3109 Multivariable Analysis Notes

Based on the 2013 autumn lectures by Dr I Petridis

MATH3109 – Multivariable Analysis.

2 October 2013
Dr. Yiannis PETRIDIS
Maths 706.

Lectures by Dr. Y Petridis. Office Hours Mon 1pm, Fri 11am. Room 504B.

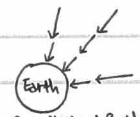
Homeworks due at 4pm on Fridays. Will be difficult problems.

Recommended text: Baby Spivak, *Calculus on manifolds*. Notation in course follows this book. Get newest edition.

3 hours of lectures a week, 4 hours from after Reading Week (incl Fri 9-10).

Course aims at unifying material covered in Methods 2. Stoke's theorem: $\int_{\partial M} \omega = \int_M d\omega$.

All classical theorems of multivariable calculus are derived from this.



gravitational field.

In this course, we discuss functions $\vec{F}: \mathbb{R}^n \to \mathbb{R}^m$. if $m=1$, this is a <u>scalar field</u>. If $m>1$, it is a <u>vector field</u>.

STOKES' THEOREM.

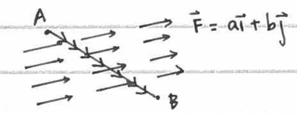
Theorem first appeared in 1850, in a letter of Lord Kelvin to Prof Stokes.

Key idea: "Differential forms are meant to be integrated". Differential forms were introduced in 1899 by Élie Cartan.

Recall the Fundamental Theorem of Calculus: $\int_a^b F'(x)\,dx = F(b) - F(a)$, where F is a function.

Instead of calling F a function, we adopt a new perspective... Consider here $F'(x)\,dx$ to be a 1-form on the singular cube $[a,b]$.

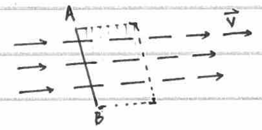
More generally, $\int_a^b g(x)\,dx$ is a 1-form in $\mathbb{R}$.



How much work is done moving a particle from A to B in this force field?

Then consider a 2D field. In $\mathbb{R}^2$, let $\vec{F}$ be a constant force field. A particle is moved from A to B.

If $\vec{AB} = x\vec{i} + y\vec{j}$, then work done is $\vec{F}\cdot\vec{AB} = ax + by = \int_A^B a\underbrace{dx}_{\substack{\text{horizontal}\\ \text{displacement}}} + b\underbrace{dy}_{\substack{\text{vertical}\\ \text{displacement}}}$.

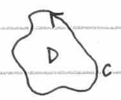


Consider constant fluid flow with $\vec{v} = v_1\vec{i} + v_2\vec{j}$. A barrier AB is placed across the flow, $\vec{AB} = x\vec{i} + y\vec{j}$.

How much fluid passes the barrier in a unit of time? We must calculate the area of the parallelogram.

This is computed by determinants – area of parallelogram $= \begin{vmatrix} v_1 & v_2 \\ x & y \end{vmatrix} = -v_2 x + v_1 y$.

Then the flow through AB in a unit of time is $-v_2 x + v_1 y = \int_A^B \overbrace{-v_2\,dx + v_1\,dy}^{\text{1-form}}$.



For a closed contour, recall Green's Theorem: $\int_C f\,dx + g\,dy = \iint_D \left(-\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x}\right) dx\,dy$.

This is also a differential form.

Moving on to $\mathbb{R}^3$, consider scalar fields $f: \mathbb{R}^3 \to \mathbb{R}$. In modern terminology, this is a 0-form on points or a 3-form $f(x,y,z)\,dx\wedge dy\wedge dz$ integrated over solids.

For vector fields, $\vec{F}: \mathbb{R}^3 \to \mathbb{R}^3$, $\vec{F} = f(x,y,z)\vec{i} + g(x,y,z)\vec{j} + h(x,y,z)\vec{k}$.

- In the case of a force field, we have 1-form $w = f\,dx + g\,dy + h\,dz$.
- In the case of a fluid flow, we have 2-form $w = f\,dy\wedge dz + g\,dz\wedge dx + h\,dx\wedge dy$.

We also have <u>operators</u>, such as

- $\nabla\cdot f = \operatorname{grad} f = \frac{\partial f}{\partial x}\vec{i} + \frac{\partial f}{\partial y}\vec{j} + \frac{\partial f}{\partial z}\vec{k}$ in classical notation. This is a 0-form $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$. (1-form)
- If $\vec{F}$ is a vector field, $\operatorname{curl}\vec{F} = \operatorname{rot}(\vec{F}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ f & g & h \end{vmatrix} = \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)\vec{i} + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\right)\vec{j} + \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)\vec{k}$

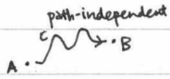
  This gives 1-form $dw = d(f\,dx + g\,dy + h\,dz) = \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)dy\wedge dz + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\right)dz\wedge dx + \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)dx\wedge dy$ (2-form).
- $\operatorname{div}\vec{F} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}$. Then $\overset{\text{3-form}}{dw} = \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}\right)dx\wedge dy\wedge dz$.

Recall that $\operatorname{curl}(\operatorname{grad} f) = 0$, if f is a potential vector field, it is irrotational i.e. $\nabla\times(\nabla\cdot f) = 0$. In modern language, $\overbrace{d(\underbrace{d\overset{\text{0-form}}{f}}_{\text{1-form}})}^{\text{2-form}} = 0$

Also, $\operatorname{div}(\operatorname{curl}\vec{F}) = \operatorname{div}(\nabla\times\vec{F}) = 0$. Here w is a 1-form, dw is a 2-form. Then $d(dw) = 0$.

Consider also the following integrals:

- Line integrals $\int_C \vec{F}\,d\vec{r}$. This is $\int_C f\,dx + g\,dy + h\,dz$.
- Surface integrals $\iint_S \vec{F}\cdot\vec{n}\,d\sigma$ This is $\int_S f\,dy\wedge dz + g\,dz\wedge dx + h\,dx\wedge dy$.
- Triple integrals over a solid T: $\iiint_T f(x,y,z)\,dx\,dy\,dz$. This is a 3-form: $\int_T f(x,y,z)\,dx\wedge dy\wedge dz$.



Recall the following theorems:

- Fundamental Theorem of Calculus: $\int_C \nabla f \cdot d\vec{r} = f(B) - f(A)$. This is the differential form $\int_C df = f(B) - f(A)$.
- Stokes' theorem. $\int_C \vec{F} \cdot d\vec{r} = \int_S \operatorname{curl} \vec{F} \cdot \vec{n}\, d\sigma$. in classical language. This is $\int_C f\,dx + g\,dy + h\,dz = \int_S \left(\frac{\partial f}{\partial y} - \frac{\partial g}{\partial x}\right) dy \wedge dx + \left(-\frac{\partial f}{\partial z} + \frac{\partial h}{\partial x}\right) dx \wedge dz + \left(\frac{\partial g}{\partial z} - \frac{\partial h}{\partial f}\right) dz \wedge dy$.
- Divergence (Gauss's) Theorem: $\int_S \vec{F} \cdot \vec{n}\, d\sigma = \iiint_T \operatorname{div} \vec{F}\, dx\,dy\,dz$, where T is a solid with boundary S.

This is a 2-form w: $\int_S f\, dy \wedge dz + g\, dz \wedge dx + h\, dx \wedge dy = \int_T dw = \int_T \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}\right) dx \wedge dy \wedge dz$. [of the form $\int_{\partial S} w = \int_S dw$].

Notation — Let $\mathbb{R}^n = \{\underset{\text{vector}}{X} = (\underbrace{X^1, X^2}_{\text{coordinates}}, \cdots, X^n)\}$ Note the use of superscripts rather than subscripts.

$|x| = \sqrt{(x^1)^2 + (x^2)^2 + \cdots + (x^n)^2}$ is the norm of a vector. We also introduce the standard basis in $\mathbb{R}^3$, $\langle e_1, e_2, \ldots, e_n \rangle$ where $e_i = (0, 0, \cdots, \underset{i^{th}\text{ slot}}{1}, \cdots 0, 0)$

Recall inner products, $\langle x, y \rangle = \sum_{i=1}^{n} x^i y^i$, $y = (y^1, y^2, \cdots, y^n)$.

Properties —

a. $|x| \geq 0$ and $|x| = 0$ iff $x = 0$

b. $|\langle x, y \rangle| \leq |x||y|$ (Cauchy-Schwartz inequality)

c. $|x + y| \leq |x| + |y|$. (Triangle Inequality)

d. $|a \cdot x| = |a| \cdot |x|$ if $a \in \mathbb{R}$

e. $\langle x, y \rangle = \langle y, x \rangle$

f. $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$

g. $\langle ax, y \rangle = \langle x, ay \rangle = a\langle x, y \rangle$ where $a \in \mathbb{R}$ $x, y \in \mathbb{R}^n$

h. $\langle x, x \rangle = |x|^2$.

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, i.e. $T(x+y) = T(x) + T(y)$ and $T(\lambda x) = \lambda \cdot T(x)$ for $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Note — Same symbol may not imply same operation! adding in $\mathbb{R}^n$ adding in $\mathbb{R}^m$ multiplication in $\mathbb{R}^n$ multiplication in $\mathbb{R}^m$.

Then we consider matrix representation of linear transformation $T \to [T]^{\mathcal{U}}_{\mathcal{B}}$. For us, it suffices to use the standard bases.

For each $e_i \in \mathbb{R}^n$, $T(e_i) = \sum_{j=1}^{m} a_{ji} e_j$. Then $[T] = (a_{ij})_{1 \leq i \leq n,\ 1 \leq j \leq m} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$. This is an $m \times n$ matrix.

Let $y = T(x)$. $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ so $x = (x^1, \cdots, x^n)$, $y = (y^1, \cdots, y^m)$. Then we get $\underbrace{\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix}}_{m \times 1} = \underbrace{[T]}_{m \times n} \underbrace{\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}}_{n \times 1}$

Matrix representations are compatible under our standard operations on linear transformations:

- if $T_1: \mathbb{R}^n \to \mathbb{R}^m$, $T_2: \mathbb{R}^n \to \mathbb{R}^m$, $[T_1 + T_2] = [T_1] + [T_2]$ where $(T_1 + T_2)(x) = T_1(x) + T_2(x)$ $\forall x \in \mathbb{R}^n$.
- $T: \mathbb{R}^n \to \mathbb{R}^m$, $S: \mathbb{R}^m \to \mathbb{R}^k$. then $[S \circ T] = \underbrace{[S]}_{k \times m} \cdot \underbrace{[T]}_{m \times n}$. This is sensible as $S \circ T: \mathbb{R}^n \to \mathbb{R}^k$ gives a $k \times n$ matrix.

<u>Functions, Limits, Continuity.</u>

Consider functions $f: \mathbb{R}^n \to \mathbb{R}^m$. Occasionally, f is defined on a subset, so $f: A \to \mathbb{R}^m$ where $A \subseteq \mathbb{R}^n$.

Then $f(x^1, \cdots, x^n) = f^1(x^1, \ldots, x^n) e_1 + f^2(x^1, \cdots, x^n) e_2 + \cdots + f^m(x^1, \cdots, x^n) e_m = (f^1(x), f^2(x), \cdots, f^m(x))$, where f^i are scalar fields $f^i: A \to \mathbb{R}$.

We also have projections: $\pi^i: \mathbb{R}^m \to \mathbb{R}$, $\pi^i(y^1, \cdots, y^m) = y^i$. Then

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "f"] \arrow[dr, "\pi^i \circ f = f^i"'] & \mathbb{R}^m \arrow[d, "\pi^i"] \\ & \mathbb{R} \end{tikzcd}$$

$\boxed{\text{Definition}}$ $\lim_{x \to a} f(x) = b$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. $0 < \underbrace{|x - a|}_{\text{norm in } \mathbb{R}^n} < \delta \Rightarrow \underbrace{|f(x) - b|}_{\text{norm in } \mathbb{R}^m} < \varepsilon$ for $f: \mathbb{R}^n \to \mathbb{R}^m$

f is continuous at a if $\lim_{x \to a} f(x) = f(a)$. f is continuous on A if it is continuous for all $a \in A$.

Remark — $\underbrace{\lim_{x \to a} f(x) = b \iff \lim_{h \to 0} f(a+h) = b}_{\text{vector functions tending to a vector}} \iff \lim_{h \to 0} \underbrace{|f(a+h) - b|}_{\text{real number.}} = 0$.

$\boxed{\text{Theorem}}$ Assume $f, g: \mathbb{R}^n \to \mathbb{R}^m$ and $\lim_{x \to a} f(x) = b$, $\lim_{x \to a} g(x) = c$. Then

(a) $\lim_{x \to a} (f(x) + g(x)) = b + c$, (b) $\lim_{x \to a} (\lambda \cdot f(x)) = \lambda \cdot b$, $\lambda \in \mathbb{R}$ (c) $\lim_{x \to a} f(x) \cdot g(x) = b \cdot c$ $(\in \mathbb{R}!)$ dot products in $\mathbb{R}^m$, yields a real number.

(d) $\lim_{x \to a} |f(x)| = |b|$, where $|\cdot|$ is a norm in $\mathbb{R}^m$.

Proof — (d) NTP: $\forall \varepsilon > 0$ $\exists \delta > 0$, $0 < |x - a| < \delta \Rightarrow ||f(x)| - |b|| < \varepsilon$. We have $\lim_{x \to a} f(x) = b \iff \forall \varepsilon > 0$, $0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon$.

Thus, by triangle inequality, $||f(x)| - |b|| \leq |f(x) - b| < \varepsilon$, so we can use same δ. // q.e.d.

(c) Try $f(x) \cdot g(x) \to b \cdot c$ Then consider difference, $f(x) \cdot g(x) - b \cdot c = f(x) \cdot g(x) - f(x) \cdot c + f(x) \cdot c - b \cdot c = f(x) \cdot (g(x) - c) + (f(x) - b) \cdot c$

Apply absolute values, $|f(x) \cdot g(x) - b \cdot c| = |f(x) \cdot (g(x) - c) + (f(x) - b) \cdot c| \overset{TI}{\leq} |f(x) \cdot (g(x) - c)| + |(f(x) - b) \cdot c| \overset{CSI}{\leq} |f(x)||g(x) - c| + |f(x) - b| \cdot |c|$

Here, $|\cdot|$ is norm in $\mathbb{R}^m$. Since $\lim_{x \to a} f(x) = b$, $f(x)$ is bounded close to b.

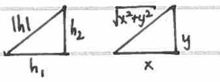
Thus, $|f(x)||g(x) - c| \xrightarrow{x \to a} 0$.

4 October 2013
Dr. Yiannis PETRIDIS
Maths 706.

General Remarks: • $F:\mathbb{R}^n\to\mathbb{R}^m$, $f=(f^1,\dots,f^m)$ • f is continuous $\Leftrightarrow$ f^i is continuous at a for $i=1,\dots,m$

• A polynomial in $x^1,\dots,x^n$ is a (finite) linear combination of monomials in terms of the form $(x^1)^{i_1}(x^2)^{i_2}\dots(x^n)^{i_n}$

where $i_1,\dots,i_n\in\mathbb{N}\cup\{0\}$. Polynomials are continuous. Rational functions $R(x)=\frac{P(x)}{Q(x)}$ where P,Q are polynomials are continuous $\forall x$ where $q(x)\neq 0$.



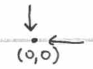
Ex. Show that $f(x,y)=\begin{cases}\frac{xy}{\sqrt{x^2+y^2}} & \text{for } (x,y)\neq 0\\ 0 & \text{for } (x,y)=(0,0)\end{cases}$

Soln. $|f(x,y)-0|=\frac{|xy|}{\sqrt{x^2+y^2}}\leq\frac{\sqrt{x^2+y^2}\cdot\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}=\sqrt{x^2+y^2}=|(x,y)|\to 0$ as $(x,y)\to 0$. $\Rightarrow$ continuous at $(0,0)$.

Ex. Let $f(x,y)=\frac{x^2-y^2}{x^2+y^2}$ for $(x,y)\neq(0,0)$. Prove that $\lim_{(x,y)\to(0,0)}f(x,y)$ does not exist.

Soln. Assume $\lim_{(x,y)\to(0,0)}f(x,y)=\ell$. i.e. $\forall\varepsilon>0\ \exists\delta>0$ s.t. $0<|(x,y)-(0,0)|<\delta\Rightarrow|f(x,y)-\ell|<\varepsilon$.

$$\left.\begin{aligned}&\text{Let } (x,y)=(x,0). && f(x,0)=\frac{x^2-0^2}{x^2+0^2}=\frac{x^2}{x^2}=1\Rightarrow|1-\ell|<\varepsilon\Rightarrow\ell=1.\\ &\text{Let } (x,y)=(0,y) && f(0,y)=\frac{0^2-y^2}{0^2+y^2}=\frac{-y^2}{y^2}=-1\Rightarrow|-1-\ell|<\varepsilon\Rightarrow\ell=-1\end{aligned}\right\}\ \ell \text{ does not exist}\Rightarrow\text{limit does not exist}_{//}\text{ q.e.d.}$$



Remark – Moreover, we could approach along any direction and get a different ℓ. For instance, let $y=mx$ for arbitrary x.

$$f(x,y)=f(x,mx)=\frac{x^2-(mx)^2}{x^2+(mx)^2}=\frac{1-m^2}{1+m^2},\ \text{which depends on the slope}$$

Note – Even if the limit along all straight lines is the same, limit may not exist. We still must consider all curves (e.g. parabolae).

