $w \in T_xU$ is computed as follows

$$\begin{aligned}
\nabla_{\mathrm{w}}\mathrm{v}(\mathrm{x}) &= \frac{d}{dt}\bigg|_{t=0} \sum_i v_i(x_1 + tw_1, x_2 + tw_2, x_3 + tw_3)\mathrm{e}_i \\
&= \sum_{i,j} \partial_j v_i(\mathrm{x}) w_j \mathrm{e}_i \\
&= \underbrace{\left(\sum_{i,j} \partial_j v_i(\mathrm{x})\, \mathrm{e}_i \otimes \mathrm{e}_j\right)}_{\nabla v(x)} \cdot \underbrace{\left(\sum_k w_k \mathrm{e}_k\right)}_{w} \\
&= \sum_{i,j,k} \partial_j v_j(x) w_k (e_j \cdot e_k) e_i \\
&= \sum \partial_j v_j(x) w_k (\delta_{jk}) e_i
\end{aligned}$$

Note that $e_j \cdot e_k = \delta_{jk}$ because we using the *standard basis*. It follows that that the gradient of the vector field v has the followig cartesian coordinate representation

$$\boxed{\nabla \mathrm{v}(\mathrm{x}) = \sum \partial_j v_i(\mathrm{x})\, \underbrace{\mathrm{e}_i \otimes \mathrm{e}_j}_{\text{order } 2}}$$

The calculation also shows that the components of the gradient of a vector field with respect to the Cartesian coordinate system can be represented in matrix form as follows

$$[\nabla \mathrm{v}(\mathrm{x})] = \begin{bmatrix} \partial_1 v_1(\mathrm{x}) & \partial_2 v_1(\mathrm{x}) & \partial_3 v_1(\mathrm{x}) \\ \partial_1 v_2(\mathrm{x}) & \partial_2 v_2(\mathrm{x}) & \partial_3 v_2(\mathrm{x}) \\ \partial_1 v_3(\mathrm{x}) & \partial_2 v_3(\mathrm{x}) & \partial_3 v_3(\mathrm{x}) \end{bmatrix}.$$

Let us now extend our definition of covarient derivative of a vector field to that of tensor fields.

Definition 101 (Tensor field version)
Suppose that $A : U \to \otimes^k TU$ is a smooth tensor field of order k on U. The **covariant derivative** $\nabla_v A$ of A at $x \in U$ along $v \in T_x U$ is defined as the kth order tensor $\nabla_v A(x) \in \otimes^k T_x U$ such that

$$\nabla_v \mathrm{A}(\mathrm{x}) = \frac{d}{dt}\bigg|_{t=0} \mathrm{A}(\mathrm{x} + t\mathrm{v}).$$

Definition 102 (Tensor field version)
The **gradient** ∇A of the k-tensor field A is defined as the tensor field $\nabla A : U \to \otimes^{k+1} TU$ of order $k+1$ on U such that at every $x \in U$

$$\nabla \mathrm{A}(\mathrm{x}) \cdot \mathrm{v} = \nabla_\mathrm{v} \mathrm{A}(\mathrm{x})$$

for every $v \in T_x U$

As before the catesian coordinate representation of $\nabla A(x)$ can be worked out to be(just compare this with the 2nd order version above)

$$\boxed{\nabla \mathrm{A}(\mathrm{x}) = \sum \partial_j A_{i_1 \ldots i_k}(\mathrm{x})\, \underbrace{\mathrm{e}_{i_1} \otimes \ldots \mathrm{e}_{i_k} \otimes \mathrm{e}_j}_{\text{order} k+1}}$$

Remark 103. *Observe that gradient always increases the order by* 1 *while dot product which is a contraction reduces order by* 2

7.4 divergence

Definition 104
Suppose that $A : U \to \otimes^k TU$ is a smooth tensor field of order k where $k \geq 1$ over an open subset $U \subseteq \mathbb{R}^3$. The **divergence** $\nabla \cdot \mathrm{A}$ of A is defined as the smooth tensor field $\nabla \cdot A : U \to \otimes^{k-1} TU$ of order $k-1$ such that for any $x \in U$

$$\nabla \cdot \mathrm{A}(\mathrm{x}) = \mathcal{C}_{k,k+1} \nabla \mathrm{A}(\mathrm{x}).$$

Consider an order k tensor A. Note that the gradient $\nabla A : U \to \otimes^{k+1} TU$ of A is a tensor field of order $k+1$ on U whence its contraction is a tensor field of order $k-1$ as required(since divergence minus 2 order from k+1 order ∇A so we have k-1). Note also since A is a *tensor field* of order k there are k possible contractions of ∇A each of these defines a distinct notion of divergence and the convention adopted here is to call $(k, k+1)$. This chosen contraction has some purpose as you will see later on.

Remark 105. *we denote* div A *to denote the the divergence* $\nabla \cdot A$ *of the tensor field* A

Remark 106. *you should also notice that in total a divergence reduces order by 1*

Definition 107 (Divergence Cartesian)
The order k smooth tensor field A In terms of Cartesian coordinate system on $\mathbb{R}^3$ has its **divergence** $\nabla \cdot \mathrm{A}$ is given by

$$\nabla \cdot \mathrm{A}(\mathrm{x}) = \sum \partial_a A_{i_1 \dots i_{k-1} a}(\mathrm{x})\, \mathrm{e}_{i_1} \otimes \dots \otimes \mathrm{e}_{i_{k-1}}$$

which is order $k-1$

Example 108
Suppose that $v : U \to TU$ is a smooth vector field on U. The divergence of v is a scalar field on U that is easily computed as follows for any $x \in U$

$$\nabla \cdot \mathrm{v}(\mathrm{x}) = \mathcal{C}_{1,2} \nabla \mathrm{v}(\mathrm{x}) = \sum \partial_i v_i(\mathrm{x})$$

Example 109
Similarly given a smooth second order tensor field $A : U \to \otimes^2 TU$ on U ,its divergence is a vector field on U that is obtained as follows: for any $x \in U$

$$\nabla \cdot \mathrm{A}(\mathrm{x}) = \mathcal{C}_{2,3} \nabla \mathrm{A}(\mathrm{x}) = \sum \partial_j A_{ij}(\mathrm{x}) \mathrm{e}_i.$$

compare this with 99 notice it is the same just without the w_j. Well that is because the latter $\nabla A \cdot w$ while this is $\nabla \cdot A$. Both are contractions which will result on both having 1 order less. The difference is you could see $\nabla \cdot A$ as $\nabla A \cdot e_k$

7.5 curl

Definition 110
Suppose that $v : U \to \otimes^k TU$ is a smooth vector field over an open subset $U \subseteq \mathbb{R}^3$. The **curl**$\nabla \times v$ of v is defined as the smooth vector field $\nabla \times v : U \to \times TU$ such that for any $x \in U$

$$\mathrm{w} \cdot \nabla \times \mathrm{v}(\mathrm{x}) = \nabla \cdot (\mathrm{v} \times \mathrm{w})(\mathrm{x})$$

In the definition above $w : U \to TU$ is a constant vector field on U: this can be thought of as the set $\bigcup_{x \in U}(x, w)$. Since w is a constant vector field its dependence on x will be supressed.

Definition 111
The vector field $v \times w : U \to \otimes TU$ is defined as follows: for any $x \in U$,

$$v \times w(x) = v(x) \times w$$

Example 112

As a quick digression ,to recap as an exercise we first compute the *covariant derivative* of the vector field $v \times w : U \to TU$ first. So now consider any $x \in U$ and $z \in T_x U$

$$\begin{aligned}(\mathrm{w} \cdot \nabla \times \mathrm{v}(\mathrm{x})) \cdot \mathrm{z} = \nabla(\mathrm{v} \times \mathrm{w})(\mathrm{x}) \cdot \mathrm{z} &= \frac{d}{dt}\bigg|_{t=0} \mathrm{v}(\mathrm{x} + t\mathrm{z}) \times \mathrm{w} \\ &= \frac{d}{dt}\bigg|_{t=0} \sum \epsilon_{ijk} v_j(x_1 + tz_1, x_2 + tz_2, x_3 + tz_3)\, w_k \mathrm{e}_i \\ &= \underbrace{\left(\sum \epsilon_{ijk} \partial_a v_j(\mathrm{x}) w_k \mathrm{e}_i \otimes \mathrm{e}_a\right)}_{\nabla(v \times w)} \cdot \underbrace{\left(\sum z_b \mathrm{e}_b\right)}_{z}\end{aligned}$$

This result can be seen from

$$v \times w = \sum \varepsilon_{ijk} v_j(x) w_k e_i$$

and hence

$$\nabla(v \times w) = \sum \partial_a \varepsilon_{ijk} v_j(x) w_k e_i \otimes e_a$$

by definition of gradient which essentially turns k tensor to $k+1$ tensors if you recall(in contrast divergence which turn to $k-1$).

