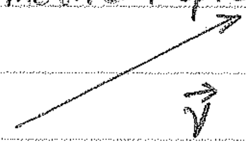
VECTORS IN $\mathbb{R}^N$.

Definition. A (real) vector in $\mathbb{R}^N$ is a quantity which has both magnitude and direction. The magnitude is a nonnegative number, and the directions are described by rays in $\mathbb{R}^N$

Geometric Representation of Vectors:



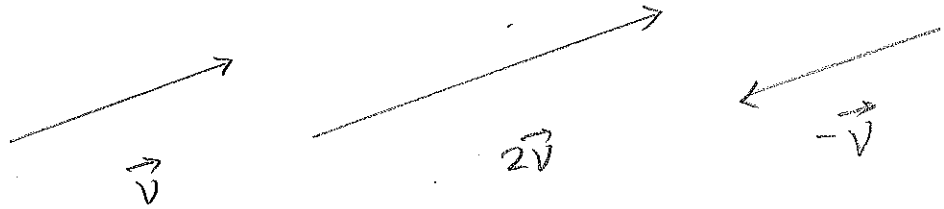
$\vec{v}$

An oriented line segment in $\mathbb{R}^N$ represents a vector $\vec{v}$. The length of the line segments represents the magnitude of $\vec{v}$ and the orientation represents the direction of $\vec{v}$.

Note that a vector is determined by its magnitude and direction, but not its position. In other words, if we can translate a vector to coincide with another, then these two vectors are equal.

Now we are going to review the basic algebraic operations on vectors.

- Scalar Multiplication: If $\vec{v}$ is a vector and c is a real number (called a scalar), we can define the scalar multiplication $c\vec{v}$, which is a vector, as follows:
  The magnitude of $c\vec{v}$ is $|c|$ times the magnitude of $\vec{v}$, and for the direction:
  ① If $c>0$, $c\vec{v}$ has the same direction as that of $\vec{v}$
  ② If $c=0$, $c\vec{v}=\vec{0}$, no direction
  ③ If $c<0$, $c\vec{v}$ has the opposite direction as that of $\vec{v}$



And we write $-\vec{v}$ for $(-1)\vec{v}$

- Vector Addition:

If $\vec{u}$ and $\vec{v}$ are vectors positioned so the initial point of $\vec{v}$ is at the terminal point of $\vec{u}$, then $\vec{u}+\vec{v}$ is the vector from the initial point of $\vec{u}$ to the terminal point of $\vec{v}$

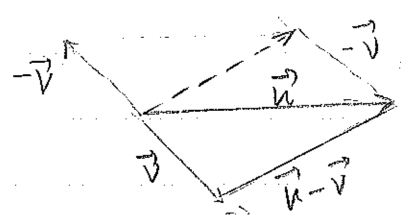
Note that vector addition is commutative: $\vec{u}+\vec{v}=\vec{v}+\vec{u}$
and associative: $(\vec{u}+\vec{v})+\vec{w}=\vec{u}+(\vec{v}+\vec{w})$.
There's also a distributive law of scalars on vectors:

$$c(\vec{u}+\vec{v})=c\vec{u}+c\vec{v}$$

The above laws can be verified using some geometry.

- Vector Subtraction

If $\vec{u}$ and $\vec{v}$ are vectors, then define $\vec{u}-\vec{v}$ to be $\vec{u}+(-\vec{v})$



• Vector Representation in Coordinates: (We'll take $\mathbb{R}^3$ as example, but the principle extends naturally to all $\mathbb{R}^N$)

In a coordinate system, for any point $P=(x,y,z)$, we can construct the vector with initial point the origin $O=(0,0,0)$, and terminal point P. We denote this vector $\overrightarrow{OP}$ and call it the position vector of the point P

Observe that the assignment of position vector to point P gives an identification of (3-dimensional) vectors and points in $\mathbb{R}^3$. Another observation is that for each vector $\vec{v}$ and any point $A=(a_1,a_2,a_3)\in\mathbb{R}^3$, there is a point $B=(b_1,b_2,b_3)$ such that $\vec{v}$ is equivalent to $\overrightarrow{AB}$. We say $\overrightarrow{AB}$ is a representation of the vector $\vec{v}$.

In a coordinate system, if $\vec{v}$ is equivalent to the position vector of $P=(a,b,c)$, we can write $\vec{v}=(a,b,c)$

The coordinate system makes vector algebra simple to compute:

Theorem. If $\vec{u}=(a_1,b_1,c_1)$, $\vec{v}=(a_2,b_2,c_2)$ and $\lambda\in\mathbb{R}$, then:
$\vec{u}+\vec{v}=(a_1+a_2,\ b_1+b_2,\ c_1+c_2)$, $\vec{u}-\vec{v}=(a_1-a_2,\ b_1-b_2,\ c_1-c_2)$
$\lambda\vec{u}=(\lambda a_1,\lambda b_1,\lambda c_1)$

Theorem: If $\vec{u}=(a,b,c)$, then $|\vec{u}|=\sqrt{a^2+b^2+c^2}$

Example. If $\vec{u}=(3,2,5)$, $\vec{v}=(4,1,3)$, then

$$\vec{u}-2\vec{v}=(3,2,5)-2\cdot(4,1,3)=(3,2,5)-(8,2,6)$$
$$=(-5,0,-1)$$

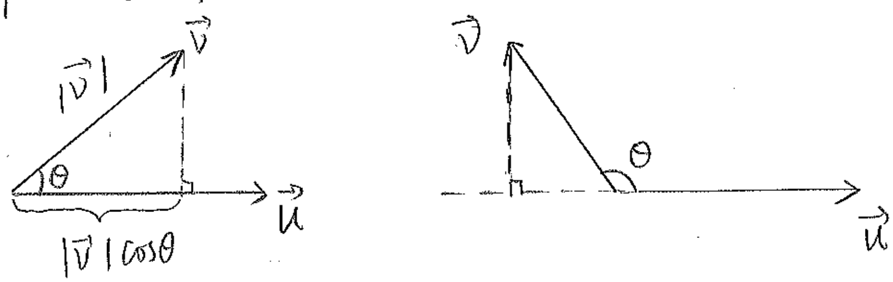
Definition. If a vector has length 1, we call it a unit vector.

For example, $\vec{i}=(1,0,0)$, $\vec{j}=(0,1,0)$, $\vec{k}=(0,0,1)$

Observe that if $\vec{v}$ is a nonzero vector, then there's a unit vector which has the same direction as $\vec{v}$, and the unit vector is $\frac{1}{|\vec{v}|}\vec{v}$.

Example. $\vec{v}=(2,-2,-1)$. $|\vec{v}|=\sqrt{2^2+(-2)^2+(-1)^2}=3$, so the unit vector in the direction of $\vec{v}$ is $(\frac{2}{3},-\frac{2}{3},-\frac{1}{3})$.

Definition. If θ is the angle between two vectors $\vec{u}$ and $\vec{v}$, we define the dot product of $\vec{u}$ and $\vec{v}$ to be $\vec{u}\cdot\vec{v}=|\vec{u}||\vec{v}|\cos\theta$

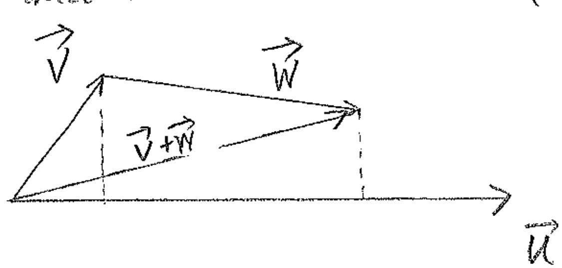


By this definition, we can immediately see some of the properties of dot product:

(i) $\vec{u}\cdot\vec{v}=\vec{v}\cdot\vec{u}$ $\forall \vec{u},\vec{v}$.
(ii) $\vec{u}\cdot\vec{0}=0$ $\forall \vec{u}$.
(iii) If $\theta=\frac{\pi}{2}$, i.e. $\vec{u}\perp\vec{v}$, then $\vec{u}\cdot\vec{v}=0$
(iv) $c(\vec{u}\cdot\vec{v})=(c\vec{u})\cdot\vec{v}=\vec{u}\cdot(c\vec{v})$ $\forall \vec{u},\vec{v}$, $\forall c\in\mathbb{R}$

Lemma. Dot Product is distributive: $\vec{u}\cdot(\vec{v}+\vec{w})=\vec{u}\cdot\vec{v}+\vec{u}\cdot\vec{w}$

Proof.



From Linear Algebra, we know that we can choose a standard basis of $\mathbb{R}^3$, $\{\vec{i}, \vec{j}, \vec{k}\}$ such that $|\vec{i}| = |\vec{j}| = |\vec{k}| = 1$,

they're pairwisely perpendicular to each other, and they're the unit vectors in the direction of the Cartisian Coordinate axis. i.e in coordinates $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ and $\vec{k} = (0, 0, 1)$

Each vector $\vec{u} = (a, b, c)$ can therefore be decomposed into

$$\begin{aligned}\vec{u} &= (a, 0, 0) + (0, b, 0) + (0, 0, c) \\ &= a(1,0,0) + b(0,1,0) + c(0,0,1) \\ &= a\vec{i} + b\vec{j} + c\vec{k}\end{aligned}$$

Lemma If $\vec{u} = (a_1, b_1, c_1)$, $\vec{v} = (a_2, b_2, c_2)$, then $\vec{u} \cdot \vec{v} = a_1a_2 + b_1b_2 + c_1c_2$

Proof

$$\begin{aligned}\vec{u} \cdot \vec{v} &= (a_1\vec{i} + b_1\vec{j} + c_1\vec{k}) \cdot (a_2\vec{i} + b_2\vec{j} + c_2\vec{k}) \\ &= (a_1\vec{i} \cdot a_2\vec{i}) + (a_1\vec{i} \cdot b_2\vec{j}) + (a_1\vec{i} \cdot c_2\vec{k}) \\ &\quad + (b_1\vec{j} \cdot a_2\vec{i}) + (b_1\vec{j} \cdot b_2\vec{j}) + (b_1\vec{j} \cdot c_2\vec{k}) \\ &\quad + (c_1\vec{k} \cdot a_2\vec{i}) + (c_1\vec{k} \cdot b_2\vec{j}) + (c_1\vec{k} \cdot c_2\vec{k}) \\ &= a_1a_2 + b_1b_2 + c_1c_2\end{aligned}$$

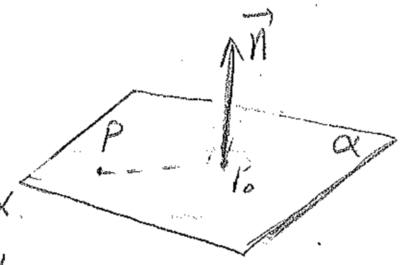
(since $\vec{i} \cdot \vec{j} = \vec{j} \cdot \vec{k} = \vec{k} \cdot \vec{i} = 0$)

Lemma. $|\vec{v}|^2 = \vec{v} \cdot \vec{v}$

Proof. $\vec{v} \cdot \vec{v} = |\vec{v}||\vec{v}| \cos 0 = |\vec{v}|^2 \cdot 1 = |\vec{v}|^2$

Corollary If $\vec{v} = (a, b, c)$, then $|\vec{v}| = \sqrt{a^2 + b^2 + c^2}$

An important application of dot product is in geometry. We can describe a plane in $\mathbb{R}^3$ by an equation that involves dot product:



Pick a vector $\vec{n} = (a, b, c)$ perpendicular to the plane α.
Pick a point $P_0 = (x_0, y_0, z_0) \in \alpha$.
Now for any $P = (x, y, z)$, $P \in \alpha$ if and only if $\overrightarrow{P_0P} \perp \vec{n}$. (We call $\vec{n}$ a normal vector of α).

Algebraically, $\overrightarrow{P_0P} \perp \vec{n}$ is equivalent to
$$(x-x_0, y-y_0, z-z_0) \cdot (a, b, c) = 0.$$

i.e. $a(x-x_0) + b(y-y_0) + c(z-z_0) = 0$

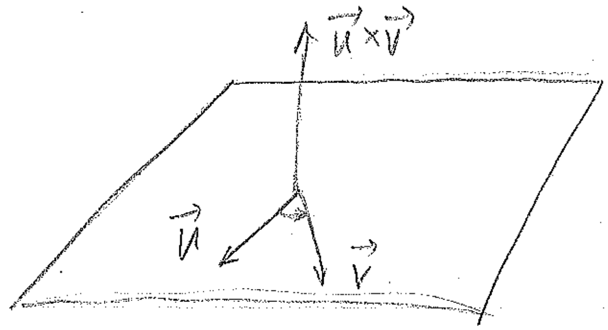
or $\quad ax + by + cz = ax_0 + by_0 + cz_0$

Example The plane that passes through $(1, 2, 3)$ with a normal vector $(5, -2, 1)$ has equation
$$5(x-1) - 2(x-2) + (x-1) = 0.$$

The dot product assigns a real number (i.e. scalar) to a given pair of vectors, as we've seen in its definition. Now we're going to consider another form of vector product, called cross product.

Definition. $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^3$. Define the cross product $\vec{u}\times\vec{v}$ to be the vector whose magnitude is $|\vec{u}|\cdot|\vec{v}|\sin\theta$, whose direction is determined by the "right hand rule". (θ is the angle between $\vec{u}$ and $\vec{v}$)

Example.



$\vec{u}\times\vec{v}$ is a vector perpendicular to both $\vec{u}$ and $\vec{v}$.

Proposition. $\vec{u}$, $\vec{v}$, $\vec{w}$ are vectors in $\mathbb{R}^3$

(i) $\vec{u}\times\vec{v} = -\vec{v}\times\vec{u}$

(ii) $\vec{u}\times\vec{0} = \vec{0}$

(iii) $\vec{u}\times\vec{u} = \vec{0}$

(iv) $\vec{u} \parallel \vec{v} \Rightarrow \vec{u}\times\vec{v} = \vec{0}$

(v) $|\vec{u}\times\vec{v}|$ equals to the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$

(vi) $\vec{u}\times(\vec{v}+\vec{w}) = \vec{u}\times\vec{v} + \vec{u}\times\vec{w}$

Proof. (i) By the right hand rule, if the order of the two vectors is switched, the direction of the cross product will be changed to the opposite.

(ii) $|\vec{u}\times\vec{0}| = |\vec{u}|\cdot|\vec{0}|\cdot\sin\theta = 0$, so $\vec{u}\times\vec{0} = \vec{0}$

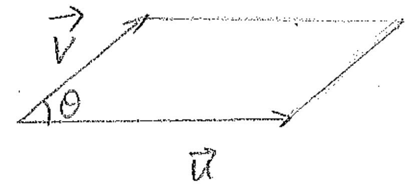
(iii) by (i), let $\vec{v} = \vec{u}$, then $\vec{u}\times\vec{u} = -\vec{u}\times\vec{u}$, so $\vec{u}\times\vec{u} = \vec{0}$

(iv). If $\vec{u} \parallel \vec{v}$, then $\theta = 0$ or $\theta = \pi$, $\sin 0 = \sin \pi = 0$

$$|\vec{u} \times \vec{v}| = |\vec{u}| \cdot |\vec{v}| \sin\theta = 0$$

(v). By the Sine Theorem, the area of the parallelgram is

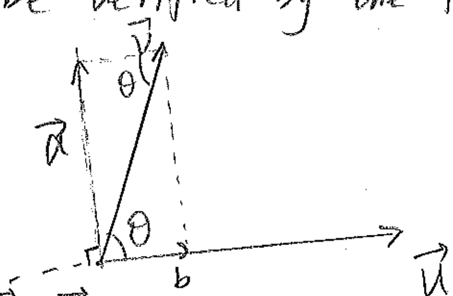
$$2 \cdot \frac{|\vec{u}| \cdot |\vec{v}| \sin\theta}{2} = |\vec{u} \times \vec{v}|$$



(vi). To prove (vi), we need a Lemma:

Lemma. The cross product of $\vec{u}$ and $\vec{v}$ only depends on the Component of $\vec{u}$ perpendicular to $\vec{u}$.

This Lemma can be verified by the following picture:



we decompose $\vec{v} = \vec{a} + \vec{b}$ such that $\vec{a} \perp \vec{u}$ and $\vec{b} \parallel \vec{u}$.

Then $\vec{u} \times \vec{v} = \vec{u} \times \vec{a}$.

Now with the help of this Lemma, we can prove

(vi) $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$:

Let $\vec{v} = \vec{a} + \vec{b}$, $\vec{w} = \vec{c} + \vec{d}$, such that

$\vec{a} \perp \vec{u}$, $\vec{b} \parallel \vec{u}$, $\vec{c} \perp \vec{u}$, $\vec{d} \parallel \vec{u}$.

by the Lemma, we know that

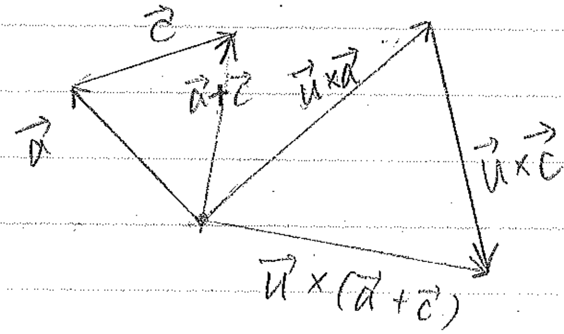
$$\vec{u}\times(\vec{v}+\vec{w}) = \vec{u}\times(\vec{a}+\vec{b}+\vec{c}+\vec{d}) = \vec{u}\times((\vec{a}+\vec{c})+(\vec{b}+\vec{d}))$$
$$= \vec{u}\times(\vec{a}+\vec{c})$$

$$\vec{u}\times\vec{v}+\vec{u}\times\vec{w} = \vec{u}\times\vec{a}+\vec{u}\times\vec{c}$$

So it suffices to verify $\vec{u}\times(\vec{a}+\vec{c}) = \vec{u}\times\vec{a}+\vec{u}\times\vec{c}$ for $\vec{a}\perp\vec{u}$, $\vec{c}\perp\vec{u}$. This can be done as follows:

Consider $\vec{u}$ to be pointing into the page, so $\vec{a}$ and $\vec{c}$ are parallel to the page.

$\vec{u}\times\vec{a}$ is the vector of magnitude $|\vec{u}|\cdot|\vec{a}|$, in the direction of $\vec{a}$ rotating clockwise $90°$,
$\vec{u}\times\vec{c}$ is the vector of magnitude $|\vec{u}|\cdot|\vec{b}|$, in the direction of $\vec{c}$ rotating clockwise $90°$.
$\vec{u}\times(\vec{a}+\vec{c})$ is the vector of magnitude $|\vec{u}|\cdot|\vec{a}+\vec{c}|$, in the direction of $\vec{a}+\vec{c}$ rotating clockwise $90°$.
Then geometrically we see very clearly that $\vec{u}\times(\vec{a}+\vec{c})$ coincides with $\vec{u}\times\vec{a}+\vec{u}\times\vec{c}$



Now with the help of the distributive law, we can develop an algebraic way of computation.

Again we take the standard basis $\{\vec{i}, \vec{j}, \vec{k}\}$.
Observe that $\vec{i}\times\vec{j}=\vec{k}$, $\vec{j}\times\vec{k}=\vec{i}$, $\vec{k}\times\vec{i}=\vec{j}$.

If $\vec{u}=(x_1, y_1, z_1)$, $\vec{v}=(x_2, y_2, z_2)$, then
$\vec{u}=x_1\vec{i}+y_1\vec{j}+z_1\vec{k}$, $\vec{v}=x_2\vec{i}+y_2\vec{j}+z_2\vec{k}$.

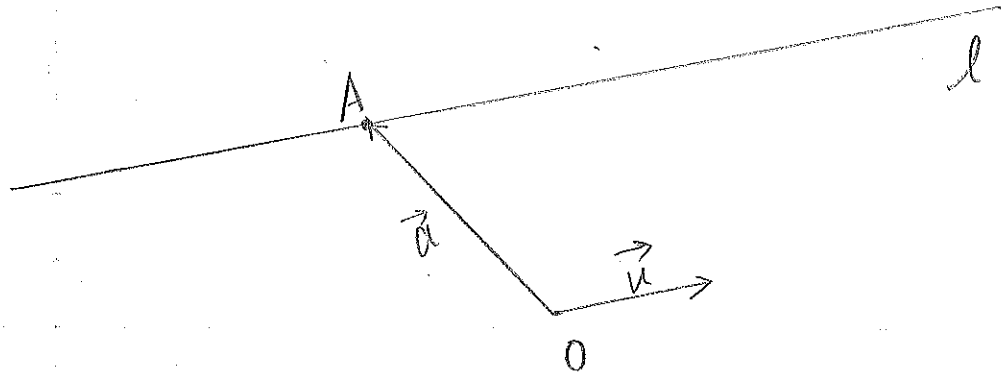
$$\begin{aligned}\vec{u}\times\vec{v} &= (x_1\vec{i}+y_1\vec{j}+z_1\vec{k})\times(x_2\vec{i}+y_2\vec{j}+z_2\vec{k}) \\ &= (x_1\vec{i})\times(y_2\vec{j})+(x_1\vec{i})\times(z_2\vec{k})+(y_1\vec{j})\times(x_2\vec{i}) \\ &\quad +(y_1\vec{j})\times(z_2\vec{k})+(z_1\vec{k})\times(x_2\vec{i})+(z_1\vec{k})\times(y_2\vec{j}) \\ &= (y_1z_2-z_1y_2)\vec{i}+(z_1x_2-x_1z_2)\vec{j}+(x_1y_2-y_1x_2)\vec{k}\end{aligned}$$

An interesting observation is that above indicates:

$$\vec{u}\times\vec{v}=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2\end{vmatrix}$$

The cross product also has applications in geometry: it can be used to describe straight lines in $\mathbb{R}^3$.

Given a straight line $l\subseteq\mathbb{R}^3$, it can be described as $\vec{r}=\vec{a}+\lambda\vec{u}$, where $\vec{a}$ is the position vector of a Point $A=(a_1, a_2, a_3)$ on l, and $\vec{u}\parallel l$.
λ is the parameter.



We wish to obtain a form of equation without the parameter λ, so we use the cross product:

$$\vec{r} \times \vec{u} = (\vec{a} + \lambda \vec{u}) \times \vec{u} = \vec{a} \times \vec{u}$$

i.e. $\vec{r} \times \vec{u} = \vec{a} \times \vec{u}$

Note $\vec{a}$ and $\vec{u}$ are some fixed vectors, it follows $\vec{a} \times \vec{u} = \vec{b}$ is some fixed vector.

so the equation of l can be written as

$$\vec{r} \times \vec{u} = \vec{b}.$$

Now we are going to see another product of vectors: Scalar Triple Product.