What happens when we try to compute iterated limits? Using previous example $f(x,y)=\frac{x^2-y^2}{x^2+y^2}$,

• $\lim_{x\to0}\left(\lim_{y\to0}f(x,y)\right)\overset{x\neq0}{=}\lim_{x\to0}\left(\lim_{y\to0}\frac{x^2-y^2}{x^2+y^2}\right)=\lim_{x\to0}\frac{x^2-0^2}{x^2+0^2}=\lim_{x\to0}1=1$ • $\lim_{y\to0}\left(\lim_{x\to0}f(x,y)\right)\overset{y\neq0}{=}\lim_{y\to0}\left(\lim_{x\to0}\frac{x^2-y^2}{x^2+y^2}\right)=\lim_{y\to0}(-1)=-1$.

It is thus important to note that numerous pathologies of multivariable functions exist.

Lemma. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then $\exists M>0$ s.t. $\boxed{|T(x)|\leq M|x|\quad\forall x\in\mathbb{R}^n}$

Remark — In functional analysis, we say T is a bounded linear operator.

Proof – $x=x^1e_1+x^2e_2+\dots+x^ne_n$. $T(x)=T\left(\sum_{j=1}^n x^je_j\right)=\sum_{j=1}^n x^jT(e_j)$ by linearity. Taking norms and using triangle inequality,

$$|T(x)|=\left|\sum_{j=1}^n x^jT(e_j)\right|\leq\sum_{j=1}^n|x^j||T(e_j)|=\sum_{j=1}^n|x^j||T(e_j)|\leq\sum_{j=1}^n|x||T(e_j)|=\left(\sum_{j=1}^n|T(e_j)|\right)|x|.\ \text{Let } M=\sum_{j=1}^n|T(e_j)| \text{ which is finite}_{//}\text{ q.e.d.}$$

Theorem. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then it is continuous for all $a\in\mathbb{R}^n$

Proof – $|T(a+h)-T(a)|\overset{\text{linear}}{=}|T(a+h-a)|=|T(h)|\leq M|h|$. Given $\varepsilon>0$, take $\delta=\frac{\varepsilon}{M}$, then $|h|<\delta\Rightarrow|T(a+h)-T(a)|<M|h|<M\cdot\frac{\varepsilon}{M}=\varepsilon$

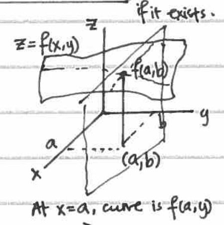
$\therefore$ T is continuous at $a\in\mathbb{R}^n$$_{//}$ q.e.d.

Theorem. If f is continuous at a and g is continuous at $f(a)$, then $g\circ f$ is continuous at a.

Proof – omitted, similar to MATH1101.

Definition. Let $f:\mathbb{R}^n\to\mathbb{R}$, $a\in\mathbb{R}^n$. We define the partial derivative $\frac{\partial f}{\partial x^i}(a)=\lim_{\substack{h\to0\\h\in\mathbb{R}}}\frac{f(a^1,a^2,\dots,a^{i-1},a^i+h,a^{i+1},\dots,a^n)-f(a^1,\dots,a^{i-1},a^i,a^{i+1},\dots,a^n)}{h}$ if it exists.

In $\mathbb{R}^2$, $\frac{\partial f}{\partial x}(a,b)=\lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a}$ and $\frac{\partial f}{\partial y}(a,b)=\lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b}$. [Notation: $\frac{\partial f}{\partial x}(a,b)=f_x(a,b)=D_xf(a,b)$].



Geometric meaning: Consider $x=a$, which is a plane intersecting the surface at a curve $f(a,y)$.

The slope of the tangent line at $f(a,b)$ is $\frac{\partial f}{\partial y}(a,b)$.

Likewise, for the $y=b$ plane, the slope of the tangent line is $\frac{\partial f}{\partial x}(a,b)$

Ex. Consider $f(x,y)=\begin{cases}\frac{x^2-y^2}{x^2+y^2} & (x,y)\neq(0,0)\\ 1 & (x,y)=(0,0)\end{cases}$. Calculate partial derivatives at $(0,0)$ w.r.t. x and y.

Soln. $\frac{\partial f}{\partial x}(0,0)=\lim_{x\to0}\frac{f(x,0)-f(0,0)}{x-0}=\lim_{x\to0}\frac{\frac{x^2-0^2}{x^2+0^2}-1}{x}=\lim_{x\to0}\frac{1-1}{x}=\lim_{x\to0}0=0$.

$\frac{\partial f}{\partial y}(0,0)=\lim_{y\to0}\frac{f(0,y)-f(0,0)}{y-0}=\lim_{y\to0}\frac{\frac{0^2-y^2}{0^2+y^2}-1}{y}=\lim_{y\to0}\frac{-1-1}{y}=\lim_{y\to0}\frac{-2}{y}$ does not exist.

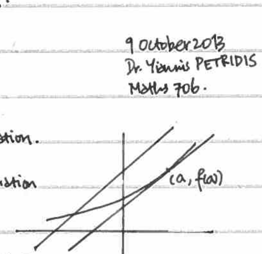
We know that $f'(a)=\lim_{h\to0}\frac{f(a+h)-f(a)}{h}$. However, this definition makes no sense if $f:\mathbb{R}^n\to\mathbb{R}^m$ as it is not well-defined.

9 October 2013
Dr. Yiannis PETRIDIS
Maths 706.

Recall that if $f:\mathbb{R}^n\to\mathbb{R}^m$, the derivative is given by working on $\frac{f(a+h)-f(a)}{h}$.

For $f:\mathbb{R}\to\mathbb{R}$, $f'(a)=\lim_{h\to0}\frac{f(a+h)-f(a)}{h}\Rightarrow 0=\lim_{h\to0}\frac{f(a+h)-f(a)-hf'(a)}{h}$. For $x-a=h$,

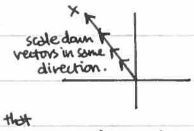
$y=f(a)+(x-a)f'(a)$ gives the equation of the tangent line at $(a,f(a))$. Then $h\to f(a)+f'(a)h$ is an affine transformation.



Definition. We say $f:\mathbb{R}^n\to\mathbb{R}^m$ (or $f:U\to\mathbb{R}^m$, U open in $\mathbb{R}^n$, $a\in U$) is differentiable at a if we can find a linear transformation $\lambda:\mathbb{R}^n\to\mathbb{R}^m$ s.t. $\lim_{h\to0}\frac{|f(a+h)-f(a)-\lambda(h)|}{|h|}=0$.

The linear transformation λ is called the (total) derivative of f at a. Notation: $Df(a)=\lambda\ \forall h\in\mathbb{R}^n$, $Df(a)(h)=\lambda(h)$.

The matrix representation of λ w.r.t. standard bases of $\mathbb{R}^n$, $\mathbb{R}^m$ is called the Jacobian matrix $f'(a)\in M_{m\times n}$.



$\boxed{\text{Theorem}}$ If f is differentiable at a and the definition works for two linear transformations $\lambda, \mu : \mathbb{R}^n \to \mathbb{R}^m$, then $\lambda = \mu$.

Proof – Since λ, μ are linear, $\lambda(0) = \mu(0) = 0$. So NTP: that $\lambda(x) = \mu(x)$ $\forall x \in \mathbb{R}^n \setminus \{0\}$. By definition, we have that
$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = \lim_{h \to 0} \frac{|f(a+h) - f(a) - \mu(h)|}{|h|} = 0. \quad \text{Consider } \frac{|\lambda(h) - \mu(h)|}{|h|} = \frac{|\lambda(h) + f(a) - f(a+h) + f(a+h) - f(a) - \mu(h)|}{h}$$
$$\overset{(\Delta)}{\le} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} + \frac{|f(a+h) - f(a) - \mu(h)|}{|h|} \xrightarrow{h \to 0} 0 + 0 = 0. \therefore \lim_{h\to 0} \frac{|\lambda(h) - \mu(h)|}{|h|} = 0.$$ Take $h = t \cdot x$ with $t \to 0$ s.t. $h \to 0$.
Then $$0 = \lim_{t\to 0} \frac{|\lambda(tx) - \mu(tx)|}{|tx|} = \lim_{t \to 0} \frac{|t\lambda(x) - t\mu(x)|}{|t||x|} = \lim_{t\to 0} \frac{|t|}{|t|} \cdot \frac{|\lambda(x) - \mu(x)|}{|x|} = \lim_{t\to 0} \frac{|\lambda(x) - \mu(x)|}{|x|} = \frac{|\lambda(x) - \mu(x)|}{|x|}$$ (independent of t) // q.e.d.

Examples of derivatives –

1. $f: \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = k$, constant. Then $Df(a) = O : \mathbb{R}^n \to \mathbb{R}^m$, $O(h) = 0 \in \mathbb{R}^m$ $\forall h \in \mathbb{R}^n$. $f(a+h) - f(a) - Df(a)(h) = k - k - O(h) = 0 - 0 = 0$ ✓.
2. $f: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. $Df(a): \mathbb{R}^n \to \mathbb{R}^m$ is also a linear transformation. We examine numerator: $f(a+h) - f(a) - Df(a)(h) \overset{\text{linear}}{=} f(a) + f(h) - f(a) - Df(a)(h)$
$\Rightarrow f(a+h) - f(a) - Df(a)(h) = f(h) - Df(a)(h) = 0$. We can thus take $Df(a) = f : \mathbb{R}^n \to \mathbb{R}^m$.

Remark – $f: \mathbb{R} \to \mathbb{R}$ is linear, so $f(x) = c \cdot x$. Then $f'(x) = c$ = 1x1 Jacobian matrix for $Df(a)$. [i.e. $Df(a) = f$].

• Since $f'(a)$ is the matrix representation of $Df(a)$, let $h = (h^1, h^2, \ldots, h^n) \in \mathbb{R}^n$. Then we can calculate action: $Df(a)(h) = \underset{m\times n}{f'(a)} \underset{n \times 1}{\begin{pmatrix} h^1 \\ h^2 \\ \vdots \\ h^n \end{pmatrix}}$

$\boxed{\text{Theorem}}$ If $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a, then it is continuous at a.

Proof – $|f(a+h) - f(a)| = |f(a+h) - f(a) + Df(a)(h) - Df(a)(h)| \overset{(\Delta)}{\le} |f(a+h) - f(a) - Df(a)(h)| + |Df(a)(h)| = \frac{|h||f(a+h) - f(a) - Df(a)(h)|}{|h|} + |Df(a)(h)|$

Taking limits as $h \to 0$, $\lim_{h\to 0} |Df(a)(h)| \overset{\text{linear} \Rightarrow \text{continuous}}{=} |\lim_{h\to 0} Df(a)(h)| = |Df(a)(0)| \overset{\text{linear}}{=} |0| = 0 \Rightarrow \lim_{h \to 0} |f(a+h) - f(a)| = 0$.

$\boxed{\text{Theorem}}$ (Chain Rule)

Let $f: \mathbb{R}^n \to U \subseteq \mathbb{R}^m$ be differentiable at a, and let $g: U \to \mathbb{R}^p$ be differentiable at $b = f(a)$. Then $g \circ f: \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a.

Then $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$ is the derivative. In Jacobian matrices, $(g\circ f)'(a) = g'(f(a)) \cdot f'(a)$ where composition takes form of matrix multiplication.

Proof – We know that $\lim_{h\to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = \lim_{x \to a} \frac{|f(x) - f(a) - \lambda(x-a)|}{|x-a|}$. Let $\varphi(x) = f(x) - f(a) - \lambda(x-a)$, $\overset{(**)}{}$ $\varphi: \mathbb{R}^n \to \mathbb{R}^m$, $\varphi(a) = 0$.

Then $Df(a) = \lambda \Leftrightarrow \lim_{x \to a} \frac{|\varphi(x)|}{|x-a|} = 0$. Using this, let $b = f(a)$, $\lambda = Df(a)$, $\mu = Dg(b)$. NTP: $D(g\circ f)(a) = \mu \circ \lambda$.

$\overset{x}{\mathbb{R}^n} \to \overset{y}{\mathbb{R}^m} \to \mathbb{R}^p$. $g(y) - g(b) - \mu(y-b) = \overset{(*)}{\psi(y)}$, with $\lim_{y \to b} \frac{|\psi(y)|}{|y-b|} = 0$. We have $(g\circ f)(x) - (g \circ f)(a) = g(f(x)) - g(b)$

By (*), $(g\circ f)(x) - (g\circ f)(a) = \mu(f(x) - b) + \psi(f(x)) \overset{(**)}{=} \mu(\lambda(x-a) + \varphi(x)) + \psi(f(x)) \overset{\mu \text{ linear}}{=} \mu(\lambda(x-a)) + \mu(\varphi(x)) + \psi(f(x))$.

$\Rightarrow (g\circ f)(x) - (g\circ f)(a) - (\mu \circ \lambda)(x-a) = \mu(\varphi(x)) + \psi(f(x))$ NTP: $\lim_{x\to a} \frac{|\mu(\varphi(x)) + \psi(f(x))|}{|x-a|} = 0$.

$$\lim_{x\to a} \frac{|\mu(\varphi(x)) + \psi(f(x))|}{|x-a|} \le \lim_{x\to a} \frac{|\mu(\varphi(x))|}{|x-a|} + \lim_{x \to a} \frac{|\psi(f(x))|}{|x-a|} \overset{\text{lemma}}{\le} \lim_{x\to a} \frac{M|\varphi(x)|}{|x-a|} + \lim_{x\to a} \frac{|\psi(f(x))|}{|x-a|} = M \overbrace{\lim_{x\to a} \frac{|\varphi(x)|}{|x-a|}}^{0} + \lim_{x \to a} \frac{|\psi(f(x))|}{|x-a|}$$
$= \lim_{x\to a} \frac{|\psi(f(x))|}{|x-a|}$. Note that $\frac{|\psi(f(x))|}{|x-a|} = \frac{|\psi(y)|}{|x-a|} = \underbrace{\frac{|\psi(y)|}{|y-b|}}_{\to 0} \cdot \frac{|y-b|}{|x-a|}$ as $x \to a$, $f(x) = y \to b = f(a)$ since y is continuous at a.

The limit will tend to 0 only if we can show that the quantity $\frac{|y-b|}{|x-a|}$ remains bounded.
$$\frac{|y-b|}{|x-a|} = \frac{|\lambda(x-a) + \varphi(x)|}{|x-a|} \le \frac{|\lambda(x-a)|}{|x-a|} + \underbrace{\frac{|\varphi(x)|}{|x-a|}}_{\to 0} \overset{\text{lemma}}{\le} \frac{K|x-a|}{|x-a|} + 0 = K$$ ($\because \exists K$ s.t. $|\lambda(x)| \le K|x|$) // q.e.d.

$\boxed{\text{Theorem}}$ (a) Let $s: \mathbb{R}^2 \to \mathbb{R}$, $s(x,y) = x+y$. Then s is differentiable and $Ds = s$.

(b) Let $p: \mathbb{R}^2 \to \mathbb{R}$, $p(x,y) = xy$. Then p is differentiable and $Dp(a,b)(h^1, h^2) = ah^2 + bh^1$.

Proof – (a) It suffices to show s is linear: Let $(x,y), (x', y') \in \mathbb{R}^2$. $s((x,y) + (x',y')) = s(x+x', y+y') \overset{\text{def}}{=} (x+x') + (y+y') = (x+y) + (x'+y') = s(x,y) + s(x',y')$.

$s(\lambda(x,y)) = s(\lambda x, \lambda y) = \lambda x + \lambda y = \lambda(x+y) = \lambda s(x,y) \Rightarrow s$ is linear $\Rightarrow$ s is differentiable, $Ds = s$ // q.e.d.

(b) $p((a,b) + (h^1, h^2)) - p(a,b) - Dp(a,b)(h^1, h^2) = p(a+h^1, b+h^2) - ab - (ah^2 + bh^1) = (a+h^1)(b+h^2) - ab - ah^2 - bh^1$
$= ab + h^1 b + ah^2 + h^1h^2 - ab - ah^2 - bh^1 = h^1h^2$. Then $$\frac{|p(a+h^1, b+h^2) - p(a,b) - Dp(a,b)(h^1,h^2)|}{|(h^1,h^2)|} = \frac{|h^1 \cdot h^2|}{\sqrt{(h^1)^2 + (h^2)^2}} \le \frac{\sqrt{(h^1)^2 + (h^2)^2}^2}{\sqrt{(h^1)^2 + (h^2)^2}} = \sqrt{(h^1)^2 + (h^2)^2} \to 0$$ // q.e.d.