The goal of the above is for you compare with the result and realize that if you know the covariant derivative you can actually derive all the rest of the key differential quantities. In particular

$$\text{Covariant derivative} \to \text{Gradient} \to \text{Divergence} \to \text{Curl}$$

Lets demonstrate this by getting back to calculating our curl $\nabla \times v$. Instead of taking the gradient of $v \times w$ we take the divergence instead

$$\begin{aligned}\mathrm{w} \cdot \nabla \times \mathrm{v}(\mathrm{x}) = \nabla \cdot (\mathrm{v} \times \mathrm{w})(\mathrm{x}) &= \sum \epsilon_{ijk} \partial_i v_j(\mathrm{x}) w_k \\ &= \underbrace{\left(\sum_{ijk} \epsilon_{kij} \partial_i v_j(\mathrm{x})\, \mathrm{e}_k\right)}_{\nabla \times v(x)} \cdot \underbrace{\left(\sum_{l} w_l \mathrm{e}_l\right)}_{w}\end{aligned}$$

Essentially we have proven

$$\boxed{\nabla \times \mathrm{v}(\mathrm{x}) = \sum \epsilon_{ijk} \partial_j v_k(\mathrm{x})\, \mathrm{e}_i,}$$

Remark 113. *To remember this just look at how we got the first line*

$$(v \times w) = \sum \epsilon_{ijk} v_j(\mathrm{x}) w_k e_i$$

the divergence of this is clearly

$$\nabla \cdot (v \times w) = \sum \epsilon_{ijk} \partial_i v_j(\mathrm{x}) w_k$$

and then split it out from the dot product like so

$$\nabla \times v = \sum \epsilon_{ijk} \partial_j v_k(\mathrm{x}) e_i$$

Remark 114. *You should see that the curl maintains the order of the vector field*

Fact 115
to summarize for far

1. divergence: contract $(k, k+1)$ of gradient
2. curl : divergence of cross product $v \times w$ where w is some axes. And then remove "dot product dependence" of w

There is curl of tensor fields but i will skip unuless necessary.

Proposition 116
prove that

1. $\nabla \cdot \nabla \times v(x) = 0$ for any vector field v
2. $\nabla \times \nabla f(x)$ for any scalar field f

Proof. Consider that for (1) we have

$$\nabla \cdot (\sum \varepsilon_{ijk} \partial_j v_k(x) e_i) = \sum \partial_i \varepsilon_{ijk} \partial_j v_k(x) = \sum \varepsilon_{ijk} \partial_i \partial_j v_k(x) = 0$$

notice the blue terms are symmetric while the levi cevit term is anti symmetric. Therefore they will cancel out eg.

$$\partial_i \partial_j v_k - \partial_j \partial_i v_k = \partial_i \partial_j v_k - \partial_i \partial_j v_k = 0$$

consider that for (2) we have

$$\nabla \times (\sum \partial_i f(x) e_i) = \sum \varepsilon_{ijk} \partial_j \partial_k f(x) e_i = 0$$

same reason

8 curvilinear coordinates

By working with $\mathbb{R}^n$ we were able to use the **global coordinate system** defined by the identity map. However it it sometimes to work with domains that only admit a local coordinate system(local diffeomorphism - refer to `geometry of a manifold 18.905` notes for more). For example instead of $\mathbb{R}^n$ we could be working on the domain S^2, the 2D sphere etc. In such scenarios it is helpful to use other coordinate systems that allow us to simplify calculations. For example spherical coordinates in the case of the sphere are obviously easier to work with than Cartesian coordinates. So we now consider the issue of how to construct a general coordinate system

Fact 117 (Motivation: coordinate basis to construct a general coordinate system)
Consider the case of a smooth vector field $v : U \to TU$. Given any $x \in U$ the vectr $v(x) \in T_xU$ amits the following represenation with respect to the Cartesian coordinate system on $\mathbb{R}^3$:

$$\mathsf{v}(\mathsf{x}) = \sum v_i(\mathsf{x})\mathsf{e}_i.$$

The question that naturally arises is how to represent v using a **general coordinate system** on U. In particular if each tangent space T_xU is equipped with a different basis then the vector field v can be locally expressed with respect to the corresponding basis. What will become evident soon is that given a *general coordinate system* on U it is possible to construct a natural set of basis vectors for each *tangent space* called the **coordinate basis** which in turn can be used to build a systematic calculus using curivlinear coordinates.

Remark 118. *This is not the only way to define general coordinate systems. For example there is method called* **Cartan's method of moving frames***(which we will obviously not be covering here)*

That is to say whatever follows here holds in general. As long as we can find a smooth paramterization(local diffeomorphism) at a point we can establish a general coordinate system it using the coordinate basis of its tangent space. Curvillinear coordinates are just special useful applications which we will only discuss at the end.

8.1 reciprocal basis vs dual basis

Definition 119
Let $B = (b_i)_{i=1}^n$ be a basis of V. Any vector $v \in V$ admits a basis representation of the following form

$$v = \sum v^i b_i$$

Note that the components of v with respect to the basis B are represented using superscripts. The reason will be clear soon

Definition 120
The **dual space** of V written V^* is defined as the set of all linear functions defined on V:

$$V^\star = \{\omega : V \to \mathbb{R} \,|\, \omega \text{ is linear}\}.$$

Elements of V^* are called **dual vectors** or **covectors**.

Note that since V is finite dimensional the element of the dual space V^* are also continuous - this is always true if V is infinite dimensional.(Recall we have covered this in `functional anlaysis 18.103`). We see at once that for any $v \in V$ we have

$$\omega(\mathsf{v}) = \omega\left(\sum v^i \mathsf{b}_i\right) = \sum v^i \omega(\mathsf{b}_i).$$

Definition 121 (Dual Basis)
introducing the special linear maps $\beta^i \in V^*$, $i = 1, \dots, n$ as

$$\beta^i(\mathsf{v}) = v^i,$$

for any $\mathsf{v} = \sum v^i b_i \in V$ we see that

$$\omega(\mathsf{v}) = \sum \omega(\mathsf{b}_i)\beta^i(\mathsf{v}).$$

since this is true for every $v \in V$ we have

$$\omega = \sum \omega_i \beta^i,$$

where $w_i = w(b_i) \in \mathbb{R}$. It is straightforward to establish that $B^* = (\beta^i)_{i=1}^n$ is a basis of the dual space V^* and is called the **dual basis** of V^* corresponding to the basis B of V. Explicitly wee that

$$\beta^i(\mathsf{b}_j) = \delta^i_j.$$

so given an arbitary $w \in V^*$ we can express it in terms of the dual basis B^* as

$$w = \sum w_i \beta^i$$

and then its action on any $v \in V$ can see seen in

$$\begin{aligned}
\omega(\mathsf{v}) &= \sum \omega_i \beta^i \left(\sum v^j \mathsf{b}_j\right) \\
&= \sum \omega_i v^j \beta^i(\mathsf{b}_j) \\
&= \sum \omega_i v^j \delta^i_j \\
&= \sum \omega_i v^i.
\end{aligned}$$

Remark 122. *The choice of subscript and superscript notation is hcosen so that the component representation of* **natural pairings** *or* **duality pairing** *like $w(v)$ has a neat component representation involving sums that have matching subscripts and superscripts*

Remark 123. *We have proven its existence and that it is indeed a basis for the dual space before in your* `MIT 18.101` *notes*

We recall the **reciprocal basis** which is defined in an inner product space which we denote as $\hat{B} = (b^i)_{i=1}^n$ as

$$\mathsf{b}^i \cdot \mathsf{b}_j = \delta^i_j, \quad i, j = 1, \dots, n.$$

We haven't really proved its exstence rigoursly so we do so now.

Theorem 124
The reciprocal basis relative to a given basis exists and is unique for inner product space(euclidean)

Proof. Recall that a euclidean space is hilbert, that is the metric induced by the inner product is complete. Therefore consider a subspace W spanned by $\{e_2, \dots, e_n\}$(essentially the basis of V but exclude e_1). Clearly it is closed. Therefore recall from `functional analysis` this implies that $V = W \oplus W^{\perp}$ and that such a decomposition is unique. This means $\dim W^{\perp} = N - \dim W = 1$. So we may pick a non zero vector $w \in W^{\perp}$. Since $e_1 \notin W$ we have

$$e_1 \cdot w \neq 0$$

Now define

$$e^1 = \frac{1}{e_1 \cdot w} w$$

then clearly

$$e^1 \cdot e_i = \delta^1_i$$

as desired. We now repeat this process for every other e_i

Theorem 125

The reciprocal basis $\{e^1, \ldots, e^n\}$ with respect to the basis $\{e_1, \ldots, e_N\}$ of the inner product space V(euclidean) is itself a basis for V

Proof. Consider

$$\sum \lambda_i e^i = 0$$

now compute the inner product of this equation with some e_k

$$\sum \lambda_i e^i \cdot e_i = \sum \lambda_i \delta^i_k = \lambda_k = 0$$

doing this for $k = 1, 2 \ldots, n$ we clearly see the only solution for λ_i to get the trivial sum is the trivial solution so the reciprocal basis is independent. Since it also has n vectors it must also be a basis for V ☐

We now study the relationship between the *reciprocal* and *dual* basis.

Fact 126

Recall that if you consider the basis B of V and the reciprocal basis $\hat{B}$ of V and dual basis B^* of V^*. Given $\omega \in V^*$ the **Riesz representation theorem** tell us that we can find a $w \in V$ such that for any $v \in V$

$$\omega(\mathrm{v}) = \mathrm{w} \cdot \mathrm{v}.$$

as proven in your `functional analysis` notes.

Proposition 127

The identity

$$\beta^i(\mathrm{v}) = \mathrm{b}^i \cdot \mathrm{v}$$

suggests identification of the dual basis covector β^i with the reciprocal basis b^i. Show that this is well-defined

Proof. For every $\beta^i(v)$ by reisz representation theorem there exists

$$\beta^i(v) = w \cdot v$$

for some unique w for all v. Suppose that $w = b^i$. We claim that such an identification is valid. By definition of the dual basis we have to show this indentifcation satisfies

$$\beta^i(b_j) = \delta^i_j$$

and is unqiue. Indeed with $w = b^i$ we have

$$\beta^i(b_j) = b^i \cdot b_j = \delta^i_j$$

by definition of reciprocal basis and this holds for all i, j as desired. Moreover because both β^i and b^i are basis of V^* an V respectively such an indentification is unique.