Definition. Given three vectors $\vec{u}, \vec{v}, \vec{w}$ in $\mathbb{R}^3$ their scalar triple product is $\vec{u} \cdot (\vec{v} \times \vec{w})$

If $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$, $\vec{w} = (w_1, w_2, w_3)$ then

$$\vec{u} \cdot (\vec{v} \times \vec{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

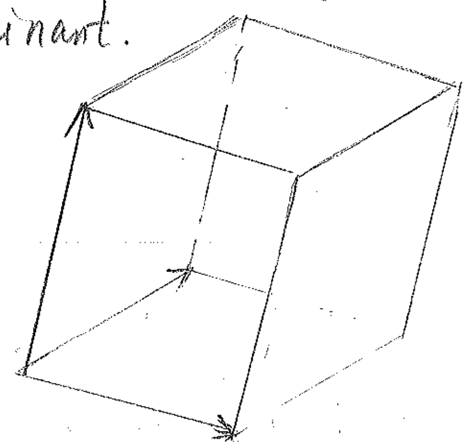
Proposition. (i) $\vec{u}\cdot(\vec{v}\times\vec{w}) = \vec{v}\cdot(\vec{w}\times\vec{u}) = \vec{w}\cdot(\vec{u}\times\vec{v})$

(ii) $\vec{u}\cdot(\vec{v}\times\vec{w}) = (\vec{u}\times\vec{v})\cdot\vec{w}$.

(iii) $\vec{u}\cdot(\vec{v}\times\vec{w}) \neq 0$ if and only if $\vec{u},\vec{v},\vec{w}$ are linearly independent.

(iv) $\vec{u}\cdot(\vec{v}\times\vec{w})$ is the signed volume of the parallelepiped formed by $\vec{a},\vec{b},\vec{c}$.

Proof. The proofs follow easily from elementary properties of determinant.



VECTOR FIELD AND LINE INTEGRAL

The concept of vector field is widely used in physics, and it also has its own interest in mathematics.

Definition. A vector field on $\mathbb{R}^n$ is a function $\vec{F}$ that assigns to each $p \in \mathbb{R}^n$ an n-dimensional vector. If we write in coordinates, a vector field is a function

$$\vec{F}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$(x_1, \dots, x_n) \longmapsto (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))$$

Example. We can define the Gravitational Field by

$$\vec{F}(x,y,z) = -\frac{MG}{|(x,y,z)|^3}(x,y,z)$$

Now we are going to study the integration of vector fields along some geometric objects. Intuitively, the integration tells us the net effect of some vector field on some geometric objects.

The first kind of integral we are interested is called the line integral.

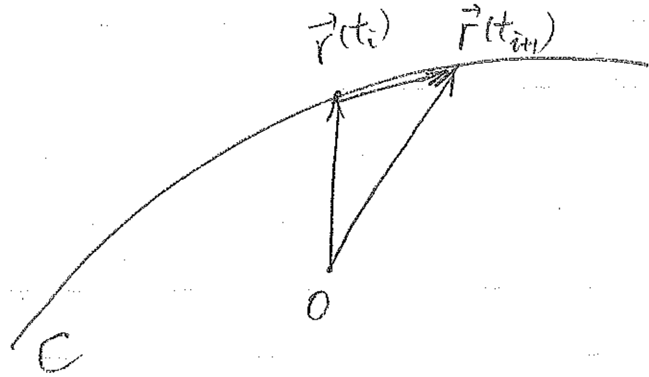
A curve in $\mathbb{R}^3$ can be described in coordinates by $\vec{r}(t) = (x(t), y(t), z(t))$

Definition. If $\vec{r}(t) = (x(t), y(t), z(t))$ is a curve, define the tangent vector of $\vec{r}(t)$ at $\vec{r}(t_0)$ to be $\vec{r}\,'(t_0) = (\frac{dx}{dt}(t_0), \frac{dy}{dt}(t_0), \frac{dz}{dt}(t_0))$

Given a vector field $\vec{F}(x,y,z)$ and a curve $C : \vec{r}(t)$, $t \in [a,b]$
we can define the line integral of $\vec{F}(x,y,z)$ along C
to be the limit of a riemann sum:

$$\int_C \vec{F} \cdot d\vec{r} = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{N} \vec{F}(\vec{r}(t_i)) \cdot \Delta \vec{r} = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{N} \vec{F}(\vec{r}(t_i)) \cdot (\vec{r}(t_i) - \vec{r}(t_{i-1}))$$

where $a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b$.



The spirit of this definition is the observation that when $\vec{r}(t_i)$ and $\vec{r}(t_{i+1})$ are close to each other, $\vec{r}(t_{i+1}) - \vec{r}(t_i)$ is a good approximation of the curve between $t = t_i$ and $t = t_{i+1}$

Similar to the integral of a function along the real line it's in general very hard to evaluate the integral by the Riemann sum. We need to develop some ways of integration:

$\frac{d\vec{r}}{dt} = \lim_{\Delta t \to 0} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{\Delta t}$, so when Δt is small,

$\vec{r}(t+\Delta t) - \vec{r}(t)$ can be approximated by $\frac{d\vec{r}}{dt} \Delta t$.

This indicates

$$\int_C \vec{F}\cdot d\vec{r} = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{N} \vec{F}(\vec{r}(t_i))\left(\vec{r}(t_i)-\vec{r}(t_{i-1})\right) = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{N} \vec{F}(\vec{r}(t_i))\cdot \frac{d\vec{r}}{dt}\Delta t_i$$

$$= \int_a^b \vec{F}(\vec{r}(t))\cdot \frac{d\vec{r}}{dt}\, dt$$

We thus obtain:

Proposition. $\int_C \vec{F}\cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t))\cdot \frac{d\vec{r}}{dt}\, dt$

Example. $\vec{F}(x,y,z) = (y, x, z)$. the curve C is parameterized by $\vec{r}(t) = (t, t^2, 2t^2)$, $t \in [0,1]$

$$\int_C \vec{F}\cdot d\vec{r} = \int_0^1 \vec{F}(\vec{r}(t))\cdot \frac{d\vec{r}}{dt}\, dt$$

$$= \int_0^1 (t^2, t, 2t^2)\cdot(1, 2t, 4t)\, dt$$

$$= \int_0^1 t^2 + 2t^2 + 8t^3\, dt$$

$$= 3$$

Definition. If C is a closed path, i.e. when the starting and end points coincide, the line integral of $\vec{F}$ along C is written as $\oint_C \vec{F}\cdot d\vec{r}$, and we call it the circulation of $\vec{F}$ around C.

Proposition. If we change the orientation of C, the line integral is:

$$\int_{-C} \vec{F}\cdot d\vec{r} = -\int_C \vec{F}\cdot d\vec{r}$$

where $-C$ is the curve obtained by reversing the orientation of C.

Proof. It follows directly from the definition, since the tangent vectors of C and $-C$ at each point are opposite to each other.

There is one more thing we need to take care of:
A geometric curve C may have more than one parameterizations, so we need to make sure different choices of parameterizations lead to the same line integral.

Suppose C can also be parameterized by s with $t = f(s)$ for some function f, i.e. $\vec{\rho}(s) = \vec{r}(f(s))$. Then

$$\int_a^b \vec{F}(\vec{\rho}(s)) \cdot \frac{d\vec{\rho}}{ds} \cdot ds = \int_a^b \vec{F}(\vec{r}(f(s))) \cdot \frac{d\vec{r}}{dt} \cdot f'(s) ds$$

$$= \int_{f(a)}^{f(b)} \vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt} \, dt$$

Proposition. The line integral is independent of the parameterization of the path C.

There're also other forms of line integrals: if we consider the scalar multiplication and cross product instead of dot product.

If $f: \mathbb{R}^3 \to \mathbb{R}$ is a scalar function, C is a curve parameterized by $\vec{r}(t)$, $t \in [a,b]$. define the line integral

$$\int_C f \, d\vec{r} = \int_a^b f \, \frac{d\vec{r}}{dt} \, dt = \left(\int_a^b f \, \frac{dx}{dt} \, dt, \int_a^b f \, \frac{dy}{dt} \, dt, \int_a^b f \, \frac{dz}{dt} \, dt \right)$$

CONSERVATIVE VECTOR FIELDS

Definition. A vector field $\vec{F}$ is conservative if it has the property that the line integral of $\vec{F}$ along any closed curve C is zero: $\oint_C \vec{F}\cdot d\vec{r} = 0$

Equivalently, a vector field $\vec{F}$ is conservative if the line integral of $\vec{F}$ along a curve only depends on the endpoints of the curve, i.e. independent of the path taken.

$\int_{C_1} \vec{F}\cdot d\vec{r} = \int_{C_2} \vec{F}\cdot d\vec{r}$ where C_1, C_2 are two curves that have the same endpoints.

Definition. Given a differentiable function $f: \mathbb{R}^3 \longrightarrow \mathbb{R}$, we define the gradient of f to be the vector field $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$

Theorem. $\vec{F}$ is a vector field on some region $D \subseteq \mathbb{R}^3$. Then $\vec{F}$ is conservative if and only if there exists a function $\phi: \mathbb{R}^3 \to \mathbb{R}$ such that $\vec{F} = \nabla\phi$

Proof. "$\Leftarrow$" If $\vec{F} = \nabla\phi$, then for any curve C from $A \in \mathbb{R}^3$ to $B \in \mathbb{R}^3$, parameterize C by $\vec{r}(t)$, $t \in [a,b]$, so that $\vec{r}(a) = A$, $\vec{r}(b) = B$. Write $\vec{r}(t) = (x(t), y(t), z(t))$

$$\int_C \vec{F}\cdot d\vec{r} = \int_C \nabla\phi(\vec{r})\cdot \frac{d\vec{r}}{dt}\,dt$$

$$= \int_a^b \frac{\partial\phi}{\partial x}(x(t),y(t),z(t))\frac{dx}{dt} + \frac{\partial\phi}{\partial y}(x(t),y(t),z(t))\frac{dy}{dt} + \frac{\partial\phi}{\partial z}(x(t),y(t),z(t))\,dt$$

$$= \int_a^b \frac{d\phi(\vec{r}(t))}{dt}\,dt$$

$$= \phi(\vec{r}(b)) - \phi(\vec{r}(a)) = \phi(B) - \phi(A)$$

so the integral only depends on the endpoints of C.

"$\Longrightarrow$". Conversely, if $\vec{F}$ is conservative on D.

define $\phi(x,y,z) = \int_C \vec{F} \cdot d\vec{r}$, where C is any path from $\vec{0}$ to (x,y,z)

We need to show $\nabla\phi = \vec{F} = (f_1, f_2, f_3)$

$$\frac{\partial \phi}{\partial x} = \lim_{h\to 0} \frac{\phi(x+h,y,z) - \phi(x,y,z)}{h}$$

$$= \lim_{h\to 0} \frac{\int_L \vec{F}\cdot d\vec{r}}{h}$$

, where L is the line segment from (x,y,z) to $(x+h, y, z)$ parameterized by: $\vec{r}(t) = (t, y, z)$, $t \in [x, x+h]$

$$= \lim_{h\to 0} \frac{\int_x^{x+h} f_1(t,y,z)\,dt}{h}$$

$$= \lim_{h\to 0} \frac{\int_0^{x+h} f_1(t,y,z)\,dt - \int_0^{x} f_1(t,y,z)\,dt}{h}$$

$$= \frac{d}{dx} \int_0^x f_1(t,y,z)\,dt$$

$$= f_1(x,y,z)$$

Similarly, you can verify $\frac{\partial\phi}{\partial y} = f_2$ and $\frac{\partial\phi}{\partial z} = f_3$

Definition. If $\vec{F} = \nabla\phi$, we say ϕ is the potential for $\vec{F}$.

Example. $\vec{F}(x,y,z) = -\frac{MG}{|(x,y,z)|^3}(x,y,z)$ is conservative with potential function $\frac{MG}{\sqrt{x^2+y^2+z^2}}$

Example. $\vec{F}(x,y) = (y,-x)$ is not conservative on $\mathbb{R}^2$:

Suppose it's conservative, $\vec{F}(x,y) = \nabla\phi$ for some $\phi: \mathbb{R}^2 \to \mathbb{R}$

then $\begin{cases} \frac{\partial\phi}{\partial x} = y \\ \frac{\partial\phi}{\partial y} = -x \end{cases} \Rightarrow \begin{cases} \frac{\partial^2\phi}{\partial y\partial x} = 1 \\ \frac{\partial^2\phi}{\partial x\partial y} = -1 \end{cases}$ Contradicts to Young's Theorem

By the previous example, we see that for a 2-dimensional vector field $\vec{F}(x,y) = (f_1(x,y), f_2(x,y))$; a necessary condition for $\vec{F}$ to be conservative is that $\frac{\partial f_1}{\partial y} = \frac{\partial f_2}{\partial x}$

We would like also to see whether this is a sufficient condition. And it turns out that we need to impose some conditions on the region.

Theorem. Let $\vec{F} = (P, Q)$ be a vector field on an open simply-connected region D. Suppose that P and Q have continuous first-order derivatives, and $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ on D, then $\vec{F}$ is conservative.

The proof of the above theorem requires the Green's Theorem. We will prove the Green's Theorem later as a special case of the Stokes' Theorem, but we would like to state the Green's Theorem here:

Theorem (Green's Theorem). Let C be a positively oriented, piecewise smooth, simple closed curve in a plane, and D the region bounded by C. If $P(x,y)$ and $Q(x,y)$ have continuous partial derivatives on an open region containing D, then

$$\oint_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$$

In the previous discussion, a "simply connected" region means a region that has no holes.

If the vector field is defined on a region which is not simply connected, the condition $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ is not sufficient to conclude $\vec{F}$ is conservative.

Example. $\vec{F}(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$

$$\frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{-(x^2+y^2)+2y^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}$$

$$\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{(x^2+y^2)-2x^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}.$$

So $\frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right)$

But $\vec{F}(x,y)$ is not conservative:

Take C to be the unit circle. $x^2+y^2=1$.

$$\oint_C \vec{F}\cdot d\vec{r} = \int_0^{2\pi} \left(\frac{-\sin\theta}{\cos^2\theta+\sin^2\theta}, \frac{\cos\theta}{\cos^2\theta+\sin^2\theta}\right)\cdot(\cos\theta, \sin\theta)'\,d\theta$$

$$= \int_0^{2\pi} (-\sin\theta, \cos\theta)\cdot(-\sin\theta, \cos\theta)\,d\theta$$

$$= \int_0^{2\pi} \sin^2\theta + \cos^2\theta\,d\theta$$

$$= 2\pi \neq 0$$

SURFACE INTEGRAL

In order to define surface integrals, we need to first find a way to describe a surface:

Definition. A parametric surface S is a function $\mathbb{R}^2 \supseteq D \longrightarrow \mathbb{R}^3$ $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$. We usually require $\vec{r}$ to be injective when restricted to the interior of D.

Example. If S is the surface of the graph of $z=f(x,y)$, we can write it as $\vec{r}(s,t) = (s, t, f(s,t))$

Example The cylinder $x^2+y^2=4$, $0 \le z \le 1$ can be parameterized by $\vec{r}(\theta, z) = (2\cos\theta, 2\sin\theta, v)$, $0 \le \theta \le 2\pi$, $0 \le v \le 1$

Example. The unit sphere $x^2+y^2+z^2=1$ can be parameterized by $\vec{r}(\theta,\varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$, $0 \le \varphi \le \pi$, $0 \le \theta \le 2\pi$

When we study curves, we need to make use of tangent lines; now we need to look into the tangent planes in order to study surfaces.

Given a surface $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$, we will find its tangent plane by finding two tangent vectors which will span the plane.

For a fixed (s_0, t_0), consider the curves:

$\vec{\alpha}(s) = \vec{r}(x(s,t_0), y(s,t_0), z(s,t_0))$

$\vec{\beta}(t) = \vec{r}(x(s_0,t), y(s_0,t), z(s_0,t))$

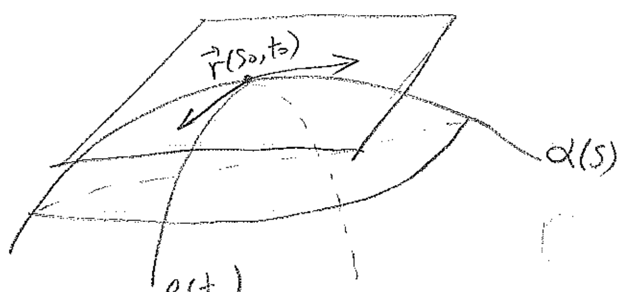
$\vec{\alpha}(s) \subseteq \vec{r}(s,t)$, $\vec{\beta}(t) \subseteq \vec{r}(s,t)$.

$$\vec{\alpha}'(s_0) = \frac{\partial \vec{r}}{\partial s}(s_0,t_0) = \left(\frac{\partial x}{\partial s}(s_0,t_0), \frac{\partial y}{\partial s}(s_0,t_0), \frac{\partial z}{\partial s}(s_0,t_0)\right)$$

$$\vec{\beta}'(t_0) = \frac{\partial \vec{r}}{\partial t}(s_0,t_0) = \left(\frac{\partial x}{\partial t}(s_0,t_0), \frac{\partial y}{\partial t}(s_0,t_0), \frac{\partial z}{\partial t}(s_0,t_0)\right)$$

both $\vec{\alpha}'(s_0)$ and $\vec{\beta}'(t_0)$ are in the tangent plane of $\vec{r}(s,t)$ at $\vec{r}(s_0,t_0)$, so if they're not parallel, $\vec{\alpha}'(s_0) \times \vec{\beta}'(t_0)$ gives a normal vector of the tangent plane at $\vec{r}(s_0,t_0)$.

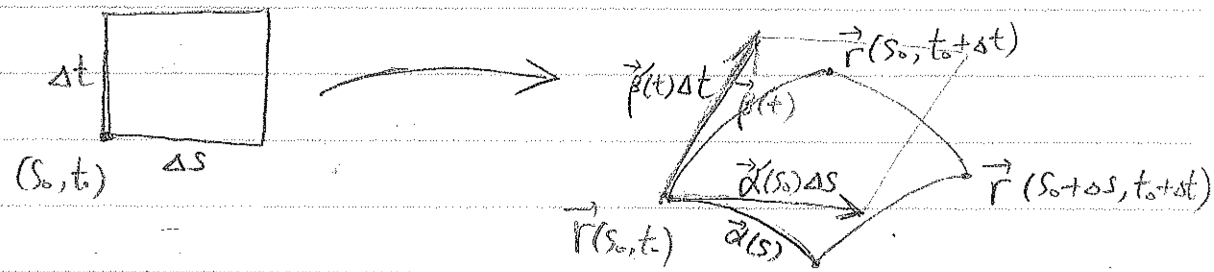
Definition. If S is parameterized by $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$ with domain $D \subseteq \mathbb{R}^2$, the area of the surface is defined by the double integral



$$A(S) = \iint_D \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

We will see this is a reasonable definition:
One way to try to define the area of a given surface is to first approximate the surface by many small pieces of planes, then the limit of the sum of the area of those small pieces as the size of pieces approaching 0 should be a good description of the area of the surface.

So we first take a small square in the st-plane with lower left corner at (s_0, t_0), then try to approximate the area of its image under $\vec{r}(s,t)$ by the tangent plane passing through $\vec{r}(s_0, t_0)$.



When Δs and Δt are small, recall that the length of $\vec{\alpha}(s)$ for $s \in [s_0, s_0+\Delta s]$ can be estimated by $|\vec{\alpha}'(s_0)| \cdot \Delta s$, and similarly the length of $\vec{\beta}(t)$ for $t \in [t_0, t_0+\Delta t]$ can be estimated by $|\vec{\beta}'(t_0)|\Delta t$, and the part of the surface is close to the tangent plane, so we can estimate the area of that piece of the surface by the area of the parallelogram bounded by $\vec{\alpha}'(s_0)\Delta s$ and $\vec{\beta}'(t)\Delta t$, which is

$$|\vec{\alpha}'(s_0)\Delta s \times \vec{\beta}'(t_0)\Delta t| = |\vec{\alpha}'(s_0) \times \vec{\beta}'(t_0)|\, \Delta s \Delta t$$
$$= \left|\frac{\partial \vec{r}}{\partial s}(s_0, t_0) \times \frac{\partial \vec{r}}{\partial t}(s_0, t_0)\right| \Delta s \Delta t$$

Finally we sum up all these small pieces by integration, to get
$$\iint_D \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

Example. The surface area of the sphere $x^2+y^2+z^2=1$ is computed as follows:
We parameterize the sphere by
$$\vec{r}(\varphi, \theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi); \ \varphi \in [0, \pi], \ \theta \in [0, 2\pi]$$

$$\frac{\partial \vec{r}}{\partial \varphi} = (\cos\varphi\cos\theta, \cos\varphi\sin\theta, -\sin\varphi)$$

$$\frac{\partial \vec{r}}{\partial \theta} = (-\sin\varphi \sin\theta,\ \sin\varphi\cos\theta,\ 0)$$

$$\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta} = (\sin^2\varphi\cos\theta,\ \sin^2\varphi\sin\theta,\ \sin\varphi\cos\varphi)$$

$$\left|\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}\right| = \sin\varphi$$

So the surface area is

$$\int_0^{2\pi}\int_0^{\pi} \sin\varphi\, d\varphi\, d\theta = \int_0^{2\pi} d\theta \int_0^{\pi} \sin\varphi\, d\varphi = 2\pi \cdot 2 = 4\pi$$

Now based on the definition of surface area, we can define different kinds of surface integrals.