Remark – If $h: \mathbb{R}^n \to \mathbb{R}$ is linear, it is called a <u>linear functional</u>: $\begin{matrix} h(x+y) = h(x) + h(y) \\ h(\lambda x) = \lambda h(x) \end{matrix} \Leftrightarrow h(\lambda x + y) = \lambda h(x) + h(y)$ $\forall x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Let $g^i : \mathbb{R}^n \to \mathbb{R}$ be linear functionals $i = 1, 2, \ldots, m$. Then we can construct linear $g: \mathbb{R}^n \to \mathbb{R}$ s.t. $g(x) = (g^1(x), g^2(x), \ldots, g^m(x)) \in \mathbb{R}^m$.

We check that g is linear: $g(\lambda x + y) = (g^1(\lambda x + y), \ldots, g^m(\lambda x + y)) \overset{\text{linearity of } g_i}{=} (\lambda g^1(x) + g^1(y), \ldots, \lambda g^m(x) + g^m(y)) = (\lambda g^1(x), \ldots, \lambda g^m(x)) + (g^1(y), \ldots, g^m(y))$
$= \lambda(g^1(x), \ldots, g^m(x)) + (g^1(y), \ldots, g^m(y)) = \lambda g(x) + g(y) \Rightarrow g$ is indeed linear.

Matrix representation of each g^i is a row vector. $[g^i] = [g^i_1, g^i_2, \ldots, g^i_n]$. Then representation of g is $[g] = \begin{bmatrix} [g^1] \\ [g^2] \\ \vdots \\ [g^m] \end{bmatrix} = \begin{pmatrix} g^1_1 & \cdots & g^1_n \\ \vdots & \ddots & \vdots \\ g^m_1 & \cdots & g^m_n \end{pmatrix}$.

$\boxed{\text{Theorem}}$ Let $f: \mathbb{R}^n \to \mathbb{R}^m$, $f = (f^1, f^2, \ldots, f^m)$. Then f is differentiable at a $\Leftrightarrow$ each f^i is differentiable at a for $i = 1, 2, \ldots, m$.

Moreover, $Df(a)(h) = (Df^1(a)(h), Df^2(a)(h), \ldots, Df^m(a)(h))$. [or in terms of Jacobians, $Df^i: \mathbb{R}^n \to \mathbb{R}$, $f'(a) = \begin{pmatrix} (f^1)'(a) \\ \vdots \\ (f^m)'(a) \end{pmatrix}$

Here, $f^i : \mathbb{R}^n \to \mathbb{R}$ are linear functionals.

Proof – ($\Rightarrow$). Assume f is differentiable. Consider linear function $\pi^i(y^1, \dots, y^m) = y^i$. Then $f^i = \pi^i \circ f$. By Chain Rule, we have $Df^i(a) = D(\pi^i \circ f)(a)$

$= D\pi^i(f(a)) \circ Df(a) = \pi^i \circ Df(a)$.

($\Leftarrow$) $f(a+h) - f(a) - Df(a)(h) = (f^1(a+h) - f^1(a), \dots, f^m(a+h) - f^m(a)) - (Df^1(a)(h), \dots, Df^m(a)(h)) = (f^1(a+h) - f^1(a) - Df^1(a)(h), \dots, f^m(a+h) - f^m(a) - Df^m(a)(h))$

Then $\dfrac{|f(a+h) - f(a) - Df(a)(h)|}{|h|} \underset{|y| \le \sum_i |y^i|}{\le} \dfrac{|f^1(a+h) - f^1(a) - Df^1(a)(h)|}{h} + \dots + \dfrac{|f^m(a+h) - f^m(a) - Df^m(a)(h)|}{h} \to 0$ // q.e.d.

11 October 2013
Torrington Pl (1-19) 115.
Dr Yiannis PETRIDIS.

$\boxed{\text{Theorem}}$ Let $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}$ be differentiable at a. Then

(1) $f+g$ is differentiable at a: $D(f+g)(a) = Df(a) + Dg(a)$ [sum of linear transformations]

(2) Let $\lambda \in \mathbb{R}$, $\lambda \cdot f$ is differentiable at a: $D(\lambda f)(a) = \lambda \cdot Df(a)$

(Product Rule)
(3) $f \cdot g: \mathbb{R}^n \to \mathbb{R}$ is differentiable at a: $D(f \cdot g)(a) = \underbrace{g(a)}_{\mathbb{R}} \cdot \underbrace{Df(a)}_{LT} + \underbrace{f(a)}_{\mathbb{R}} \cdot \underbrace{Dg(a)}_{LT}$

(Quotient Rule)
(4) If $g(a) \neq 0$, then f/g is differentiable at a: $D(f/g)(a) = \dfrac{1}{g(a)^2}[g(a) \cdot Df(a) - f(a) \cdot Dg(a)]$.

Reminder – By Chain Rule, $f: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^p$, $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$. If $s(x,y) = x+y$, then $Ds = s$ and $p(x,y) = x \cdot y \Rightarrow Dp(a,b)(h^1, h^2) = bh^2 + ah^1$.

Proof – (1) $\begin{matrix} f \searrow \\ g \nearrow \end{matrix} (f,g) \xrightarrow{s} f+g$. Then $D(f+g)(a) = D(s \circ (f,g))(a) = Ds(f(a), g(a)) \circ D(f,g)(a) = s(f(a), g(a)) \circ (Df(a), Dg(a)) = s \circ (Df(a), Dg(a)) = Df(a) + Dg(a)$ // q.e.d.

(3) $x \to (f(x), g(x)) \xrightarrow{p} f(x) \cdot g(x)$. Then $f \cdot g = p \circ (f,g): \mathbb{R}^n \to \mathbb{R}$. Then we have, for $h \in \mathbb{R}^n$, $D(f \cdot g)(a)(h) = D(p \circ (f,g))(a)(h) = [Dp((f(a), g(a)) \circ D(f,g)(a)](h)$

$= [Dp(f(a), g(a)) \circ ((Df(a), Dg(a))](h) = Dp(f(a), g(a))[(Df(a), Dg(a))(h)] = Dp(f(a), g(a))\overbrace{(Df(a)(h), Dg(a)(h))}^{\mathbb{R}^2}$

$= g(a) \cdot Df(a)(h) + f(a)\, Dg(a)(h) = [g(a) \cdot Df(a) + f(a) \cdot Dg(a)](h)$ // q.e.d.

$\boxed{\text{Theorem}}$ If $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a, then all partial derivatives $D_j f^i(a)$ exist and $f'(a) = (D_j f^i(a))$, $i = 1, \dots, m$ and $j = 1, \dots, n$.

Proof – Recall that if $f = (f^1, f^2, \dots, f^m)$, $Df(a) = (Df^1(a), Df^2(a), \dots, Df^m(a))$. $f'(a) = \begin{pmatrix} \text{---}\ (f^1)'(a)\ \text{---} \\ \vdots \end{pmatrix}$. Suppose $f: \mathbb{R}^n \to \mathbb{R}$, NTP: $D_j f(a)$ exists.

$D_j f(a) = \lim_{x^j \to a^j} \dfrac{f(a^1, a^2, \dots, a^{j-1}, x^j, a^{j+1}, \dots, a^n) - f(a^1, \dots, a^{j-1}, a^j, a^{j+1}, \dots, a^n)}{x^j - a^j}$. Let $h: \mathbb{R} \to \mathbb{R}^n$, $h(x) = (a^1, a^2, \dots, a^{j-1}, x, a^{j+1}, \dots, a^n)$

Then $f \circ h: \mathbb{R} \to \mathbb{R}$, $(f \circ h)(x) = f(a^1, \dots, x, \dots, a^n)$. Then $D_j f(a) = \dfrac{d}{dx}\underbrace{(f \circ h)}_{\mathbb{R} \to \mathbb{R}}(a^j) = (f \circ h)'(a^j) \overset{\text{chain (Jacobians)}}{=} f'(h(a^j)) \cdot h'(a^j) = f'(a) \cdot h'(a^j)$ — ⊛

$h(x) = (a^1, a^2, \dots, a^{j-1}, x, a^{j+1}, \dots, a^n)$. Here, $h^i(x) = a^i$ $\forall i$ except j, $h^j(x) = x \Rightarrow$ all components are differentiable, $Dh^i(a^j) = 0$, $(h^i)'(a^j) = 0$.

And also, $Dh^j = h^j$ (linear) $\Rightarrow (h^j)'(a^j) = I_1 \Rightarrow h'(a^j) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow$ 1 in j-component. So, from ⊛, $D_j f(a) = f'(a) \cdot e_j = j\text{-entry of } \underset{\text{(row vector)}}{f'(a)}$

$\Rightarrow f'(a) = (D_j f^i(a))$ exists // q.e.d.

$\boxed{\text{Theorem}}$ $f: \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at a ~~if differentiable at a~~ if the following hold: $\exists$ an open set U, $a \in U$ s.t. $D_j f^i(x)$ exist $\forall x \in U$ and are continuous at a.

Proof – Let $n=2$. Then $f(a^1+h^1, a^2+h^2) - f(a^1, a^2) = f(a^1+h^1, a^2+h^2) - f(a^1+h^1, a^2) + f(a^1+h^1, a^2) - f(a^1, a^2)$. By Mean Value Theorem, —⊛ Define $g(x) = (x, a^2)$, $D_1 f$ is derivative of $g(x)$.

$\exists b^1 \in (a^1, a^1+h^1)$, $f(a^1+h^1, a^2) - f(a^1, a^2) = D_1 f(b^1, a^2) \cdot h^1$. Likewise, $\exists b_2 \in (a^2, a^2+h^2)$ s.t. $f(a^1+h^1, a^2+h^2) - f(a^1+h^1, a^2) = D_2 f(a^1+h^1, b^2) \cdot h^2$.

(a^1+h^1, a^2+h^2)
(a^1, a^2) (a^1+h^1, a^2).

If f is differentiable, $D_1 f(a^1, a^2) h^1 + D_2 f(a^1, a^2) h^2$ is the function we need to consider. Then we have:

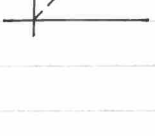
18 October 2013
Maths 706.

$\left|\dfrac{f(a+h) - f(a) - D_1 f(a^1, a^2) h^1 - D_2 f(a^1, a^2) h^2}{h}\right| = \dfrac{|(D_1 f(b^1, a^2) h^1 - D_1 f(a^1, a^2) h^1) + (D_2 f(a^1+h^1, b^2) h^2 - D_2 f(a^1, a^2) h^2)|}{|h|}$

$\overset{(\Delta)}{\le} \dfrac{|D_1 f(b^1, a^2) - D_1 f(a^1, a^2)||h^1|}{|h|} + \dfrac{|D_2 f(a^1+h^1, b^2) - D_2 f(a^1, a^2)||h^2|}{|h|} \le |D_1 f(b^1, a^2) - D_1 f(a^1, a^2)| + |D_2 f(a^1+h^1, b^2) - D_2 f(a^1, a^2)|$

By continuity, as $h \to 0$, $b^1 \to a^1$, $b^2 \to a^2$. So, expression $\to 0 + 0 = 0$, so f is differentiable // q.e.d.

$\boxed{\text{Definition}}$ Let $f: \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}^n$, $u \in \mathbb{R}^n$. Then the <u>directional derivative</u> of f at a in the direction of u is $D_u f(a) = \lim_{t \to 0} \dfrac{f(a+tu) - f(a)}{t}$



Note – Compare this to definition of partial derivative, then $D_i f(a) = D_{e_i} f(a)$.

Having all directional derivatives does not necessarily imply continuity at a.

$\boxed{\text{Ex}}$ Let $G(x,y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. Show all directional derivatives exist at 0, but $G(x,y)$ is not continuous at 0.

Soln. Let $\underline{u} = (u^1, u^2)$. Then $\dfrac{G((0,0) + t(u^1, u^2)) - G((0,0))}{t} = \dfrac{G(tu^1, tu^2)}{t}$. Consider the case where $u^2 \neq 0$. Then we get:

$\dfrac{G(tu^1, tu^2)}{t} = \dfrac{t^2 (u^1)^2 t u^2}{t(t^4 (u^1)^4 + t^2 (u^2)^2)} = \dfrac{(u^1)^2 u^2}{t^2 (u^1)^4 + (u^2)^2} \xrightarrow{t \to 0} \dfrac{(u^1)^2 u^2}{(u^2)^2} = \dfrac{(u^1)^2}{u^2} \in \mathbb{R}$ exists. However, if $u^2 = 0$, $\dfrac{G(tu^1, tu^2)}{t} = \dfrac{G(tu^1, 0)}{t} = \dfrac{0}{t} \xrightarrow{t \to 0} 0$.

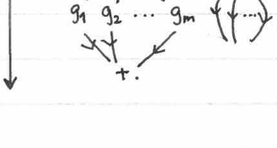
Hence all directional derivatives exist // q.e.d. However, G is not continuous at $(0,0)$. Take $y = x^2$, then $G(x, x^2) = \dfrac{x^4}{2x^4} = \dfrac{1}{2} \not\to 0 = G(0,0)$ // q.e.d.

$\boxed{\text{Definition}}$ f is continuously differentiable if $D_j f^i(x)$ exist and are continuous $\forall i, j$.

$\boxed{\text{Theorem}}$ Let $g_1, \dots, g_m: \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable at a, $f: \mathbb{R}^m \to \mathbb{R}$ be differentiable at $(g_1(a), \dots, g_m(a))$.

Define $F: \mathbb{R}^n \to \mathbb{R}$, then $F(x) = f(g_1(x), \dots, g_m(x))$. Then $D_i F(a) = \sum_{j=1}^{m} D_j f(g_1(a), \dots, g_m(a))\, D_i(g_j(a))$.

Proof – Let $F = f \circ g$ and use chain rule. Specifics omitted //.



Summing up all terms, we get the derivative.

Inverse function theorem

Let us consider case $n=1$, let $I \subset \mathbb{R}$, $f: I \to \mathbb{R}$ be continuously differentiable. $f'(a) > 0 \Rightarrow f$ is strictly increasing in some $J \subset I$ s.t. the point $a \in J \Rightarrow f$ is injective. Likewise, $f'(a) < 0 \Rightarrow f$ strictly decreasing in $J \subset I$ s.t. $a \in J \Rightarrow f$ is injective. Then $f'(a) \neq 0 \Rightarrow f^{-1}$ exists $: W \to \mathbb{R}$, $f(J) = W$.

$\boxed{\text{Theorem}}$ (Inverse function theorem)

f^{-1} is differentiable, $f^{-1}(y) = \frac{1}{f'(f^{-1}(y))}$ $\forall y \in W$.

Proof - $(f^{-1})'(y) = \lim_{h \to 0} \frac{f^{-1}(y+h) - f^{-1}(y)}{h} = \lim_{\delta \to 0} \frac{x + \delta - x}{f^{-1}(x+\delta) - f^{-1}(x)} = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$

Note - If $f: \mathbb{R}^n \to \mathbb{R}^n$, then $f'(a)$ is an $n \times n$ matrix, $f^{-1}(y) = (f'(f^{-1}(y)))^{-1} \Rightarrow \det f(a) \neq 0$.

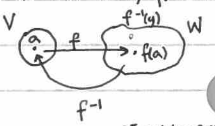
Let $f: \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable at on open set A with $a \in A$, $\det f'(a) \neq 0$. Then $\exists$ open $V \subseteq A$, $a \in V$ and open $W: f(a) \in W$ s.t. $f: V \to W$ is bijective, $f^{-1}: W \to V$ is continuous differentiable and $(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$ $\forall y \in W$.

23 October 2013
Dr. Yiannis PETRIDIS.
Maths 706.

Today's lecture focuses on the inverse function theorem — refer to Handout 2!

$\boxed{\text{Theorem}}$ (Inverse function theorem, restated).

Let $f: \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable (i.e. $D_j f^i$ exist and are continuous). $\det \underset{n \times n}{f'(a)} \neq 0$. Then $\exists$ V, W open in $\mathbb{R}^n$ with $a \in V$, $f(a) \in W$ s.t. $f: V \to W$ is bijective, f^{-1} continuously differentiable on W, s.t. $(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}$ $\forall y \in W$.



Proof - For proof of theorem and lemma, see handout 2.

25 October 2013
Dr. Yiannis PETRIDIS
Maths 706.

$\boxed{\text{Ex}}$ Consider $f(x,y) = (xy, x^2+y^2) = (z, w)$, $f: \mathbb{R}^2 \to \mathbb{R}^2$. Find $\frac{\partial x}{\partial z}, \frac{\partial x}{\partial w}, \frac{\partial y}{\partial z}, \frac{\partial y}{\partial w}$.

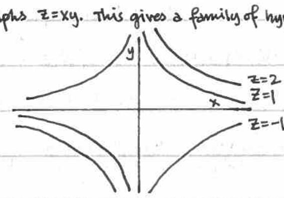
Soln. $f^1(x,y) = xy$, $f^2(x,y) = x^2+y^2$. Then $f'(x,y) = \begin{pmatrix} y & x \\ 2x & 2y \end{pmatrix}$. $\det f'(x,y) = 2y^2 - 2x^2 = 2(x+y)(y-x) \neq 0$ if $x \neq \pm y$.