Remark 128. *This is why the components of both covectors and components of the corresponding vectors with respect to the reciprocal basis are represented with subscripts(to indicate such a identification)*

8.2 coordinate basis

Definition 129

A **coordinate curve** at $x \in \mathbb{R}^3$ is a map of the form $c_i : I_\delta \to \mathbb{R}^3$ where $I_\delta = [-\delta, \delta] \subseteq \mathbb{R}$ for some $\delta > 0$ such that for any $t \in I_\delta$

$$\begin{aligned} \mathrm{c}_1(t) &= \mathrm{x}(y^1 + t, y^2, y^3), \\ \mathrm{c}_2(t) &= \mathrm{x}(y^1, y^2 + t, y^3), \\ \mathrm{c}_3(t) &= \mathrm{x}(y^1, y^2, y^3 + t), \end{aligned}$$

where $y = y(x)$. Note that $c_i(0) = x$ for $i = 1, 2, 3$.

Remark 130. *y in our context are the curvilinear coordinates while $x \in U$ where recall $U \subseteq M$ where the manifold M in our context is simply $\mathbb{R}^3$. These curves are clearly curves within the manifold on U*

Definition 131

The tangent vector $g_i(y)$ to the coordinate curve c_i at $t = 0$ is defined as the ith **coordinate tangent vector**

$$\mathrm{g}_i(\mathrm{y}) = \dot{\mathrm{c}}_i(0) = \sum \frac{\partial x^j(\mathrm{y})}{\partial y^i} \mathrm{e}_j.$$

Note that each $g_i(y) \in T_{x(y)}U$(tangent space at the point $x(y)$ on the manifold).

Theorem 132

We tangent vectors $(g_i(y))$ are linearly independent on form the **basis** of $T_{x(y)}U$. They are what we call the **coordinate basis** with respect to the **coordinate system** (U, ϕ)

Proof. Consider

$$\sum a^i \frac{\partial x^j(\mathrm{y})}{\partial y^i} \mathrm{e}_j = 0 \quad \Rightarrow \quad \sum a^i \frac{\partial x^j(\mathrm{y})}{\partial y^i} = 0, \ i = 1, 2, 3$$

Well this is because (e_j) are basis vectors which are non-zero. Now let us put this equations into matrix form. Also because we have assumed a **smooth** coordinate chart(local diffeomorphism).This means the determinant of such a matrix(known as the jacobian matrix) satisfies

$$\det \begin{bmatrix} \partial_1 x^1(\mathrm{y}) & \partial_2 x^1(\mathrm{y}) & \partial_3 x^1(\mathrm{y}) \\ \partial_1 x^2(\mathrm{y}) & \partial_2 x^2(\mathrm{y}) & \partial_3 x^2(\mathrm{y}) \\ \partial_1 x^3(\mathrm{y}) & \partial_2 x^3(\mathrm{y}) & \partial_3 x^3(\mathrm{y}) \end{bmatrix} \neq 0$$

as implied by smooth(locally). So it follows that the only solution(i.e the column vector consisting of (a_i) which this matrix multiplies such that the result of the matrix multiplication is trivial is when $a_i = 0$ for all i. This directly implies the linear independence of our defined coordinate basis as desired $\square$

Remark 133. *This is to be expected, recall this is our we defined(and proven) the basis of our tangent space too in* `18.101 Analysis on manifolds` *too!*

Given a smooth vector field $\mathrm{v} : U \to TU$ the vector $\mathrm{v}(x) \in T_pU$ can be expressed in terms of the corresponding coorindate basis $(g_i(y))$ where $y = y(x) \in V$ as follows

$$\mathrm{v}(\mathrm{x}) = \sum \hat{v}^i(\mathrm{y})\mathrm{g}_i(\mathrm{y}).$$

The change of basis rules can be used here to study how the curvilinear components $(\hat{v}^i)$ of v are related to its cartesian components (v^i) if $y = y(x)$ where $x \in U$ then

$$\sum v^i(\mathrm{x})\mathrm{e}_i = \sum \hat{v}^i(\mathrm{y})\mathrm{g}_i(\mathrm{y}) \quad \Rightarrow \quad v^i(\mathrm{x}) = \sum \frac{\partial x^i(\mathrm{y})}{\partial y^j}\hat{v}^j(\mathrm{y}).$$

Remark 134. *It is clear that in our context x^i refers to the coordinate function of $x \in U$ which are the local coordinates of U of the manifold $\mathbb{R}^3$. In contrast y^i are the coordinate functions of y of an open subset diffeomorphic to U of the manifold*

8.3 Christoffel symbols

Let us consider how our coordinate basis vectors $g_i(y)$ varies as y varies over V. So we consider

$$\mathrm{g}_i(\mathrm{y}) = \sum_k \frac{\partial x^k(\mathrm{y})}{\partial y^i}\mathrm{e}_k \quad \Rightarrow \quad \begin{aligned} \frac{\partial \mathrm{g}_i(\mathrm{y})}{\partial y^j} &= \frac{\partial}{\partial y^j}\left(\sum_\ell \frac{\partial x^\ell(\mathrm{y})}{\partial y^i}\mathrm{e}_\ell\right) \\ &= \sum_\ell \frac{\partial^2 x^l(\mathrm{y})}{\partial y^i \partial y^j}\mathrm{e}_l \\ &= \sum_{k,\ell} \frac{\partial^2 x^l(\mathrm{y})}{\partial y^i \partial y^j}\frac{\partial y^k(\mathrm{x})}{\partial x^l}\mathrm{g}_k(\mathrm{y}), \end{aligned}$$

where $x = x(y)$.

Definition 135

We introduce the **Christoffel symbols** $\Gamma^k_{ij} : V \to \mathbb{R}$ as

$$\Gamma^k_{ij}(\mathrm{y}) = \sum_\ell \frac{\partial^2 x^l(\mathrm{y})}{\partial y^i \partial y^j}\frac{\partial y^k(\mathrm{x})}{\partial x^l},$$

So using our christoffel symbols we may now express

$$\frac{\partial \mathrm{g}_i(\mathrm{y})}{\partial y^j} = \sum_k \Gamma^k_{ij}\mathrm{g}_k(\mathrm{y}).$$

The christoffel symbols thus provide a means to study how the coordinate basis vectors change when moving from one tangent space to a neighbouring one. Now dot this result by $g^k(y)$ and we see that

$$\Gamma^k_{ij}(\mathrm{y}) = \mathrm{g}^k(\mathrm{y}) \cdot \frac{\partial \mathrm{g}_i(\mathrm{y})}{\partial y^j}.$$

which clearly demonstrates in the special case of the cartesian coordinate system the christoffel symbols vanish indentically since th basis vectors don't change with respect to position. Refer to 141 for more.

8.4 metric tensor

Definition 136

The **metric tensor** $g : U \to TU \otimes TU$ where $U \subseteq \mathbb{R}^3$ is open is defined as follows: for any $x \in U$

$$\mathrm{g}(\mathrm{x}) = \sum \mathrm{e}^i \otimes \mathrm{e}_i.$$

Example 137

This is in fact the identity tensor field which in this context means

$$\mathrm{Id}(v_1, v_2) = v_1 \cdot v_2$$

Consider the basis e_i and its reciprocal basis e^j. The following are all identity tensor fields

$$\begin{aligned}
\mathrm{g}(\mathrm{x}) &= \sum \delta^i_j \mathrm{e}_i \otimes \mathrm{e}^j \\
&= \sum g_{ij}(\mathrm{y})\, \mathrm{g}^i(\mathrm{y}) \otimes \mathrm{g}^j(\mathrm{y}) \\
&= \sum g^{ij}(\mathrm{y})\, \mathrm{g}_i(\mathrm{y}) \otimes \mathrm{g}_j(\mathrm{y}) \\
&= \sum \mathrm{g}_i \otimes \mathrm{g}^i \\
&= \sum \mathrm{g}^i \otimes \mathrm{g}_i.
\end{aligned}$$

Lemma 138

$$\sum \delta^i_j (\mathrm{e}_i \cdot \mathrm{v}_1)(\mathrm{e}^j \cdot \mathrm{v}_2) = \mathrm{v}_1 \cdot \mathrm{v}_2.$$

Here, e_i and e^j are the basis and reciprocal basis vectors, respectively, and δ^i_j is the Kronecker delta.

Proof. Expanding v_1 and v_2**: Any vector v_1 can be expanded in terms of the basis vectors e_i as:

$$\mathrm{v}_1 = \sum_i (\mathrm{e}_i \cdot \mathrm{v}_1)\mathrm{e}_i.$$

Similarly, for v_2:

$$\mathrm{v}_2 = \sum_i (\mathrm{e}^i \cdot \mathrm{v}_2)\mathrm{e}_i.$$

Now substitute in and you will see why it follows. Intuitiion

$$\begin{aligned}\sum u^i w_i = \mathrm{u}\cdot\mathrm{w} = \mathrm{g}(\mathrm{x})(\mathrm{u},\mathrm{w}) &= \sum \hat{u}^i \hat{w}_i \\ &= \sum g_{ij}(\mathrm{y})\hat{u}^i\hat{w}^j \\ &= \sum g^{ij}(\mathrm{y})\hat{u}_i\hat{w}_j.\end{aligned}$$