Definition. If $f(x,y,z)$ is a function defined on a surface S parameterized by $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$ on $D \subseteq \mathbb{R}^2$, define the surface integral of f on S to be:

$$\iint_S f(x,y,z)\, dS = \iint_D f(\vec{r}(s,t)) \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

Remark. From the Riemann Sum approach, the surface integral should be defined by the following process:

$$\iint_S f(x,y,z)\, dS = \lim_{\max \Delta S \to 0} \sum_{i \in I} f(x_i^*, y_i^*, z_i^*) \cdot \Delta S_i$$

$$= \lim_{\max \Delta A \to 0} \sum_{i \in I} f(\vec{r}(s_i^*, t_i^*)) \left|\frac{\partial \vec{r}}{\partial s}(s_i^*, t_i^*) \times \frac{\partial \vec{r}}{\partial t}(s_i^*, t_i^*)\right| \Delta A$$

$$= \iint_D f(\vec{r}(s,t)) \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

Example. A thin sphere $x^2+y^2+z^2=1$ made of some mixed medals has density function $\rho(x,y,z)=x^2$. Compute its mass.

$$\vec{r}(\varphi,\theta)=(\sin\varphi\cos\theta,\ \sin\varphi\sin\theta,\ \cos\varphi)$$

$$\left|\frac{\partial\vec{r}}{\partial\varphi}\times\frac{\partial\vec{r}}{\partial\theta}\right|=\sin\varphi$$ as we computed in the previous example.

$$\text{Mass}=\iint_S \rho(x,y,z)\,dS=\int_0^{2\pi}\int_0^{\pi}(\sin\varphi\cos\theta)^2\sin\varphi\,d\varphi\,d\theta$$
$$=\int_0^{2\pi}\cos^2\theta\,d\theta\cdot\int_0^{\pi}\sin^3\varphi\,d\varphi$$
$$=\frac{4}{3}\pi$$

Example. If S is the graph of the function $z=g(x,y)$ on $D\subseteq\mathbb{R}^2$. We can parameterize S by $\vec{r}(x,y)=(x,y,g(x,y))$

$$\frac{\partial\vec{r}}{\partial x}=\left(1,0,\frac{\partial g}{\partial x}\right),\quad \frac{\partial\vec{r}}{\partial y}=\left(0,1,\frac{\partial g}{\partial y}\right)$$

$$\frac{\partial\vec{r}}{\partial x}\times\frac{\partial\vec{r}}{\partial y}=\left(-\frac{\partial g}{\partial x},-\frac{\partial g}{\partial y},1\right).$$

$$\left|\frac{\partial\vec{r}}{\partial x}\times\frac{\partial\vec{r}}{\partial y}\right|=\sqrt{\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2+1}$$

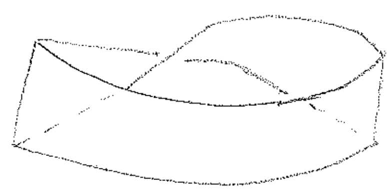
So $$\iint_S f(x,y,z)\,dS=\iint_D f(x,y,g(x,y))\sqrt{\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2+1}\,dA$$

Next we are going to define another kind of surface integral, which is to study the amound of flow passing through a surface.

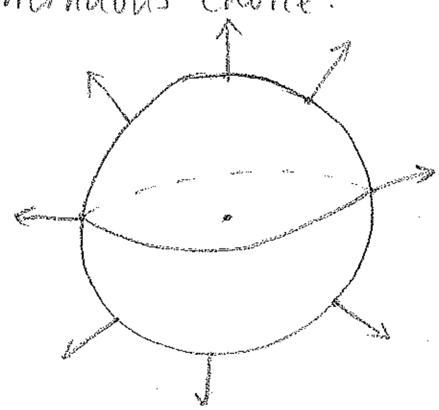
Definition. An oriented surface $\vec{S}$ is a surface with a continuous choice of unit normal vectors. $\vec{n}$

Definition. If it's possible to have a continuous choice of unit normal vectors for a surface S, we say S is orientable.

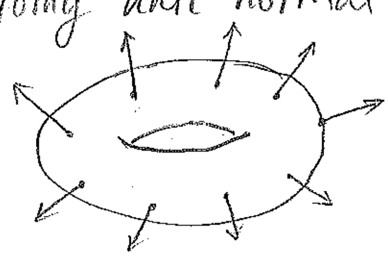
Example. The Mobius Stripe is NOT orientable:



Example. The sphere is orientable, we can choose the outgoing unit normal vector at every point, which is a continuous choice:



Remark. A convension is that when we regard a closed surface as an oriented surface, if not specified, we choose the outgoing unit normal vector as the orientation.



Definition. If $\vec{F}(x,y,z)$ is a vector field defined on a region containing an oriented surface $\vec{S}$ (a surface S with continuous choice of unit normal vectors $\vec{n}$), define the flux of $\vec{F}$ through $\vec{S}$ to be

$$\iint_S \vec{F}\cdot d\vec{S} = \iint_S \vec{F}\cdot\vec{n}\, dS$$

when S is parameterized by $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$, we know $\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}$ is a normal vector of the surface, so a choice of unit normal vector is

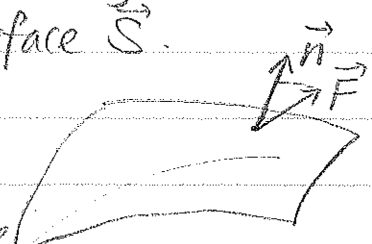
$$\vec{n} = \frac{\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}}{\left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right|}$$

We then can evaluate the integral based on this orientation:

$$\iint_S \vec{F}\cdot d\vec{S} = \iint_D \vec{F}(\vec{r}(s,t))\cdot \frac{\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}}{\left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right|} \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

$$= \iint_D \vec{F}(\vec{r}(s,t))\cdot \left(\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right) dA$$

The practical meaning of the flux is to study the amount of some flow $\vec{F}$ through some surface $\vec{S}$. The observation is only the component of $\vec{F}$ that is parallel to $\vec{n}$ (i.e. normal to the surface) passes through the surface



and if $|\vec{n}|=1$, this component is computed by $\vec{F}\cdot\vec{n} = |\vec{F}|\cos\theta$

Then the total flux is to sum up the flux at each point along the surface, via the language of integral.

This concept of flux plays an important role in several branches of physics, such as fluid dynamics, heat transfer, and electromagnetic dynamics.

Notation. When the surface is closed (i.e. bounds some volume), and $\vec{n}$ is taken to be the outgoing unit normal vector, we write the flux as

$$\oiint_S \vec{F} \cdot d\vec{S}$$

Example. Find the flux of $\vec{F}(x,y,z) = (z, y, x)$ across the unit sphere $x^2+y^2+z^2=1$: $\oiint_S \vec{F} \cdot d\vec{S}$

We can parameterize the sphere by

$\vec{r}(\varphi,\theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$, $0 \le \varphi \le \pi$, $0 \le \theta \le 2\pi$

Then

$$\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta} = (\sin^2\varphi\cos\theta, \ \sin^2\varphi\sin\theta, \ \sin\varphi\cos\varphi)$$

So

$$\frac{\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}}{\left|\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \varphi}\right|} = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$$

We see at $(1,0,0)$, $\varphi = \frac{\pi}{2}$, $\theta = 0$. So

$$\frac{\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}}{\left|\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \varphi}\right|} = (1, 0, 0),$$ which is outward going.

So $\dfrac{\frac{\partial\vec{r}}{\partial\varphi}\times\frac{\partial\vec{r}}{\partial\theta}}{\left|\frac{\partial\vec{r}}{\partial\varphi}\times\frac{\partial\vec{r}}{\partial\theta}\right|}$ gives the orientation that we want.

$$\iint_S \vec{F}\cdot d\vec{S} = \iint_D \vec{F}(\vec{r}(\varphi,\theta))\cdot\left(\frac{\partial\vec{r}}{\partial\varphi}\times\frac{\partial\vec{r}}{\partial\theta}\right)dA$$

$$= \int_0^{2\pi}\int_0^{\pi}(\cos\varphi,\ \sin\varphi\sin\theta,\ \sin\varphi\cos\theta)\cdot(\sin\varphi\cos\theta,\ \sin\varphi\sin\theta,\ \cos\varphi)\,d\varphi\,d\theta$$

$$= \int_0^{2\pi}\int_0^{\pi} 2\sin^2\varphi\cos\varphi\cos\theta + \sin^3\varphi\sin^2\theta\,d\varphi\,d\theta$$

$$= 2\int_0^{\pi}\sin^2\varphi\cos\varphi\,d\varphi\int_0^{2\pi}\cos\theta\,d\theta + \int_0^{\pi}\sin^3\varphi\,d\varphi\int_0^{2\pi}\sin^2\theta\,d\theta$$

$$= \frac{4\pi}{3}$$

Example. When $\vec{S}$ is the graph of the function $z=g(x,y)$ defined on $D\subseteq\mathbb{R}^2$ with upward orientation, observe that the parameterization $\vec{r}(x,y)=(x,y,g(x,y))$ gives the orientation $\frac{\partial\vec{r}}{\partial x}\times\frac{\partial\vec{r}}{\partial y}=\left(-\frac{\partial g}{\partial x},-\frac{\partial g}{\partial y},1\right)$ which is indeed upward, so for a vector field $\vec{F}=(P,Q,R)$.

$$\iint_S \vec{F}\cdot d\vec{S} = \iint_D \vec{F}(\vec{r}(x,y))\cdot\left(\frac{\partial\vec{r}}{\partial x}\times\frac{\partial\vec{r}}{\partial y}\right)dA$$

$$= \iint_D \left(P(\vec{r}(x,y)),\ Q(\vec{r}(x,y)),\ R(\vec{r}(x,y))\right)\cdot\left(-\frac{\partial g}{\partial x},-\frac{\partial g}{\partial y},1\right)dA$$

$$= \iint_D -\frac{\partial g}{\partial x}P(x,y,g(x,y)) - \frac{\partial g}{\partial y}Q(x,y,g(x,y)) + R(x,y,g(x,y))\,dA$$

OTHER TYPES OF INTEGRAL

Definition. Given a path C and a scalar function $\phi(x,y,z)$ defined on some region containing C, defined the line integral of ϕ along C to be

$$\int_C \phi \, d\vec{r} = \int_a^b \phi(\vec{r}(t)) \, \vec{r}\,'(t) \, dt$$

where $\vec{r}(t) : [a,b] \to \mathbb{R}^3$ is a parameterization of C

Note the above integral should leads to a vector as a result, since $\phi(\vec{r}(t)) \, \vec{r}\,'(t)$ is a vector.

Example. $\int_C x+y^2 \, d\vec{r}$ where C is the parabola $y=x^2$ in the plane $z=0$, connecting $(0,0,0)$ and $(1,1,0)$

C can be parameterized by $\vec{r}(t) = (t, t^2, 0)$, $t \in [0,1]$.

$$\begin{aligned}
\int_C x+y^2 \, d\vec{r} &= \int_0^1 (t+(t^2)^2) \, \vec{r}\,'(t) \, dt \\
&= \int_0^1 (t+t^4) \cdot (1, 2t, 0) \, dt \\
&= \int_0^1 (t+t^4, \; 2t^2+2t^5, 0) \, dt \\
&= \left(\int_0^1 t+t^4 \, dt, \; \int_0^1 2t^2+2t^5 \, dt, \; \int_0^1 0 \, dt \right) \\
&= \left(\tfrac{7}{10}, 1, 0 \right)
\end{aligned}$$

Definition. Given a path C parameterized by $\vec{r}(t)$, $t \in [a,b]$, and a vector field $\vec{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ define the cross integral of $\vec{F}$ along C to be

$$\int_C \vec{F} \times d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \times \vec{r}\,'(t)\, dt$$

Similar as the previous case, the result of this integral is also a vector.

We can also define the corresponding concepts on oriented surfaces, and an important one is the following:

Definition Given an oriented surface $\vec{S}$ with orientation $\hat{n}$, if $\vec{S}$ can be parameterized by $\vec{r}(s,t) = (x(s,t), y(s,t), z(s,t))$ with domain $(s,t) \in D$, and $\vec{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ is a vector field on $\vec{S}$, we define the cross integral of $\vec{F}$ along $\vec{S}$ to be:

$$\iint_S \vec{F} \times d\vec{S} = \iint_S \vec{F}(\vec{r}(s,t)) \times \hat{n}\, dS$$

$$= \iint_D \vec{F}(\vec{r}(s,t)) \times \hat{n} \left| \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right| dA$$

When $\hat{n}$ agrees with the orientation induced by parameterization, i.e. $\hat{n} = \dfrac{\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}}{\left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right|}$, we can further write

$$\iint_S \vec{F} \times d\vec{S} = \iint_D \vec{F}(\vec{r}(s,t)) \times \left(\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} \right) dA$$

Example. Let $\vec{F}(x,y,z) = (x,y,z)$, and $\vec{S}$ is the unit sphere $x^2+y^2+z^2=1$ with outward orientation.

$$\iint_S \vec{F} \times d\vec{S} = \iint_D \vec{F}(\vec{r}(\varphi,\theta)) \times \left(\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}\right) dA$$

$$= \iint_D (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi) \times (\sin^2\varphi\cos\theta, \sin^2\varphi\sin\theta, \sin\varphi\cos\varphi) dA$$

$$= \iint_D (0,0,0)\, dA$$

$$= (0,0,0)$$

Example. Let $\vec{F}(x,y,z) = (-y, x, 0)$, and $\vec{S}$ is the unit sphere $x^2+y^2+z^2=1$ with outward orientation.

$$\iint_S \vec{F} \times d\vec{S} = \iint_D \vec{F}(\vec{r}(\varphi,\theta)) \times \left(\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}\right) dA$$

$$= \iint_D (-\sin\varphi\sin\theta, \sin\varphi\cos\theta, 0) \times (\sin^2\varphi\cos\theta, \sin^2\varphi\sin\theta, \sin\varphi\cos\varphi) dA$$

$$= \iint_D (\sin^2\varphi\cos\varphi\cos\theta, \sin^2\varphi\cos\varphi\sin\theta, -\sin^3\varphi)\, dA$$

$$= \left(\int_0^{2\pi}\int_0^{\pi} \sin^2\varphi\cos\varphi\cos\theta \, d\varphi d\theta, \int_0^{2\pi}\int_0^{\pi} \sin^2\varphi\cos\varphi\sin\theta \, d\varphi d\theta, \int_0^{2\pi}\int_0^{\pi} -\sin^3\varphi \, d\varphi d\theta\right)$$

$$= \left(0, 0, -\frac{8}{3}\right)$$

We can also integrate a scalar function on an oriented surface:

Definition If $f(x,y,z)$ is a scalar function defined on an oriented surface S, define:

$$\iint_S f\, d\vec{S} = \iint_S f\hat{n}\, dS = \iint_D f(\vec{r}(s,t))\, \hat{n}(\vec{r}(s,t)) \left|\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right| dA$$

if $\vec{S}$ is parameterized by $\vec{r}(s,t)$ with domain D.

If $\vec{r}(s,t)$ agrees with the orientation of $\vec{S}$, we see

$$\iint_S f\, d\vec{S} = \iint_D f(\vec{r}(s,t)) \left(\frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t}\right) dA$$

Example. If $\vec{S}$ is the unit sphere $x^2+y^2+z^2=1$ with outward orientation, and $f(x,y,z)=z$, then

$$\iint_S f\, d\vec{S} = \iint_D \cos\varphi \left(\frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta}\right) dA$$

$$= \iint_D (\sin^2\varphi\cos\varphi\cos\theta,\ \sin^2\varphi\cos\varphi\sin\theta,\ \sin\varphi\cos^2\varphi)\, dA$$

$$= \left(\iint_D \sin^2\varphi\cos\varphi\cos\theta\, dA,\ \iint_D \sin^2\varphi\cos\varphi\sin\theta\, dA,\ \iint_D \sin\varphi\cos^2\varphi\, dA\right)$$

$$= \left(0, 0, \frac{4\pi}{3}\right)$$

We can reinterpret the concept of gradient by considering integral of this form:

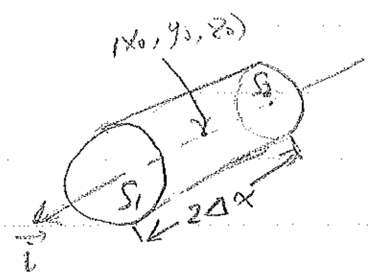
Definition $f(x,y,z)$ is a scalar function, V is some region in $\mathbb{R}^3$ with (x_0, y_0, z_0) in its interior, and $\vec{S}$ is the boundary of V with outward orientation. Then define the gradient of f at (x_0, y_0, z_0) to be the vector

$$\nabla f = \lim_{V \to 0} \frac{\oiint_S f \, d\vec{S}}{V}$$

if the limit exists. So we get a vector field ∇f, called the gradient of f.

Proposition $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$

Proof We're going to show that $\nabla f \cdot \vec{i} = \frac{\partial f}{\partial x}$ at each point.
Let V be a cylindrical region with core in the direction of $\vec{i}$.



$$\nabla f(x_0, y_0, z_0) \cdot \vec{i} = \left(\lim_{V \to 0} \frac{\oiint_S f \, d\vec{S}}{V}\right) \cdot \vec{i}$$

$$= \lim_{V \to 0} \frac{\oiint_S f \hat{n} \, dS}{V} \cdot \vec{i}$$

$$= \lim_{V \to 0} \frac{\oiint_S f \, \vec{i} \cdot \hat{n} \, dS}{V}$$

Note for the curved parts of the cylinder, $\hat{n} \perp \vec{i}$. on the two disks, S_1 & S_2, $\hat{n} \parallel \vec{i}$

$$= \lim_{V \to 0} \frac{\iint_{S_1} f \, dS - \iint_{S_2} f \, dS}{V}$$

$$= \lim_{V \to 0} \frac{\iint_{D_r} f(x_0 + \Delta x, y, z) - f(x_0 - \Delta x, y, z) \, dA}{\text{Area}(D_r) \cdot 2\Delta x}$$

where D_r is the disk on y-z plane centered at (y_0, z_0) with radius r.

$$= \lim_{V \to 0} \frac{\iint_{D_r} \frac{\partial f}{\partial x}(x_0, y, z) \cdot 2\Delta x + \varepsilon(y,z) \Delta x \, dA}{\iint_{D_r} 1 \, dA \cdot 2\Delta x}$$

$$= \left(\lim_{r \to 0} \frac{\iint_{D_r} \frac{\partial f}{\partial x}(x_0, y, z) \, dA}{\iint_{D_r} 1 \, dA} + \lim_{\Delta x \to 0} \frac{\iint_{D_r} \varepsilon(y,z) \, dA}{2 \iint_{D_r} 1 \, dA} \right)$$

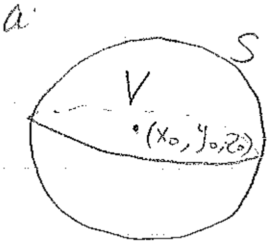
$$= \frac{\partial f}{\partial x}(x_0, y_0, z_0)$$

CURL

Definition. If $\vec{F}(x,y,z)$ is a vector field defined in a neighbourhood of a point (x_0, y_0, z_0), V is a region in $\mathbb{R}^3$ such that (x_0, y_0, z_0) is an interior point of V, with boundary of V to be a surface S, outward oriented. Then define the curl of $\vec{F}$ at (x_0, y_0, z_0) to be the vector

$$\operatorname{Curl}\vec{F}(x_0,y_0,z_0) = -\lim_{Vol(V)\to 0} \frac{\oiint_S \vec{F}\times d\vec{S}}{V}$$

if the limit exists. We then also get a corresponding vector field $\operatorname{Curl}\vec{F}$ for the given vector field $\vec{F}$.



Proposition. If $\vec{F}$ and $\vec{G}$ are vector fields, $\lambda, \mu \in \mathbb{R}$, then $\operatorname{Curl}(\lambda\vec{F} + \mu\vec{G}) = \lambda \operatorname{Curl}(\vec{F}) + \mu \operatorname{Curl}(\vec{G})$.

Proof. It follows directly from the definition of curl and the fact that cross product is distributive.

Proposition. If $f(x,y,z)$ is a scalar function and $\vec{u}$ is a constant vector field, then $\operatorname{Curl}(f\vec{u}) = \nabla f \times \vec{u}$

Proof.
$$\operatorname{Curl}(f\vec{u}) = -\lim_{Vol(V)\to 0} \frac{\oiint_S f\vec{u}\times d\vec{S}}{V} = -\lim_{Vol(V)\to 0} \frac{\oiint_S f\vec{u}\times \hat{n}\, dS}{V}$$

$$= -\lim_{Vol(V)\to 0} \vec{u} \times \frac{\oiint_S f\hat{n}\, dS}{V}$$

$$= -\vec{u}\times\nabla f = \nabla f \times \vec{u}$$

Proposition. If $\vec{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ is a vector field, then $\operatorname{Curl}\vec{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$

Proof. $\vec{F} = P\vec{i} + Q\vec{j} + R\vec{k}$, so by the previous propositions.

$$\begin{aligned}
\operatorname{Curl}\vec{F} &= \operatorname{Curl}(P\vec{i} + Q\vec{j} + R\vec{k}) \\
&= \operatorname{Curl}(P\vec{i}) + \operatorname{Curl}(Q\vec{j}) + \operatorname{Curl}(R\vec{k}) \\
&= \nabla P \times \vec{i} + \nabla Q \times \vec{j} + \nabla R \times \vec{k} \\
&= \left(\frac{\partial P}{\partial x}, \frac{\partial P}{\partial y}, \frac{\partial P}{\partial z}\right) \times (1,0,0) + \left(\frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial y}, \frac{\partial Q}{\partial z}\right) \times (0,1,0) + \\
&\quad \left(\frac{\partial R}{\partial x}, \frac{\partial R}{\partial y}, \frac{\partial R}{\partial z}\right) \times (0,0,1) \\
&= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)
\end{aligned}$$

Remark. A good way for memorizing the above result is that

$$\operatorname{Curl}\vec{F} = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \vec{\nabla} \times \vec{F}$$

Proposition $\operatorname{Curl}(\nabla f) = \vec{0}$ for any smooth function f.

Proof. $\operatorname{Curl}(\nabla f) = \operatorname{Curl}\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$

$$= \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right), \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right), \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right)\right)$$

Corollary. If $\vec{F}$ is a conservative vector field, then $\operatorname{curl}\vec{F} = \vec{0}$

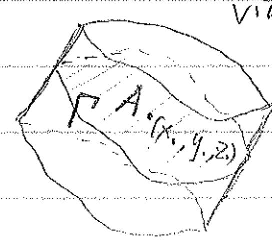
Next we are going to find another interpretation of curl, which will be useful in the proof of Stokes' Theorem.