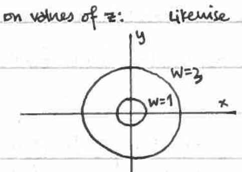
We could substitute $y = \frac{z}{x}$ to get $x^2 + \frac{z^2}{x^2} = w \Rightarrow x^4 - wx^2 + z^2 = 0 \Rightarrow x^2 = \frac{w \pm \sqrt{w^2 - 4z^2}}{2}$, $x = \pm\sqrt{\frac{w \pm \sqrt{w^2-4z^2}}{2}}$. From here, we could technically calculate $\frac{\partial x}{\partial z}, \frac{\partial x}{\partial w}$; but this is troublesome. So instead, we use the inverse function theorem: $(f^{-1})'(z,w) = (f'(x,y))^{-1}$ if $(z,w) = f(x,y)$.

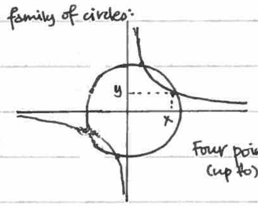
Then $\begin{pmatrix} y & x \\ 2x & 2y \end{pmatrix}^{-1} = \frac{1}{\det f'(x,y)} \begin{pmatrix} 2y & -x \\ -2x & y \end{pmatrix} = \frac{1}{2(y^2-x^2)} \begin{pmatrix} 2y & -x \\ -2x & y \end{pmatrix} = \begin{pmatrix} \partial(f^{-1})^1/\partial z & \partial(f^{-1})^1/\partial w \\ \partial(f^{-1})^2/\partial z & \partial(f^{-1})^2/\partial w \end{pmatrix} = \begin{pmatrix} \partial x/\partial z & \partial x/\partial w \\ \partial y/\partial z & \partial y/\partial w \end{pmatrix}$ //

Note - This answer suffices — there is no need to express answers in terms of z and w in practice.

Consider graphs $z = xy$. This gives a family of hyperbolae based on values of z: Likewise $w = x^2+y^2$ gives a family of circles:



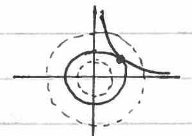




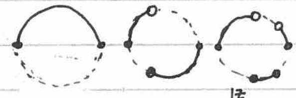
Four points of intersection (up to)

For our previous example, this gives us a geometric interpretation. (z,w) are given by points of intersection. If we jiggle both curves a little bit, we get another solution for (x,y) in the neighbourhood of our original solution.

However, our graph need not appear this way: for instance, if we started with one point of intersection, jiggling the curve could result in two or no solutions $\Rightarrow$ no unique preimage. This occurs, as mentioned, at $y = x$ or $y = \pm x$.

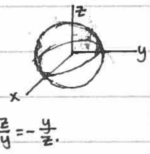


Consider $x^2+y^2=1$. How many functions can we derive from it? Infinitely many. However, only two are continuous.



Let (a,b) be a point on the circle. $b > 0$: Then $y = \sqrt{1-x^2}$ and $\exists I \in (-1,1)$ $\forall x \in I$ $\exists$ unique y s.t. $x^2+y^2=1$, and this function is $y = \sqrt{1-x^2}$. If $b < 0$: $y = -\sqrt{1-x^2}$. If we let $y = g(x)$ to find a functional representation for f,

$x^2 + (g(x))^2 = 1 \Rightarrow 2x + 2g(x)\frac{dg}{dx} = 0 \Rightarrow \frac{dg}{dx} = -\frac{2x}{2g(x)} = \frac{-x}{g(x)}$: If $g(x) = \sqrt{1-x^2}$, $\frac{dg}{dx} = \frac{-x}{\sqrt{1-x^2}}$, if $g(x) = -\sqrt{1-x^2}$, $\frac{dg}{dx} = \frac{-x}{-\sqrt{1-x^2}} = \frac{x}{\sqrt{1-x^2}}$



We can extend this notion to a sphere, $z = \pm\sqrt{1-x^2-y^2} = g(x,y)$ (could be either sign) w.r.t. x, $2x + 2z\frac{\partial z}{\partial x} = 0$, $\frac{\partial z}{\partial x} = -\frac{x}{z}$; w.r.t. y, $\frac{\partial z}{\partial y} = -\frac{y}{z}$.

Then consider $y^2 + x + z^2 - e^z - 4 = 0$. We cannot express $z = g(x,y)$ explicitly, but we try it implicitly and seek values for $\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}$.

w.r.t. x, $1 + 2z\frac{\partial z}{\partial x} - e^z \frac{\partial z}{\partial x} = 0 \Rightarrow \frac{\partial z}{\partial x}(2z - e^z) = -1 \Rightarrow \frac{\partial z}{\partial x} = \frac{-1}{2z - e^z}$. w.r.t. y, $2y + 0 + 2z\frac{\partial z}{\partial y} - e^z\frac{\partial z}{\partial y} - 0 = 0 \Rightarrow \frac{\partial z}{\partial y} = \frac{-2y}{2z - e^z} = \frac{-\partial F/\partial y}{\partial F/\partial z}$.

For the general situation, we consider a system of equations: $f^1(\underbrace{x^1, \dots, x^n}_{\text{parameters}}, \underbrace{y^1, \dots, y^m}_{\text{variables}})$, $f^2(x^1, \dots, x^n, y^1, \dots, y^m)$, $\dots$, $f^m(x^1, \dots, x^n, y^1, \dots, y^m)$.

Then we solve for $y^1, y^2, \dots, y^m$ depending on $(x^1, \dots, x^n)$. We start with a point $(a,b) = (a^1, a^2, \dots, a^n, b^1, \dots, b^m)$ satisfying the system. For $(x^1, \dots, x^n)$ close to $(a^1, \dots, a^n)$, when can we find a solution of the system depending differentially on $x^1, \dots, x^n$? $\begin{cases} g^1(x^1, \dots, x^n) \\ \vdots \\ g^m(x^1, \dots, x^n) \end{cases}$. $g(x^1, x^2, \dots, x^n) = (g^1(\ ), g^2(\ ), \dots, g^m(\ ))$.

$g: \mathbb{R}^n \to \mathbb{R}^m$, $f(x, g(x)) = 0$.

3109-06.

$\boxed{\text{Theorem}}$ (Implicit Function Theorem)

$f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$ is continuously differentiable in U, $(a,b)\in U$, $a\in\mathbb{R}^n$, $b\in\mathbb{R}^m$, $f(a,b)=0$. If $M=(D_{j+n}f^i(a,b))_{i=1,\dots,m,\ j=1,\dots,m}$ is not singular (i.e. $\det M\neq 0$).

Then $\exists$ open set $A\subseteq\mathbb{R}^n$, $a\in A$, $B\subseteq\mathbb{R}^m$, $b\in B$. $\forall x\in A$ $\exists$ unique $y\in B$ with $f(x,y)=0$. Call $y=g(x)$ s.t. $g:A\to B$ then g is differentiable.

Proof – Refer to Handout 3 (and annotations).

30 October 2013
Dr. Yiannis PETRIDIS
Maths 706.

First attempt at implicit solutions to functions came from Newton (1669); implicit differentiation introduced by Leibniz (1684), furthered by Lagrange (1770) to solve $w=f(z)$ for holomorphic, Cauchy did it with power series (later). Finally formally stated by Dini (1876).

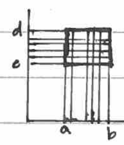
Let $f(x,y)=0$, $y=g(x)$, $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$, $g=(g^1,\dots,g^m)$. i.e. $f^i(x^1,\dots,x^n,g^1(x^1,\dots,x^n),\dots,g^m(x^1,\dots,x^n))=0$.

$\left(\frac{\partial g^1}{\partial x^1},\frac{\partial g^2}{\partial x^1},\dots,\frac{\partial g^m}{\partial x^1}\right)$.

Differentiate this w.r.t. x^1: $\frac{\partial f^i}{\partial x^1}\cdot 1+\frac{\partial f^i}{\partial x^2}\cdot 0+\dots+\frac{\partial f^i}{\partial x^n}\cdot 0+\frac{\partial f^i}{\partial y^1}\cdot\frac{\partial y^1}{\partial x^1}+\dots+\frac{\partial f^i}{\partial y^m}\frac{\partial y^m}{\partial x^1}=0$. f has m components, m equations. Need to solve for m unknowns

This yields a linear system. $\Rightarrow \begin{pmatrix}\partial f^1/\partial y^1 & \cdots & \partial f^1/\partial y^m\\ \vdots & & \vdots\\ \partial f^m/\partial y^1 & \cdots & \partial f^m/\partial y^m\end{pmatrix}\begin{pmatrix}\partial g^1/\partial x^1\\ \vdots\\ \partial g^m/\partial x^1\end{pmatrix}=0 \Rightarrow M\begin{pmatrix}\partial g^1/\partial x^1\\ \vdots\\ \partial g^m/\partial x^1\end{pmatrix}=0.$ $\det M\neq 0\Rightarrow$ solution exists.

Integration.



Consider a function $f:[a,b]\to\mathbb{R}$. Recall that a partition of $[a,b]$, $a=t_0<t_1<\dots<t_n=b$. On $\mathbb{R}^n$, we have a partition of $[a_1,b_1]$ called P_1, of $[a_2,b_2]$ called $P_2,\dots,$ of $[a_n,b_n]$ called P_n. Then the partition $P=(P_1,\dots,P_n)$. If $P_i=\{t_{i1},t_{i2},\dots,t_{i\,n(i)}\}$, then the partition P gives rectangles in $\mathbb{R}^n$ of the form $S=[t_{1j},t_{1(j+1)}]\times[t_{2k},t_{2(k+1)}]\times\dots\times[t_{ns},t_{n(s+1)}]$. The subrectangles of the partition will be denoted by S, $f:A\to\mathbb{R}$ bounded.

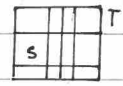
Set $m_S(f)=\inf_{x\in S}f(x)$, $M_S(f)=\sup_{x\in S}f(x)$. Also, volume of rectangle is $v(S)=|t_{1j}-t_{1(j+1)}||t_{2(k+1)}-t_2|\cdots|t_{n(s+1)}-t_{ns}|$.

We define Darboux/Riemann sums: $U(f,P)=\sum_S M_S(f)\cdot v(S)$, $L(f,P)=\sum_S m_S(f)\cdot v(S)\le U(f,P)$. Then we conduct a refinement: let P' be a refinement of P:

$\boxed{\text{Definition}}$ P' is a refinement of P if every subrectangle S of P' is contained in a subrectangle T of P.

$\boxed{\text{Lemma}}$ If P' is a refinement of P, $L(f,P)\le L(f,P')$ and $U(f,P)\ge U(f,P')$.

Proof – (for upper, lower sums are analogous). Let T denote subrectangles of P, S denote subrectangles of P'. Let S be a subrectangle of P', then $\exists$ a T, subrectangle of P s.t. $S\subseteq T$. Then $\sup_{x\in S}f(x)=M_S(f)\le M_T(f)=\sup_{x\in T}f(x)$. $\Rightarrow M_S(f)v(S)\le M_T(f)v(S)$. Then fix T and sum up over all $S\subseteq T$: $\sum_{S\subseteq T}M_S(f)v(S)\le\sum_{S\subseteq T}M_T(f)v(S)=M_T(f)\sum_{S\subseteq T}v(S)=M_T(f)v(T)$. Then we sum up over all T s.t. we get $\sum_T\sum_{S\subseteq T}M_S(f)v(S)\le\sum_T M_T(f)v(T)=U(f,P)\Rightarrow$ LHS $=\sum_{S\subseteq P}M_S(f)v(S)=U(f,P')\Rightarrow U(f,P)\ge U(f,P')$, q.e.d.



$\boxed{\text{Corollary}}$ If P_1,P_2 are any partitions, $L(f,P_1)\le U(f,P_2)$.

Proof – Take P' to be a common refinement. Then $L(f,P_1)\overset{\text{(lemma)}}{\le}L(f,P')\le U(f,P')\le U(f,P_2)$.

$\boxed{\text{Definition}}$ Define $\overline{\int}_A f=\inf_P U(f,P)$ to be the upper Riemann integral, $\underline{\int}_A=\sup_P L(f,P)$ to be the lower Riemann integral.

If $\overline{\int}_A f=\underline{\int}_A f$, we say that f is Riemann integrable.

$\boxed{\text{Theorem}}$ (Riemann's criterion)

f is integrable $\Leftrightarrow$ $\forall\varepsilon>0$, $\exists$ partition P s.t. $U(f,P)-L(f,P)<\varepsilon$.

Proof – ($\Leftarrow$) Let $U(f,P)-L(f,P)<\varepsilon$. $\inf_P[U(f,P)-L(f,P)]=\inf_P U(f,P)-\sup_P L(f,P)<\varepsilon\ \forall\varepsilon>0$. [Fix $\varepsilon>0$. We have P s.t. $U(f,P)-L(f,P)<\varepsilon$. $L(f,P)\le\underline{\int}_A f$, $U(f,P)\ge\overline{\int}_A f\Rightarrow\overline{\int}_A f-\underline{\int}_A f<\varepsilon$].

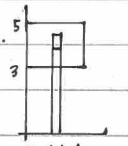
$\therefore$ Since this quantity is non-negative, $\overline{\int}_A f-\underline{\int}_A f=0\Rightarrow\overline{\int}_A f=\underline{\int}_A f\Rightarrow f$ is integrable.

($\Rightarrow$) $\overline{\int}_A f=\inf_P U(f,P)$. Then $\overline{\int}_A f+\frac{\varepsilon}{2}$ is not a lower bound for $U(f,P)\Rightarrow\exists$ partition P_1 s.t. $U(f,P_1)<\overline{\int}_A f+\frac{\varepsilon}{2}$. Fix $\varepsilon>0$, assuming f is integrable. $\underline{\int}_A f=\overline{\int}_A f=\int_A f$. $\underline{\int}_A f=\sup L(f,P)$, then $\underline{\int}_A f-\frac{\varepsilon}{2}$ is not an upper bound for $L(f,P)$. $\exists P_2$ partition s.t. $\underline{\int}_A f-\frac{\varepsilon}{2}<L(f,P_2)\le L(f,P)\le U(f,P)\le U(f,P_1)<\overline{\int}_A f+\frac{\varepsilon}{2}$. Take common refinement P of P_1 and P_2. i.e. we have found partition P s.t. $\underline{\int}_A f-\frac{\varepsilon}{2}\le L(f,P)\le U(f,P)\le\int_A f+\frac{\varepsilon}{2}\Rightarrow U(f,P)-L(f,P)<\varepsilon$, q.e.d.

Not all functions are integrable! For instance, define $f:[0,1]\times[3,5]\to\mathbb{R}$. Let $f(x,y)=\begin{cases}0, & x\text{ is irrational}\\ 1, & x\text{ is rational}\end{cases}$. Then f is not integrable on $[0,1]\times[3,5]$.

For this example, $m_T(f)=0$ and $M_T(f)=1$. Here, we simply consider interval (projected) $x\in[t_j,t_{j+1})$. Then $\exists x$ rational and x irrational.

Then $U(f,P)=1\times[(1-0)(5-3)]=2$, $L(f,P)=0$.



$\boxed{\text{Theorem}}$ (Fubini's theorem, 2 dimensions).

Let $A\subseteq\mathbb{R}^n$, $B\subseteq\mathbb{R}^m$ be rectangles, $f:A\times B\to\mathbb{R}$ be integrable. We define $\mathcal{L}(x)=\underline{\int}_B f(x,y)\,dy$ and $\mathcal{U}(x)=\overline{\int}_B f(x,y)\,dy$. [These always exist, but might not agree]

Then $\mathcal{L},\mathcal{U}:A\to\mathbb{R}$ are Riemann integrable and $\int_{A\times B}f=\int_A\mathcal{L}(x)dx=\int_A\mathcal{U}(x)\,dx=\int_A\left(\underline{\int}_B f(x,y)\,dy\right)dx=\int_A\left(\overline{\int}_B f(x,y)\,dy\right)dx$.

Proof – We need a partition of $A\times B$, $P=(P_1,P_2)$. Here, let P_1 be a partition of A, P_2 be a partition of B; S_A be a subrectangle for P_1, S_B be a subrectangle for P_2.

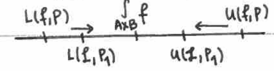
Consider a rectangle of this partition $S=S_A\times S_B$. Fix $x\in S_A$. Then we get the vertical segment $\{x\}\times S_B$. Naturally, $\inf_{\{x\}\times S_B}f\ge\inf_{S_A\times S_B}f\Rightarrow m_{\{x\}\times S_B}f\ge m_{S_A\times S_B}f$.