Theorem 139

Denoting $J(y)$ to be the determinant of the matrix whose (i,j)th entry is $J_{ij}(y)$ and $g(y)$ the determinant of the matrix whose (i,j)th entry is $g_{ij}(y)$ it follows from

$$g_{ij}(\mathrm{y}) = \mathrm{g}_i(y)\cdot\mathrm{g}_j(y) = \sum \frac{\partial x_k(\mathrm{y})}{\partial y^i}\frac{\partial x_k(\mathrm{y})}{\partial y^j} = \sum J_{ki}(\mathrm{y})J_{kj}(\mathrm{y}).$$

that

$$g(\mathrm{y}) = J(\mathrm{y})^2 \quad \Rightarrow \quad J(\mathrm{y}) = \sqrt{g(\mathrm{y})}.$$

$$g^{ij}(\mathrm{y}) = \sum \frac{\partial y^i(\mathrm{x})}{\partial x^l}\frac{\partial y^j(\mathrm{x})}{\partial x^l}.$$

$$\sum g^{ij}(\mathrm{y})g_{jk}(\mathrm{y}) = \delta^i_k = \sum g_{jk}(\mathrm{y})g^{ki}(\mathrm{y}).$$

Proof. Consider that

$$g(y) = \det([(g)_{ij}]) = \det\left([(J)_{ij}]^T[(J)_{ij}]\right) = \det^2\left([(J)_{ij}]\right) = J^2(y)$$

where $[(J)_{ij}]^T[(J)_{ij}]$ since we are taking dot product, see that this matrix multiplication(row times column) directly represents $\sum J_{ki}(\mathrm{y})J_{kj}(\mathrm{y})$.Then by properties of determinant the rest of the relation follow

8.5 gradient

Definition 140 (Covariant derivative)

Let $v : U \to TU$ be a smooth vector field over an open suset $U \subseteq \mathbb{R}^3$. Recall from the earlier discussion that all key differential quantities are related to v are obtained from the covariant derivative of v.Given a coordinate system (U,ϕ) the curvilinear coordinate representation of the covariant derivative of v is obtained as follows

$$\begin{aligned}\nabla_{\mathrm{w}}\mathrm{v}(\mathrm{x}) &= \left.\frac{d}{dt}\right|_{t=0}\mathrm{v}(\mathrm{x}+t\mathrm{w}) \\ &= \left.\frac{d}{dt}\right|_{t=0}\sum_i \hat{v}^i(\mathrm{y}+t\hat{\mathrm{w}})\mathrm{g}_i(\mathrm{y}+t\hat{\mathrm{w}}) \\ &= \sum_{i,j}\left(\partial_j\hat{v}^i(\mathrm{y})\hat{w}^j\mathrm{g}_i(\mathrm{y}) + \hat{v}^i(\mathrm{y})\partial_j\mathrm{g}_i(\mathrm{y})\hat{w}^j\right) \\ &= \sum_{i,j}\left(\partial_j\hat{v}^i(\mathrm{y}) + \sum_k\Gamma^k_{ij}(\mathrm{y})\hat{v}^i(\mathrm{y})\right)\hat{w}^j\mathrm{g}_i(\mathrm{y}).\end{aligned}$$

Immediately you will notice the difference with cartesian coordinates. In particular for cartesian the red highlighted part $\mathrm{g}_i(\mathrm{y}) = 0$. See the example below for why

Example 141

Well consider this then

1. Curvilinear Basis Vectors($g_i(y)$): In a curvilinear coordinate system, the basis vectors g_1, g_2, g_3 vary with the position in the space and are defined in terms of the curvilinear coordinates (y^1, y^2, y^3). For example, in cylindrical coordinates, the basis vectors might look like this:

- $g_1 = (\cos(y^2), \sin(y^2), 0)$ (radial direction),
- $g_2 = (-\sin(y^2), \cos(y^2), 0)$ (angular direction)
- $g_3 = (0, 0, 1)$ (vertical direction).

2. Cartesian Basis Vectors(e_i): In a Cartesian coordinate system, the basis vectors are constant with respect to their coordinates y:

- $e_1 = (1, 0, 0)$
- $e_2 = (0, 1, 0)$
- $e_3 = (0, 0, 1)$

(think about/picture it, no matter the position in $\mathbb{R}^3$ you know x, y, z axes all point the same respective directions...)

Now, calculating the Partial Derivative:

1. For g_1:

$$\partial_{y^2} g_1(y) = \partial_{y^2}(\cos(y^2), \sin(y^2), 0) = (-\sin(y^2), \cos(y^2), 0).$$

This shows that g_1 changes as y^2 changes.

2. For g_2:

$$\partial_{y^2} g_2(y) = \partial_{y^2}(-\sin(y^2), \cos(y^2), 0) = (-\cos(y^2), -\sin(y^2), 0).$$

This similarly indicates that g_2 also changes with y^2. And clearly for Cartesian the the partial derivatives are all going to be zero.

So in fact we have just shown that our coordinate basis is indeed a general coordinate system(in particular we just showed that how our familiar cartesian coordinates can be derived from these equations!)

Definition 142

We introduce a special notation the denote the components of the covariant derivative of v

$$\hat{v}^i_{|j}(y) = \partial_j \hat{v}^i(y) + \sum_k \Gamma^k_{ij}(y)\hat{v}^i(y).$$

Using this, our expression for the covariant derivative can be rewritten more compactly as follows

$$\sum_i \hat{v}^i_{|j}(y)\hat{w}^j g_i(y) = \left(\sum_{i,j} \hat{v}^i_{|j}(y) g_i(y) \otimes g^j(y)\right) \cdot \left(\sum_l \hat{w}^l g_l(y)\right).$$

which implies our gradient in coordinate basis can be written as

$$\boxed{\nabla v(x) = \sum_{i,j} \hat{v}^i_{|j}(y) g_i(y) \otimes g^j(y)}$$

The foregoing calculatins can be extended to tensor fields too.

Example 143

For example in the case of the 2nd order tensor field $A : U \to \otimes^2 TU$ is consider here. The covariant derivative of A at $x \in U$ along $w \in T_xU$ is computed as follows

$$\begin{aligned}\nabla_{\mathrm{w}}\mathrm{A}(\mathrm{x}) &= \left.\frac{d}{dt}\right|_{t=0} \hat{A}^{ij}(\mathrm{y}+t\hat{\mathrm{w}})\,\mathrm{g}_i(\mathrm{y}+t\hat{\mathrm{w}})\otimes\mathrm{g}_j(\mathrm{y}+t\hat{\mathrm{w}})\\ &= \sum\left(\partial_k\hat{A}^{ij}(\mathrm{y})+\Gamma^i_{ak}(\mathrm{y})\hat{A}^{aj}(\mathrm{y})+\Gamma^j_{bk}(\mathrm{y})\hat{A}^{ib}(\mathrm{y})\right)\hat{w}^k\,\mathrm{g}_i(\mathrm{y})\otimes\mathrm{g}_j(\mathrm{y}).\end{aligned}$$

Note that the product rule for differentiation extends to tensor products too. By definition $A(x)\otimes B(x) = A(x)B(x)$ so clearly it holds. As before we introduce the special notation for components of the covariant derivative of A:

$$\hat{A}^{ij}_{|k}(\mathrm{y}) = \sum\left(\partial_k\hat{A}^{ij}(\mathrm{y})+\Gamma^i_{ak}(\mathrm{y})\hat{A}^{aj}(\mathrm{y})+\Gamma^j_{bk}(\mathrm{y})\hat{A}^{ib}(\mathrm{y})\right).$$

The gradient of A can be easily seem from the above calculation as

$$\nabla\mathrm{A}(\mathrm{x}) = \sum\hat{A}^{ij}_{|k}(\mathrm{y})\mathrm{g}_i(\mathrm{y})\otimes\mathrm{g}_j(\mathrm{y})\otimes\mathrm{g}^k(\mathrm{y}).$$

As an imediate application let us compute the covariant derivative of the metric tensor $g : U \to \otimes^2 TU$. It follows from the definition of covariant derivative thtaa since $g(x) = \sum e^i \otimes e_i$ does not vary as x varies over U,

$$\nabla g(x) = 0$$

8.6 sperical coordinate system

Coord

$$\begin{aligned}\mathrm{g}_1(\mathrm{y}) &= \cos y^2\mathrm{e}_1+\sin y^2\mathrm{e}_2 = \mathrm{e}_r(r,\theta,z),\\ \mathrm{g}_2(\mathrm{y}) &= -y^1\sin y^2\mathrm{e}_1+y^1\cos y^2\mathrm{e}_2 = r\mathrm{e}_\theta(r,\theta,z),\\ \mathrm{g}_3(\mathrm{y}) &= \mathrm{e}_3 = \mathrm{e}_z(r,\theta,z).\end{aligned}$$

then

$$\begin{aligned}\nabla\mathrm{v} = &\frac{\partial v_r}{\partial r}\mathrm{e}_r\otimes\mathrm{e}_r+\left(\frac{1}{r}\frac{\partial v_r}{\partial\theta}-\frac{v_\theta}{r}\right)\mathrm{e}_r\otimes\mathrm{e}_\theta+\frac{\partial v_r}{\partial z}\mathrm{e}_r\otimes\mathrm{e}_z\\ &+\frac{\partial v_\theta}{\partial r}\mathrm{e}_\theta\otimes\mathrm{e}_r+\left(\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{v_r}{r}\right)\mathrm{e}_\theta\otimes\mathrm{e}_\theta+\frac{\partial v_\theta}{\partial z}\mathrm{e}_\theta\otimes\mathrm{e}_z\\ &+\frac{\partial v_z}{\partial r}\mathrm{e}_z\otimes\mathrm{e}_r+\frac{1}{r}\frac{\partial v_z}{\partial\theta}\mathrm{e}_z\otimes\mathrm{e}_\theta+\frac{\partial v_z}{\partial z}\mathrm{e}_z\otimes\mathrm{e}_z.\end{aligned}$$

so we have the divergence

$$\nabla\cdot\mathrm{v} = \frac{1}{r}\frac{\partial}{\partial r}(rv_r)+\frac{1}{r}\frac{\partial v_\theta}{\partial\theta}+\frac{\partial v_z}{\partial z}.$$

and the curl

$$\nabla\times\mathrm{v} = \left(\frac{1}{r}\frac{\partial v_z}{\partial\theta}-\frac{\partial v_\theta}{\partial z}\right)\mathrm{e}_r+\left(\frac{\partial v_r}{\partial z}-\frac{\partial v_z}{\partial r}\right)\mathrm{e}_\theta+\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_\theta)-\frac{\partial v_r}{\partial\theta}\right)\mathrm{e}_z.$$

9 advanced tensor analysis

9.1 intro

Example 144

In Cartesian coordinates the velocity component of the vector is easily given by

$$\vec{V} = (\dot{x}, \dot{y}, \dot{z}) = \dot{x}\hat{e}_x + \dot{y}\hat{e}_y + \dot{z}\hat{e}_z$$

However if we wish to do the same in spherical polar coordinates like for example

$$\vec{V} = (\dot{r}, \dot{\theta}, \dot{\varphi}) = \dot{r}\hat{e}_r + \dot{\theta}\hat{e}_\theta, \dot{\varphi}\hat{e}_\varphi$$

this might not seem correct. First of all they basis vectors are not even the same anymore.