Theorem: If $\vec{F}$ is a vector field, $\hat{u}$ is a constant unit vector. A is a flat surface with boundary curve Γ, and (x_0, y_0, z_0) is a point in the interior of A and $\hat{u} \perp A$. Then:

$$\hat{u} \cdot \text{Curl}\,\vec{F}(x_0, y_0, z_0) = \lim_{A \to 0} \frac{1}{A} \oint_{\Gamma} \vec{F} \cdot d\vec{r}$$

where the orientation of Γ agrees with $\hat{u}$ via right-hand rule

We will here only provide a sketched proof. The reader may refer to Barry Spain's Vector Analysis book to fill in more technical details



Proof: We thicken A to be a solid of thickness $2l$, with A the middle layer. We call this solid $A \times I$. (I means the interval $[-l, l]$)

$$\partial(A \times I) = \partial A \times I \cup A \times \partial I$$

$$= \underbrace{\Gamma \times I}_{\text{side}} \cup \underbrace{A \times \{-l\}}_{\text{bottom}} \cup \underbrace{A \times \{l\}}_{\text{top}}$$

$$\hat{u} \cdot \text{curl}\,\vec{F}(x_0, y_0, z_0) = \hat{u} \cdot \left(- \lim_{A \times I \to 0} \frac{1}{A \times I} \oiint_{\partial(A \times I)} \vec{F} \times d\vec{S} \right)$$

$$= \lim_{A \times I \to 0} \frac{1}{A \times I} \oiint_{\partial(A \times I)} \vec{F} \cdot (\hat{u} \times \hat{n})\, dS$$

$$= \lim_{A \times I \to 0} \frac{1}{A \times I} \oiint_{\Gamma \times I} \vec{F} \cdot \hat{T}\, dS$$

let $\hat{T} = \hat{u} \times \hat{n}$, then $\{\hat{u}, \hat{n}, \hat{T}\}$ forms a right-handed orthonormal basis ...

$$= \lim_{A\to 0} \frac{1}{A} \lim_{l\to 0} \frac{1}{2l} \cdot \int_{-l}^{l} \oint_{\Gamma\times\{h\}} \vec{F}\cdot\hat{T}\, ds\, dh$$

$$= \lim_{A\to 0} \frac{1}{A} \cdot \frac{1}{2l} \cdot 2l \oint_{\Gamma} \vec{F}\cdot d\vec{r}$$

$$= \lim_{A\to 0} \frac{1}{A} \oint_{\Gamma} \vec{F}\cdot d\vec{r}$$

This new formula for curl shows us clearly that $\operatorname{curl}\vec{F}$ tells us about how much the vector field $\vec{F}$ is spinning with respect to each direction.

Remark An interesting remark is that we can make use of this presentation of curl to conclude the Corollary: if $\vec{F}$ is a conservative vector field, then $\operatorname{curl}\vec{F}=\vec{0}$ (which we have proved earlier by other methods)

For any $\hat{u}$, we see if $\vec{F}$ is conservative,

$$\hat{u}\cdot\operatorname{curl}\vec{F} = \lim_{A\to 0}\frac{1}{A}\oint_{\Gamma}\vec{F}\cdot d\vec{r} = \lim_{A\to 0}\frac{1}{A}\cdot 0 = 0$$

So $\operatorname{curl}\vec{F}=\vec{0}$.

If we make use of the Stokes' Theorem, we can also show that if $\vec{F}$ is defined on an open simply-connected region in $\mathbb{R}^3$, then $\operatorname{curl}\vec{F}=0$ implies $\vec{F}$ is conservative.

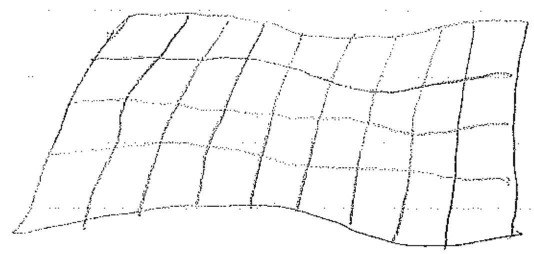
STOKES' THEOREM.

Theorem. Suppose $\vec{F}(x,y,z)$ is a vector field defined in a region containing a surface $\vec{S}$ and $\vec{F}$ has continuous partial derivatives, then:

$$\iint_S \operatorname{curl}\vec{F}\cdot d\vec{S} = \oint_{\partial S} \vec{F}\cdot d\vec{r}$$

where ∂S is the boundary curve of $\vec{S}$ with counterclockwise orientation if seen from the side of $\vec{S}$ that $\hat{n}$ is pointing to, i.e. the orientation of $\vec{S}$ induces the orientation of ∂S by the right-hand rule.

Proof $\iint_S \operatorname{curl}\vec{F}\cdot d\vec{S} = \iint_S \operatorname{curl}\vec{F}\cdot\hat{n}\, dS$



$$= \lim_{\max\Delta S\to 0} \sum_{\Delta S} \operatorname{curl}\vec{F}\cdot\hat{n}\,\Delta S$$

$$\overset{(*)}{=} \lim_{\max\Delta S\to 0} \sum_{\Delta S} \frac{1}{\Delta S}\oint_{\partial\Delta S} \vec{F}\cdot d\vec{r}\,\Delta S$$

$$= \lim_{\max\Delta S\to 0} \sum_{\Delta S} \oint_{\partial\Delta S} \vec{F}\cdot d\vec{r}$$

(Note that when the line integrals are summed up, those on the sides shared by two pieces of ΔS will be cancelled)

$$= \oint_{\partial S} \vec{F}\cdot d\vec{r}$$

Remark. A rigorous argument for $(*)$ is a standard one in analysis:

$\hat{n}\cdot\operatorname{curl}\vec{F} = \lim_{A\to 0} \oint_\Gamma \frac{1}{A}\oint_\Gamma \vec{F}\cdot d\vec{r}$ implies there exists $\varepsilon(A) > 0$

such that $\hat{n}\cdot\operatorname{curl}\vec{F}\cdot A = \oint_\Gamma \vec{F}\cdot d\vec{r} + \varepsilon(A)A$ and $\varepsilon(A)\to 0$ as $A\to 0$. Then

$$\sum_{\Delta S} \operatorname{curl}\vec{F}\cdot\hat{n}\,\Delta S = \sum_{\partial\Delta S}\left(\oint_\Gamma \vec{F}\cdot d\vec{r} + \varepsilon(\Delta S)\Delta S\right) = \oint_{\partial S}\vec{F}\cdot d\vec{r} + \sum_{\partial\Delta S}\varepsilon(\Delta S)\Delta S$$

Taking the limit as $\max \Delta S \to 0$, we get

$$\iint_S \operatorname{curl} \vec{F} \cdot \hat{n}\, dS = \oint_{\partial S} \vec{F} \cdot d\vec{r} + \lim_{\max \Delta S \to 0} \mathcal{E}(\Delta S)\, \Delta S$$

and $\displaystyle \lim_{\max \Delta S \to 0} \sum \mathcal{E}(\Delta S) \Delta S \leq \lim_{\max \Delta S \to 0} \left(\max_{\Delta S} \mathcal{E}(\Delta S)\right) \cdot \sum_{\Delta S} \Delta S = 0$ since

$\max \mathcal{E}(\Delta S) \to 0$ and $\sum_{\Delta S} \Delta S$ is the area of S, so bounded.

Corollary. If S is a closed surface, $\vec{F}$ is a vector field on S, then

$$\oiint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = 0$$

Proof. If S is closed, then $\partial S = \phi$. so by Stokes' Theorem:

$$\oiint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = \int_{\phi} \vec{F} \cdot d\vec{r} = 0$$

Corollary (Green's Theorem) $\vec{F}(x,y) = (P(x,y), Q(x,y))$ is a vector field on a curve C, with S the region enclosed by C. If $\vec{F}$ has continuous partial derivatives, then

$$\iint_S \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\, dA = \oint_C \vec{F} \cdot d\vec{r}$$

Proof. We regard $\vec{F}$ as a vector field in $\mathbb{R}^3$: $\vec{F}(x,y,z) = (P(x,y), Q(x,y), 0)$

Then $\operatorname{Curl} \vec{F} = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & 0 \end{vmatrix} = \left(0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$

By Stokes' Theorem:

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = \iint_S \left(0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \cdot (0,0,1)\, dA$$

$$= \iint_S \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\, dA$$

Proposition. If $\vec{F}(x,y,z)$ is a vector field defined on a simply-connected region $R \subseteq \mathbb{R}^3$, then $\text{curl}\vec{F} = \vec{0}$ on R implies $\vec{F}$ is a conservative vector field

Remark. We can interpret "simply-connected" region in $\mathbb{R}^3$ to be a region R that for any closed curve $C \subseteq R$, C bounds some surface $S \subseteq R$ that has no hole

Proof. For any closed curve $C \subset R$, C bounds some surface $S \subseteq R$ that has no hole. So by Stokes' Theorem,

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S \text{curl}\vec{F} \cdot d\vec{S} = 0$$

Now we also obtain a geometric proof of the fact that $\text{curl}(\nabla f) = \vec{0}$:

For any surface $\vec{S}$, $\iint_S \text{curl}(\nabla f) \cdot d\vec{S} = \oint_{\partial S} \nabla f \cdot d\vec{r} = 0$

since ∇f is always a conservative vector field.

Because the above equality holds for any $\vec{S}$, we conclude the integrand $\text{curl}(\nabla f) = \vec{0}$

Review of Triple Integral

We will first have a brief review of triple integral before discussing about divergence.
Recall that if E is some solid in $\mathbb{R}^3$ and f is a function defined on E, we define the triple integral of f along E to be the Riemann Sum

$$\iiint_E f\,dV = \lim_{\max \Delta V_i \to 0} \sum_{\Delta V_i} f(x_i^*, y_i^*, z_i^*)\Delta V_i$$

A good way of understanding triple integral is to regard f as the density function, then $\iiint_E f\,dV$ gives the mass of E

The evaluation of $\iiint_E f\,dV$ is in general complicated, but in some special cases we can translate it into the iterated form. For example, if E is of the form

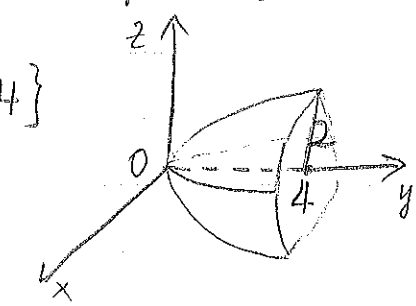
$$E = \{(x,y,z)\in\mathbb{R}^3 \mid a \le x \le b,\ g_1(x) \le y \le g_2(x),\ u_1(x,y) \le z \le u_2(x,y)\}$$

then $$\iiint_E f\,dV = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz\,dy\,dx = \iint_D \int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz\,dA$$

where $D = \{(x,y)\in\mathbb{R}^2 \mid a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$

Example. Evaluate $\iiint_E \sqrt{x^2+z^2}\,dV$ where E is the region bounded by the paraboloid $y = x^2+z^2$ and the plane $y=4$.

$E = \{(x,y,z)\in\mathbb{R}^3 \mid (x,z)\in D,\ x^2+z^2 \le y \le 4\}$
where $D = \{(x,z)\in\mathbb{R}^2 \mid x^2+z^2 \le 4\}$



So the integral is

$$\iiint_E \sqrt{x^2+z^2}\, dV = \iint_D \left(\int_{x^2+z^2}^{4} \sqrt{x^2+z^2}\, dy\right) dA$$

$$= \iint_D \sqrt{x^2+z^2}\,(4-x^2-z^2)\, dA$$

$$= \int_0^{2\pi}\int_0^2 r\,(4-r^2)\, r\, dr\, d\theta$$

$$= \frac{128}{15}\pi$$

Example Compute the volume of a ball of radius r.

We may take the ball $x^2+y^2+z^2=r^2$, call it B.

Then the volume is:

$$\iiint_B 1\, dV = \int_0^r \int_0^\pi \int_0^{2\pi} \rho^2 \sin\phi\, d\theta\, d\phi\, d\rho$$

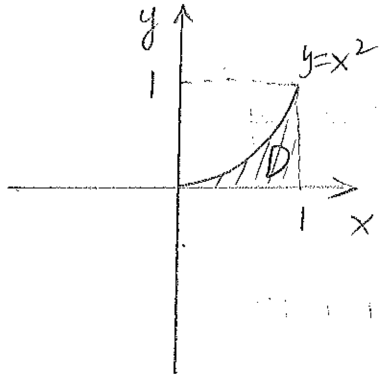
$$= \int_0^1 \rho^2\, d\rho \cdot \int_0^\pi \sin\phi\, d\phi \cdot \int_0^{2\pi} 1\, d\theta$$

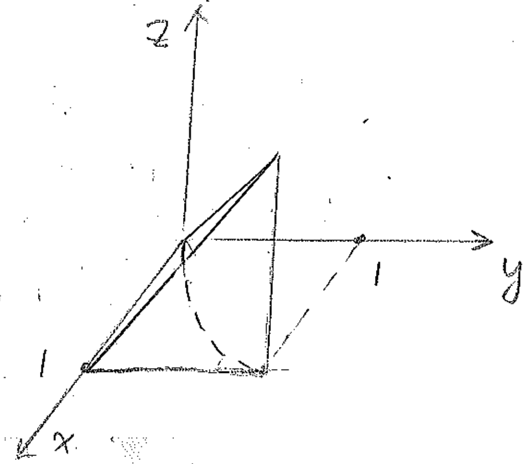
$$= \frac{r^3}{3} \cdot 2 \cdot 2\pi$$

$$= \frac{4}{3}\pi r^3$$

Example Rewrite $\int_0^1\int_0^{x^2}\int_0^y f(x,y,z)\,dz\,dy\,dx$ into the form $\iiint f(x,y,z)\,dx\,dz\,dy$

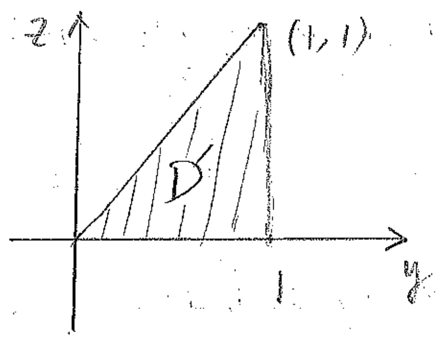
First, we need to find the solid represented by the integral





$E=\{(x,y,z)\in\mathbb{R}^3 \mid 0\le x\le 1,\ 0\le y\le x^2,\ 0\le z\le y\}$

The projection of E on y-z plane is



$D'=\{(y,z)\in\mathbb{R}^2 \mid 0\le y\le 1,\ 0\le z\le y\}$

$E=\{(x,y,z)\in\mathbb{R}^3 \mid (y,z)\in D',\ \sqrt{y}\le x\le 1\}$

So $\iiint_E f\,dV=\int_0^1\int_0^y\int_{\sqrt{y}}^1 f(x,y,z)\,dx\,dz\,dy$

DIVERGENCE

Definition $\vec{F}(x,y,z)$ is a vector field defined in a neighbourhood of $(x_0, y_0, z_0) \in \mathbb{R}^3$. V is a solid with (x_0, y_0, z_0) in its interior, $\partial V = \vec{S}$ is outward oriented. Define the divergence of $\vec{F}(x,y,z)$ at (x_0, y_0, z_0) to be:

$$\operatorname{div}\vec{F}(x_0, y_0, z_0) = \lim_{V\to 0} \frac{1}{V} \oiint_S \vec{F}\cdot d\vec{S}$$

So as (x_0, y_0, z_0) varies around, we obtain a scalar function $\operatorname{div}\vec{F}$.

Proposition (i) If λ and μ are constants, $\vec{F}$ and $\vec{G}$ are vector fields, then

$$\operatorname{div}(\lambda\vec{F} + \mu\vec{G}) = \lambda \operatorname{div}(\vec{F}) + \mu \operatorname{div}(\vec{G})$$

(ii) If f is a scalar function and $\vec{u}$ is a constant vector, then

$$\operatorname{div}(f\vec{u}) = \vec{u}\cdot\nabla f$$

Proof. (i) Follows directly from the distributive law of dot product

(ii)
$$\operatorname{div}(f\vec{u}) = \lim_{V\to 0} \frac{1}{V} \oiint_S (f\vec{u})\cdot d\vec{S} = \lim_{V\to 0} \frac{1}{V}\, \vec{u}\cdot \oiint_S f\, d\vec{S}$$

$$= \vec{u}\cdot \lim_{V\to 0} \frac{1}{V} \oiint_S f\, d\vec{S}$$

$$= \vec{u}\cdot\nabla f$$

Proposition. If $\vec{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))$ is a vector field, then $\operatorname{div}\vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$

So formally we write $\operatorname{div}\vec{F} = \vec{\nabla}.\vec{F}$

Proof $\vec{F} = P\hat{i} + Q\hat{j} + R\hat{k}$

$$\begin{aligned}
\operatorname{div}\vec{F} &= \operatorname{div}(P\hat{i} + Q\hat{j} + R\hat{k}) \\
&= \operatorname{div}(P\hat{i}) + \operatorname{div}(Q\hat{j}) + \operatorname{div}(R\hat{k}) \\
&= \nabla P.\hat{i} + \nabla Q.\hat{j} + \nabla R.\hat{k} \\
&= \left(\frac{\partial P}{\partial x}, \frac{\partial P}{\partial y}, \frac{\partial P}{\partial z}\right).(1,0,0) + \left(\frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial y}, \frac{\partial Q}{\partial z}\right).(0,1,0) \\
&\quad + \left(\frac{\partial R}{\partial x}, \frac{\partial R}{\partial y}, \frac{\partial R}{\partial z}\right).(0,0,1) \\
&= \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}
\end{aligned}$$

Proposition. $\operatorname{div}(\operatorname{curl}\vec{F}) = 0$

Proof $\operatorname{div}(\operatorname{curl}\vec{F}) = \operatorname{div}(R_y - Q_z, P_z - R_x, Q_x - P_y)$

$$\begin{aligned}
&= R_{yx} - Q_{zx} + P_{zy} - R_{xy} + Q_{xz} - P_{yz} \\
&= 0
\end{aligned}$$

Definition. If f is a scalar function, define the Laplacian of f to be $\Delta f = \vec{\nabla}.(\vec{\nabla} f) = \operatorname{div}(\nabla f)$

$$= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

Exercise. Show that $\Delta(fg) = f\Delta g + g\Delta f + 2\nabla f.\nabla g$

GAUSS' THEOREM (DIVERGENCE THEOREM)

Theorem. V is a solid with boundary $\partial V = \vec{S}$ outward oriented. $\vec{F}$ is a vector field defined on a region containing V. Then:

$$\iiint_V \operatorname{div}\vec{F}\, dV = \oiint_S \vec{F}\cdot d\vec{S}$$

Proof. By the definition of $\operatorname{div}\vec{F} = \lim_{V\to 0} \frac{1}{V} \oiint_S \vec{F}\cdot d\vec{S}$

For each (x_0, y_0, z_0), $\operatorname{div}\vec{F}(x_0, y_0, z_0) = \frac{1}{V}\oiint_S \vec{F}\cdot d\vec{S} + \varepsilon(V)$

and $\varepsilon(V) \longrightarrow 0$ as $V \longrightarrow 0$.

so $\qquad \operatorname{div}\vec{F}(x,y,z)\, V = \oiint_S \vec{F}\cdot d\vec{S} + V\cdot\varepsilon(V)$

$$\iiint_V \operatorname{div}\vec{F}\, dV = \lim_{\max \Delta V_i \to 0} \sum_{\Delta V_i} \operatorname{div}\vec{F}(x_i^*, y_i^*, z_i^*)\,\Delta V_i$$

$$= \lim_{\max \Delta V_i \to 0} \sum_{\Delta V_i} \oiint_{\Delta V_i} \vec{F}\cdot d\vec{S} + \Delta V_i \cdot \varepsilon(\Delta V_i)$$

$(*)$

$$= \sum_{\Delta V_i} \oiint_{\Delta V_i} \vec{F}\cdot d\vec{S}$$

$$= \oiint_S \vec{F}\cdot d\vec{S}$$

$(*)$ is because $\left|\sum_{\Delta V_i} \Delta V_i \cdot \varepsilon(\Delta V_i)\right| \le \left|\max \varepsilon(\Delta V_i)\right| \sum_{\Delta V_i} \Delta V_i$

$$= \left|\max \varepsilon(\Delta V_i)\right| \cdot V \longrightarrow 0$$

Corollary. $\iiint_V \nabla f \, dV = \oiint_{\partial V} f \, d\vec{S}$. for a solid V.

Proof. For any constant vector $\vec{u}$, apply Gauss' Theorem to $f\vec{u}$.

$$\iiint_V \operatorname{div}(f\vec{u})\, dV = \oiint_{\partial V} (f\vec{u}) \cdot d\vec{S}$$

$$\iiint_V \vec{u} \cdot \nabla f \, dV = \iint_{\partial V} \vec{u} \cdot (f\hat{n}) \, dS$$

$$\vec{u} \cdot \iiint_V \nabla f \, dV = \vec{u} \cdot \oiint_{\partial V} f \, d\vec{S}$$

Since this holds for all $\vec{u} \in \mathbb{R}^3$, we conclude

$$\iiint_V \nabla f \, dV = \oiint_{\partial V} f \, d\vec{S}$$

Corollary. If $\vec{S}$ is a closed surface, $\oiint_S 1 \, d\vec{S} = \vec{0}$

Corollary. If $\vec{S}$ is a closed surface, $\iint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = 0$ for any vector field $\vec{F}$.

Proof. $\oiint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = \iiint_V \operatorname{div}(\operatorname{curl} \vec{F})\, dV = \iiint_V 0 \, dV = 0$

Example. Evaluate $\oiint_S (3x, 2y, z) \cdot d\vec{S}$ where $\vec{S}$ is the unit sphere with centre $(0,0,0)$.

$$\oiint_S (3x, 2y, z) \cdot d\vec{S} = \iiint_B \frac{\partial(3x)}{\partial x} + \frac{\partial(2y)}{\partial y} + \frac{\partial(z)}{\partial z}\, dV$$

$$= \iiint_B 6\, dV$$

$$= 6 \iiint_B 1\, dV$$

$$= 6 \cdot \frac{4}{3}\pi$$

$$= 8\pi$$

Remark. The Green's Theorem can also be viewed as a special case of Gauss' Theorem:

If we parameterize a simple closed curve C on $\mathbb{R}^2$ by $\vec{r}(t) = (x(t), y(t))$, then let $\hat{n}(t) = \frac{1}{|\vec{r}'(t)|} \cdot (y'(t), -x'(t))$, $\hat{n}(t)$ is the unit normal vector of C at $\vec{r}(t)$ pointing outward.

$S = s(t) = \int_a^t |\vec{r}'(\tau)|\, d\tau$ is the arclength, then the Green's Theorem implies:

$$\oint_C \vec{F} \cdot \hat{n}\, ds = \int_a^b \vec{F}(\vec{r}(t)) \cdot \hat{n}(t)\, |\vec{r}'(t)|\, dt$$

$$= \int_a^b \frac{1}{|\vec{r}'(t)|} \cdot \left(P(\vec{r}(t))\, y'(t) - Q(\vec{r}(t))\, x'(t)\right) \cdot |\vec{r}'(t)|\, dt$$

$$= \oint_C P\, dy - Q\, dx$$

$$= \oint_C (-Q, P) \cdot d\vec{r} = \iint_D P_x + Q_y\, dA = \iint_D \operatorname{div}\vec{F}\, dA$$

SUFFIX NOTATION

When we write the equations involving vectors, it often happen that we are doing the same thing on each of the coordinates. For example, if $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ then $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$.