Then $\sum_{S_B}m_{\{x\}\times S_B}f\cdot v(S_B)\ge\underbrace{\sum_{S_B}m_{S_A\times S_B}f\cdot v(S_B)}_{①}\Rightarrow$ LHS is $L(g_x,P_2)$. $L(g_x,P_2)\le\mathcal{L}(x)$ because $\underline{\int}_B=\sup$. These inequalities are true $\forall x$ on S_A. However, RHS of inequality in ① is x-independent, so it is a lower bound. $\mathcal{L}(x) \geq \sum_{S_B} m_{S_A \times S_B}(f)\, v(S_B)\ \forall x \in S_A$. $\inf_{x \in S_A} \mathcal{L}(x) \geq \sum_{S_B} m_{S_A \times S_B}(f) \cdot v(S_B) \Rightarrow$

$[\inf_{x \in S_A} \mathcal{L}(x)]\, v(S_A) \geq \sum_{S_B} m_{S_A \times S_B}(f) \cdot v(S_B) \cdot v(S_A) \Rightarrow \sum_{S_A} [\inf_{x \in S_A} \mathcal{L}(x)]\, v(S_A) \geq \sum_{S_A} \sum_{S_B} m_{S_A \times S_B}(f)\, \overbrace{v(S_B)\, v(S_A)}^{v(S_A \times S_B)} \Rightarrow L(\mathcal{L}(x), P_1) \geq L(f, P)$

Using similar approach for upper sums — we get $L(f,P) \leq L(\mathcal{L}, P_1) \leq U(\mathcal{L}, P_1) \leq U(\mathcal{U}, P_1)$ [since if $f \leq g$, $U(f,P) \leq U(g,P)$] $\leq U(f,P)$.

As f is integrable, we apply Riemann's criterion: Given $\varepsilon > 0$, $\exists P$ s.t. $U(f,P) - L(f,P) < \varepsilon \Rightarrow U(\mathcal{L}, P_1) - L(\mathcal{L}, P_1) < \varepsilon$. Since $\int_{A \times B} f$ is the unique number trapped between all $L(f,P)$, $U(f,P)$; and $\int_A \mathcal{L}$ is the unique number trapped between $L(\mathcal{L}, P_1)$ and $U(\mathcal{L}, P_1)$.

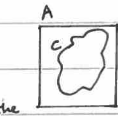


$\Rightarrow \int_{A \times B} f = \int_A \mathcal{L}$ // q.e.d. Similar approach can be used to show $\int_{A \times B} f = \int_A \mathcal{U}$.

remarks — $\left.\begin{array}{l} \mathcal{L}(x) = \underline{\int}_B f(x,y)\, dy \\ \mathcal{U}(x) = \overline{\int}_B f(x,y)\, dy \end{array}\right\}$ always exist i.e. $\forall x \in A$. If $f(\cdot, y)$ is integrable in y, then $\mathcal{L}(x) = \mathcal{U}(x)$.

1. If this holds $\forall x \in A$, $\mathcal{L}(x) = \mathcal{U}(x)\ \forall x \in A$ and they both $= \int_B f(x,y)\, dy$. Then write conclusion of Fubini's Theorem: $\int_{A \times B} f = \int_A \mathcal{L} = \int_A \left(\int_B f(x,y)\, dy \right) dx$.
2. We can consider $\int_{A \times B} f = \int_B \left(\underline{\int}_A f(x,y)\, dx \right) dy = \int_B \left(\overline{\int}_A f(x,y)\, dx \right) dy$ [Proof is the same, just using horizontal segments].

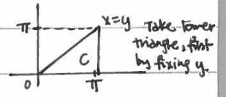
   If $\int_A f(x,y)\, dx$ exists $\forall y \in B$ then we have the familiar formula: $\int_{A \times B} f = \int_B \left(\int_A f(x,y)\, dx \right) dy$.



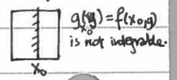
3. If C is a "nice" subset of $\mathbb{R}^n$ that is bounded, we can put it inside a rectangle A. We define χ_C (or sometimes denoted $\mathbb{1}_C$) as the characteristic function, where $\chi_C(a) = \begin{cases} 1 & \text{if } a \in C \\ 0 & \text{if } a \notin C \end{cases}$. Define $\int_C f = \int_A \chi_C \cdot f$

   For example, $\int_0^\pi \left(\int_y^\pi \frac{\sin x}{x}\, dx \right) dy \overset{\text{Fubini}}{=} \int_{[0,\pi] \times [0,\pi]}^{x=y} \mathbb{1}_C(x,y) \cdot \frac{\sin x}{x} \overset{\text{iterated Fubini}}{=} \int_0^\pi \left(\int_0^\pi \mathbb{1}_C(x,y) \frac{\sin x}{x}\, dy \right) dx$ we fix x, so we treat it last.

   $= \int_0^\pi \left(\int_0^x \frac{\sin x}{x}\, dy \right) dx$ [inner integral is a constant of length x] $= \int_0^\pi \sin x\, dx = [-\cos x]_0^\pi = 2$.



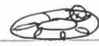
4. It is possible that $\int_B f(x,y)\, dy$ does not exist $\forall x \in A$ e.g. fix $x_0 \in [0,1]$, set $f(x,y) = \begin{cases} 0 & \text{if } x \neq x_0 \\ 1 & \text{if } x = x_0,\ y \in \mathbb{Q} \\ 0 & \text{if } x = x_0,\ y \notin \mathbb{Q} \end{cases}$.

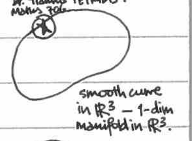


13 November
Dr. Yiannis PETRIDIS.
Maths 7.

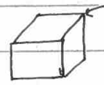
Manifolds in $\mathbb{R}^n$.

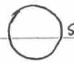
The sphere is a 2D-manifold in $\mathbb{R}^3$. This is not the only such manifold — for instance, the torus is also a 2D manifold in $\mathbb{R}^3$



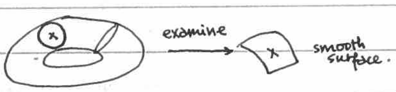


Not every surface is a manifold, such as the following:

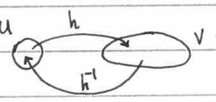




Examine a point x on the surface of a manifold — e.g. a torus: we surround it by a ball (3D), which is "smooth" — can be "flattened out".

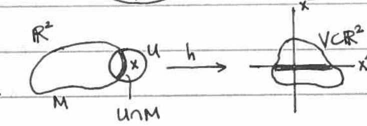


Remark — From here on, "differentiable" means all partials of all orders exist and are continuous.

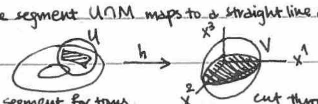


$\boxed{\text{Definition}}$ Let U, V be open sets in $\mathbb{R}^n$. $h: U \to V$ is called a ~~diffeomorphism~~ diffeomorphism if h^{-1} exists and h, h^{-1} are differentiable.

Consider a curve in $\mathbb{R}^2$, M, which is a 1D manifold. Pick an x with an open set U s.t. $x \in U$. Consider $U \cap M$, the intersection of the curve with open set U. U maps to V, and the segment $U \cap M$ maps to a straight line along x^1 axis.



There is an analogous treatment for higher definitions:



$\boxed{\text{Definition}}$ A k-dimensional manifold M in $\mathbb{R}^n$ is a set of points in $\mathbb{R}^n$ s.t. the following condition holds for all $x \in M$:

(Condition M). $\exists$ open sets $U, V \subset \mathbb{R}^n$ s.t. $x \in U$, $\exists$ diffeomorphism $h: U \to V$ s.t. $h(U \cap M) = V \cap \{y \in \mathbb{R}^n \mid y^{k+1} = y^{k+2} = \dots = y^n = 0\}$.

Remark — In V for 2D example, $x^2 = 0$; in V for 3D example, $x^3 = 0$ [no x^3 component]

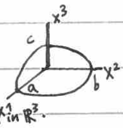
$\boxed{\text{Theorem}}$ Let $A \subseteq \mathbb{R}^n$, $g: A \to \mathbb{R}^p$, $n \geq p$. Assume that $g'(x)$ $p \times n$ has rank p for all x with $g(x) = 0$. Then $g^{-1}(0)$ is a $(n-p)$-dimensional manifold in $\mathbb{R}^n$ (g is called submersion).

Remark — Recall that $g'(x)$ is the matrix representation of $Dg(x): \mathbb{R}^n \to \mathbb{R}^p$. $\operatorname{rank} g'(x) = \dim \operatorname{Im}[Dg(x)]$. $\operatorname{rk} g'(x) = p$ means $\operatorname{Im}(Dg(x)) = \mathbb{R}^p$, so $Dg(x)$ is surjective.

Examples of manifolds —

1. $S^2 = \{(x^1, x^2, x^3) \in \mathbb{R}^3 : (x^1)^2 + (x^2)^2 + (x^3)^2 = 1\}$. $g: \mathbb{R}^3 \to \mathbb{R}$ $\left\{ \begin{array}{l} p=1 \\ n=3 \end{array} \right\}$ Then let $g(x^1, x^2, x^3) = (x^1)^2 + (x^2)^2 + (x^3)^2 - 1$ s.t. $g^{-1}(0) = S^2$. $g'(x) = (2x^1\ 2x^2\ 2x^3)$.

   $g'(x)$ either has rank 0 or 1 [c.f. 0 or 1 linearly independent rows]. $g'(x) \begin{cases} = (0,0,0) \Rightarrow \operatorname{rank} g'(x) = 0 \\ \neq (0,0,0) \Rightarrow \operatorname{rank} g'(x) = 1. \end{cases}$ $(2x^1, 2x^2, 2x^3) = 0 \Leftrightarrow x^1, x^2, x^3 = 0$. However, $(0,0,0) \notin S^2 \Rightarrow \operatorname{rk} g'(x) = 1$. By theorem, $\forall x \in g^{-1}(0)$, $\operatorname{rk} g'(x) = 1 \Rightarrow g^{-1}(0) = S^2$ is a $3-1=2$-dimensional manifold in $\mathbb{R}^3$.



2. Ellipsoid is 2-D manifold in $\mathbb{R}^3$. $\frac{(x^1)^2}{a^2} + \frac{(x^2)^2}{b^2} + \frac{(x^3)^2}{c^2} = 1 \Rightarrow$ Let $g: \mathbb{R}^3 \to \mathbb{R}$, $g(x^1, x^2, x^3) = \frac{(x^1)^2}{a^2} + \frac{(x^2)^2}{b^2} + \frac{(x^3)^2}{c^2} - 1$ s.t. ellipsoid is $g^{-1}(0)$.

   $g'(x^1, x^2, x^3) = \left(\frac{2x^1}{a^2}, \frac{2x^2}{b^2}, \frac{2x^3}{c^2} \right) = (0,0,0) \Leftrightarrow x^1 = x^2 = x^3 = 0$. However $(0,0,0) \notin$ ellipsoid $\Rightarrow \forall (x^1, x^2, x^3) \in g^{-1}(0)$, $\operatorname{rk} g'(x) = 1 \Rightarrow g^{-1}(0)$ is 2D-manifold in $\mathbb{R}^3$.

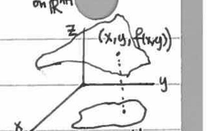
3. Sphere of dimension n in $\mathbb{R}^{n+1}$. $S^n = \{(x^1, \dots, x^{n+1}) : \sum_{i=1}^{n+1} (x^i)^2 = 1\}$. $g: \mathbb{R}^{n+1} \to \mathbb{R}$, $g(x^1, \dots, x^{n+1}) = (x^1)^2 + (x^2)^2 + \dots + (x^{n+1})^2 - 1$, $S^n = g^{-1}(0)$.

   $g'(x^1, \dots, x^{n+1}) = (2x^1, \dots, 2x^{n+1}) \neq 0$ as long as $(x^1, \dots, x^{n+1}) \neq 0$, which is not on S^n. Thus $\operatorname{rk} g'(x^1, \dots, x^{n+1}) = 1$ on $g^{-1}(0) \Rightarrow S^n$ is a n-dim manifold on $\mathbb{R}^{n+1}$ (ambient space).

4. (Monge patch): $f(x,y) = z$. Assume that f is continuously differentiable. Then graph of f is a 2D manifold in $\mathbb{R}^3$.

   Then $g(x,y,z) = f(x,y) - z$, $g: \mathbb{R}^3 \to \mathbb{R}$. Then graph is given by $g^{-1}(0)$. $g'(x,y,z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1 \right)$ which is always non-zero.

   $\therefore \operatorname{rk} g'(x,y,z) = 1 \Rightarrow$ Monge patch is $3-1=2$-dimensional manifold in $\mathbb{R}^3$.

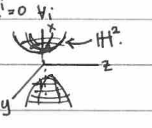


3109-08.

5. Hyperbolic n-space: $\mathbb{H}^n = \{(x^1, x^2, \dots, x^{n+1}) \in \mathbb{R}^{n+1}$ with $x^1 > 0$ and $(x^1)^2 - [(x^2)^2 + (x^3)^2 + \dots + (x^{n+1})^2] = 1\}$. is a n-dimensional manifold in $\mathbb{R}^{n+1}$. Then we have $g: \mathbb{R}^{n+1} \to \mathbb{R}$.

$g(x^1, x^2, \dots, x^{n+1}) = (x^1)^2 - [(x^2)^2 + (x^3)^2 + \dots + (x^{n+1})^2 - 1]$, $\mathbb{H}^n = g^{-1}(0) \cap \{(x^1, \dots, x^{n+1}), x^1 > 0\}$. $g'(x^1, \dots, x^{n+1}) = (2x^1, -2x^2, -2x^3, \dots, -2x^{n+1}) = 0 \Leftrightarrow x^i = 0 \ \forall i$

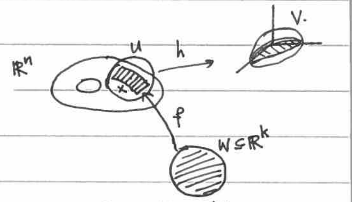
However, $(0, \dots, 0) \notin \mathbb{H}^n$. Particular example: $n=2$ $(x^1, x^2, x^3) = (x, y, z)$, $\mathbb{H}^2 = x^2 - y^2 - z^2 = 1$ (two-sheeted hyperboloid). Set $y=0$, then $x^2 - z^2 = 1$



This is a hyperbola. We repeat this by fixing x, then we get the two-sheeted hyperboloid. [Note that for $\mathbb{H}^2$, $x > 0$, so only the top half is $\mathbb{H}^2$.].

Remark – We can also have 1-sheeted hyperboloids of form $x^2 + y^2 - z^2 = 1$, which are also manifolds by same logic.

Instead of thinking of manifolds as "flattened out" regions, we can think of them as parametrisations (or charts).

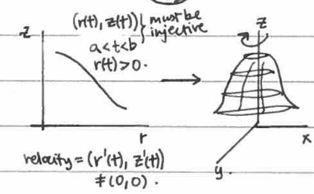


Theorem M is a k-dimensional manifold if $\forall x \in M$ the following condition holds:

(Condition C) $\exists$ open sets $W \subseteq \mathbb{R}^k$, $U \subseteq \mathbb{R}^n$, $x \in U$. Then $\exists f: W \to U$ injective s.t.

(a) $f(W) = M \cap U$ (b) $f'(y)$ has rank k $\forall y \in W$ (c) $f^{-1}: f(W) \to W$ is continuous.

Examples –



1. (Surface of revolution). Using the diagram on the right, we can parametrise the surface of revolution by the formula:

$f(t, \theta) = (r(t)\cos\theta, r(t)\sin\theta, z(t)) \in \mathbb{R}^3$. $r(t)^2 = x^2 + y^2$. [injective, no self-intersection] We set that $0 < \theta < 2\pi$ to maintain injectivity of f. [surface does not "quite" close].

NTP: $f(\underbrace{(a,b) \times (0, 2\pi)}_{W})$ is a 2D-manifold in $\mathbb{R}^3$. To do this, we check our conditions:

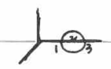
(a) $f(W) = U \cap M$, $M = \operatorname{Im} f$, $U = \mathbb{R}^3$ ✓ (b) NTP: $f'(y)$ has rank 2 $\forall y \in W$. We note $f'(t, \theta) = \begin{pmatrix} r'(t)\cos\theta & -r(t)\sin\theta \\ r'(t)\sin\theta & r(t)\cos\theta \\ z'(t) & 0 \end{pmatrix}$

To find the rank, calculate the largest square submatrix with determinant $\neq 0$. Try $\begin{vmatrix} r'(t)\cos\theta & -r(t)\sin\theta \\ r'(t)\sin\theta & r(t)\cos\theta \end{vmatrix} = r'(t)r(t)[\cos^2\theta + \sin^2\theta] = r'(t)r(t)$. We know $r(t) \neq 0$,

so $r'(t)r(t) \neq 0$ if $r'(t) \neq 0$. Is it? Yes, in most instances: Case 1: $r'(t) \neq 0$, then we have submatrix of rank 2 satisfying conditions. Otherwise, we have "bad" alternative –

Case 2: $r'(t) = 0$. $f'(t, \theta) = \begin{pmatrix} 0 & -r(t)\sin\theta \\ 0 & r(t)\cos\theta \\ z'(t) & 0 \end{pmatrix}$. Then $\begin{vmatrix} 0 & -r(t)\sin\theta \\ z'(t) & 0 \end{vmatrix} = z'(t)r(t)\sin\theta$ and $\begin{vmatrix} 0 & r(t)\cos\theta \\ z'(t) & 0 \end{vmatrix} = z'(t)r(t)\cos\theta$. $r(t) > 0$ and $z'(t) \neq 0$ since $r'(t) = 0$.