Now suppose

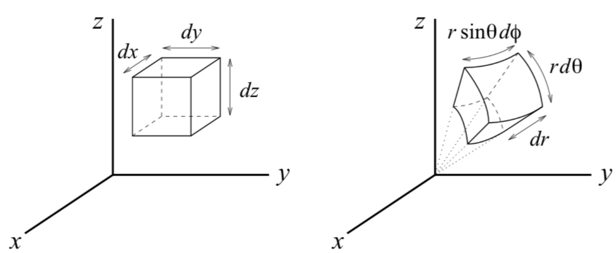

Figure 2: Differential volume in cartesian(left) and spherical(right)

Now suppose we accounted by this by considering the "differential volume" in which case we get

$$\vec{V} = (\dot{r}, r\dot{\theta}, r\sin\theta\dot{\varphi}) = \dot{r}\hat{e}_r + r\dot{\theta}\hat{e}_\theta, r\sin\theta\dot{\varphi}\hat{e}_\varphi$$

But clearly such a mode a definition is very ambiguous. In what follows we will learn that the velocity vector we got above via the "differential volume" method is what is known as the **physical form**. In **tensor calculus** we learn there are 2 other forms known as the **covariant,contrvariant** or **mixed**.

Remark 145. *It turns out that this vector*

$$\vec{V} = (\dot{r}, \dot{\theta}, \dot{\varphi}) = \dot{r}\hat{e}_r + \dot{\theta}\hat{e}_\theta, \dot{\varphi}\hat{e}_\varphi$$

is covariant while the physical form is neither covariant or contravairant

Definition 146

The components of a **covariant vector** transform like a gradient and obey the transformation law

$$\tilde{A}_i = \sum_{j=1}^{n} \frac{\partial x^j}{\partial \tilde{x}^i} A_j$$

Definition 147

The components of a **contravariant vector** transform like coordinate differential and obey the transformation law

$$\tilde{A}^i = \sum_{j=1}^{n} \frac{\partial \tilde{x}^i}{\partial x^j} A^j$$

We will use **einstein summation convention** in what follows

Example 148

The rank of a tensor refers to the number of indices (or subscripts/superscripts) required to uniquely specify a component of the tensor. For example the scalar is rank 0. The vector is rank 1. The matrix is rank 2. The cube matrix is rank 3 etc. Essentially you can see it as a "dimension"

Consider the following general tensor transformation below

Definition 149

A tensor of **dimension** m(each index varies from 1 to m) and **rank** n(number of indices) is an entity that under arbitrary coordinate transformation tansforms as

$$\tilde{T}^{k_1 \dots k_q}_{i_1 \dots i_p} = \underbrace{\frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}}}_{\text{covariant}} \underbrace{\frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}}}_{\text{contravariant}} T^{l_1 \dots l_q}_{j_1 \dots j_p}$$

where $p + q = n$(look at each tensor) and the indices $i_1, \dots, i_p$ and $j_1, \dots, j_p$ are the **covariant indices** and $k_1, \dots, k_q$ and $l_1, \dots, l_q$ are the **contravariant indices**. Inidices that apear just once in a term such as ($i_1, \dots, i_p$ and $k_1, \dots, k_q$) are called **free indices** whle indices appearing twice - one covariant and one contravariant such as ($j_1, \dots, j_q$ and $l_1, \dots, l_q$) are called **dummy indices**

Remark 150. *Notice in all cases the contravariant indices are found in subscripts on T or $\tilde{T}$ while the coariant indices are found in the superscripts. The free indices are not reflected in T and opposite for $\tilde{T}$*

For a tensor relationship to be valid , each term whether on the left or right side of the equation must have the same free indices each in the same position. For example if a certain free index is covariant then it must be covariant in all terms and vice versa.

Fact 151

If $q = 0$ then the tensor is said to **covariant** while if $p = 0$ the tensor is said to be **contravariant**

Definition 152

A rank 2 **dyad**, D results from taking the **dyadic product** of two vectors(rank 1 tensors) $\overrightarrow{A}$ and $\overrightarrow{B}$ as follows

$$D_{ij} = A_i B_j, \quad D_i^j = A_i B^j, \quad D_j^i = A^i B_j, \quad D^{ij} = A^i B^j$$

Theorem 153

The sum or difference of two like-tensors is a tensor of the same type

Proof. Let $R = S + T$ then

$$
\begin{aligned}
\tilde{R}^{k_1\dots k_q}_{i_1\dots i_p} &= \tilde{S}^{k_1\dots k_q}_{i_1\dots i_p} + \tilde{T}^{k_1\dots k_q}_{i_1\dots i_p} \\
&= \underbrace{\frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}}}_{\text{covariant}} \underbrace{\frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}}}_{\text{contravariant}} S^{l_1\dots l_q}_{j_1\dots j_p} \\
&+ \underbrace{\frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}}}_{\text{covariant}} \underbrace{\frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}}}_{\text{contravariant}} T^{l_1\dots l_q}_{j_1\dots j_p} \\
&= \underbrace{\frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}}}_{\text{covariant}} \underbrace{\frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}}}_{\text{contravariant}} R^{l_1\dots l_q}_{j_1\dots j_p}
\end{aligned}
$$

Definition 154

A **tensor contraction** occurs when one of a tensor's **free covariant indices** is set equal to one of its **free contravariant indices**. In this case a sum if perfromed on the now repeated indices and the result is a tensor with two fewer *free indices*

Example 155

For example if T^j_{ij} is a contraction on the 2nd and 3rd indices of the rank 3 tensor T^k_{ij} we would write

$$T^j_{ij} = T_i$$

And we now have a rank 1 tensor T_i

Theorem 156

A contraction of a rank 2 tensor(its **trace**) is a scalar whose value is independent of the coordinate system chosen. Such a scalar is referred to as a rank 0 tensor

Proof. Let $T = T^i_i$ be the trace of the tensor T. If $\tilde{T}^l_k$ is a tensor in a coordinate system $\tilde{x}^k$ then its trace transforms to coordinate system x^i accoridng to

$$\tilde{T}^k_k = \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^k}{\partial x^j} T^j_i = \delta^i_j T^j_i = T^i_i = T$$

9.2 the metric

Definition 157
In an arbitary m-dimensional coordinate system, x^i the differential displacement vector is

$$d\overrightarrow{r} = (h_{(1)}dx^1, \ldots, h_{(m)}dx^m) = \sum_{i=1}^{m} h_{(i)}dx^i\hat{e}_{(i)}$$

The lenght of the vector $d\overrightarrow{r}$ is given by the dot product of $d\overrightarrow{r}$ with itself(recall differential geometry unit arc lenghth)

$$(dr^2) = \sum_i \sum_j h_{(i)}h_{(j)}\hat{e}_{(i)} \cdot \hat{e}_{(j)}dx^i dx^j$$

where $\hat{e}_{(i)}$ are they physical(not covariant) *unit* vectors and $h_{(i)} = h_{(i)}(x^1, \ldots, x^m)$ are the scale factors which give each omponent the apporopriate units of lenght

That is to say $e_i \neq e_{(i)}$

Definition 158
The **metric** g_{ij} is given by

$$g_{ij} = h_{(i)}h_{(j)}\hat{e}_{(i)} \cdot \hat{e}_{(j)}$$

which by definition of dot prouct is symmetric unde the interchange of its indices : $g_{ij} = g_{ji}$

So we rewrite 157 as

$$(dr)^2 = g_{ij}dx^i dx^j$$

Theorem 159
The metric is a rank 2 covariant tensor

Proof. Because $(dr)^2$ is a distance between two physical points it mut be invariant(hence the 1st equality) under coordinate transformations. Thus consider $(dr)^2$ in the coordinate system $\tilde{x}^k$ and x^i:

$$(dr)^2 = \tilde{g}_{kl}d\tilde{x}^k\tilde{x}^l = g_{ij}dx^i dx^j = g_{ij}\frac{\partial x^i}{\partial \tilde{x}^k}d\tilde{x}^k\frac{\partial x^j}{\partial \tilde{x}^l}d\tilde{x}^l = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}g_{ij}d\tilde{x}^k d\tilde{x}^l \quad \Rightarrow \left(\tilde{g}_{kl} - \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}g_{ij}\right)d\tilde{x}^k d\tilde{x}^l = 0$$

which must be true $\forall d\tilde{x}^k d\tilde{x}^l$. This can only be true if and only if

$$\tilde{g}_{kl} = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}g_{ij}$$

Definition 160
The **conjugate metric** g^{kl} is the inverse to the metric tensor and therefore satisfies:

$$g^{kp}g_{ip} = g_{ip}g^{kp} = \delta_i^k$$

Well obviously think about it from a matrix point of a view. If inverse exists then left and right inverses must exist. Also it is easy to prove that g^{ij} is rank 2 **contravariant** tensor