It seems to be tedius to repeat the same thing 3 times. So we now introduce another way of representing vector operations, called the suffix notation.

If $\vec{u} = (u_1, u_2, u_3)$ is a vector, we represent it by u_i where it's understood that $i = 1, 2, 3$. Then the equation $\vec{u} + \vec{v} = \vec{w}$ can be written as

$$u_i + v_i = w_i .$$

The suffix "i" is called a free suffix. We say it's "free" because if we replace all the "i" by another letter, say "j", we refer to the same equation. But, you must make sure that the same letter to be used for each term of the expression.

Next we consider another type of operation, the dot product. We know if $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ then

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{3} u_i v_i$$

So in suffix notation, we denote the dot product of $\vec{u}$ and $\vec{v}$ by:

$$u_i v_i$$

The repeated suffix means we should take the sum of the terms as i goes over from 1 to 3. We call it the Summation Convention: Whenever a suffix is repeated in a single term in an expression, we are to take the sum as the suffix goes from 1 to 3.
We call this kind of suffix a "dummy suffix".

- A principle is that any suffix should appear no more than twice in each term of an expression.

Now we can do more complicated examples: $(\vec{a}.\vec{b})(\vec{c}.\vec{d})$
We know $\vec{a}.\vec{b}$ is a scalar, and by suffix notation, it can be expressed as $a_i b_i$; $\vec{c}.\vec{d}$ is also a scalar, and it can be expressed as $c_j d_j$. So we can write

$$(\vec{a}.\vec{b})(\vec{c}.\vec{d}) = a_i b_i c_j d_j$$

and it indeed means

$$(\vec{a}.\vec{b})(\vec{c}.\vec{d}) = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_i c_j d_j$$

Note we cannot choose the same suffix for $\vec{a}.\vec{b}$ and $\vec{c}.\vec{d}$, since each suffix can appear at most twice in each term. But we can reorder the letters:

$a_i c_j b_i d_j$ is the same as $a_i b_i c_j d_j$.

Example Write the suffix notation expression $a_j b_i c_j$ in ordinary vector notation

$$a_j b_i c_j = a_j c_j b_i$$

So the equation stands for the vector whose i-th component is $\sum_{j=1}^{3} a_j c_j b_i = (\vec{a}.\vec{c}) b_i$. thus the equation is $(\vec{a}.\vec{c})\vec{b}$

Example. Write the vector equation $\vec{u} + (\vec{a}.\vec{b})\vec{v} = |\vec{a}|^2(\vec{b}.\vec{v})\vec{a}$ in suffix notation

First, the i-th component of this vector equation

$$u_i + (\vec{a}.\vec{b}) v_i = (\vec{a}.\vec{a})(\vec{b}.\vec{v}) a_i$$

Now replace each dot product by suffix notation, we get:

$$u_i + a_j b_j v_i = a_j a_j b_k v_k a_i$$

Example. We are going to show $tr(AB) = tr(BA)$ where A and B are two $n \times n$ matrices with real entries.

If $C = AB$, we know the ij-th entry of C is $C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$, so by suffix notation, we can write

$$C_{ij} = A_{ik} B_{kj}$$

(Since k appear twice in the product $A_{ik}B_{kj}$, we take the sum of the terms $\sum_{k=1}^{n} A_{ik} B_{kj}$ by the rule)

Recall the trace of a matrix C is

$Tr(C) = \sum_{j=1}^{n} C_{jj}$. we can write in suffix notation as C_{jj}

The trace of AB is $C_{jj} = A_{jk} B_{kj}$

Similarly, the trace of BA is $(BA)_{jj} = B_{jk} A_{kj}$

So we see $Tr(BA)$ is identified with $Tr(AB)$ by relabelling

Definition. We define the Kronecker delta δ_{ij} to be the suffix notation of the $n \times n$ identity matrix. i.e.

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Lemma. $\delta_{ij} = \delta_{ji}$

The δ_{ij} has lots of applications in computations with suffix notation.

Proposition (i) $\delta_{ij} a_i = a_j$
(ii) $\vec{a} \cdot \vec{b}$ can be written as $\delta_{ij} a_i b_j$ in suffix notation

Proof. (i) $\delta_{ij} a_i = \delta_{1j} a_1 + \delta_{2j} a_2 + \delta_{3j} a_3 = a_j$

(ii) $\delta_{ij} a_i b_j = \sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij} a_i b_j = a_1 b_1 + a_2 b_2 + a_3 b_3 = \vec{a} \cdot \vec{b}$

Proposition $\delta_{ij}\delta_{jk} = \delta_{ik}$

Proof: $\delta_{ij}\delta_{jk} = \delta_{i1}\delta_{1k} + \delta_{i2}\delta_{2k} + \delta_{i3}\delta_{3k}$

If $i \neq k$, then either $i \neq 1$ or $k \neq 1$, we get either $\delta_{i1} = 0$ or $\delta_{1k} = 0$

so $\delta_{i1}\delta_{1k} = 0$

Similarly, $\delta_{i2}\delta_{2k} = 0$, $\delta_{i3}\delta_{3k} = 0$

we get in this case $\delta_{ij}\delta_{jk} = 0$

If $i = k$, $\delta_{ij}\delta_{jk} = \delta_{k1}\delta_{1k} + \delta_{k2}\delta_{2k} + \delta_{k3}\delta_{3k} = \delta_{kk}\cdot\delta_{kk} = 1$

We finish the proof.

Definition. The alternating tensor ε_{ijk} is defined to be

$$\varepsilon_{ijk} = \begin{cases} 0 & \text{if any of } i, j, k \text{ are equal} \\ +1 & \text{if } (i,j,k) = (1,2,3), (2,3,1) \text{ or } (3,1,2) \\ -1 & \text{if } (i,j,k) = (1,3,2), (3,2,1) \text{ or } (2,1,3) \end{cases}$$

Remark: when $\{i,j,k\} = \{1,2,3\}$, ε_{ijk} is indeed the signature (also called parity) of the permutation $\begin{pmatrix} 1 & 2 & 3 \\ i & j & k \end{pmatrix} \in S_3$.

Proposition. (i) $\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij}$

(ii) $\varepsilon_{ijk} = -\varepsilon_{jik}$

Proof: Follow directly from definition

One application of the alternating tensor is to express cross product of vectors in suffix notation:

Proposition. $(\vec{a}\times\vec{b})_i = \varepsilon_{ijk} a_j b_k$

Proof. We know $\vec{a}\times\vec{b} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \sum_{\sigma\in S_3} \operatorname{sgn}(\sigma)\, \vec{e}_{\sigma(1)} a_{\sigma(2)} b_{\sigma(3)}$

$$= \sum_{\substack{i,j,k \\ \text{distinct}}} \varepsilon_{ijk} a_j b_k \vec{e}_i.$$

So the coefficient for $\vec{e}_i$ is $\sum_{j\neq i} \sum_{\substack{k\neq i \\ k\neq j}} \varepsilon_{ijk} a_j b_k = \sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon_{ijk} a_j b_k$

since whenever two of i,j,k are the same $\varepsilon_{ijk}=0$.

Example. Write $\vec{a}\cdot(\vec{b}\times\vec{c})$ in suffix notation

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = a_i(\vec{b}\times\vec{c})_i = a_i\varepsilon_{ijk} b_j c_k = \varepsilon_{ijk} a_i b_j c_k$$

The above expression tells us:

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = \varepsilon_{ijk} a_i b_j c_k = \varepsilon_{jki} b_j c_k a_i = \vec{b}\cdot(\vec{c}\times\vec{a})$$

Exercise. What is $\varepsilon_{ijk}\varepsilon_{ijk}$?

Proposition. $\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$

Proof. You may refer to the textbook for the proof.

Example. Prove $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\cdot\vec{c})\vec{b} - (\vec{a}\cdot\vec{b})\vec{c}$

$$
\begin{aligned}
(\vec{a}\times(\vec{b}\times\vec{c}))_i &= \varepsilon_{ijk}\, a_j\, (\vec{b}\times\vec{c})_k \\
&= \varepsilon_{ijk}\, a_j\, \varepsilon_{klm}\, b_l\, c_m \\
&= \varepsilon_{kij}\,\varepsilon_{klm}\, a_j\, b_l\, c_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, a_j\, b_l\, c_m \\
&= (\delta_{jm}a_j)(\delta_{il}b_l)c_m - (\delta_{jl}a_j)\, b_l\, (\delta_{im}c_m) \\
&= a_m c_m b_i - a_l b_l c_i \\
&= ((\vec{a}\cdot\vec{c})\vec{b} - (\vec{a}\cdot\vec{b})\vec{c})_i
\end{aligned}
$$

SUFFIX NOTATION IN VECTOR CALCULUS

We can make use of the suffix notation to represent the differential operator $\vec{\nabla} = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3})$ by

$$\vec{\nabla}_i = \frac{\partial}{\partial x_i}$$

So the gradient of f, $\vec{\nabla} f$ in suffix notation is

$$(\vec{\nabla} f)_i = \frac{\partial f}{\partial x_i}$$

the curl of $\vec{u}$, $\vec{\nabla} \times \vec{u}$ in suffix notation is

$$(\vec{\nabla} \times \vec{u})_i = \varepsilon_{ijk} \frac{\partial}{\partial x_j} u_k = \varepsilon_{ijk} \frac{\partial u_k}{\partial x_j}$$

the divergence of $\vec{u}$, $\vec{\nabla} \cdot \vec{u}$ in suffix notation is

$$\vec{\nabla} \cdot \vec{u} = \frac{\partial}{\partial x_i} u_i = \frac{\partial u_i}{\partial x_i}$$

Example. $\vec{r}(x_1, x_2, x_3) = (x_1, x_2, x_3)$. Let $r = |\vec{r}| = (\vec{r} \cdot \vec{r})^{\frac{1}{2}}$.

First, we know that $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$ by the rules of differentiation.

$$[\vec{\nabla} r]_i = \frac{\partial r}{\partial x_i} = \frac{\partial}{\partial x_i} (\vec{r} \cdot \vec{r})^{\frac{1}{2}} = \frac{\partial}{\partial x_i} (r_j \cdot r_j)^{\frac{1}{2}} = \frac{\partial}{\partial x_i} (x_j \cdot x_j)^{\frac{1}{2}}$$

$$= \frac{1}{2} (x_j \cdot x_j)^{-\frac{1}{2}} \cdot \frac{\partial}{\partial x_i} (x_j \cdot x_j)$$

$$= \frac{1}{2r} \cdot 2 x_j \frac{\partial x_j}{\partial x_i}$$

$$= \frac{1}{r} x_j \delta_{ij}$$

$$= \frac{1}{r} x_i$$

· This implies

$$\vec{\nabla} r = \frac{\vec{r}}{r}$$

Next let's try to compute the curl and divergence of $\vec{r}$.

$$(\vec{\nabla}\times\vec{r})_i = \varepsilon_{ijk}\frac{\partial r_k}{\partial x_j} = \varepsilon_{ijk}\frac{\partial x_k}{\partial x_j} = \varepsilon_{ijk}\delta_{jk} = 0$$

$$\text{so} \quad \vec{\nabla}\times\vec{r} = \vec{0}$$

$$\vec{\nabla}\cdot\vec{r} = \frac{\partial r_i}{\partial x_i} = \frac{\partial x_i}{\partial x_i} = \delta_{ii} = 3$$

The suffix notation can be used to verify the identities of vector calculus.

Example. Show $\mathrm{curl}(\nabla f) = \vec{0}$ and $\mathrm{div}(\mathrm{curl}\,\vec{u}) = 0$

$$[\vec{\nabla}\times(\vec{\nabla}f)]_i = \varepsilon_{ijk}\frac{\partial}{\partial x_j}(\vec{\nabla}f)_k = \varepsilon_{ijk}\frac{\partial}{\partial x_j}\cdot\frac{\partial f}{\partial x_k}$$

$$= -\varepsilon_{ikj}\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_k}$$

$$= -\varepsilon_{ikj}\frac{\partial}{\partial x_k}\frac{\partial f}{\partial x_j}$$

$$= -[\vec{\nabla}\times(\vec{\nabla}f)]_i$$

We conclude $\vec{\nabla}\times(\vec{\nabla}f) = \vec{0}$

$$\vec{\nabla}\cdot(\vec{\nabla}\times\vec{u}) = \frac{\partial}{\partial x_i}(\vec{\nabla}\times\vec{u})_i = \frac{\partial}{\partial x_i}\varepsilon_{ijk}\frac{\partial u_k}{\partial x_j}$$

$$= \varepsilon_{ijk}\frac{\partial}{\partial x_i}\frac{\partial x_k}{\partial x_j}$$

$$= -\varepsilon_{jik}\frac{\partial}{\partial x_j}\cdot\frac{\partial x_k}{\partial x_i}$$

$$= -\vec{\nabla}\cdot(\vec{\nabla}\times\vec{u})$$

We conclude $\vec{\nabla}\cdot(\vec{\nabla}\times\vec{u}) = 0$

Example. Show that $\vec{\nabla}\times(\vec{\nabla}\times\vec{u}) = \vec{\nabla}(\vec{\nabla}\cdot\vec{u}) - \vec{\nabla}^2\vec{u}$

$$[\vec{\nabla}\times(\vec{\nabla}\times\vec{u})]_i = \varepsilon_{ijk}\frac{\partial}{\partial x_j}(\vec{\nabla}\times\vec{u})_k$$

$$= \varepsilon_{ijk}\frac{\partial}{\partial x_j}\varepsilon_{klm}\frac{\partial u_m}{\partial x_l}$$

$$= \varepsilon_{ijk}\varepsilon_{klm}\frac{\partial^2 u_m}{\partial x_j\partial x_l}$$

$$= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\frac{\partial^2 u_m}{\partial x_j\partial x_l}$$

$$= \delta_{il}\cdot\delta_{jm}\frac{\partial^2 u_m}{\partial x_j\partial x_l} - \delta_{im}\delta_{jl}\frac{\partial^2 u_m}{\partial x_j\partial x_l}$$

$$= \frac{\partial^2 u_j}{\partial x_j\partial x_i} - \frac{\partial^2 u_i}{\partial x_j\partial x_j}$$

$$= \frac{\partial}{\partial x_i}\left(\frac{\partial u_j}{\partial x_j}\right) - \frac{\partial^2 u_i}{\partial x_j^2}$$

$$= [\vec{\nabla}(\vec{\nabla}\cdot\vec{u})]_i - [\vec{\nabla}^2\vec{u}]_i$$

Example. Show that $\vec{\nabla}\cdot(\vec{u}\times\vec{v}) = (\vec{\nabla}\times\vec{u})\cdot\vec{v} - (\vec{\nabla}\times\vec{v})\cdot\vec{u}$

$$\vec{\nabla}\cdot(\vec{u}\times\vec{v}) = \frac{\partial}{\partial x_i}(\vec{u}\times\vec{v})_i = \frac{\partial}{\partial x_i}\varepsilon_{ijk}u_j v_k$$

$$= \varepsilon_{ijk}\frac{\partial u_j}{\partial x_i}\cdot v_k + \varepsilon_{ijk}\frac{\partial v_k}{\partial x_i}u_j$$

$$= \varepsilon_{kij}\frac{\partial u_j}{\partial x_i}v_k - \varepsilon_{jik}\frac{\partial v_k}{\partial x_i}u_j$$

$$= [\vec{\nabla}\times\vec{u}]_k\cdot v_k - [\vec{\nabla}\times\vec{v}]_j\cdot u_j$$

$$= (\vec{\nabla}\times\vec{u})\cdot\vec{v} - (\vec{\nabla}\times\vec{v})\cdot\vec{u}$$

Definition Given a vector $\vec{u}$, define the operator $\vec{u}.\vec{\nabla}$ to be

$$\vec{u}.\vec{\nabla} = u_j \frac{\partial}{\partial x_j}$$, which acts in the following way:

$$\vec{u}.\vec{\nabla} f = u_j \frac{\partial f}{\partial x_j} = \vec{u}.(\vec{\nabla} f)$$

$$\vec{u}.\vec{\nabla}\vec{V} = u_j \frac{\partial V_i}{\partial x_j} = (\vec{u}.(\vec{\nabla} V_1), \vec{u}.(\vec{\nabla} V_2), \vec{u}.(\vec{\nabla} V_3))$$

Example Show that $\vec{\nabla}(\vec{u}.\vec{V}) = \vec{u}\times(\vec{\nabla}\times\vec{V}) + \vec{V}\times(\vec{\nabla}\times\vec{u}) + \vec{u}.\vec{\nabla}\vec{V} + \vec{V}.\vec{\nabla}\vec{u}$

$$[\vec{u}\times(\vec{\nabla}\times\vec{V})]_i = \varepsilon_{ijk} u_j (\vec{\nabla}\times\vec{V})_k = \varepsilon_{ijk} u_j \varepsilon_{klm} \frac{\partial V_m}{\partial x_l}$$

$$= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, u_j \frac{\partial V_m}{\partial x_l}$$

$$= \left(\delta_{il}\delta_{jm} u_j \frac{\partial V_m}{\partial x_l}\right) - \left(\delta_{im}\delta_{jl} u_j \frac{\partial V_m}{\partial x_l}\right)$$

$$= u_j \frac{\partial V_j}{\partial x_i} - u_j \frac{\partial V_i}{\partial x_j}$$

similarly, we can show

$$[\vec{V}\times(\vec{\nabla}\times\vec{u})]_i = V_j \frac{\partial u_j}{\partial x_i} - V_j \frac{\partial u_i}{\partial x_j}$$

so $$[\vec{u}\times(\vec{\nabla}\times\vec{V})]_i + [\vec{V}\times(\vec{\nabla}\times\vec{u})]_i = u_j \frac{\partial V_j}{\partial x_i} - u_j \frac{\partial V_i}{\partial x_j} + V_j \frac{\partial u_j}{\partial x_i} - V_j \frac{\partial u_i}{\partial x_j}$$

$$= \left(u_j \frac{\partial V_j}{\partial x_i} + V_j \frac{\partial u_j}{\partial x_i}\right) - u_j \frac{\partial V_i}{\partial x_j} - V_j \frac{\partial u_i}{\partial x_j}$$

$$= \frac{\partial}{\partial x_i}(u_j V_j) - u_j \frac{\partial V_i}{\partial x_j} - V_j \frac{\partial u_i}{\partial x_j}$$

$$= [\vec{\nabla}(\vec{u}.\vec{V})]_i - [\vec{u}.\vec{\nabla}\vec{V}]_i - [\vec{V}.\vec{\nabla}\vec{u}]_i$$

DUAL SPACES AND BILINEAR FORMS

Definition. V is a vector space over a field $\mathbb{F}$. A map $T: V \longrightarrow \mathbb{F}$ is defined to be a linear map if $\forall \vec{u}, \vec{v} \in V$, $\lambda, \mu \in \mathbb{F} \Rightarrow T(\lambda\vec{u}+\mu\vec{v}) = \lambda T(\vec{u}) + \mu T(\vec{v})$

Given a vector space V over a field $\mathbb{F}$. Denote V^* to be the set of all linear maps $V \longrightarrow \mathbb{F}$. We can define an addition and a scalar multiplication on V^*:

(i) (Addition) $\forall T_1, T_2 \in V^*$, T_1+T_2 is the linear map

$$(T_1+T_2)(\vec{v}) = T_1(\vec{v}) + T_2(\vec{v}) \quad \forall v \in V.$$

(ii) (Scalar Multiplication) $\forall T \in V^*, \lambda \in \mathbb{F}$, λT is the linear map $(\lambda T)(\vec{v}) = \lambda T(\vec{v}) \quad \forall v \in V$.

Definition. The dual vector space of V is the vector V^* with addition and scalar multiplication defined as above.

In case when the vector space is of finite dimension, there's a good way to represent linear maps:

One way to understand a linear map is to regard it as a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^1$. so it will be represented by a $1 \times n$ matrix (i.e a row vector $(a_1, \ldots, a_n)$. If $\varphi = (a_1, \ldots, a_n)$, then

$$\varphi(\vec{v}) = (a_1, \ldots, a_n)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = a_1 v_1 + \cdots + a_n v_n,$$

and $a_i = \varphi(\vec{e_i})$. This observation motivates the following:

Lemma. If V is a finite dimensional vector space over $\mathbb{F}$, then V^* is also a finite dimensional vector space over $\mathbb{F}$, and $\dim V^* = \dim V$

Proof. If we take a basis $\{v_1, v_2, \dots, v_n\}$ for V, where $n = \dim V$. define $\varphi_i \in V^*$ to be the linear map satisfying.

$$\varphi_i(\vec{v}_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

Then we'll show $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ forms a basis of V^* and we call it the dual basis of $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$

First, we show $\varphi_1, \dots, \varphi_n$ are linearly independent:
If $\lambda_1\varphi_1 + \cdots + \lambda_n\varphi_n = 0$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{F}$.
For any $\vec{e}_i$, $(\lambda_1\varphi_1 + \cdots + \lambda_n\varphi_n)(\vec{v}_i) = 0$

$$\Rightarrow \quad \lambda_i \varphi_i(\vec{v}_i) = 0$$

$$\Rightarrow \quad \lambda_i = 0 \quad \text{since } \varphi_i(\vec{v}_i) = 1$$

So $\lambda_1 = \cdots = \lambda_n = 0$. we see $\varphi_1, \dots, \varphi_n$ are linearly independent.