Since $\sin\theta$ and $\cos\theta$ are never simultaneously 0, one of the determinants is $\neq 0$. i.e. $\operatorname{rk} f'(t, \theta) = 2$. q.e.d.



2. (Torus as a surface of revolution). We can parametrise this as $(r-2)^2 + z^2 = 1$. Alternatively, this is also: $f(\theta, \varphi) = ((2+\cos\varphi)\cos\theta, (2+\cos\varphi)\sin\theta, \sin\varphi)$, with $0 < \theta < 2\pi$,

$-\pi < \varphi < \pi$. $(r-2)^2 + z^2 = \cos^2\varphi + \sin^2\varphi = 1$. $f'(\theta, \varphi) = \begin{pmatrix} -(2+\cos\varphi)\sin\theta & -\sin\varphi\cos\theta \\ (2+\cos\varphi)\cos\theta & -\sin\varphi\sin\theta \\ 0 & \cos\varphi \end{pmatrix}$. We examine the respective determinants of submatrices:

$\begin{vmatrix} -(2+\cos\varphi)\sin\theta & -\sin\varphi\cos\theta \\ (2+\cos\varphi)\cos\theta & -\sin\varphi\sin\theta \end{vmatrix} = -(2+\cos\varphi)(\sin\varphi)\begin{vmatrix} \sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{vmatrix} = \underbrace{(2+\cos\varphi)}_{\neq 0}(\sin\varphi) = 0$ when $\sin\varphi = 0$. So if $\sin\varphi > 0$, $f'(\theta, \varphi) = \begin{pmatrix} \cancel{-3\sin\theta} & 0 \\ \cancel{2\cos\theta} & 0 \\ 0 & 1 \end{pmatrix}$

Then we get the two determinants $-(2+\cos\varphi)\sin\theta\cos\varphi$, $(2+\cos\varphi)\cos\theta\cos\varphi$. $\cos\varphi \neq 0$. So determinants are never simultaneously 0, so $\operatorname{rk} f'(\theta, \varphi) = 0$.

15 November 2013.
Dr Yiannis PETRIDIS
Maths 706.

Returning to the surface of revolution, we try to reclaim (t, θ) from $f(t, \theta) = (x, y, z)$. Then $(t, \theta) = f^{-1}(x, y, z)$. Set $x = r(t)\cos\theta$, $y = r(t)\sin\theta$, $z = z(t)$.

Then $\frac{y}{x} = \tan\theta \Rightarrow \theta = \arctan\frac{y}{x}$ or $\theta = \operatorname{arccot}\frac{x}{y}$. Moreover, $x^2 + y^2 = r(t)^2$, $r(t) = \sqrt{x^2 + y^2} > 0$. $z(t) = z$. Since $(r(t), z(t))$ is bijective onto its image,

Given $(x, y, z) \in M$, $(r(t), z(t))$ can be inverted to give t: $t = z^{-1}$ or r^{-1}.

For Ex 2 (torus), we try to invert our parametrisation f. $W = (-\pi, \pi) \times (-\pi, \pi)$. Why is $f^{-1}: f(W) \to W$ continuous? $x = (2+\cos\varphi)\cos\theta$, $y = (2+\cos\varphi)\sin\theta$, $z = \sin\varphi$.

$\theta = \arctan\frac{y}{x}$ or $\operatorname{arccot}\frac{x}{y}$. $\varphi = \arcsin z$. [φ may not be uniquely defined as sin is not injective on the interval.

We now move towards some theory: binding some concepts together.

Theorem If $g: \mathbb{R}^n \to \mathbb{R}^p$ is continuously differentiable, $M = g^{-1}(0)$ $\forall x \in M$ rank $g'(x) = p$. Then M is an $(n-p)$-dimensional manifold in $\mathbb{R}^n$

Proof – Refer to Handouts 3 and 4.

20 November 2013
Dr. Yiannis PETRIDIS.
Maths 706.

Let V be a finite dimensional vector space over $\mathbb{R}$, $f: V \to \mathbb{R}$ a linear functional. Then $f(\lambda x + y) = \lambda f(x) + f(y)$ $\forall x, y \in V$, $\lambda \in \mathbb{R}$.

Definition The dual space of V is $V^* = \{f: V \to \mathbb{R}$ linear functionals$\}$.

If $f, g: V \to \mathbb{R}$ are linear functionals, define $f+g: V \to \mathbb{R}$, $(f+g)(x) = f(x) + g(x)$ $\forall x \in V$. If $\lambda \in \mathbb{R}$, $(\lambda f)(x) = \lambda \cdot f(x)$.

Theorem V^* is a vector space with the operations defined above.

Proof – (partial) $(f+g)(\lambda x + y) \overset{\text{def}}{=} f(\lambda x + y) + g(\lambda x + y) \overset{\text{linearity}}{=} \lambda f(x) + f(y) + \lambda g(x) + g(y) = \lambda(f(x) + g(x)) + (f(y) + g(y)) = \lambda(f+g)(x) + (f+g)(y)$. etc. q.e.d.

Proposition $\dim V^* = \dim V$.

Proof – To construct a basis for V^*, start with $\{v_1, \dots, v_n\}$ as a basis for V. $x \in V$, $x = x^1 v_1 + \dots + x^n v_n$. Define linear functionals $\varphi^i(x) = x^i$ $i = 1, 2, \dots, n$. Then $\{\varphi^i\}$ is a basis for V.

(1) φ^i is a linear functional: $\varphi^i(\lambda x + y) = \varphi^i((\lambda x^1 + y^1)v_1, (\lambda x^2 + y^2)v_2, \dots, (\lambda x^n + y^n)v_n) = \lambda x^i + y^i = \lambda\varphi^i(x) + \varphi^i(y)$, so φ^i is a linear functional. [$\lambda \in \mathbb{R}$]

(2) $\{\varphi^i\}$ forms a LI set: $\sum_{i=1}^n a_i x^i = 0$ linear functional. Plug $x = v_1$, $\sum_{i=1}^n a_i \varphi^i(v_1) = 0 \Rightarrow a_1 \cdot 1 + 0 = 0 \Rightarrow a_1 = 0$. More generally, $\varphi^i(v_j) = \delta_{ij}$, so each $a_i = 0 \Rightarrow$ set is LI.

(3) $\{\varphi^i\}$ spans V^*: Given $f \in V^*$, we claim $\exists\{b_i\}$ s.t. $f = \sum_{i=1}^n b_i \varphi^i$. Then $f(v_j) = \sum_{i=1}^n b_i \varphi^i(v_j) = \sum_{i=1}^n b_i \delta_{ij} = b_j$. Then apply for $j = 1, \dots, n$. Then obviously,

$f(v) = (b_1\varphi^1 + \dots + b_n\varphi^n)(v)$ holds for $\forall$ a basis element. $\Rightarrow$ they agree for any $v \in V$, so $f = b_1\varphi^1 + \dots + b_n\varphi^n \Rightarrow \{\varphi^i\}$ spans V^*. Thus, $\{\varphi^i\}$ forms basis for V^*. q.e.d.

Let V be a finite dimensional vector space over $\mathbb{R}$. Define the k-fold product $V^k = \overbrace{V \times \cdots \times V}^{k \text{ factors}}$ by $(w_1, w_2, \dots, w_k)$ where $w_1, w_2, \dots, w_k \in V$.

This has operations $(w_1, \dots, w_k) + (z_1, \dots, z_k) = (w_1 + z_1, \dots, w_k + z_k)$, $\lambda \in \mathbb{R}$ then $\lambda(w_1, \dots, w_k) = (\lambda w_1, \dots, \lambda w_k)$.

$\boxed{\text{Definition}}$ A function $T: V^k \to \mathbb{R}$ is called multilinear if $T(v_1, v_2, \dots, v_{i-1}, v_i + v_i', v_{i+1}, \dots, v_k) = T(v_1, v_2, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_k) + T(v_1, v_2, \dots, v_{i-1}, v_i', v_{i+1}, \dots, v_k)$ $\forall i = 1, \dots, k$ and $\lambda \in \mathbb{R}$.

$T(v_1, v_2, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_k) = \lambda T(v_1, v_2, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_k)$.

The set of k-multilinear maps is $\mathcal{J}^k(V) = \{T: V^k \to \mathbb{R}\}$. Each $T: V^k \to \mathbb{R}$ is called a k-tensor.

Let $T, S: V^k \to \mathbb{R}$ be multilinear, then $T+S: V^k \to \mathbb{R}$ is defined as $(T+S)(v_1, \dots, v_k) = T(v_1, \dots, v_k) + S(v_1, \dots, v_k)$, if $\lambda \in \mathbb{R}$, $(\lambda \cdot T)(v_1, \dots, v_k) = \lambda \cdot T(v_1, \dots, v_k)$.

$\boxed{\text{Definition}}$ Let $S \in \mathcal{J}^k(V)$, $T \in \mathcal{J}^l(V)$. Define $S \otimes T: V^{k+l} \to \mathbb{R}$ by $(S \otimes T)(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}) = S(v_1, \dots, v_k) \cdot T(v_{k+1}, \dots, v_{k+l})$.

Note – We see (in Homework 7) that $S \otimes T \in \mathcal{J}^{k+l}(V)$. Also, observe that $S \otimes T \neq T \otimes S$. $(T \otimes S)(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}) = T(v_1, \dots, v_l) \cdot S(v_{l+1}, \dots, v_{k+l})$.

$\boxed{\text{Proposition}}$ (Properties of Tensor Products).

(1) $(S_1 + S_2) \otimes T = (S_1 \otimes T) + (S_2 \otimes T)$ (2) $S \otimes (T_1 + T_2) = S \otimes T_1 + S \otimes T_2$. where $S_1, S_2, S \in \mathcal{J}^k(V)$, $T, T_1, T_2 \in \mathcal{J}^l(V)$

(3) $a \in \mathbb{R}$, $(aS) \otimes T = S \otimes (aT) = a(S \otimes T)$. (4) $(S \otimes T) \otimes U = S \otimes (T \otimes U)$, $S \in \mathcal{J}^k(V)$, $T \in \mathcal{J}^l(V)$, $U \in \mathcal{J}^m(V)$

(5) $\mathcal{J}^1(V) = V^*$.

Proof – (Part) (4): Let $(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}, v_{k+l+1}, \dots, v_{k+l+m}) \in V$. Then $[(S \otimes T) \otimes U](v) = (S \otimes T)(v_1, \dots, v_{k+l}) \cdot U(v_{k+l+1}, \dots, v_{k+l+m}) =$

$S(v_1, \dots, v_k) \cdot T(v_{k+1}, \dots, v_{k+l}) \cdot U(v_{k+l+1}, \dots, v_{k+l+m}) \overset{\text{associativity of real numbers}}{=} S(v_1, \dots, v_k) \cdot (T \otimes U)(v_{k+1}, \dots, v_{k+l}, v_{k+l+1}, \dots, v_{k+l+m}) = (S \otimes (T \otimes U))(v_1, \dots, v_{k+l+m})$ // q.e.d.

$\boxed{\text{Theorem}}$ $\dim \mathcal{J}^k(V) = n^k$.

Proof – Let $\{i_1, \dots, i_k\} \subset \{1, 2, \dots, n\}$, $\{v_1, v_2, \dots, v_n\}$ is a basis for V, $\{\varphi^1, \dots, \varphi^n\}$ is a basis for V^*. Then we claim that $\{\underset{\mathcal{J}^1(V)}{\varphi^{i_1}} \otimes \varphi^{i_2} \otimes \dots \otimes \underset{\mathcal{J}^1(V)}{\varphi^{i_k}}, 1 \le i_1, \dots, i_k \le n\}$ is a basis for $\mathcal{J}^k(V)$.

To prove LI, let $\sum_{i_1, \dots, i_k = 1}^{n} a_{i_1 i_2 \dots i_k} \varphi^{i_1} \otimes \varphi^{i_2} \otimes \dots \otimes \varphi^{i_k} = 0$. Take $(v_{j_1}, v_{j_2}, \dots, v_{j_k})$ and apply functional to it: $\sum_{i_1, \dots, i_k}^{n} a_{i_1 i_2 \dots i_k} \varphi^{i_1}(v_{j_1}) \cdot \varphi^{i_2}(v_{j_2}) \dots \varphi^{i_k}(v_{j_k}) = 0$

LHS reduces to $\sum a_{i_1} a_{i_2} \dots a_{i_k} \delta_{i_1 j_1} \delta_{i_2 j_2} \dots \delta_{i_k j_k} = a_{j_1} a_{j_2} \dots a_{j_k} = 0 \Rightarrow$ each coefficient is 0 $\Rightarrow$ set is LI.

To prove spanning, NTP: $\varphi^{i_1} \otimes \dots \otimes \varphi^{i_k}$ span $\mathcal{J}^k(V)$ as $1 \le i_1, i_2, \dots, i_k \le n$. Take any $T \in \mathcal{J}^k(V)$, write $T = \sum_{i_r = 1}^{n} c_{i_1 \dots i_k} \varphi^{i_1} \otimes \dots \otimes \varphi^{i_k}$. Plug in the k-tuple $(v_{j_1}, v_{j_2}, \dots, v_{j_k})$, then $T(v_{j_1}, v_{j_2}, \dots, v_{j_k}) = \sum_{i_r = 1}^{n} c_{i_1 \dots i_k} \delta_{i_1 j_1} \delta_{i_2 j_2} \dots \delta_{i_k j_k} = c_{j_1 j_2 \dots j_k}$. We still need to prove that such an expression is legitimate.

Plug any vectors $w_1, \dots, w_k$ on both sides. Let $w_1 = a^{11} v_1 + a^{12} v_2 + \dots + a^{1n} v_n$, $w_2 = a^{21} v_1 + \dots + a^{2n} v_n$, $\dots$, $w_k = a^{k1} v_1 + \dots + a^{kn} v_n$.

$\text{LHS} = T(w_1, \dots, w_k) = T\left(\sum_{j_1=1}^{n} a^{1 j_1} v_{j_1}, \sum_{j_2=1}^{n} a^{2 j_2} v_{j_2}, \dots, \sum_{j_k=1}^{n} a^{k j_k}\right) = \sum_{j_1, \dots, j_k = 1}^{n} a^{1 j_1} a^{2 j_2} \dots a^{k j_k} T(v_{j_1}, \dots, v_{j_k}) = \sum_{j_1, \dots, j_k}^{n} a^{1 j_1} \dots a^{k j_k} \cdot c_{j_1 \dots j_k}$.

$\text{RHS} = \sum_{i_1, \dots, i_k = 1}^{n} c_{i_1 \dots i_k} \varphi^{i_1} \otimes \dots \otimes \varphi^{i_k}(w_1, \dots, w_k) = \sum_{i_1, \dots, i_k = 1}^{n} c_{i_1 \dots i_k} a^{1 i_1} a^{2 i_2} \dots a^{k i_k}$. Call $i = j$, and we are done // q.e.d.

$\boxed{\text{Definition}}$ A $T \in \mathcal{J}^k(V)$ is called a symmetric k-tensor if $T(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_k) = T(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_k)$ $\forall i, j$.

Note – For $k=2$, these correspond to $T(v_1, v_2) = T(v_2, v_1) \Rightarrow$ inner products if positive definite i.e. $T(v, v) \ge 0$ $\forall v$.

A $T \in \mathcal{J}^k(V)$ is called alternating if $T(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_k) = -T(v_1, v_2, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_k)$.

Note – An example of these are k-determinants.

Recall that f is an even function if $f(x) = f(-x)$, an odd function if $f(x) = -f(-x)$. For instance then, if $f(x) = x^2 + x^3$, it is the sum of an even and odd function. More generally, what about any f, e.g. $f(x) = \cos(\sin(x^5 + \cos x))$? Let us have $f(x) = \underbrace{\frac{f(x) + f(-x)}{2}}_{g(x)} - \underbrace{\frac{f(x) - f(-x)}{2}}_{h(x)}$. Then $g(x) = \frac{f(x) + f(-x)}{2}$ is even, $h(x) = \frac{f(x) - f(-x)}{2}$ is odd.