Definition 161

A **conjugate tensor** is the result of multiplying a tensor with the metric then contracting one of the indices of the metric with one of the indices of the tensor

Example 162

Consider

$$T^{k\quad\ \ j_1\ldots j_q}_{i_1\ldots i_{r-1}\,i_{r+1}\ldots i_p} = g^{ki_r}T^{j_1\ldots j_q}_{i_1\ldots i_p}$$

This is known as **raising an index**. See on the RHS that i_r on g is equal to one of the contravariant indices on T. So upon multiplication to get the LHS where i_r has disappeared altogether. The remaining k in the superscript was contributed by g.

$$T_{k\quad\ \ j_1\ldots j_q}^{i_1\ldots i_{r-1}\,i_{r+1}\ldots i_p} = g_{ki_r}T^{i_1\ldots i_p}_{j_1\ldots j_q}$$

The same logic applies here. This is known as **lowering an index**

This is expressed below respectively

$$A^j = g^{ij}A_i \quad A_i = g_{ij}A^j$$

Example 163

Suppose we have some random coordinate system x^i and we know that it is represented in cartesian coordinates by $\chi^k(x^i)$ then we may write

$$(dr)^2 = \delta_{kl}d\chi^k d\chi^l = \delta_{kl}\left(\frac{\partial\chi^k}{\partial x^i}dx^i\right)\left(\frac{\partial\chi^l}{\partial x^j}dx^j\right) = \left(\delta_{kl}\frac{\partial\chi^k}{\partial x^i}\frac{\partial\chi^l}{\partial x^j}\right)dx^i dx^j$$

which then implies

$$g_{ij} = \delta_{kl}\frac{\partial\chi^k}{\partial x^i}\frac{\partial\chi^l}{\partial x^j}$$

and its inverse is given by

$$g^{ij} = \delta_{kl}\frac{\partial x^i}{\partial\chi^k}\frac{\partial x^j}{\partial\chi^l}$$

Definition 164

We write for any vector $\overrightarrow{A}$:

$$\overrightarrow{A} = \sum_{i=1}^{m} h_{(i)}A^i\hat{e}_{(i)} = \sum_{i=1}^{m} A_{(i)}\hat{e}_{(i)}$$

where we have defined

$$A_{(i)} = h_{(i)}A^i \qquad A^i = \frac{1}{h_{(i)}}A_{(i)} \tag{11}$$

We essentially we have supposed there exist a covariant form of this vector

You will see why soon. Essentially the scale factor the amount to scale the unit physical basis such that it becomes a covariant basis. Then due to in-variance of representation it follows that the physical component must be inversely scaled to get the contra-variant component

Remark 165. *As a reminder we put brackets (i) in the subscript to indicate there is no sum over i with einstein summation convention because it is not a covariant vector. Rather it is just a physical component.* $e_i \neq e_{(i)}$

Remark 166. *A physical of a vector field has the same units of the field. For example a physical component of velocity has units ms^{-1}, for force its N, for electric field is Vm^{-1} etc*

Substituting $A^i = g^{ij}A_j$ into the left of 11 and multiplying the other by g_{ij} we get a relationship between physical and covariant components

$$A_{(i)} = h_{(i)}g^{ij}A_j; \quad A_j = \sum_i \frac{g_{ij}}{h_{(i)}}A_{(i)}$$

Using the above we write down the relationships between physical components of higher rank tensors and their contravariant mixed and covariant foms. For rank 2 tensors these are

$$T_{(ij)} = h_{(i)}h_{(j)}T^{ij} = h_{(i)}h_{(j)}g^{ik}T^j_k \dots$$

Definition 167

Let $\vec{r}_x$ be a displacement vector whose componenets are express in terms of the coordinate system x^i then the covariant basis vector e_i is defined to be

$$e_i = \frac{d\vec{r}}{dx^i}$$

It is easy to see that e_i is a covariant vector by considering its transfrmation to a new coordinte system $\tilde{x}^j$

$$\tilde{e}_j = \frac{d\vec{r}}{d\tilde{x}_j} = \frac{\partial x^i}{\partial \tilde{x}^j}\frac{d\vec{r}}{dx^i} = \frac{\partial x^i}{\partial \tilde{x}^j}e_i$$

Note that i dropped the x in $\vec{r_x}$ because it should be representation invariant.

Remark 168. *This is contrast to the physical basis $\hat{e}_{(i)}$ which is the normalized version of e_i. This is because $h_{(i)}$ depends on the system it doesn't transform directly like above*

Now recalling our relationship between covariant components and physical components we defined above in 164

$$e_i = h_{(i)}\hat{e}_{(i)} \quad \hat{e}_{(i)} = \frac{1}{h_{(i)}}e_i$$

Once again we can then get the relationships similar to previously by multiplying the left of the above equation with g^{ij} and replacing e_i in the second with $g_{ij}e^j$ where we get

$$g^{ij}e_i = e^j = \sum_i g^{ij}h_{(i)}\hat{e}_{(i)} \qquad \hat{e}_{(i)} = \frac{g_{ij}}{h_{(i)}}e^j$$

Theorem 169

Regardless of whether the coordinate sysetm is orthogonal $e_i \cdot e^j = \delta^j_i$

Proof. Consider

$$e_i \cdot e^j = h_{(i)}\hat{e}_{(i)} \cdot \sum_k g^{kj}h_{(k)}\hat{e}_{(k)} = \sum_k g^{kj}h_{(i)}h_{(k)}\hat{e}_{(i)}\hat{e}_{(k)} = g^{kj}g_{ik} = \delta^i_j$$

Corollary 170

$$e_i \cdot e_J = g_{ij} \quad e^i \cdot e^j = g^{ij}$$

Proof. Simply compare our results above with 158

Now here comes the power of using covariant basis vectors

Example 171

Recall that we had expressed our vector in physical basis via

$$\vec{A} = \sum_i A_{(i)} \hat{e}_{(i)}$$

we now know this is equal to

$$\vec{A} = \sum_i h_{(i)} A^i \frac{1}{h_{(i)}} e_i = A^i e_i$$

Similarly we could have substituted to express in terms of contra-variant basis vectors instead via

$$\vec{A} = \sum_i h_{(i)} g^{ij} A_j \frac{g_{ik}}{h_{(i)}} e^k = g^{ij} g_{ik} A_j e^k = \delta^j_k A_j e^k = A_j e^j$$

Now notice we can always require the covariant component from the contra variant representation via

$$\hat{A} \cdot e_i = A_j e^j \cdot e_i = A_i$$

just like you would do if you had an orthonormal basis!

9.3 scalar and inner products

Definition 172

The **covariant scalar product** and **contravariant scalar product** of two rank 1 tensor A and B are defined as

$$g^{ij} A_i B_j$$

and

$$g_{ij} A^i B^j$$

respectively

In fact we could prove in this case the two products are actually the same value. Simply consider

$$g^{ij} A_i B_j = A^j B_j \qquad g_{ij} A^i B^j = A^i B_i$$

However such a product is *not invariant* simply consider

$$\sum_k \tilde{A}_k \tilde{B}_k = \sum_k \frac{\partial x^i}{\partial \tilde{x}^k} A_i \frac{\partial x^j}{\partial \tilde{x}^k} B_j = \sum_k \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial x^j}{\partial \tilde{x}^k} A_i B_j \neq \sum_i A_i B_i$$

Definition 173

The **covariant** and **contravariant** scalar products of two rank 2 tensors S and T are defined as $g^{ik} g^{jl} S_{kl} T_{ij}$ and $g_{ik} g_{jl} S^{kl} T^{ij}$ respectively

Similary the to the 1 rank case the result is a scalar and the same.

What about the case of unequal tensor ranks?

Example 174
Let us now contract a rank 2 tensor T with a rank 1 tensor A. In tensor notation there are sixteen ways to do this...but we can arrange them into 2 groups:

$$T^{ij}A_j = T^i_j A^j = g^{ik}T^j_k A_j = g^{ik}T_{kj}A^j = (T \cdot A)^i$$

$$A_i T^{ij} = A^i T^j_i = g^{ik}A_i T^i_k j = g^{ik}A^i T_{ik} = (A \cdot T)^j$$

In both cases the end result is a rank 1 tensor. This is in fact an example of **contravariant inner product** between tensor of rank 1 and 2

Definition 175
The **inner product** between two tensors of any rank is the contraction of the **inner indices** namely the last index of the first tensor and the first index of the last tensor

Example 176 (1. Contracting Over One Index)
When contracting over one index, we sum over a single pair of matching indices from the two tensors. The resulting tensor's rank is $3 + 5 - 2 = 6$.

$$C_{ijlkm} = \sum_p A_{ijp} B_{plkm}.$$

- A_{ijp}: Rank-3 tensor.
- B_{plkm}: Rank-5 tensor.
- Result: C_{ijlkm}, a rank-6 tensor.

Example 177 (2. Contracting Over Two Indices)
When contracting over two indices, we sum over two pairs of matching indices from the two tensors. The resulting tensor's rank is $3 + 5 - 4 = 4$.

$$C_{iklm} = \sum_{j,p} A_{ijp} B_{pjklm}.$$

- A_{ijp}: Rank-3 tensor.
- B_{pjklm}: Rank-5 tensor.
- Result: C_{iklm}, a rank-4 tensor.

Example 178 (3. Contracting Over Three Indices)
When contracting over all three indices of the rank-3 tensor with three matching indices of the rank-5 tensor, the result is a rank-2 tensor. The resulting tensor's rank is $3 + 5 - 6 = 2$.

$$C_{lm} = \sum_{i,j,k} A_{ijk} B_{ijklm}.$$

- A_{ijk}: Rank-3 tensor.
- B_{ijklm}: Rank-5 tensor.
- Result: C_{lm}, a rank-2 tensor.