Next, we show $\operatorname{span}\{\varphi_1, \dots, \varphi_n\} = V^*$:
For any $\varphi \in V^*$, let $\mu_i = \varphi(\vec{v}_i)$, then we see

$$\varphi(\vec{v}_i) = \mu_i = (\mu_1\varphi_1 + \cdots + \mu_n\varphi_n)(\vec{v}_i)$$

i.e. φ and $\mu_1\varphi_1 + \cdots + \mu_n\varphi_n$ agree on the basis $\{\vec{v}_1, \dots, \vec{v}_n\}$, so by linearity, $\varphi = \mu_1\varphi_1 + \cdots + \mu_n\varphi_n$

We conclude $\{\varphi_1, \dots, \varphi_n\}$ is a basis for V^*

Definition. V is a vector space over $\mathbb{F}$. A bilinear form on V is a map
$$V \times V \xrightarrow{\langle\ ,\ \rangle} \mathbb{F}$$
$$(\vec{u}, \vec{v}) \longmapsto \langle \vec{u}, \vec{v} \rangle$$
such that (i) $\forall\ \lambda, \mu \in \mathbb{F},\ \vec{u}, \vec{v}, \vec{w} \in V$
$$\langle \lambda\vec{u} + \mu\vec{v}, \vec{w} \rangle = \lambda\langle \vec{u}, \vec{w}\rangle + \mu\langle \vec{v}, \vec{w}\rangle$$
(ii) $\forall\ \lambda, \mu \in \mathbb{F},\ \vec{u}, \vec{v}, \vec{w} \in V$.
$$\langle \vec{w}, \lambda\vec{u} + \mu\vec{v} \rangle = \lambda\langle \vec{w}, \vec{u}\rangle + \mu\langle \vec{w}, \vec{v}\rangle$$

Example. The dot product on $\mathbb{R}^n$ is a bilinear form:
$$(\lambda\vec{u} + \mu\vec{v}) \cdot \vec{w} = \lambda\vec{u} \cdot \vec{w} + \mu\vec{v} \cdot \vec{w}$$
$$\vec{w} \cdot (\lambda\vec{u} + \mu\vec{v}) = \lambda\vec{w} \cdot \vec{u} + \mu\vec{w} \cdot \vec{v}$$

Definition
- A bilinear form is symmetric if $\forall\ \vec{u}, \vec{v} \in V$, $\langle \vec{u}, \vec{v}\rangle = \langle \vec{v}, \vec{u}\rangle$
- A bilinear form is nondegenerate if $\langle \vec{u}, \vec{v}\rangle = 0\ \forall\ \vec{v} \in V$ implies $\vec{u} = \vec{0}$.

There is a linear transformation $V \longrightarrow V^*$ corresponding to a given bilinear form $\langle\ ,\ \rangle$ on V:
$$V \longrightarrow V^*$$
$$\vec{u} \longmapsto \langle \vec{u}, \cdot \rangle$$

Proposition The map $V \longrightarrow V^*$ induced by a bilinear form $\langle\ ,\ \rangle$ is injective if and only if $\langle\ ,\ \rangle$ is nondegenerate

Proof. Follows directly from the definition

A good way to represent a bilinear form is to use matrices.

Given a bilinear form $< \ , \ >$ on a finite dimensional vector space V over $\mathbb{F}$, let $n = \dim V$.
Take a basis of V: $\{\vec{v}_1, \ldots, \vec{v}_n\}$
let $a_{ij} = <\vec{v}_i, \vec{v}_j>$, we form an $n\times n$ matrix $A = (a_{ij})$
For any $\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$ and $\vec{y} = y_1\vec{v}_1 + \cdots + y_n\vec{v}_n$.
We can see

$$<\vec{x}, \vec{y}> = (x_1, \ldots, x_n)\, A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

by verifying for $<\vec{v}_i, \vec{v}_j>$ then extend by linearity:

$$<\vec{v}_i, \vec{v}_j> = a_{ij} = (0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-th}}}{1}, 0, \ldots, 0)\, A \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j\text{-th}$$

So once we fix a basis of V, there's a one-to-one correspondence between bilinear forms on V and $n\times n$ matrices with entries in $\mathbb{F}$

So now we've seen a new point of view for vectors and matrices: When we fix a basis for an n-dimensional vector space V over a field $\mathbb{F}$, a vector $(a_1, \ldots, a_n)$ describes a linear map $V \to \mathbb{F}$ and a matrix (a_{ij}) describes a bilinear form (map) $V\times V \to \mathbb{F}$. This point of view indicates we can generalize the concepts of vectors and matrices by considering multilinear maps:

$$\underbrace{V\times V\times\cdots\times V}_{m \text{ copies}} \longrightarrow \mathbb{F}.$$

TENSORS AS MULTILINEAR MAPS

Definition. V is a vector space over a field $\mathbb{F}$. m is a positive integer. Define an m-tensor on V to be a multilinear map $\underbrace{V \times V \times \cdots \times V}_{m \text{ copies}} \xrightarrow{T} \mathbb{F}$

i.e. for each $1 \leq i \leq m$,

$$T(\vec{v}_1, \ldots, \lambda \vec{v}_i + \mu \vec{v}_i', \ldots, \vec{v}_n) = \lambda T(\vec{v}_1, \ldots, \vec{v}_i, \ldots \vec{v}_n) + \mu T(\vec{v}_1, \ldots \vec{v}_i', \ldots \vec{v}_n)$$

Definition. The set of m-tensors on V is denoted as $T^m(V)$.

Proposition. $T^m(V)$ is a $\mathbb{F}$-vector space, with addition:

$$(S+T)(\vec{v}_1, \ldots, \vec{v}_n) = S(\vec{v}_1, \ldots, \vec{v}_n) + T(\vec{v}_1, \ldots, \vec{v}_n)$$

scalar multiplication: $(\lambda S)(\vec{v}_1, \ldots, \vec{v}_n) = \lambda S(\vec{v}_1, \ldots, \vec{v}_n)$

Proof. Exercise

Definition. Given $S \in T^k(V)$ and $T \in T^l(V)$, define the tensor product $S \otimes T \in T^{k+l}(V)$ to be the multilinear map

$$S \otimes T(\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_{k+l}) = S(\vec{v}_1, \ldots, \vec{v}_k) \cdot T(\vec{v}_{k+1}, \ldots, \vec{v}_{k+l})$$

Proposition. The following rules hold for tensors:

$$(S_1 + S_2) \otimes T = S_1 \otimes T + S_2 \otimes T$$
$$T \otimes (S_1 + S_2) = T \otimes S_1 + T \otimes S_2$$
$$(aS) \otimes T = S \otimes (aT) = a(S \otimes T)$$
$$(S \otimes T) \otimes U = S \otimes (T \otimes U)$$

Remark · Note in general, $S\otimes T \neq T\otimes S$.

· Since tensor product is associative, we can write $S\otimes T\otimes U$ to denote either $(S\otimes T)\otimes U$ or $S\otimes(T\otimes U)$

Theorem. Let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be a basis of V, and $\{\varphi_1, \dots, \varphi_n\}$ be the dual basis, (i.e. $\varphi_i(\vec{v}_j) = \delta_{ij}$). Then the set of all k-fold tensor products

$$\varphi_{i_1}\otimes\cdots\otimes\varphi_{i_k} \quad 1\le i_1, \dots, i_k \le n$$

is a basis for $T^k(V)$, so $\dim T^k(V) = n^k$.

Proof. First we prove these n^k k-tensor products span $T^k(V)$:

$$\begin{aligned}\varphi_{i_1}\otimes\cdots\otimes\varphi_{i_k}(\vec{v}_{j_1}, \dots, \vec{v}_{j_k}) &= \varphi_{i_1}(\vec{v}_{j_1})\cdot\ \cdots\ \cdot\varphi_{i_k}(\vec{v}_{j_k}) \\ &= \delta_{i_1j_1}\cdots\delta_{i_kj_k} \\ &= \begin{cases} 1, & \text{if } i_1=j_1, \dots, i_k=j_k \\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Now for $\vec{w}_1, \dots, \vec{w}_k$ in V, since $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis of V,

$$\vec{w}_i = \sum_{j=1}^{n} a_{ij}\vec{v}_j$$

For any $T\in T^k(V)$.

$$\begin{aligned}T(\vec{w}_1, \dots, \vec{w}_k) &= T\Big(\sum_{j_1=1}^{n} a_{1j_1}\vec{v}_{j_1}, \dots, \sum_{j_k=1}^{n} a_{kj_k}\vec{v}_{j_k}\Big) \\ &= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} a_{1j_1}\cdots a_{kj_k}\, T(\vec{v}_{j_1}, \dots, \vec{v}_{j_k}) \\ &= \sum_{j_1=1}^{n}\cdots\sum_{j_k=1}^{n} T(\vec{v}_{j_1}, \dots, \vec{v}_{j_k})\cdot\varphi_{j_1}\otimes\cdots\otimes\varphi_{j_k}(\vec{w}_1, \dots, \vec{w}_k)\end{aligned}$$

So $T = \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} T(\vec{v}_{j_1}, \ldots, \vec{v}_{j_k}) \varphi_{j_1} \otimes \cdots \otimes \varphi_{j_k}$

i.e. $T \in \operatorname{Span}\{\varphi_{j_1} \otimes \cdots \otimes \varphi_{j_k}\}$, $1 \le j_1, \ldots, j_k \le n$

Next we prove $\{\varphi_{j_1} \otimes \cdots \otimes \varphi_{j_k}\}$ are linearly independent.

If $\exists\, a_{i_1, \ldots, i_k} \in \mathbb{F}$, $1 \le i_1, \ldots, i_k \le n$ such that

$\sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} a_{i_1, \ldots, i_k} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}$ is the zero map.

Then $\sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} a_{i_1, \ldots, i_k} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k} (\vec{v}_{j_1}, \ldots, \vec{v}_{j_k}) = 0$

i.e. $a_{j_1, \ldots, j_k} = 0$

We conclude $\{\varphi_{j_1} \otimes \cdots \otimes \varphi_{j_k}\}$ are linearly independent.

So we conclude $\{\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}\}$ $1 \le i_1, \ldots, i_k \le n$ form a basis of $\mathcal{T}^k(V)$, and $\dim \mathcal{T}^k(V) = n^k$

Example. A bilinear form $\langle\, , \rangle$ on V is a 2-tensor on V. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for V, and A is the $n \times n$ matrix representing $\langle\, , \rangle$ under this basis, then $\langle\, , \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \delta_i \otimes \delta_j$, where $\delta_1, \ldots, \delta_n$ is the dual basis in V^*, and $A_{ij} = \langle \vec{v}_i, \vec{v}_j \rangle$

Definition. A bilinear form on a vector space V is positive definite if $\forall \vec{v} \in V$ and $\vec{v} \neq \vec{0}$, $\langle \vec{v}, \vec{v} \rangle > 0$

Theorem. (Gram-Schmidt) If $\langle\ ,\ \rangle$ is a symmetric, positive definite bilinear form on V, then there exists an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ of V with respect to $\langle\ ,\ \rangle$. i.e. $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$

Proof. Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis of V.

Let $\vec{w}_1' = \vec{w}_1$

$$\vec{w}_2' = \vec{w}_2 - \frac{\langle \vec{w}_1', \vec{w}_2 \rangle}{\langle \vec{w}_1', \vec{w}_1' \rangle} \vec{w}_1'$$

$$\vec{w}_3' = \vec{w}_3 - \frac{\langle \vec{w}_1', \vec{w}_3 \rangle}{\langle \vec{w}_1', \vec{w}_1' \rangle} \vec{w}_1' - \frac{\langle \vec{w}_2', \vec{w}_3 \rangle}{\langle \vec{w}_2', \vec{w}_2' \rangle} \vec{w}_2'$$

$$\vdots$$

$$\vec{w}_n' = \vec{w}_n - \frac{\langle \vec{w}_1', \vec{w}_n \rangle}{\langle \vec{w}_1', \vec{w}_1' \rangle} \vec{w}_1' - \dots - \frac{\langle \vec{w}_{n-1}', \vec{w}_n \rangle}{\langle \vec{w}_{n-1}', \vec{w}_{n-1}' \rangle} \vec{w}_{n-1}'$$

We can show that if $i \neq j$, $\langle \vec{w}_i', \vec{w}_j' \rangle = 0$ by induction

First, $\langle \vec{w}_1', \vec{w}_2' \rangle = \langle \vec{w}_1, \vec{w}_2 \rangle - \frac{\langle \vec{w}_1', \vec{w}_2 \rangle}{\langle \vec{w}_1', \vec{w}_1' \rangle} \langle \vec{w}_1, \vec{w}_1' \rangle$

$$= \langle \vec{w}_1, \vec{w}_2 \rangle - \langle \vec{w}_1, \vec{w}_2 \rangle$$

$$= 0$$

Next assume for $1 \le i < j \le k$, $\langle \vec{w}_i', \vec{w}_j' \rangle = 0$

then for $k+1$,

$$\langle \vec{w}_{k+1}', \vec{w}_i' \rangle = \left\langle \vec{w}_{k+1} - \frac{\langle \vec{w}_1', \vec{w}_{k+1} \rangle}{\langle \vec{w}_1', \vec{w}_1' \rangle} \vec{w}_1' - \dots - \frac{\langle \vec{w}_k', \vec{w}_{k+1} \rangle}{\langle \vec{w}_k', \vec{w}_k' \rangle} \vec{w}_k', \vec{w}_i' \right\rangle$$

$$= \langle \vec{w}_{k+1}, \vec{w}_i' \rangle - \frac{\langle \vec{w}_i', \vec{w}_{k+1} \rangle}{\langle \vec{w}_i', \vec{w}_i' \rangle} \cdot \langle \vec{w}_i', \vec{w}_i' \rangle = 0$$

At last, we do a "normalization":

Let $\vec{v}_i = \dfrac{\vec{w}_i'}{\sqrt{\langle \vec{w}_i', \vec{w}_i' \rangle}}$ for $i = 1, \dots, n$.

$$\langle \vec{v}_i, \vec{v}_i \rangle = \frac{\langle \vec{w}_i', \vec{w}_i' \rangle}{\langle \vec{w}_i', \vec{w}_i' \rangle} = 1$$

$$\langle \vec{v}_i, \vec{v}_j \rangle = \frac{\langle \vec{w}_i', \vec{w}_j' \rangle}{\sqrt{\langle \vec{w}_i', \vec{w}_i' \rangle} \cdot \sqrt{\langle \vec{w}_j', \vec{w}_j' \rangle}} = 0 \quad \text{if } i \neq j$$

so $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$.

And this also implies $\vec{v}_1, \dots, \vec{v}_n$ are linearly independent:
if $a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = \vec{0}$, then for any $\vec{v}_i$,

$$0 = \langle a_1\vec{v}_1 + \cdots + a_n\vec{v}_n, \vec{v}_i \rangle = a_i \langle \vec{v}_i, \vec{v}_i \rangle = a_i$$

and we know V is of dimension n, so $\{\vec{v}_1, \dots, \vec{v}_n\}$ form a basis of V.

Example. The determinant of $n \times n$ matrices can be interpreted as an n-tensor on a vector space $\mathbb{R}^n$:
Let $\{\vec{e}_1, \dots, \vec{e}_n\}$ be the standard basis of $\mathbb{R}^n$

$$\mathbb{R}^n \times \cdots \times \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$\left(\sum_{j=1}^n a_{1j}\vec{e}_j, \dots, \sum_{j=1}^n a_{nj}\vec{e}_j\right) \longmapsto \det(a_{ij})$$

Definition. $T \in T^k(V)$ is alternating if for any $1 \leq i < j \leq k$,
$$T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_k) = -T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i, \dots, \vec{w}_k)$$

Proposition. $T \in T^k(V)$. T is alternating if and only if

$T(\vec{w}_1, \dots, \vec{w}_k) = 0$ whenever some $\vec{w}_i = \vec{w}_j$ for $i \neq j$

Proof. If T is an alternating tensor,

$(\vec{w}_1, \dots, \vec{w}_k) \in V \times \cdots \times V$ such that $\vec{w}_i = \vec{w}_j$ for some $i \neq j$. then

$$T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_k) = -T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i, \dots, \vec{w}_k)$$

$$= -T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_k)$$

so $T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_k) = 0$

Conversely, if for any $(\vec{w}_1, \dots, \vec{w}_k) \in V \times \cdots \times V$ such that $\vec{w}_i = \vec{w}_j$ for some $i \neq j$, $T(\vec{w}_1, \dots, \vec{w}_k) = 0$.

then

$$T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_n)$$

$$= T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_n) + T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_i, \dots, \vec{w}_n)$$

$$= T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_n)$$

Similarly. $T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i, \dots, \vec{w}_n)$

$$= T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_n)$$

So $T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_j, \dots, \vec{w}_n) + T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i, \dots, \vec{w}_n)$

$$= T(\vec{w}_1, \dots, \vec{w}_i, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_n) + T(\vec{w}_1, \dots, \vec{w}_j, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_n)$$

$$= T(\vec{w}_1, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_i + \vec{w}_j, \dots, \vec{w}_n)$$

$$= 0$$

Corollary. If $T \in T^k(V)$ is an alternating tensor, $\{\vec{w}_1, \dots, \vec{w}_k\}$ is a set of k linearly depent vectors in V. then

$$T(\vec{w}_1, \dots, \vec{w}_k) = 0$$

Proof. Will be homework.

Given any $T \in T^k(V)$, there is a way to construct an alternating k-tensor, $Alt(T)$, as follows:

$$Alt(T)(\vec{v}_1, \dots, \vec{v}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} sgn\,\sigma \cdot T(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(k)})$$

where $\sigma \in S_k$ is a permutation of k letters, and $sgn(\sigma)$ is the signature of σ. i.e. if we decompose σ as a product of m transpositions, $sgn(\sigma) = (-1)^m$

Definition. The set of all alternating k tensors form a subspace of $T^k(V)$, and it's denoted as $\Lambda^k(V)$.

Theorem. (i) If $T \in T^k(V)$, then $AlT(T) \in \Lambda^k(V)$

(ii) If $T \in \Lambda^k(V)$, then $AlT(T) = T$

Proof.

$$AlT(T)(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_j, \dots, \vec{v}_k)$$

$$= \frac{1}{k!} \sum_{\sigma \in S_k} sgn\,\sigma \cdot T(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(i)}, \dots, \vec{v}_{\sigma(j)}, \dots, \vec{v}_{\sigma(k)})$$

$$= \frac{1}{k!} \sum_{\sigma(ij) \in S_k} sgn((ij)\sigma)\, T(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(j)}, \dots, \vec{v}_{\sigma(i)}, \dots, \vec{v}_{\sigma(k)})$$

$$= -\frac{1}{k!} \sum_{\sigma \in S_k} sgn\,\sigma \cdot T(\vec{v}_{\sigma(1)}, \dots, \vec{v}_{\sigma(j)}, \dots, \vec{v}_{\sigma(i)}, \dots, \vec{v}_{\sigma(k)})$$

$$= -AlT(T)(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_i, \dots, \vec{v}_k)$$

(ii). If $T \in \Lambda^k(V)$,

$$\mathrm{Alt}(T)(\vec{v}_1, \ldots, \vec{v}_k)$$

$$= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}\sigma \cdot T(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(k)})$$

$$= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}\sigma \cdot \operatorname{sgn}\sigma \, T(\vec{v}_1, \ldots, \vec{v}_k)$$

$$= \frac{1}{k!} \cdot k! \cdot T(\vec{v}_1, \ldots, \vec{v}_k)$$

$$= T(\vec{v}_1, \ldots, \vec{v}_k)$$

Exercise. In the above proof, we made use of the fact: $T \in \Lambda^k(V)$, $T(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(k)}) = \operatorname{sgn}(\sigma) T(\vec{v}_1, \ldots, \vec{v}_k)$. Prove this fact.

Definition. The wedge product $\Lambda^k(V) \times \Lambda^l(V) \xrightarrow{\wedge} \Lambda^{k+l}(V)$ is defined as follows: if $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^l(V)$, then $\omega \wedge \eta = \frac{(k+l)!}{k!\,l!} \mathrm{Alt}(\omega \otimes \eta)$

Proposition. The following properties hold for wedge product:

(i). $(\omega_1 + \omega_2) \wedge \eta = \omega_1 \wedge \eta + \omega_2 \wedge \eta$

(ii). $\omega \wedge (\eta_1 + \eta_2) = \omega \wedge \eta_1 + \omega \wedge \eta_2$

(iii) $a\omega \wedge \eta = \omega \wedge (a\eta) = a(\omega \wedge \eta) \quad (a \in \mathbb{F})$

(iv). $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$, if $\omega \in \Lambda^k(V)$ and $\eta \in \Lambda^l(V)$

(v). $(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k!\,l!\,m!} \mathrm{Alt}(\omega \otimes \eta \otimes \theta)$

if $\omega \in \Lambda^k(V)$, $\eta \in \Lambda^l(V)$, $\theta \in \Lambda^m(V)$.

(vi) $\omega_1 \wedge \cdots \wedge \omega_k = \frac{(l_1 + \cdots + l_k)!}{l_1! \cdots l_k!} \mathrm{Alt}(\omega_1 \otimes \cdots \otimes \omega_k)$

if $\omega_1 \in \Lambda^{l_1}(V), \ldots, \omega_k \in \Lambda^{l_k}(V)$

Theorem. If V has basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ and V^* has dual basis $\{\varphi_1, \dots, \varphi_n\}$, then $\{\varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}\}$, $1 \le i_1 < i_2 < \dots < i_k \le n$ forms a basis for $\Lambda^k(V)$. So $\dim \Lambda^k(V) = \frac{n!}{k!(n-k)!}$

Proof. If $\omega \in \Lambda^k(V)$, we can write it as

$$\omega = \sum_{i_1, \dots, i_k} a_{i_1 \dots i_k} \varphi_{i_1} \otimes \dots \otimes \varphi_{i_k}$$

$$\omega = \mathrm{Alt}(\omega) = \sum_{i_1, \dots, i_k} a_{i_1 \dots i_k} \mathrm{Alt}(\varphi_{i_1} \otimes \dots \otimes \varphi_{i_k})$$

and each $\mathrm{Alt}(\varphi_{i_1} \otimes \dots \otimes \varphi_{i_k}) = \begin{cases} \frac{1}{k!} \varphi_{i_1} \wedge \dots \wedge \varphi_{i_k} & \text{if } i_1, \dots, i_k \text{ distinct} \\ 0 & \text{otherwise} \end{cases}$

So $\{\varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}\}$ span $\Lambda^k(V)$.

We also need to show the linearly independence:

If $\sum c_{i_1, \dots, i_k} \varphi_{i_1} \wedge \dots \wedge \varphi_{i_k} = 0$,

Observe $\varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}(\vec{v}_{j_1}, \dots, \vec{v}_{j_k}) = \begin{cases} 1 & \text{if } i_1 = j_1, \dots, i_k = j_k \\ 0 & \text{otherwise} \end{cases}$
(Exercise)

$(j_1 < \dots < j_k)$.

So we can get linearly independence by applying

$\sum c_{i_1, \dots, i_k} \varphi_{i_1} \wedge \dots \wedge \varphi_{i_k}$ to each $(\vec{v}_{j_1}, \dots, \vec{v}_{j_k})$

Corollary If $\dim V = n$, then $\dim \Lambda^n(V) = 1$.