$\sigma: x \mapsto -x$, then $\sigma^2 = 1$. $g \circ \sigma(x) = g(-x) = g(x)$, $h(-x) = -h(x)$. Thus, $g(x) = \frac{f(x) + f(\sigma x)}{2}$, $h(x) = \frac{f(x) - f(\sigma x)}{2}$. Suppose we start with any k-tensor $T \in \mathcal{J}^k(V)$, which is neither symmetric nor alternating. We can obtain such tensors (symmetric or alternating) out of them. Consider permutation $\sigma \in S_k$, $\sigma: \begin{pmatrix} 1 & 2 & \cdots & k \\ \sigma(1) & \sigma(2) & \cdots & \sigma(k) \end{pmatrix}$.

Then the tensor $\frac{1}{k!} \sum_{\sigma \in S_k} T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = (\text{Sym}\, T)(v_1, \dots, v_k)$ and if T was symmetric it agrees with this definition. Consider $\text{sgn}: S_k \to \{\pm 1\}$, then we have that given $T \in \mathcal{J}^k(V)$, we define $\text{Alt}\, T \in \mathcal{J}^k(V)$ by $(\text{Alt}\, T)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma)\, T(v_{\sigma(1)}, \dots, v_{\sigma(k)})$.

$\boxed{\text{Theorem}}$ (a) If $T \in \mathcal{J}^k(V)$, then $\text{Alt}\, T$ is alternating. (or $\text{Alt}(T) \in \Lambda^k(V)$).

(b) Let $\Lambda^k(V)$ be the set of alternating k-tensors. Then $\Lambda^k(V)$ is a subspace of $\mathcal{J}^k(V)$. [Proof – left as exercise in Homework 7] Nomenclature – [$w \in \Lambda^k(V) \Rightarrow w$ is of degree k].

(c) If $w \in \Lambda^k(V)$, $\text{Alt}(w) = w$. (d) $\text{Alt}(\text{Alt}\, T) = \text{Alt}\, T$.

22 November 2013
Dr. Yiannis PETRIDIS
Maths 706.

Note – In the bilinear case, $T \in \mathcal{J}^2(V) \Rightarrow w_1, w_2 \in V$. $S(w_1, w_2) = \frac{T(w_1, w_2) + T(w_2, w_1)}{2}$ is symmetric, $\text{Alt}(T)(w_1, w_2) = \frac{T(w_1, w_2) - T(w_2, w_1)}{2}$ is alternating.

Proof – (d) follows from (c): Let $w = \text{Alt}(T) \in \Lambda^k(V)$, so $\text{Alt}(w) = w \Rightarrow \text{Alt}(\text{Alt}\, T) = \text{Alt}\, T$ // q.e.d. Part (a) – NTP: fix i, j, $\text{Alt}(T)(w_1, \dots, \overset{i}{w_j}, \dots, \overset{j}{w_i}, \dots, w_k) = -\text{Alt}(w_1, \dots, w_k)$.

$\text{LHS} = \text{Alt}(T)(w_1, \dots, \overset{i}{w_j}, \dots, \overset{j}{w_i}, \dots, w_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma)\, T(w_{\sigma(1)}, w_{\sigma(2)}, \dots, \overset{i}{w_{\sigma(j)}}, \dots, \overset{j}{w_{\sigma(i)}}, \dots, w_{\sigma(k)})$. Define $\sigma' = \sigma \cdot (i\, j)$, $\exists$ 1-1 correspondence $S_k \to S_k$ by $\sigma \to \sigma(i\, j)$ (bijective). Then we get that

$\text{LHS} = \frac{1}{k!} \sum_{\sigma \in S_k} -\text{sgn}(\sigma')\, T(w_{\sigma'(1)}, w_{\sigma'(2)}, \dots, w_{\sigma'(i)}, \dots, w_{\sigma'(j)}, \dots, w_{\sigma'(k)})$ $\because$ $\sigma \in S_k \Rightarrow \sigma' \in S_k$, $\text{sgn}(\sigma') = -\text{sgn}(\sigma) \Rightarrow -\text{Alt}(T)(w_1, \dots, w_i, \dots, w_j, \dots, w_k)$ // q.e.d.

Part (c): If $w \in \Lambda^k(V)$, then $w(w_1, \dots, w_k) = -w(w_1, \dots, \overset{i}{w_j}, \dots, \overset{j}{w_i}, \dots, w_k) = w(w_2, w_1, \dots, \overset{i}{w_j}, \dots, \overset{j}{w_i}, \dots, w_k)$. Thus, if $\sigma \in S_k$, we get that

$w(w_{\sigma(1)}, w_{\sigma(2)}, \dots, w_{\sigma(k)}) = \text{sgn}(\sigma) \cdot w(w_1, \dots, w_k)$. $w \in \Lambda^k(V)$, $\text{Alt}(w)(w_1, \dots, w_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma)\, w(w_{\sigma(1)}, \dots, w_{\sigma(k)}) = \frac{1}{k!} \sum_{\sigma \in S_k} [\text{sgn}(\sigma)]^2 w(w_1, \dots, w_k) = w(w_1, \dots, w_k)$ // q.e.d.

3109-10.

Let $w \in \Lambda^k(V)$ and $\eta \in \Lambda^\ell(V)$. $w \otimes \eta \in \mathcal{J}^{k+\ell}(V)$ is a tensor, but does not have to be alternating. We define the wedge of w and η to be $\boxed{w \wedge \eta = \frac{(k+\ell)!}{k!\,\ell!} \mathrm{Alt}(w \otimes \eta)} \in \Lambda^{k+\ell}(V)$.

Proposition If $w_1, w_2, w \in \Lambda^k(V)$, $\eta_1, \eta_2, \eta \in \Lambda^\ell(V)$. Then the following properties hold:

(a) $(w_1 + w_2) \wedge \eta = w_1 \wedge \eta + w_2 \wedge \eta$ (b) $w \wedge (\eta_1 + \eta_2) = w \wedge \eta_1 + w \wedge \eta_2$

(c) If $a \in \mathbb{R}$, $(aw) \wedge \eta = a(w \wedge \eta) = w \wedge (a\eta)$ (d) $w \wedge \eta = (-1)^{k\ell} \eta \wedge w$

Proof – Part (a), (d) in homework, (b) similar to (a), (c) omitted.

Let V, W be finite dimensional vector spaces, $V \xrightarrow{f} W$ where f is a linear transformation. If $T: W \to \mathbb{R}$ is a linear functional on V, then $T \circ f$ is a linear functional on V. Thus, if $T \in W^*$, then $T \circ f \in V^*$. For notation, we write $T \circ f = f^*(T)$, which gives a mapping $f^*: W^* \to V^*$, the pullback of T. We note also that $W^* = \mathcal{J}^1(W)$, $V^* = \mathcal{J}^1(V)$. We can generalise this to tensors: If $T \in \mathcal{J}^k(W)$ then $f^*(T) \in \mathcal{J}^k(V)$. Then $f^*(T)(v_1, v_2, \dots, v_k) = T(f(v_1), f(v_2), \dots, f(v_k))$

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "T\circ f = f^* \text{ (pullback)}"'] & W \arrow[d, "T"] \\ & \mathbb{R} \end{tikzcd}$$

Theorem (a) $f^*(T) \in \mathcal{J}^k(V)$ (b) $f^*(T \otimes S) = f^*(T) \otimes f^*(S)$ (c) $f^*(w \wedge \eta) = f^*(w) \wedge f^*(\eta)$

Proof – (b) left as exercise, (c) in Homework 7. (a): Need to show $f^*(T)$ is k-multilinear on V. Take $v_1, \dots, v_k, v_i' + \lambda \in \mathbb{R}$. Then we have that:

$$f^*(T)(v_1, \dots, \lambda v_i + v_i', \dots, v_k) = T(f(v_1), f(v_2), \dots, f(\lambda v_i + v_i'), \dots, f(v_k)) = T(f(v_1), \dots, \lambda f(v_i) + f(v_i'), \dots, f(v_k)) \overset{T \in \mathcal{J}^k(V)}{=} \lambda(T(f(v_1), \dots, f(v_k))) + T(f(v_1), \dots, f(v_i'), \dots, f(v_k))$$

$$= \lambda f^*(T)(v_1, \dots, v_k) + f^*(T)(v_1, \dots, v_i', \dots, v_k) \Rightarrow f^*(T) \in \mathcal{J}^k(V) \text{ q.e.d.}$$

Theorem (a) If $S \in \mathcal{J}^k(V)$, $T \in \mathcal{J}^\ell(V)$ and $\mathrm{Alt}(S) = 0$, then $\mathrm{Alt}(S \otimes T) = 0$ and $\mathrm{Alt}(T \otimes S) = 0$.

(b) $\mathrm{Alt}(\mathrm{Alt}(w \otimes \eta) \otimes \theta) = \mathrm{Alt}(w \otimes \eta \otimes \theta) = \mathrm{Alt}(w \otimes \mathrm{Alt}(\eta \otimes \theta))$ where $w \in \mathcal{J}^k(V)$, $\eta \in \mathcal{J}^\ell(V)$ and $\theta \in \mathcal{J}^m(V)$.

(c) If $w \in \Lambda^k(V)$, $\eta \in \Lambda^\ell(V)$ and $\theta \in \Lambda^m(V)$, then $(w \wedge \eta) \wedge \theta = \frac{(k+\ell+m)!}{k!\,\ell!\,m!} \mathrm{Alt}(w \otimes \eta \otimes \theta) = w \wedge (\eta \wedge \theta)$

Proof – (a): Given that $\mathrm{Alt}(S) = 0$. Thus for $w_1, \dots, w_k \in V$, $\mathrm{Alt}(S)(w_1, \dots, w_k) = \frac{1}{k!} \sum_{\tau \in S_k} \overset{\mathrm{sgn}(\tau)}{} S(w_{\tau(1)}, w_{\tau(2)}, \dots, w_{\tau(k)}) = 0$. Then $\mathrm{Alt}(S \otimes T) = \frac{1}{(k+\ell)!} \sum_{\sigma \in S_{k+\ell}} \mathrm{sgn}(\sigma)(S \otimes T)(w_{\sigma(1)}, \dots, w_{\sigma(k+\ell)})$

We split these into parts: let $G = \{\sigma \in S_k : \sigma(k+1) = k+1, \sigma(k+2) = k+2, \dots, \sigma(k+\ell) = k+\ell\}$, $G \le S_{k+\ell}$. Then these terms contribute: $\sum_{\sigma \in G} \mathrm{sgn}(\sigma) S(w_{\sigma(1)}, \dots, w_{\sigma(k)}) T(w_{k+1}, \dots, w_{k+\ell})$

$= \overbrace{T(w_{k+1}, \dots, w_{k+\ell})}^{\text{constant}} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) S(w_{\sigma(1)}, \dots, w_{\sigma(k)})$ since $G \cong S_k$, so $= T(w_{k+1}, \dots, w_{k+\ell}) \mathrm{Alt}(S)(w) = 0$. Consider coset decomposition of $S_{k+\ell}$ by subgroup G.

Let $\sigma_0 \in S_{k+\ell}$ but $\sigma_0 \notin G$ s.t. $G\sigma_0 \neq G$. Then $\sigma_0(w_1, \dots, w_k, w_{k+1}, \dots, w_{k+\ell}) = (\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_k, \tilde{w}_{k+1}, \dots, \tilde{w}_{k+\ell})$. Then elements of $G\sigma_0$ have the form $\sigma \cdot \sigma_0$, with $\sigma \in G$, i.e. $\sigma\sigma_0(w_1, \dots, w_k, w_{k+1}, \dots, w_{k+\ell}) = (\tilde{w}_{\sigma(1)}, \tilde{w}_{\sigma(2)}, \dots, \tilde{w}_{\sigma(k)}, \underbrace{\tilde{w}_{k+1}, \dots, \tilde{w}_{k+\ell}}_{\text{last } \ell \text{ entries fixed}})$. Then we sum over $\sigma\sigma_0 \in G\sigma_0$, i.e. $\sigma \in G$,

$\sum_{\sigma \in G} \mathrm{sgn}(\sigma\sigma_0) S(\tilde{w}_{\sigma(1)}, \tilde{w}_{\sigma(2)}, \dots, \tilde{w}_{\sigma(k)}) \cdot \overbrace{T(\tilde{w}_{k+1}, \tilde{w}_{k+2}, \dots, \tilde{w}_{k+\ell})}^{\text{constant}} = \mathrm{sgn}(\sigma_0) T(\tilde{w}_{k+1}, \tilde{w}_{k+2}, \dots, \tilde{w}_{k+\ell}) \sum_{\sigma \in G} \mathrm{sgn}(\sigma) S(\tilde{w}_{\sigma(1)}, \dots, \tilde{w}_{\sigma(k)}) = 0$. Thus, the complete sum of both parts gives $\mathrm{Alt}(S \otimes T) = 0 + 0 = 0$ q.e.d.

(b): Take $S = \mathrm{Alt}(w \otimes \eta) - w \otimes \eta$, $T = \theta$. $\mathrm{Alt}(S) = \mathrm{Alt}(\mathrm{Alt}(w \otimes \eta) - w \otimes \eta) = \mathrm{Alt}(w \otimes \eta) - \mathrm{Alt}(w \otimes \eta) = 0$. Apply part (a), then $\mathrm{Alt}[(\mathrm{Alt}(w \otimes \eta) - w \otimes \eta) \otimes \theta] = 0$.

$\mathrm{Alt}[\mathrm{Alt}(w \otimes \eta) \otimes \theta - (w \otimes \eta) \otimes \theta] = \mathrm{Alt}[\mathrm{Alt}(w \otimes \eta) \otimes \theta] - \mathrm{Alt}((w \otimes \eta) \otimes \theta) = 0 \Rightarrow \mathrm{Alt}[\mathrm{Alt}(w \otimes \eta) \otimes \theta] = \mathrm{Alt}(w \otimes \eta \otimes \theta)$ by associativity q.e.d.

(c): $\underbrace{(w \wedge \eta)}_{\in \Lambda^{k+\ell}(V)} \wedge \theta = \frac{((k+\ell)+m)!}{(k+\ell)!\,m!} \mathrm{Alt}((w \wedge \eta) \otimes \theta) = \frac{(k+\ell+m)!}{(k+\ell)!\,m!} \mathrm{Alt}\left[\frac{k+\ell!}{k!\,\ell!} \mathrm{Alt}(w \otimes \eta) \otimes \theta\right] = \frac{(k+\ell+m)!}{k!\,\ell!\,m!} \mathrm{Alt}[\mathrm{Alt}(w \otimes \eta) \otimes \theta] = \frac{(k+\ell+m)!}{k!\,\ell!\,m!} \mathrm{Alt}(w \otimes \eta \otimes \theta)$ q.e.d.

Theorem Let V have a basis $\{v_1, \dots, v_n\}$ ($\dim V = n$), and let V^* have the dual basis $\{\varphi^1, \dots, \varphi^n\}$ where $\varphi^i(v_j) = \delta_{ij}$. Then a basis for $\Lambda^k(V)$ is given by $\varphi^{i_1} \wedge \varphi^{i_2} \wedge \dots \wedge \varphi^{i_k}$ where $1 \le i_1 \le i_2 \le \dots \le i_k \le n$, $\dim \Lambda^k(V) = \binom{n}{k}$.

Remark – If $k > n$, $\dim(\Lambda^k(V)) = 0$.

Proof – Suppose $w \in \Lambda^k(V)$. In particular, $\Lambda^k(V) \le \mathcal{J}^k(V)$, so $w = \sum_{i_1, i_2, \dots, i_k = 1}^{n} a_{i_1 i_2 \dots i_k} \varphi^{i_1} \otimes \varphi^{i_2} \otimes \dots \otimes \varphi^{i_k}$. Apply Alt operator on both sides. LHS $= \mathrm{Alt}(w) = w$. Then $w = \sum_{i_1, \dots, i_k = 1}^{n} a_{i_1 \dots i_k} \mathrm{Alt}(\varphi^{i_1} \otimes \dots \otimes \varphi^{i_k}) = \sum_{i_1, \dots, i_k = 1}^{n} a_{i_1 \dots i_k} c_{i_1 \dots i_k} \varphi^{i_1} \wedge \varphi^{i_2} \wedge \varphi^{i_3} \wedge \dots \wedge \varphi^{i_k}$. Notice that this allows repeats, i.e. $i_1 = i_2$ for instance.

Consider $(\varphi^i \wedge \varphi^i)(w_1, w_2) = \frac{2!}{1!\,1!} \mathrm{Alt}(\varphi^i \otimes \varphi^i)(w_1, w_2) = \frac{2!}{1!\,1!} \cdot \frac{1}{2!}((\varphi^i \otimes \varphi^i)(w_1, w_2) - (\varphi^i \otimes \varphi^i)\overbrace{(w_2, w_1)}^{\text{swapped}}) = \varphi^i(w_1)\varphi^i(w_2) - \varphi^i(w_2)\varphi^i(w_1) = 0$, so $\varphi^i \wedge \varphi^i = 0$.