9.4 invariance of tensor product

it was briefly mentioned in the definition of tensors that it has to invariant under coordinate transformations. That is , in terms of a new basis the the tensor must still have the same value. Now

Example 179

$$U^k_{ij} = V_i$$

is invalid since each term does not have the same number of indices. To rememdy this we can contract the it on two its indices like so

$$U^j_{ij} = V_i$$

Hence a general rule of thumb to ensure invariance is as follows

Theorem 180 (Quotient Rule)
If A and B are tensors and if the expression $A = BT$ is invariant under coordinate transformation then T is a tensor

Proof. The proof for general rank is more cumbersome and no more enlighted so we show a representative case involving 2 rank 1 tensors A and B

$$\tilde{B}_l \tilde{T}^l_k = \tilde{A}_k = \frac{\partial x^i}{\partial \tilde{x}^k} A_i = \frac{\partial x^i}{\partial \tilde{x}^k} B_j T^j_i = \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^l}{\partial x^j} \tilde{B}_l T^j_i$$

This implies

$$\tilde{B}_l \left(\tilde{T}^l_k - \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^l}{\partial x^j} T^j_i \right) = 0$$

This is possible if and only if the parenthissis is zero and hence T^j_i transforms as a rank 2 mixed tensors.

Fact 181
By now you should have observed the general method to investigate the tensor properties of some relation. We simply need to transform its components into the desired set of coordinates and then factorize out the uncessary terms(B for the above example) to analyze our desired(T for the above example). Its a very natural method just by directly obseving the definition of tensors

9.5 tensor derivatives

Notice that for a vector $\overrightarrow{A}$ we have

$$\frac{\partial \overrightarrow{A}}{\partial x^j} = \frac{\partial A^i e_i}{\partial x^j} = \frac{\partial A^i}{\partial x^j} e_i + \frac{\partial e_i}{\partial x^j} A^i$$

Which is basically product rule. If e_i is one of the m covariant basis vectors spanning an m-dimensional space then since $\frac{\partial e_i}{\partial x^j}$ must clearly be a vector within that same m-dimensional space and therefore can be expressed as a linear combination of the m basis vectors. This motivates the definition below

Definition 182

The **christoffel symbols of the second kind** Γ^k_{ij} are the components of the vector $\frac{\partial e_i}{\partial x^j}$ relative to the basis e_k. Thus

$$\Gamma^k_{ij} e_k$$

To get an expression for Γ^l_{ij} by itself we take the dot product with e^l to get

$$\frac{\partial e_i}{\partial x^j} \cdot e^l = \Gamma^k_{ij} e_k \cdot e^l = \Gamma^l_{ij}$$

Theorem 183

The **christoffel symbol of the second kind** is symmetric in its lower indices

Proof. Recall

$$\frac{\partial e_i}{\partial x^j} \cdot e^l = \Gamma^l_{ij} = \frac{\partial}{\partial x_j} \frac{\partial \overrightarrow{r}_x}{\partial x^i} \cdot e^l = \frac{\partial}{\partial x_i} \frac{\partial \overrightarrow{r}_x}{\partial x^j} \cdot e^l = \frac{\partial e_j}{\partial x^i} \cdot e^l = \Gamma^l_{ji}$$

Definition 184

The **christoffel symbol of the first kind** Γ_{ijk} are given by

$$\Gamma_{ijk} = g_{lk} \Gamma^l_{ij} \quad \Gamma^l_{ij} = g^{lk} \Gamma_{ijk}$$

It is easy to show given above that it is symmetric in the first 2 indices Using the above 2 equations we can write

$$\Gamma_{ijk} = g_{lk} \Gamma^l_{ij} = g_{lk} e^l \cdot \frac{\partial e_i}{\partial x^j} = e_k \cdot \frac{\partial e_i}{\partial x^j}$$

Notice in the 3rd equality we have contracted g_{lk} with e^l to get the composite vector e_k. Now recall that $g_{ij} = e_i \cdot e_j$ and thus

$$\frac{\partial g_{ij}}{\partial x^k} = e_i \cdot \frac{\partial e_j}{\partial x^k} + e_j \cdot \frac{\partial e_i}{\partial x^k} = \Gamma_{jki} + \Gamma_{ikj}$$

where recall product rule on dot products.

Fact 185
It turns out there is a useful identity that we can get to simplify future calculations by exploiting the symmetry of its indices.
Let us permute the indices above twice getting

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma_{ijk} + \Gamma_{kji} \quad \text{and} \quad \frac{\partial g_{jk}}{\partial x^i} = \Gamma_{kij} + \Gamma_{jik}$$

By combining the above equations and solving simulataneously it is possible to derive

$$\boxed{\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right) \Rightarrow \quad \Gamma^l_{ij} = \frac{g^{lk}}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right)}$$

Fact 186
For a general 3-D coordinate system there are 27 Chirstoffel symbols of eachkind(15 of them independent - accounting for symmetry in the bottom symbols) each with as many as nin terms. However for orthogonal coordinates where $g_{ij} \propto \delta_{ij}$ and $g_{ii} = h^2_{(i)} = \frac{1}{g^{ii}}$ things get much simpler. In that case the christoffel symbols fall into three categories as follows(without summation convention)

Proposition 187
christoffel sybmols are not tensors.

Proof. To determine how christoffel symbols transform under cooordinate transformations from x^i to $\tilde{x}^p$(that's right we are gonna investigate its tensor properties) Consider

$$\frac{\partial \tilde{g}_{pq}}{\partial \tilde{x}^r} = \frac{\partial}{\partial \tilde{x}^r}\left(g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\right) = \frac{\partial g_{ij}}{\partial x^k}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^r \partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial^2 x^j}{\partial \tilde{x}^r \partial \tilde{x}^q}$$

Recognize the term in the parenthesis of the 1st equality is simply the metric transformation(recall g_{ij} is rank 2 covariant tensor). The 2nd equality is just product rule on a product of 3 terms. The purpose of this? Well we want to make use 185(again as a foreshadow, this relation will prove useful later) that's why we didn't attempt to analyze using the definition directly(well not that it reveals anything...dot product with a partial derivative huh??). Well in order to use our relation clearly we have to permute our indices accordingly in which case we get

$$\frac{\partial \tilde{g}_{qr}}{\partial \tilde{x}^p} = \frac{\partial g_{jk}}{\partial x^i}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{jk}\frac{\partial^2 x^j}{\partial \tilde{x}^p \partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{jk}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial^2 x^k}{\partial \tilde{x}^p \partial \tilde{x}^r}$$

and

$$\frac{\partial \tilde{g}_{pq}}{\partial \tilde{x}^r} = \frac{\partial g_{ij}}{\partial x^k}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^r \partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial^2 x^j}{\partial \tilde{x}^r \partial \tilde{x}^q}$$

and blah blah blah noting that the metric is symmetirc $g_{ij} = g_{ji}$ the whole thing simplifies to

$$\tilde{T}_{pqr}\Gamma_{ijk}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^p}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{ij}\frac{\partial x^j}{\partial \tilde{x}^r}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^q}$$

Oh shit, this is clearly not a well defined tensor. The indices on LHS and RHS aren't equal due to that 2nd derivative on the LHS. Similarly it can also be shown the same can be said for **Christoffel symbols** of the second kind. ☐

So whats the point of them?To repeat yet again, we will see soon as foreshadowed. In fact because of their utility later we will allow christofeel symbols to participate in einstein summation, which is why we did not enclose them in brackets like we did for physical components.

9.6 covariant derivative

With our newfound relations using christoffel symbols we substitute it into

$$\frac{\partial \overrightarrow{A}}{\partial x^j} = \frac{\partial A^i e_i}{\partial x^j} = \frac{\partial A^i}{\partial x^j} e_i + \frac{\partial e_i}{\partial x^j} A^i$$

from earlier to obtain

$$\frac{\partial \overrightarrow{A}}{\partial x^j} = \frac{\partial A^i}{\partial x^j} e_i + \Gamma^k_{ij} e_k A^i = \left(\frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk} A^k \right) e_i$$

This again motivates another definition below

Definition 188

The **covariant derivative** of a **contravariant** vector A^i is given by

$$\nabla_j A^i \equiv \frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk} A^k$$

So we may now write our equation for $\frac{\partial \overrightarrow{A}}{\partial x^j}$ more compactly by

$$\frac{\partial \overrightarrow{A}}{\partial x^j} = \nabla_j A^i$$

Proposition 189

The covariant derivative of a contravariant vector is a mixed rank 2 tensor

Proof. Like how you normally determine a tensor nature we consider the transformation of the *convariant derivative of a contravariant vector* from the coordinate system x^i to $\tilde{x}^p$.

$$\begin{aligned} \tilde{\nabla}_q \tilde{A}^p &= \frac{\partial \tilde{A}^p}{\partial \tilde{x}^q} + \tilde{\Gamma}^p_{qr} \tilde{A}^r \\ &= \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial A^i}{\partial x^j} \dots \end{aligned}$$

Definition 190

The **covarient derivative** of a **covariant vector** A_i is given by

$$\nabla_j A_i \equiv \frac{\partial A_i}{\partial x^j} - \Gamma^k_{ij} A_k$$

where we define

$$\frac{\partial e^i}{\partial x^j} = -\Gamma^i_{jk} e^k$$

Again reminder that we allowed christoffel symbols to participate in einstein summation despite them not begin tensors.