Theorem. Let $\vec{v}_1, \ldots, \vec{v}_n$ be a basis for $\mathbb{R}^n$. $T \in \Lambda^n(\mathbb{R}^n)$.
If $\vec{w}_i = \sum_{j=1}^{n} a_{ij}\vec{v}_j$ are n vectors in V, then:

$$T(\vec{w}_1, \ldots, \vec{w}_n) = \det(a_{ij}) \cdot T(\vec{v}_1, \ldots, \vec{v}_n)$$

Proof. Define $S \in \Lambda^n(\mathbb{R}^n)$ by

$$S\left(\sum_{j=1}^{n} a_{1j}\vec{e}_j, \ldots, \sum_{j=1}^{n} a_{nj}\vec{e}_j\right) = T\left(\sum_{j=1}^{n} a_{1j}\vec{v}_j, \ldots, \sum_{j=1}^{n} a_{nj}\vec{v}_j\right)$$

where $\{\vec{e}_1, \ldots, \vec{e}_n\}$ is the standard basis for $\mathbb{R}^n$.

Since $\det \in \Lambda^n(\mathbb{R}^n)$ is nonzero, $\dim \Lambda^n(\mathbb{R}^n) = 1$.
so $\exists\, \lambda \in \mathbb{R}$ such that
$$S = \lambda \det$$

Then
$$\begin{aligned} T(\vec{w}_1, \ldots, \vec{w}_n) &= S\left(\sum_{j=1}^{n} a_{1j}\vec{e}_j, \ldots, \sum_{j=1}^{n} a_{nj}\vec{e}_j\right) \\ &= \lambda \det\left(\sum_{j=1}^{n} a_{1j}\vec{e}_j, \ldots, \sum_{j=1}^{n} a_{nj}\vec{e}_j\right) \\ &= \lambda \det(a_{ij}) \end{aligned}$$

$$T(\vec{v}_1, \ldots, \vec{v}_n) = S(\vec{e}_1, \ldots, \vec{e}_n) = \lambda \det(\delta_{ij}) = \lambda$$

So $T(\vec{w}_1, \ldots, \vec{w}_n) = \det(a_{ij})\, T(\vec{v}_1, \ldots, \vec{v}_n)$

Corollary. Take the standard basis $\{\vec{e}_1, \dots, \vec{e}_n\}$ of $\mathbb{R}^n$. $\{\delta_1, \dots, \delta_n\}$ is its dual basis on V^*

Then $\det = \delta_1 \wedge \delta_2 \wedge \dots \wedge \delta_n$

Remark. By the discussion above, we have another way to define det:

The determinant map for $\mathbb{R}^n$ is the unique alternating n-tensor whose value on $(\vec{e}_1, \dots, \vec{e}_n)$ is 1. where $\{\vec{e}_1, \dots, \vec{e}_n\}$ is the standard basis for $\mathbb{R}^n$.

Example. Recall when we studies suffix notations, we defined the "alternating tensor" ε_{ijk}.

If we take the standard basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ on $\mathbb{R}^3$. $\{\varphi_1, \varphi_2, \varphi_3\}$ is the dual basis. Then we have shown before

$$\varepsilon_{ijk} u_i v_j w_k = \vec{u} \cdot (\vec{v} \times \vec{w}) = \det(\vec{u}, \vec{v}, \vec{w})$$

So ε_{ijk} indeed represents the determinant on $\mathbb{R}^3$, and this expression above tells us that

$$\det = \sum_{0 \leq i,j,k \leq 3} \varepsilon_{ijk} \varphi_i \otimes \varphi_j \otimes \varphi_k$$

DIFFERENTIAL FORMS (Brief Introduction)

Recall that given a smooth function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$,
at $p \in \mathbb{R}^n$, we can compute the directional derivative along the direction $\vec{v} = (v_1, \ldots, v_n)$:

$$D_{\vec{v}}(f) = \vec{v} \cdot \vec{\nabla} f = v_1 \frac{\partial f}{\partial x_1}(p) + \cdots + v_n \frac{\partial f}{\partial x_n}(p)$$

Since this holds for all smooth functions, we can define the operator

$$D_{\vec{v}} = \vec{v} \cdot \vec{\nabla} = v_1 \frac{\partial}{\partial x_1} + \cdots + v_n \frac{\partial}{\partial x_n}$$

This expression indicates we can use a new notation for the standard basis of $\mathbb{R}^n$: $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$.

i.e. $\frac{\partial}{\partial x_i} = (0, \ldots, \underset{\substack{\uparrow \\ i\text{th}}}{1}, \ldots 0)$

Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, there is a map

$$\begin{aligned} \mathbb{R}^n &\xrightarrow{df} (\mathbb{R}^n)^* \\ p &\longmapsto df(p) \end{aligned}$$

$(\mathbb{R}^n)^*$ is the dual space of $\mathbb{R}^n$

where $df(p)$ acts on $v_1 \frac{\partial}{\partial x_1} + \cdots + v_n \frac{\partial}{\partial x_n} \in \mathbb{R}^n$ by

$$df(p)\left(v_1 \frac{\partial}{\partial x_1} + \cdots + v_n \frac{\partial}{\partial x_n}\right) = v_1 \frac{\partial f}{\partial x_1}(p) + \cdots + v_n \frac{\partial f}{\partial x_n}(p)$$

Let x_i be the function $x_i : \mathbb{R}^n \longrightarrow \mathbb{R}$

$$(p_1, \ldots, p_n) \longmapsto p_i.$$

Then $dx_i(p)\left(\frac{\partial}{\partial x_j}\right) = \frac{\partial x_i}{\partial x_j}(p) = \delta_{ij} \quad \forall p \in \mathbb{R}^n$

so $\{dx_1(p), \ldots, dx_n(p)\}$ forms the dual basis of $\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$ for any $p \in \mathbb{R}^n$

This implies for each smooth function $f: \mathbb{R}^n \to \mathbb{R}$

$$df(p) = \frac{\partial f}{\partial x_1}(p)\, dx_1(p) + \cdots + \frac{\partial f}{\partial x_n}(p)\, dx_n(p) \quad \forall p \in \mathbb{R}^n$$

So $\boxed{df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n}$, called the differential of f

In other words, the differential of f tells us how to write df as a linear combination of the basis $\{dx_1, \ldots, dx_n\}$ at each point.

Definition A differential k-form on $\mathbb{R}^n$ is a map $\omega: \mathbb{R}^n \to \Lambda^k(\mathbb{R}^n)$

Example df is a differential 1-form.

A differential k-form ω can be written as

$$\omega(p) = \sum_{1 \le i_1 < \cdots < i_k \le n} \omega_{i_1 \cdots i_k}(p)\, dx_{i_1}(p) \wedge \cdots dx_{i_k}(p)$$

i.e. $\omega = \sum_{1 \le i_1 < \cdots < i_k \le n} \omega_{i_1 \cdots i_k}\, dx_{i_1} \wedge \cdots \wedge dx_{i_k}$.

Definition ω is a continuous/differentiable/smooth k-form if all the $\omega_{i_1 \cdots i_k}$ are continuous/differentiable/smooth functions.

TENSORS IN CARTESIAN COORDINATES

The Cartesian Coordinates is the most popular one for the Euclidean space $\mathbb{R}^n$. Once we have fixed a cartesian coordinates (equivalently, a set of orthonormal basis), we can express an k-tensor as n^k numbers $\{a_{i_1 \dots i_k}\}_{1 \le i_1, \dots, i_k \le n}$. But, if we build another Cartesian Coordinates, we may get another set of n^k numbers $\{b_{i_1 \dots i_k}\}_{1 \le i_1, \dots, i_k \le n}$

So now we are facing the question: how will the n^k numbers change if the coordinate is changed?

Definition. If $(\vec{u}, \vec{v}) = \vec{u}^T \vec{v}$ is the dot product on $\mathbb{R}^n$, a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if for any $\vec{u}, \vec{v} \in \mathbb{R}^n$, $(T(\vec{u}), T(\vec{v})) = (\vec{u}, \vec{v})$

Definition An $n \times n$ real matrix A is orthogonal if $A^{-1} = A^T$.

Lemma. If $\{\vec{e}_1, \dots, \vec{e}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if the matrix of T with respect to this basis is an orthogonal matrix.

Proof. If $T: \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal linear transformation, Its matrix with respect to the basis is A, then $T(\vec{u}) = A\vec{u}$. so:

$$\vec{u}^T \vec{v} = (\vec{u}, \vec{v}) = (T(\vec{u}), T(\vec{v})) = (A\vec{u})^T (A\vec{v}) = \vec{u}^T A^T A \vec{v}$$

holds for all $\vec{u}$ & $\vec{v} \in \mathbb{R}^n$. we conclude $A^T A = I_n$, the identity matrix, so $A^{-1} = A^T$

Conversely, if the matrix A of a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal matrix, then for any $\vec{u}, \vec{v} \in \mathbb{R}^n$.

$$\begin{aligned}(T(\vec{u}), T(\vec{v})) = (T(\vec{u}))^T (T(\vec{v})) = (A\vec{u})^T (A\vec{v}) &= u^T A^T A v \\ &= \vec{u}^T \vec{v} \\ &= (\vec{u}, \vec{v})\end{aligned}$$

So T is orthogonal linear transformation.

Definition The set of all $n \times n$ orthogonal matrices is denoted by $O_n(\mathbb{R})$, called the orthogonal linear group.

Proposition If $\{\vec{v}_1, \dots, \vec{v}_n\}$ and $\{\vec{v}_1', \dots, \vec{v}_n'\}$ are two sets of basis of $\mathbb{R}^n$ with transition matrix $A \in O_n(\mathbb{R})$, i.e. $A = (a_{ij}) \in O_n(\mathbb{R})$ and $\vec{v}_i' = \sum_{j=1}^{n} a_{ij} \vec{v}_j$. Then the coefficients of a k-tensor $T \in \mathcal{T}^k(V)$ with respect to these two basis are related in suffix notation by:

$$C'_{j_1 \cdots j_k} = a_{j_1 i_1} \cdots a_{j_k i_k} C_{i_1 \cdots i_k}$$

Remark Tensors can be expressed in suffix notation just like vectors and matrices. Fixing $\{\vec{e}_1, \dots, \vec{e}_n\}$ the cartesian coordinates, we know a k-tensor $T \in \mathcal{T}^k(\mathbb{R}^n)$ can be written as $\sum_i C_{i_1 \cdots i_k} \varphi_{i_1} \otimes \cdots \varphi_{i_k}$, so we denote it by its component $C_{i_1 \cdots i_k}$ in suffix notation

Proof. Let $\{\varphi_1, \dots, \varphi_n\}$ be the dual basis of $\{\vec{V_1}, \dots, \vec{V_n}\}$
$\{\varphi_1', \dots, \varphi_n'\}$ be the dual basis of $\{\vec{V_1}', \dots, \vec{V_n}'\}$

Then $\varphi_i(\vec{V_k}') = \varphi_i(\sum_{j=1}^{n} a_{kj} \vec{V_j}) = a_{ki}$.

we see $\varphi_i = \sum_{k=1}^{n} a_{ki} \varphi_k'$

$$\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k} = (\sum_{j_1=1}^{n} a_{j_1 i_1} \varphi_{j_1}') \otimes \cdots \otimes (\sum_{j_k=1}^{n} a_{j_k i_k} \varphi_{j_k}')$$
$$= \sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} a_{j_1 i_1} \cdots a_{j_k i_k} \varphi_{j_1}' \otimes \cdots \otimes \varphi_{j_k}'$$

If $T \in T^k(V)$ such that $C_{i_1 \cdots i_k}$ are the coefficients with respect to $\{\varphi_1 \otimes \cdots \otimes \varphi_k\}$, and $C_{j_1 \cdots j_k}'$ are the coefficients with respect to $\{\varphi_1' \otimes \cdots \otimes \varphi_k'\}$.

Then: $T = \sum_{1 \le i_1, \dots, i_k \le n} C_{i_1 \cdots i_k} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}$

$$= \sum_{1 \le i_1, \dots, i_k \le n} C_{i_1 \cdots i_k} (\sum_{j_1=1}^{n} \cdots \sum_{j_k=1}^{n} a_{j_1 i_1} \cdots a_{j_k i_k}) \varphi_{j_1}' \otimes \cdots \otimes \varphi_{j_k}'$$

$$= \sum_{1 \le j_1, \dots, j_k \le n} \sum_{1 \le i_1, \dots, i_k \le n} C_{i_1 \cdots i_k} a_{j_1 i_1} \cdots a_{j_k i_k} \varphi_{j_1}' \otimes \cdots \otimes \varphi_{j_k}'$$

So we get $C_{j_1 \cdots j_k}' = \sum_{1 \le i_1, \dots, i_k \le n} a_{j_1 i_1} \cdots a_{j_k i_k} C_{i_1 \cdots i_k}$

In suffix notation:

$$C_{j_1 \cdots j_k}' = a_{j_1 i_1} \cdots a_{j_k i_k} C_{i_1 \cdots i_k}$$

Example. We know that the dot product is represented in suffix notation as δ_{ij}. Now if we change the coordinates by $A \in O_n(\mathbb{R})$, then the suffix notation for the dot product in the new coordinates is

$$\delta'_{i'j'} = a_{i'i}\, a_{j'j}\, \delta_{ij} = a_{i'i}\, a_{j'i} = \delta_{i'j'} \quad \text{since } A \in O_n(\mathbb{R})$$

So we see the matrix of dot product with respect to any orthonormal basis is the identity matrix I_n

Example. ε_{ijk} is the suffix notation for the 3-tensor $\det \in \Lambda^3(\mathbb{R}^3)$. If we change the coordinates by $A \in O_n(\mathbb{R})$; recall we have proved the Theorem (on Page 74), a special case is:

$$T(\vec{v}_1{}', \vec{v}_2{}', \vec{v}_3{}') = \det(A)\, T(\vec{v}_1, \vec{v}_2, \vec{v}_3) \quad \text{if } T \in \Lambda^3(\mathbb{R}^3)$$

So we see under a change of basis:

$$\varepsilon'_{ijk} = \det(A)\, \varepsilon_{ijk}$$

If $\det(A) = 1$, we will get $\varepsilon'_{ijk} = \varepsilon_{ijk}$.
If $\det(A) = -1$, we will get $\varepsilon'_{ijk} = -\varepsilon_{ijk}$.

But $A \in O_n(\mathbb{R})$, these are the only two possibilities since

$$A \in O_n(\mathbb{R}) \Rightarrow A \cdot A^T = I \Rightarrow \det(A)\det(A^T) = 1$$
$$\Rightarrow \det(A) \cdot \det(A) = 1$$
$$\Rightarrow \det(A) = \pm 1$$

Definition The special orthogonal linear group is

$$SO_n(\mathbb{R}) = \{A \in O_n(\mathbb{R}) \mid \det(A) = 1\}$$

Remark. (1) Elements in $SO_n(\mathbb{R})$ corresponds to rotations in $\mathbb{R}^n$ fixing the origin.

(2) A: transition matrix. A is called orientation preserving if $\det(A) > 0$; it's called orientation reversing if $\det(A) < 0$.

(3) $\{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\{\vec{v}_1{}', \ldots, \vec{v}_n{}'\}$ are two sets of basis for $\mathbb{R}^n$. We say they have the same orientation if the transition matrix is orientation preserving, and we say they have the opposite orientation if the transition matrix is orientation reversing.

So in the previous example, we see the alternating tensor ε_{ijk} has the same expression under the change of basis by a rotation.

Example. If $T \in T^2(V)$ such that in suffix notation with respect to some orthonormal basis $T_{ij} = T_{ji}$. Then in any other orthonormal basis, the suffix notation also satisfies $T'_{i'j'} = T'_{j'i'}$

Let $A \in O_n(\mathbb{R})$ be the change of basis matrix then $T'_{i'j'} = a_{i'i} a_{j'j} T_{ij} = a_{i'i} a_{j'j} T_{ji} = a_{j'j} a_{i'i} T_{ji} = T'_{j'i'}$

Definition. A tensor $T \in T^k(V)$ is called isotropic if its components are the same in each Cartesian coordinates

Example • The dot product is an isotropic 2-tensor
• The determinand for $\mathbb{R}^3$ is an isotropic 3-tensor

Proposition The zero tensor is the only isotropic 1-tensor on $\mathbb{R}^3$

Proof. Suppose with repect to some coordinates, the isotopic 1-tensor is denoted by T_i.
Now Let the transformation matrix be $A = \begin{pmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 1 \end{pmatrix} \in SO_3(\mathbb{R})$
Then

$$T'_j = a_{ji} T_i$$

so $T'_1 = a_{11} T_1 + a_{12} T_2 + a_{13} T_3 = T_2$
$T'_2 = a_{21} T_1 + a_{22} T_2 + a_{23} T_3 = -T_1$

If T is isotropic then $T'_1 = T_1$, $T'_2 = T_2 \Rightarrow T_1 = T_2$, $-T_1 = T_2$
$\Rightarrow T_1 = T_2 = 0$

Similarly by an appropriate choice of A, we can also show $T_3 = 0$

Proposition. If T is an isotropic 2-tensor on $\mathbb{R}^3$, then $T = \lambda\delta_{ij}$ in any Cartesian coordinates

Proof. For T an isotropic 2-tensor, if in some Cartesian coordinates it's T_{ij}, then take the transformation matrix to be

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO_3(\mathbb{R})$$

$$T'_{i'j'} = a_{i'i}a_{j'j}T_{ij} = a_{i'i}T_{ij}a_{j'j} = a_{i'i}T_{ij}a^t_{jj'} = (ATA^t)_{i'j'}$$

$$ATA^t = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} T_{22} & -T_{21} & T_{23} \\ -T_{12} & T_{11} & -T_{13} \\ T_{32} & -T_{31} & T_{33} \end{pmatrix}$$

Since T is isotropic, $T'_{i'j'} = T_{i'j'}$, so

$$T_{11} = T_{22}, \quad -T_{21} = T_{12}, \quad -T_{31} = T_{13}.$$

Similarly, we can finally show $T_{11} = T_{22} = T_{33}$, and $T_{ij} = 0$ if $i \neq j$. Let $\lambda = T_{11}$, we thus see

$$T_{ij} = \lambda\delta_{ij}$$

Proposition. If T is an isotropic 3-tensor on $\mathbb{R}^3$, then $T = \lambda\varepsilon_{ijk}$ in any Cartesian coordinates.

Proof. For T an isotropic 3-tensor, if in some Cartesian coordinates it's T_{ijk}, then again take the transformation matrix to be $A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$$T'_{i'j'k'} = a_{i'i} a_{j'j} a_{k'k} T_{ijk}$$

We get $T'_{123} = \sum a_{1i} a_{2j} a_{3k} T_{ijk} = -T_{213}$

and T is isotropic $\Rightarrow T_{123} = T'_{123} = -T_{213}$

Similarly we can find the other relations to finish the proof.

Theorem If $T \in T^k(\mathbb{R}^3)$ and $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, then in suffix notation
$T(\vec{u}, \vec{v}, \vec{w}) = T_{ijk} u_i v_j w_k$

Proof. $T(\vec{u}, \vec{v}, \vec{w}) = T\left(\sum_{i=1}^{3} u_i \vec{e}_i, \sum_{j=1}^{3} v_j \vec{e}_j, \sum_{k=1}^{3} w_k \vec{e}_k\right)$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} u_i v_j w_k T(\vec{e}_i, \vec{e}_j, \vec{e}_k)$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} u_i v_j w_k T_{ijk}$$

So in suffix notation, $T(\vec{u}, \vec{v}, \vec{w})$ is $T_{ijk} u_i v_j w_k$.

Theorem (The Quotient Rule). If a suffix equation $a_i = T_{ij} b_j$ holds in all Cartesian coordinates, and for any 1-tensor b_j, a_i is also a 1-tensor. then T_{ij} represents a 2-tensor.

Proof: Let L be the change of basis matrix

$$a'_i = L_{ik} a_k = L_{ik} T_{kj} b_j$$

$$b_j = (L^{-1})_{jm} b_m \Rightarrow b_j = L_{mj} b'_m, \quad \text{since } L^{-1} = L^T$$

Then $a'_i = L_{ik} T_{kj} b_j = L_{ik} T_{kj} L_{mj} b'_m$

The assumption $a_i = T_{ij} b_j$ holds for all coordinates implies $a'_{im} = T'_{im} b'_m$

so $\quad T'_{im} = L_{ik} T_{kj} L_{mj} = L_{ik} L_{mj} T_{kj}$ for any L.

We conclude T_{ij} is a 2-tensor

INVERSE FUNCTION THEOREM

Definition. $F: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ given by $F(x_1,\dots,x_n) = (f_1(x_1,\dots,x_n),\dots,f_m(x_1,\dots,x_n))$ is called continuously differentiable at $a \in \mathbb{R}^n$ if all $\frac{\partial f_i}{\partial x_j}$ exist in a neighbourhood of a and continuous at a.

Lemma. $A \subseteq \mathbb{R}^n$ is a rectangle. $F: A \to \mathbb{R}^n$ is continuously differentiable on A. If there is $M > 0$ such that $\forall\, x \in A \setminus \partial A$ $\left|\frac{\partial f_i}{\partial x_j}(x)\right| \le M$, then:

$$|F(x) - F(y)| \le n^2 M |x - y| \quad \text{for all } x, y \in A$$

Proof. $\forall\, x, y \in A$.

$$f_i(y) - f_i(x) = \sum_{j=1}^{n} \left[f_i(y_1,\dots,y_j,x_{j+1},\dots,x_n) - f_i(y_1,\dots,y_{j-1},x_j,\dots,x_n) \right]$$

$$= \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(y_1,\dots,y_{j-1},z_j,x_{j+1},\dots,x_n) \cdot (y_j - x_j)$$

for some z_j between x_j & y_j, by Mean Value Theorem.

$$|F(x) - F(y)| \le \sum_{i=1}^{n} |f_i(y) - f_i(x)|$$

$$\le \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \frac{\partial f_i}{\partial x_j}(y_1,\dots,y_{j-1},z_j,x_{j+1},\dots,x_n) \right| \cdot |y_j - x_j|$$

$$\le \sum_{j=1}^{n} nM \cdot |y_j - x_j| \le nM \cdot n\,|y - x| = n^2 M\,|y - x|$$

Definition. If $F:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, define the Jacobian of F at a to be the $m \times n$ matrix

$$DF(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{m\times n}$$

Theorem. (Inverse Function Theorem)
If $F:\mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuously differentiable in an open set containing $a \in \mathbb{R}^n$, and $\det DF(a) \neq 0$, then there exists an open set V containing a and an open set W containing $f(a)$ such that $F: V \rightarrow W$ has inverse $F^{-1}: W \rightarrow V$ which is also differentiable and $\forall\, y \in W$

$$DF^{-1}(y) = [DF(F^{-1}(y))]^{-1}$$

Proof. Let L denote the linear transformation represented by $DF(a) = \left(\frac{\partial f_i}{\partial x_j}\right)$. $L:\mathbb{R}^n \rightarrow \mathbb{R}^n$ is invertible since $\det DF(a) \neq 0$

Then $D(L^{-1} \circ F)(a) = D(L^{-1})\, DF(a)$ (by the Chain Rule)

$$= L^{-1}\, DF(a)$$

$$= DF(a)^{-1} \cdot DF(a)$$

$$= I_n$$

If we can prove the Theorem is true for $L^{-1} \circ F$, then since L is invertible, we'll get the Theorem is true for F

The above observation implies we can assume $DF(a) = I_n$ to prove the Theorem.