In the same way we proved that $\nabla_j A^i$ is a rank 2 mixed tensor in 189 one can show that $\nabla_j A_i$ is *rank 2 covariant tensor*

Theorem 191

The covariant derivatives of the contravariant and covariant basis vectors are zero

Proof. Recall 188. We then get

$$\nabla_j e^i = \partial_j e^i + \Gamma^i_{jk} e^k$$

and

$$\nabla_j e_i = \partial_j e_i - \Gamma^i_{jk} e^k$$

It is clear to see by definition they are both zero.

Remark 192. *This is why we defined our christoffel symbols(recall they are just coefficients that the derivative of a contra/covariant basis vector can take because being in the same space they must be spanned by the same basis vectors) like above. It is explicitly so that we can get such a result and following results like Ricci theorem(see below). In fact if you study Riemann geometry you will know that such a choice was deliberate, so as to define parallel transport equations*

Theorem 193 (Ricci)

The covariant derivative of the metric and its inverse vanish

9.7 Connexion to vector calculus

Definition 194

Consider the contraction of the covariant derivative of a contravariant vector

$$\nabla_i A^i = \partial_i A^i + \Gamma^i_{ij} A^j$$

Then using 185 we have

$$\begin{aligned}
\Gamma^i_{ij} &= \frac{1}{2} g^{ki}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \\
&= \frac{1}{2} g^{ki}\partial_i g_{jk} + g^{ki}\partial_j g_{ki} - g^{ik}\partial_i g_{kj} (\text{swap dummy indices}) \\
&= \frac{1}{2} g^{ki}\partial_i g_{jk} + g^{ki}\partial_j g_{ki} - g^{ik}\partial_i g_{jk} (\text{symmetry of metric dummy in 3rd term}) \\
&= \frac{1}{2} g^{ki}\partial_i g_{jk}
\end{aligned}$$

Theorem 195 (jacobi formula)

Let Q_{ij} be a rank 2 tensor of dimension m(and thus represented by an $m \times m$ matrix) and let $Q = \det(Q_{ij})$ be its determinant. Let Q^{ji} be the cofactor of Q_{ij} and thus the matrix (Q^{ji}) is the transpose of the cofactors of Q_{ij}. Then

$$\partial_k Q = Q^{ji} \partial_k Q_{ij}$$

Proof. We claim that Q can be thought of as a function of the matrix elements Q_{ij} so we have

$$\partial_k Q = \frac{\partial Q}{\partial Q_{ij}} \partial_k Q_{ij}$$

We notice that

$$Q = \sum_{k=1}^{m} Q_{ik} Q^{ki} \quad \text{for any } i = 1, \dots, m \text{ no sum on } i$$

which means

$$\frac{\partial Q}{\partial Q_{ij}} = \frac{\partial}{\partial Q_{ij}} \sum_{k=1}^{m} Q_{ik} Q^{ki} = \sum_{k=1}^{m} \left(\frac{\partial Q_{ik}}{\partial Q_{ij}} Q^{ki} + Q_{ik} \frac{\partial Q^{ki}}{\partial Q_{ij}} \right)$$

Now $\frac{\partial Q_{ik}}{\partial Q_{ij}} = \delta_k^j$ since the matrix elements are independent of each other and that $\frac{\partial Q^{ki}}{\partial Q_{ij}} = 0$ since the cofactor Q^{ki} includes all matrix elements other than those in column i and row k and must therefore be independent of Q_{ij} so we have

$$\frac{\partial Q}{\partial Q_{ij}} = \sum_{k=1}^{m} \delta_k^j Q^{ki} = Q^{ji}$$

Therefore subbing this into our fnction of Q at the start we have the proposition as desired $\square$

With this jacobi formula now let $Q_{ij} = g_{ij}$ be the metric so we have

$$\partial_k g = G^{ji} \partial_k g_{ij}$$

where G^{ij} is the cofactor of g_{ij}. By a variation of cramer rule the invergse of g_{ij} which we denote as g^{ji} can be found by

$$g^{ji} = \frac{1}{g} G^{ji} \quad \Rightarrow \quad G^{ji} = g g^{ij}$$

where g is the determinant of the metric. Then putting this alotgoether we have

$$\partial_k g = g g^{ij} \partial_k g_{ij} \quad \Rightarrow \quad g^{jk} \partial_i g_{jk} = \frac{1}{g} \partial_i g$$

Finally we substitue our relation into our covariant derivative contration from earlier 194 to obtain

$$\nabla_i A^i = \partial_i A^i + A^i \frac{1}{2g} \partial_i g = \partial_i A^i + A^i \frac{1}{\sqrt{g}} \partial_i \sqrt{g} = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} A^i)$$

the 2nd equality can be seen by

$$\left(\frac{1}{\sqrt{g}} \right) (\partial_i \sqrt{g}) = \left(\frac{1}{\sqrt{g}} \right) \left(\frac{1}{2} \frac{1}{\sqrt{g}} \partial_i g \right) = \frac{1}{2g} \partial_i g$$

. The last equality is just product rule.

Fact 196

For *orthogonal coordinates* only g_{ij} is diagonal(obviously). Here is the interesting thing, since recall 167

$$e_i = h_{(i)}\hat{e}_{(i)} \quad \hat{e}_{(i)} = \frac{1}{h_{(i)}}e_i$$

by definition where recall $\hat{e}_{(i)}$ are the *unit physical basis vectors* while $g_{ii} = h_{(i)}h_{(i)}\hat{e}_i \cdot \hat{e}_j = e_i \cdot e_j$. It is clear to see that the metric will look like

$$\begin{pmatrix} h^2_{(1)} & 0 & 0 \\ 0 & h^2_{(2)} & 0 \\ 0 & 0 & h^2_{(3)} \end{pmatrix}$$

Hence we have $\sqrt{g} = h_{(1)}h_{(2)}h_{(3)}$ so putting everything together(converting all to physical components) we have

$$\nabla_i A^i = \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \sum_{i=1}^{3} \partial_i \left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}} A_{(i)} \right) = \nabla \cdot \overrightarrow{A}$$

...to be continued i think you got to general idea already and should know theory euclidean tensor calculus and orthogonal curvilinear systems in full by now.

Fact 197

In summary let f, $\overrightarrow{A}$ and T be an arbituary scalar, vector and contravariant rank 2 tensor function of the *orthogonal coordinates* then we have

$$\begin{aligned}
\nabla f &= \left(\frac{1}{h_{(1)}}\partial_1 f, \frac{1}{h_{(2)}}\partial_2 f, \frac{1}{h_{(3)}}\partial_3 f \right) \\
\nabla^2 f &= \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \sum_{i=1}^{3} \partial_i \left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h^2_{(i)}} \partial_i f \right) \\
\nabla \cdot \overrightarrow{A} &= \frac{1}{h_{(1)}h_{(2)}h_{(3)}} \sum_{i=1}^{3} \partial_i \left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}} A_{(i)} \right) \\
\nabla \times \overrightarrow{A} f &= \left(\frac{1}{h_{(2)}h_{(3)}} \left(\partial_2(h_{(3)}A_{(3)}) - \partial_3(h_{(2)}A_{(2)}) \right), \right. \\
\nabla \times \overrightarrow{A} &= \left(\frac{1}{h_{(2)}h_{(3)}} \left(\partial_2(h_{(3)}A_{(3)}) - \partial_3(h_{(2)}A_{(2)}) \right), \right. \\
& \frac{1}{h_{(3)}h_{(1)}} \left(\partial_3(h_{(1)}A_{(1)}) - \partial_1(h_{(3)}A_{(3)}) \right), \\
& \left. \frac{1}{h_{(1)}h_{(2)}} \left(\partial_1(h_{(2)}A_{(2)}) - \partial_2(h_{(1)}A_{(1)}) \right) \right) \\
(\nabla \overrightarrow{A})_{ij} &= \begin{cases} \partial_i \left(\frac{A_{(i)}}{h_{(i)}} + \frac{1}{h_{(i)}} \overrightarrow{A} \cdot \nabla h_{(i)} \right) & i = j \\ \frac{1}{h_{(i)}} \left(\partial_i A_{(j)} - \frac{A_{(i)}}{h_{(j)}} \partial_j h_{(i)} \right) & i \neq j \end{cases}
\end{aligned}$$

Recall that we have defined physical components to be the the direct differentiation of each component with the new component so we calculate each $h_{(i)}$ by

$$\sqrt{\left(\frac{\partial x_1}{\partial \hat{x}_i}\right)^2 + \left(\frac{\partial x_2}{\partial \hat{x}_i}\right)^2 + \left(\frac{\partial x_3}{\partial \hat{x}_i}\right)^2}$$

Fact 198

Note the angle conventions

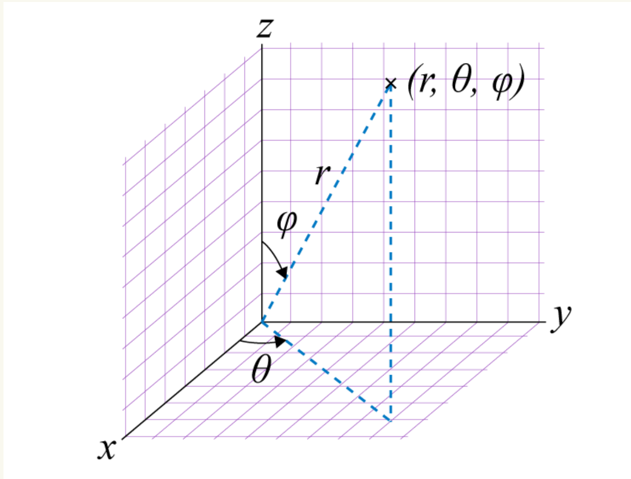

Figure 3: the polar angle φ is clockwise positive while the azimuthal angle θ is anticlockwise positive