F is continuously differentiable at $a \in \mathbb{R}^n$ implies

$$\lim_{x \to a} \frac{|F(x) - F(a) - DF(a)(x-a)|}{|x-a|} = 0$$

i.e. $$\lim_{x \to a} \frac{|F(x) - F(a) - (x-a)|}{|x-a|} = 0$$

So for $\varepsilon = \frac{1}{2}$, $\exists\ \delta > 0$ such that

$$0 < |x-a| < \delta \Rightarrow \frac{|F(x) - F(a) - (x-a)|}{|x-a|} < \frac{1}{2}$$

Suppose $0 < |b-a| < \delta$ and $F(b) = F(a)$, then

$$\frac{|\cancel{F(b) - F(a)} - (b-a)|}{|b-a|} = \frac{|b-a|}{|b-a|} = 1 > \frac{1}{2} \text{. Contradiction}$$

So for any $0 < |x-a| < \delta$, $F(x) \neq F(a)$

We can take even smaller δ', to get $0 < |x-a| < \delta'$

Such that:

(1) $0 < |x-a| < \delta' \Rightarrow F(x) \neq F(a)$

(2) $\det DF(x) \neq 0 \quad \forall\ |x-a| < \delta'$ (since $\det DF$ is continuous)

(3) $\left|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right| < \frac{1}{2n^2}$ (since F is continuously differentiable at a)
on $|x-a| < \delta'$

Take U to be a rectangle inside $|x-a| < \delta'$ and $a \in U$.
Let $G(x) = F(x) - x$, and apply the Lemma, we get
for any $x, x' \in U$.

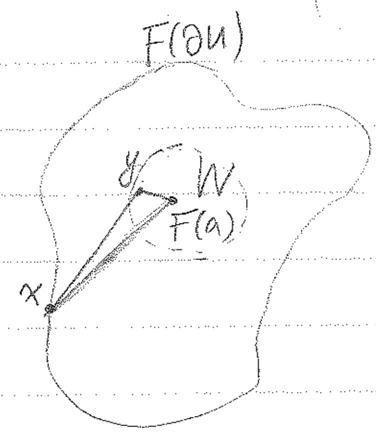
$$|G(x) - G(x')| \leq n^2 \cdot \frac{1}{2n^2} |x - x'| = \frac{1}{2}|x - x'|$$

So $|x - x'| - |F(x) - F(x')| \leq |G(x) - G(x')| \leq \frac{1}{2}|x - x'|$

$$\Rightarrow \quad |x - x'| \leq 2|F(x) - F(x')| \text{ on } U \quad \cdots (4)$$

Since $F(x)\neq F(a)\ \forall\, x\in U-\{a\}$, we see $F(\partial U)\not\ni F(a)$

Let $d=\min\limits_{x\in\partial U}|F(x)-F(a)|$.



Define $W=\{y\in\mathbb{R}^n \mid |y-F(a)|<\frac{d}{2}\}$

For any $y\in W$, any $x\in\partial U$, we get

$$
\begin{aligned}
|y-F(x)| &= |y-F(a)+F(a)-F(x)| \\
&\geqslant |F(a)-F(x)|-|y-F(a)| \\
&> d-\frac{d}{2} \\
&= \frac{d}{2} \\
&> |y-F(a)|
\end{aligned}
$$

i.e. $|y-F(a)|<|y-F(x)|\quad \forall\, y\in W,\ \forall\, x\in\partial U$. ----(5)

Now we are going to show for any $y\in W$, $\exists!\ x\in U$ such that $F(x)=y$:

Existence:

Let $g: U\to\mathbb{R}$ be $g(x)=|y-F(x)|^2=\sum_{i=1}^{n}(y_i-f_i(x))^2$

U is a rectangle, hence compact. the continuous function $g(x)$ has a minimum x_0 on U

If $x\in\partial U$, by (5), $g(a)=|y-F(a)|^2<|y-F(x)|^2$, so $x_*\notin\partial U$, x_* is in the interior of U

We know if an interior point is a minimum, then it's a critical point. so

$$\frac{\partial g}{\partial x_j}(x_*)=0 \ \forall \ j=1,\dots,n$$

$$\text{i.e. } 2\sum_{i=1}^{n} (y_i - f_i(x_*)) \cdot \frac{\partial f_i}{\partial x_j}(x_*) = 0 \ \forall \ j=1,\dots,n.$$

The matrix form is

$$2\underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_*) & \cdots & \frac{\partial f_1}{\partial x_n}(x_*) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_*) & \cdots & \frac{\partial f_n}{\partial x_n}(x_*) \end{pmatrix}}_{DF(x_0)} \begin{pmatrix} y_1 - f_1(x_*) \\ \vdots \\ y_n - f_n(x_*) \end{pmatrix} = 0$$

Since $\det DF(x_*) \neq 0$ by (2), we get
$y_i = f_i(x_*)$ $\forall$ $i=1,\dots,n$.

Uniqueness: follows directly from (4).

So we have proved $\forall y \in W$, $\exists! \, x \in U \setminus \partial U$ such that $y = F(x)$.

Let $V = U \cap F^{-1}(W)$, we just showed $F: V \to W$ is a bijection, so it has inverse $F^{-1}: W \to V$.

(4) can be rewritten as $|F^{-1}(y) - F^{-1}(y')| \leq 2|y-y'|$ $\forall$ $y, y' \in W$.
So we see F^{-1} is continuous on W.

We need to show at last F^{-1} is differentiable on W:

For any $x_0 \in V$, since F is differentiable at x_0,

we can write $F(x) = F(x_0) + DF(x_0)(x - x_0) + \varphi(x - x_0)$

such that $\displaystyle\lim_{x \to x_0} \frac{|\varphi(x - x_0)|}{|x - x_0|} = 0$

$$[DF(x_0)]^{-1} \cdot (F(x) - F(x_0)) = [DF(x_0)]^{-1}[DF(x_0)(x - x_0) + \varphi(x - x_0)]$$

$$[DF(x_0)]^{-1}(F(x) - F(x_0)) = x - x_0 + [DF(x_0)]^{-1}\varphi(x - x_0)$$

Now for any $y_0 \in W$, if we denote $x_0 = F^{-1}(y_0)$, then $\forall y \in W$:

$$[DF(x_0)]^{-1}(F(F^{-1}(y)) - F(F^{-1}(y_0))) = F^{-1}(y) - F^{-1}(y_0) + [DF(x_0)]^{-1}\varphi(F^{-1}(y) - F^{-1}(y_0))$$

$$\Rightarrow [DF(x_0)]^{-1}(y - y_0) = F^{-1}(y) - F^{-1}(y_0) + [DF(x_0)]^{-1}\varphi(F^{-1}(y) - F^{-1}(y_0))$$

$$\Rightarrow F^{-1}(y) = F^{-1}(y_0) + [DF(x_0)]^{-1}(y - y_0) - [DF(x_0)]^{-1}\varphi(F^{-1}(y) - F^{-1}(y_0))$$

Note that $\displaystyle\lim_{y \to y_0} \frac{|[DF(x_0)]^{-1}\varphi(F^{-1}(y) - F^{-1}(y_0))|}{|y - y_0|}$

$$= \lim_{y \to y_0} [DF(x_0)]^{-1} \cdot \frac{|\varphi(F^{-1}(y) - F^{-1}(y_0))|}{|F^{-1}(y) - F^{-1}(y_0)|} \cdot \frac{|F^{-1}(y) - F^{-1}(y_0)|}{|y - y_0|}$$

$$\leq 2[DF(x_0)]^{-1} \lim_{y \to y_0} \frac{|\varphi(F^{-1}(y) - F^{-1}(y_0))|}{|F^{-1}(y) - F^{-1}(y_0)|}$$

$$= 0$$

since F^{-1} is continuous and $\displaystyle\lim_{x \to x_0} \frac{|\varphi(x - x_0)|}{|x - x_0|} = 0$.

We conclude F^{-1} is differentiable at $y_0 \in W$, and

$DF^{-1}(y_0) = [DF(F^{-1}(y_0))]^{-1}$, i.e. $DF^{-1}(F(x_0)) = [DF(x_0)]^{-1}$

Remark. When $\det DF(a) = 0$, F may also be invertible near a.

For example, $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. $f'(0) = 0$, but f has inverse $f^{-1}(x) = \sqrt[3]{x}$ near 0.

Example. The conversion between polar coordinates & Cartesian coordinates for $\mathbb{R}^2$ is given by:

$$F: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(r, \theta) \longmapsto (r\cos\theta, r\sin\theta)$$

$$DF = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

$$\det DF = r\cos^2\theta + r\sin^2\theta = r$$

We see that $\det DF \neq 0$ whenever $r \neq 0$, i.e. the point is not the origin. So the Inverse Function Theorem tells us that for any point $P \neq \vec{0}$ in $\mathbb{R}^2$, there's a neighbourhood of P on which the Cartesian coordinates are in one-to-one correspondence with the polar coordinates, i.e. two points have the same cartesian coordinates iff they have the same polar coordinates

IMPLICIT FUNCTION THEOREM

Theorem. (Implicit Function Theorem)

$F: \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$ is continuously differentiable in an open set containing $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$, and $F(a,b)=0$.

Let $M = \left(\frac{\partial f^i}{\partial x_{n+j}}(a,b) \right)$ be the $m \times m$ matrix.

If $\det M \neq 0$, then there exists open set $a \in A \subseteq \mathbb{R}^n$ and open set $b \in B \subseteq \mathbb{R}^m$ such that there is a unique function

$$G: A \longrightarrow B$$

such that $F(x, G(x)) = 0$, and G is differentiable.

Proof. Define

$$\begin{aligned} \Phi: \mathbb{R}^n \times \mathbb{R}^m &\longrightarrow \mathbb{R}^n \times \mathbb{R}^m \\ (x, y) &\longmapsto (x, F(x,y)) \end{aligned}$$

$$D\Phi(a,b) = \left(\begin{array}{c|c} I_n & O \\ \hline * & M \end{array} \right)$$

$$\det D\Phi(a,b) = \det I_n \cdot \det M \neq 0$$

The Inverse Function Theorem implies there is an open set $(a,b) \in A \times B \subseteq \mathbb{R}^n \times \mathbb{R}^m$ and open set $\Phi(a,b) = (a, F(a,b)) = (a,0) \in W \subseteq \mathbb{R}^n \times \mathbb{R}^m$ such that $\Phi: A \times B \longrightarrow W$ has differentiable inverse $\Psi: W \longrightarrow A \times B$.

We can write $\Psi(x,y) = (x, K(x,y))$ for some differentiable function $K : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$

Define $\pi : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$

$$(x, y) \longmapsto y$$

The $\pi \circ \Phi(x,y) = \pi((x, F(x,y))) = F(x,y)$

So:

$$F(x, K(x,y)) = F(\Psi(x,y)) = \pi \circ \Phi \circ \Psi(x,y) = \pi(x,y) = y$$

This implies $F(x, K(x,0)) = 0$

Let $G(x) = K(x,0)$, we get $F(x, G(x)) = 0$

Corollary. The derivative of the implicit function G can be computed by implicit differentiation:

$$0 = \frac{\partial f_i}{\partial x_j}(x, g(x)) + \sum_{\alpha=1}^{m} \frac{\partial f_i}{\partial x_{n+\alpha}}(x, g(x)) \cdot \frac{\partial g_\alpha}{\partial x_j}(x)$$

In matrix Form, if

$$DF(x, g(x)) = \underbrace{\Big[\, DF_x}_{n} \,\Big|\, \underbrace{DF_y \,\Big]}_{m} \Big\} m$$

Then $DF_x(x, G(x)) + DF_y(x, G(x)) \cdot DG(x) = 0$

$$DG(x) = -DF_y(x, G(x))^{-1} \cdot DF_x(x, G(x))$$

Example. $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is $F(x,y) = x^2 + y^2 - 1$

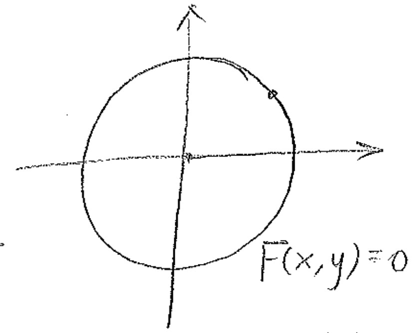
$$DF(x,y) = \left[\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right] = [2x, 2y]$$

$M = [2y] \neq 0$ for any (a,b) such that $a^2+b^2-1=0$ and $b \neq 0$

so y is a function of x near (a,b) on this level set $F(x,y)=0$, i.e. $y = G(x)$ such that $F(x, G(x)) = 0$

The derivative is computed by

$$DG(x) = -[2y]^{-1} \cdot [2x] = -\frac{x}{y} = -\frac{x}{G(x)}$$



Note that this agrees with what we have learned about implicit differentiation in Calculus.

Lemma. $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a smooth function. If $\nabla F(x_0, y_0) \neq \vec{0}$ for some $(x_0, y_0) \in \mathbb{R}^2$, then the level set $F(x,y) = F(x_0, y_0)$ locally is a differentiable curve near (x_0, y_0)

Proof. If $\nabla F(x_0, y_0) = \left(\frac{\partial F}{\partial x}(x_0, y_0), \frac{\partial F}{\partial y}(x_0, y_0)\right) \neq \vec{0}$, then at least one of the coordinates is nonzero. Without loss of generality we may assume $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$.

Then by the implicit function theorem, there is a function $g : \mathbb{R} \to \mathbb{R}$ such that $F(x, g(x)) = F(x_0, y_0)$ for x near x_0

This implies the level set $F(x,y)=F(x_0,y_0)$ near (x_0,y_0) is the graph of the function $y=g(x)$, hence it's the curve

$\vec{r}(t)=(t,g(t))$. near $t=x_0$.

Proposition. $F:\mathbb{R}^2\longrightarrow\mathbb{R}$ is a smooth function and $F(x,y)=C$ is a level set. If $\nabla F\neq\vec{0}$ on $F(x,y)=C$ everywhere, then $F(x,y)=C$ represents a differentiable curve in $\mathbb{R}^2$

Proof. By the Lemma, we see locally it's a curve everywhere so it's a curve itself.

Similar arguments lead to the following results:

Proposition. $F:\mathbb{R}^3\longrightarrow\mathbb{R}$ is a smooth function, and $F(x,y,z)=C$ is a level set. If $\nabla F(x,y,z)\neq\vec{0}$ on $F(x,y,z)=C$ everywhere, then $F(x,y,z)=C$ represents a differentiable surface in $\mathbb{R}^2$

Example. In $\mathbb{R}^2$, $F(x,y)=\frac{x^2}{a^2}+\frac{y^2}{b^2}$, $(ab\neq 0)$. Then

For any $C>0$: If $F(x,y)=C$, then $(x,y)\neq(0,0)$. So $\nabla F=(\frac{2x}{a^2},\frac{2y}{b^2})\neq 0$ on $F(x,y)=C$. we get $F(x,y)=C$ is a differentiable curve.

Indeed, $F(x,y)=C$ represents an ellipse in this case.

Example. $F(x,y,z)=x^2+y^2+z^2$. $F(x,y,z)=C>0$ is a level set on which $\nabla F(x,y,z)=(2x,2y,2z)\neq 0$. so

$F(x,y,z)=C$ is a differentiable surface.

Indeed we know it represents a sphere

Exercise. $F(x,y)=xy$. $F(x,y)=C$ is a differentiable curve if $C\neq 0$. And what happened for $C=0$?

We can use the formula in the Implicit Function Theorem to compute the tangent line / tangent surface

Example. Find the tangent line of $\frac{x^2}{4}+\frac{y^2}{9}=1$ at $(1,\frac{3}{2}\sqrt{3})$

$F(x,y)=\frac{x^2}{4}+\frac{y^2}{9}$ $\nabla F(x,y)=(\frac{x}{2},\frac{2}{9}y)$

we see $\frac{2}{9}\cdot\frac{3}{2}\sqrt{3}\neq 0$. so locally $y=g(x)$ on $F(x,y)=1$, i.e. $F(x,g(x))=1$

$$\text{then } g'(x)=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}=-\frac{\frac{x}{2}}{\frac{2}{9}y}=-\frac{9x}{4y}$$

$g'(1)=-\frac{\sqrt{3}}{2}$, the tangent line is

$$y-\frac{3}{2}\sqrt{3}=-\frac{\sqrt{3}}{2}(x-1)$$

Theorem. $F: \mathbb{R}^3 \to \mathbb{R}$ is a continuously differentiable function. If $F^{-1}(c)$ defines a surface, i.e. $\nabla F(x,y,z) \neq \vec{0}$ for any $F(x,y,z)=c$, then $\nabla F(x,y,z)$ is a normal vector field on the surface.

Proof. Without loss of generality, assume $\frac{\partial F}{\partial z}(x_0,y_0,z_0) \neq 0$. Then by Implicit Function Theorem, there's $\mathbb{R}^2 \supseteq A \xrightarrow{G} B \subseteq \mathbb{R}$ $(x_0,y_0) \in A$, $z_0 \in B$ such that $F(x,y,G(x,y)) = c$ on A. So locally the surface is the graph of $z = G(x,y)$.

$$\nabla G(x_0,y_0) = -\frac{1}{\frac{\partial F}{\partial z}(x_0,y_0,z_0)} \cdot \left[\frac{\partial F}{\partial x}(x_0,y_0,z_0), \frac{\partial F}{\partial y}(x_0,y_0,z_0)\right]$$

We know $\left(-\frac{\partial G}{\partial x}(x_0,y_0,z_0), -\frac{\partial G}{\partial y}(x_0,y_0,z_0), 1\right)$ is a normal vector to the surface at (x_0,y_0,z_0), and

$$\left(-\frac{\partial G}{\partial x}(x_0,y_0,z_0), -\frac{\partial G}{\partial y}(x_0,y_0,z_0), 1\right) = \frac{1}{\frac{\partial F}{\partial z}(x_0,y_0,z_0)} \nabla F(x_0,y_0,z_0)$$

So $\nabla F(x_0,y_0,z_0)$ is also a normal vector.

Example. Find the equation of the tangent plane of $x^2+2y^2+3z^2=6$ at $(1,1,1)$.

Let $F(x,y,z) = x^2+2y^2+3z^2$.

$\nabla F(x,y,z) = (2x, 4y, 6z)$, so $\nabla F(1,1,1) = (2,4,6)$

The equation of the tangent plane is

$$2(x-1)+4(y-1)+6(z-1)=0$$

Similar to the discussions about curves in $\mathbb{R}^2$ and surfaces in $\mathbb{R}^3$, we can also study curves in $\mathbb{R}^3$ by the implicit function theorem.

Proposition. $F: \mathbb{R}^3 \to \mathbb{R}^2$ is a smooth map. $F(x,y,z) = (C_1, C_2)$ a level set on which DF is of rank 2 everywhere, then $F(x,y,z) = (C_1, C_2)$ represents a differentiable curve in $\mathbb{R}^3$

Example. $F: \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $F(x,y,z) = (xy, yz + xz)$

$$DF = \begin{pmatrix} y & x & 0 \\ z & z & x+y \end{pmatrix}$$

We see $DF(1,1,1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}$ with

$\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ invertible, so near $(1,1,1)$ there's

function $G: \mathbb{R} \to \mathbb{R}^2$ such that $F(x, G(x)) = (1, 2)$

So near $(1,1,1)$ $F(x,y,z) = (1,2)$ is the curve $\vec{r}(t) = (t, G(t))$ for t near 1, and we can compute $\vec{r}\,'(1)$ by $\vec{r}\,'(t) = (1, G'(t))$

$$G'(1) = -\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

So $\vec{r}\,'(1) = (1, -1, 0)$

APPLICATIONS IN ECONOMICS

Example. The demand D for a goods is a functions of the price P before tax and the sales tax is t per unit:

$$D = f(t, p)$$

Suppose S is the supply, which is a function of price P:

$$S = g(p)$$

The equilibrium price p is a function of the sales tax t, implicitly defined by

$$f(t, p) = g(p)$$

Let $F(t, p) = f(t, p) - g(p) = 0$

$DF = \left[\frac{\partial f}{\partial t} \quad \frac{\partial f}{\partial p} - \frac{dg}{dp} \right]$, so by the implicit function theorem the equilibrium price p is a function of t if $\frac{\partial f}{\partial p} - \frac{dg}{dp} \neq 0$

In reality, $\frac{\partial f}{\partial p} < 0$ and $\frac{dg}{dp} > 0$, so $\frac{\partial f}{\partial p} - \frac{\partial g}{\partial p} \neq 0$.

$$\frac{dP}{dt} = -\frac{\frac{\partial f}{\partial t}}{\frac{\partial f}{\partial p} - \frac{dg}{dp}} < 0,$$ which implies a increase in sales tax will make the equilibrium of price before tax decrease.

Example. A firm produces $Q = f(L)$ units of goods when using L units of labour. Assume $f'(L) > 0$, $f''(L) < 0$.

If the firm gets P dollars per unit produced and pays w dollars for a unit of labour, Then the profit function is given by:

$$\pi(L) = \underbrace{Pf(L)}_{\text{revenue}} - \underbrace{wL}_{\text{cost}}$$

Let L^* be the units of Labour that maximizes $\pi(L)$. We know

$$\pi'(L^*) = 0$$

i.e. $$pf'(L^*) - w = 0$$

Let $F(p, w, L^*) = pf'(L^*) - w = 0$

$$\nabla F = [f'(L^*), -1, pf''(L^*)]$$

Since $pf''(L^*) \neq 0$, L^* is a function of p & w.

$L^* = G(p, w)$, $\quad pf'(G(p,w)) - w = 0$

$$\left[\frac{\partial L^*}{\partial p}, \frac{\partial L^*}{\partial w}\right] = -\frac{1}{pf''(L^*)}[f'(L^*), -1]$$

$$= \left[-\frac{f'(L^*)}{pf''(L^*)}, \frac{1}{pf''(L^*)}\right]$$

We see as price increases, L^* increases.

as salary increases, L^* decreases