Thus, $w = \sum_{i_1, \dots, i_k \text{ distinct}}^{n} a_{i_1 \dots i_k} c_{i_1 \dots i_k} (\varphi^{i_1} \wedge \varphi^{i_2} \wedge \dots \wedge \varphi^{i_k})$. For linear independence, note that $(\varphi^{i_1} \wedge \varphi^{i_2} \wedge \dots \wedge \varphi^{i_k})(v_{j_1}, v_{j_2}, \dots, v_{j_k}) = \delta_{i_1 j_1} \delta_{i_2 j_2} \dots \delta_{i_k j_k}$ $\because$ $\mathrm{Alt}(\varphi^{i_1} \otimes \varphi^{i_2} \otimes \dots \otimes \varphi^{i_k})(v_{j_1}, \dots, v_{j_k}) = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \varphi^{i_1}(v_{\sigma(j_1)}) \varphi^{i_2}(v_{\sigma(j_2)}) \dots \varphi^{i_k}(v_{\sigma(j_k)})$, which contributes a term only if $i_r = j_r$ $\forall r$. q.e.d.

Example – Let $n = 3 = \dim V$, take $V = \mathbb{R}^3$, $\{v_1 = e_1, v_2 = e_2, v_3 = e_3\}$ be a basis for V, $\{\varphi^1, \varphi^2, \varphi^3\}$ is the dual basis. Then if $k = 1$, $\dim \Lambda^1(\mathbb{R}^3) = \binom{3}{1} = 3$, $\dim \mathcal{J}^1(\mathbb{R}^3) = \dim(\mathbb{R}^3)^*$

Then $\varphi^1, \varphi^2, \varphi^3$ is the basis for $\Lambda^1(\mathbb{R}^3)$. Then try $k = 2$ (alternating 2-tensors). $\dim \Lambda^2(\mathbb{R}^3) = \binom{3}{2} = 3$. Recall $\varphi^i \wedge \varphi^i = 0 \Rightarrow \boxed{\varphi^1 \wedge \varphi^1 = \varphi^2 \wedge \varphi^2 = \varphi^3 \wedge \varphi^3 = 0.}$

$\Rightarrow$ basis for $\Lambda^2(\mathbb{R}^3)$ is $\{\varphi^1 \wedge \varphi^2, \varphi^1 \wedge \varphi^3, \varphi^1 \wedge \varphi^2\}$ (exponents in increasing order). Consider $(\varphi^1 \wedge \varphi^2)(w_1, w_2) \overset{\text{def}}{=} \frac{2!}{1!\,1!} \mathrm{Alt}(\varphi^1 \otimes \varphi^2)(w_1, w_2)$

$= \frac{2!}{1!\,1!} \frac{1!}{2!} \sum_{\sigma \in S_2} (\varphi^1 \otimes \varphi^2)(w_{\sigma(1)}, w_{\sigma(2)}) \mathrm{sgn}(\sigma) = (\varphi^1 \otimes \varphi^2)(w_1, w_2) - (\varphi^1 \otimes \varphi^2)(w_2, w_1) = \varphi^1(w_1)\varphi^2(w_2) - \varphi^1(w_2)\varphi^2(w_1) \Rightarrow \varphi^1 \wedge \varphi^2 = \varphi^1 \otimes \varphi^2 - \varphi^2 \otimes \varphi^1$.

Similarly, $\varphi^1 \wedge \varphi^3 = \varphi^1 \otimes \varphi^3 - \varphi^3 \otimes \varphi^1$ and $\varphi^2 \wedge \varphi^3 = \varphi^2 \otimes \varphi^3 - \varphi^3 \otimes \varphi^2$. We clearly see that $\boxed{\varphi^2 \wedge \varphi^1 = -\varphi^1 \wedge \varphi^2, \ \varphi^3 \wedge \varphi^1 = -\varphi^1 \wedge \varphi^3, \ \varphi^3 \wedge \varphi^1 = -\varphi^1 \wedge \varphi^3}$

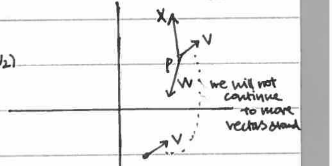
Then if $k = 3$, $\dim \Lambda^3(\mathbb{R}^3) = \binom{3}{3} = 1 \Rightarrow 1 \le 2 \le 3 \Rightarrow$ basis element is $\varphi^1 \wedge \varphi^2 \wedge \varphi^3$. We evaluate its significance: $(\varphi^1 \wedge \varphi^2 \wedge \varphi^3)(w_1, w_2, w_3) = \frac{3!}{1!\,1!\,1!} \mathrm{Alt}(\varphi^1 \otimes \varphi^2 \otimes \varphi^3)(w_1, w_2, w_3)$

$= \frac{3!}{1!\,1!\,1!} \frac{1!}{3!} \sum_{\sigma \in S_3} (\varphi^1 \otimes \varphi^2 \otimes \varphi^3)(w_{\sigma(1)}, w_{\sigma(2)}, w_{\sigma(3)}) \mathrm{sgn}(\sigma) = \sum_{\sigma \in S_3} \mathrm{sgn}(\sigma) \varphi^1(w_{\sigma(1)}) \varphi^2(w_{\sigma(2)}) \varphi^3(w_{\sigma(3)}) = \underbrace{\varphi^1(w_1)\varphi^2(w_2)\varphi^3(w_3)}_{\sigma = \mathrm{Id}} \underbrace{- \varphi^1(w_2)\varphi^2(w_1)\varphi^3(w_3)}_{\sigma = (1\,2)}$

$\underbrace{- \varphi^1(w_3)\varphi^2(w_2)\varphi^3(w_1)}_{\sigma = (1\,3)} \underbrace{- \varphi^1(w_1)\varphi^2(w_3)\varphi^3(w_2)}_{\sigma = (2\,3)} \underbrace{+ \varphi^1(w_2)\varphi^2(w_3)\varphi^3(w_1)}_{\sigma = (1\,2\,3)} \underbrace{+ \varphi^1(w_3)\varphi^2(w_1)\varphi^3(w_2)}_{\sigma = (1\,3\,2)} \dots$ Thus in conclusion, we have that $\varphi^1 \wedge \varphi^2 \wedge \varphi^3$ is equivalent to $\varphi^1 \otimes \varphi^2 \otimes \varphi^3 - \varphi^2 \otimes \varphi^1 \otimes \varphi^3 - \varphi^3 \otimes \varphi^2 \otimes \varphi^1 - \varphi^1 \otimes \varphi^3 \otimes \varphi^2 + \varphi^2 \otimes \varphi^3 \otimes \varphi^1 + \varphi^3 \otimes \varphi^1 \otimes \varphi^2$

27 November 2013
Dr. Yiannis PETRIDIS.
Maths 7b6.

Remark – If $\dim V = n$, $\dim \Lambda^n(V) = \binom{n}{n} = 1$. Let $w_1, w_2, \dots, w_n$ be vectors in V. $w_1 = \begin{pmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{pmatrix}$, $w_2 = \begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{pmatrix}, \dots, w_n = \begin{pmatrix} w_{1n} \\ w_{2n} \\ \vdots \\ w_{nn} \end{pmatrix}$. Then consider the $n \times n$ matrix $\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$. Then $\det(w_1, w_2, \dots, w_n)$ is an alternating tensor. Notably, $\det I_n = 1$ is a basis for the 1D-space.

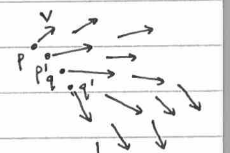


Let $p \in \mathbb{R}^n$. Let $\boxed{\mathbb{R}^n_p = \{(\underset{\text{tied, initial point}}{p}, \underset{\text{standard vector.}}{v}),\ v \in \mathbb{R}^n\}}$ (i.e. set of all vectors beginning from $p \in \mathbb{R}^n$. This is a vector space, with $(p, v_1) + (p, v_2) = (p, v_1 + v_2)$ and $a \in \mathbb{R} \Rightarrow a(p, v_1) = (p, av_1)$. This space $\mathbb{R}^n_p$ is called the <u>tangent space</u> of $\mathbb{R}^n$ at p.

Note – If $p \neq q$, $(p, v_1) + (q, v_2)$ is not well-defined.

Notation – We write $(p, v) = v_p$ [i.e. tied at point p].

Clearly, $\mathbb{R}^n_p \cong \mathbb{R}^n$, so $\mathbb{R}^n_p$ has basis $\{(e_1)_p, (e_2)_p, \dots, (e_n)_p\}$. Take a vector at p. Then $v_p = a^1(e_1)_p + a^2(e_2)_p + \dots + a^n(e_n)_p$, $a^i \in \mathbb{R}$.

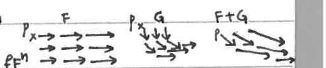


A <u>vector field</u> F on $\mathbb{R}^n$ (or a subset $A \subseteq \mathbb{R}^n$) is a function $F : \underset{\mathbb{R}^n}{\overset{}{A}} \longrightarrow \bigcup_{p \in \mathbb{R}^n} \mathbb{R}^n_p$ s.t. $F(p) \in \mathbb{R}^n_p$ [i.e. choose a vector at each point].

Then $F(p) = F^1(p)(e_1)_p + F^2(p)(e_2)_p + \dots + F^n(p)(e_n)_p$, where for a fixed p, $F^i(p) \in \mathbb{R}$. As p varies, we get $F^1(p), \dots, F^n(p)$ to be scalar functions.

The vector field is called continuous (respectively differentiable) if the functions $F^1, F^2, \dots, F^n : \mathbb{R}^n \to \mathbb{R}$ are continuous (respectively differentiable).

Recall that we can sum vector fields by adding components respectively. We can also multiply a vector field F with a scalar function $f : \mathbb{R}^n \to \mathbb{R}$. Let $F(p) \in \mathbb{R}^n_p$, $f(p) \in \mathbb{R}$, $f(p) \cdot F(p) \in \mathbb{R}^n_p$. Then $p \mapsto f(p) \cdot F(p)$ $\forall p \in \mathbb{R}^n$ defines a vector field with components $fF^1, fF^2, \dots, fF^n$.



If F is a vector field, <u>divergence</u> $\operatorname{div} F : \mathbb{R}^n \to \mathbb{R}$ (scalar) given by $(\operatorname{div} F)(p) = \frac{\partial F^1}{\partial x^1}(p) + \frac{\partial F^2}{\partial x^2}(p) + \dots + \frac{\partial F^n}{\partial x^n}(p) = \sum_{i=1}^n D_i F^i(p)$. If F is on $\mathbb{R}^3$, we define <u>curl</u> of F by

$\operatorname{curl}(F)(p) = \begin{vmatrix} i & j & k \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ F^1 & F^2 & F^3 \end{vmatrix} = \left(\frac{\partial F^3}{\partial y} - \frac{\partial F^2}{\partial z}\right)\hat{i} - \left(\frac{\partial F^3}{\partial x} - \frac{\partial F^1}{\partial z}\right)\hat{j} + \left(\frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y}\right)\hat{k}$ = or rather, $\left(\frac{\partial F^3}{\partial y}(p) - \frac{\partial F^2}{\partial z}(p)\right)(e_1)_p - \left(\frac{\partial F^3}{\partial x}(p) - \frac{\partial F^1}{\partial z}(p)\right)(e_2)_p + \left(\frac{\partial F^2}{\partial x}(p) - \frac{\partial F^1}{\partial y}(p)\right)(e_3)_p$.

This gives a new vector field.

For our considerations, take $V = \mathbb{R}^n_p$, with basis $\{(e_1)_p, \dots, (e_n)_p\}$. Let $w(p) \in \Lambda^k(\mathbb{R}^n_p)$. Then a <u>k-form</u> on $\mathbb{R}^n$ is a choice of $w(p) \in \Lambda^k(\mathbb{R}^n_p)$ for every $p \in \mathbb{R}^n$.

At p, what does $w(p)$ look like? $w(p) \in \Lambda^k(\mathbb{R}^n_p)$. Then $\sum_{1 \le i_1 < \dots < i_k = n} w_{i_1 \dots i_k}(p) \varphi^{i_1}(p) \wedge \varphi^{i_2}(p) \wedge \dots \wedge \varphi^{i_k}(p)$ where $\varphi^i(p)$ is the dual basis of $\{(e_1)_p, (e_2)_p, \dots, (e_n)_p\}$.

$w(p)$ is determined by $w_{i_1 \dots i_k}(p)$, so we have functions $p \mapsto w_{i_1} \dots w_{i_k}(p)$ for $i_1 < i_2 < \dots < i_k$, $i_1, i_2, \dots, i_k \in \{1, 2, \dots, n\}$.

A continuous (respectively differential) k-form on $\mathbb{R}^n$ is a k-form where all these functions $w_{i_1 i_2 \dots i_k}(p)$ are continuous (respectively differentiable) on $\mathbb{R}^n$.

Given w, η two differential k-forms on $\mathbb{R}^n$, we can define $w + \eta$ to be a differential k-form on $\mathbb{R}^n$ by $w(p), \eta(p) \in \Lambda^k(\mathbb{R}^n_p)$, $(w + \eta)(p) = w(p) + \eta(p) \in \Lambda^k(\mathbb{R}^n_p)$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar field, then we can define $f \cdot w$ to be a differential k-form on $\mathbb{R}^n$ by $\underbrace{f(p)}_{\in \mathbb{R}} \cdot \underbrace{w(p)}_{\in \Lambda^k(\mathbb{R}^n_p)} = (f \cdot w)(p)$.

Given $f : \mathbb{R}^n \to \mathbb{R}$ differentiable, fix p. Then $Df(p) : \mathbb{R}^n \to \mathbb{R}$ is a linear map/functional, so $Df(p) \in (\mathbb{R}^n_p)^* = J_1(\mathbb{R}^n_p) = \Lambda^1(\mathbb{R}^n_p) \Rightarrow Df$ is an alternating 1-tensor, so $p \to Df(p)$ is a differential 1-form on $\mathbb{R}^n$. We denote it by df, $(df)(p) = Df(p)$. i.e. $df(p)(v_p) = Df(p)(v) \in \mathbb{R}$ with $v_p = (p, v)$.

1-form, so 1 slot

Let $i \in \{1, \dots, n\}$. $\pi^i : \mathbb{R}^n \to \mathbb{R}$ $\pi^i(x^1, \dots, x^n) = x^i$. Since π^i is linear, $D\pi^i(p) = \pi^i$. So $D\pi^i(p)(v) = \pi^i(v) = v^i = \varphi^i(v)$. By definition, $d\pi^i(p)(v_p) = D\pi^i(p)(v)$. Thus, $\boxed{\varphi = d\pi^i}$

Notation – $\pi^i = x^i$, $d\pi^i = dx^i$, so that $dx^i = \varphi^i$. So $w(p) = \sum_{i_1 < i_2 < \dots < i_k} w_{i_1 i_2 \dots i_k}(p)\, dx^{i_1}(p) \wedge dx^{i_2}(p) \wedge \dots \wedge dx^{i_k}(p)$. Thus, $\boxed{w = \sum_{i_1 < \dots < i_k} w_{i_1 i_2 \dots i_k}\, dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}}$

We then classify differential forms –

For $\mathbb{R}^1$, $\dim \Lambda^k(\mathbb{R}^1_p) = \binom{1}{k} \neq 0 \Leftrightarrow k = 0$ or 1. If $k = 0$, only 0-form is $w = f(x)$. If $k = 1$, $w = f(x)\,dx$ is a 1-form.

For $\mathbb{R}^2$, $\dim \Lambda^k(\mathbb{R}^2_p) = \binom{2}{k} \neq 0 \Leftrightarrow k = 0, 1, 2$. We normally use $(x, y) = (x^1, x^2)$ for this case. If $k = 0$, $w = f(x, y)$ is only 0-form. If $k = 1$, $\dim \Lambda^1(\mathbb{R}^2_p) = 2$. Then we have $w = P(x, y)\,dx + Q(x, y)\,dy$ is the 1-form. If $k = 2$, $w = f(x, y)\,dx \wedge dy$. Note that $dx \wedge dx = 0$, $dy \wedge dy = 0$, $dx \wedge dy = -dy \wedge dx$.

For $\mathbb{R}^3$, $\dim \Lambda^k(\mathbb{R}^3_p) = \binom{3}{k} \neq 0 \Leftrightarrow k = 0, 1, 2, 3$. If $k = 0$, $w = f(x, y, z)$ is 0-form. [Here, $(x, y, z) = (x^1, x^2, x^3)$]. If $k = 1$, $w = f(x, y, z)\,dx + g(x, y, z)\,dy + h(x, y, z)\,dz$ is 1-form.

If $k = 2$, $w = P(x, y, z)\,dx \wedge dy + Q(x, y, z)\,dx \wedge dz + R(x, y, z)\,dy \wedge dz$ with $dx \wedge dx = dy \wedge dy = dz \wedge dz = 0$, $dx \wedge dy = -dy \wedge dx$, $dx \wedge dz = -dz \wedge dx$, $dy \wedge dz = -dz \wedge dy$.

Thus, w is equivalent also to $w = -R(x, y, z)\,dz \wedge dy - P(x, y, z)\,dy \wedge dx + Q(x, y, z)\,dx \wedge dz = f_1(x, y, z)\,dz \wedge dy + f_2(x, y, z)\,dy \wedge dx + f_3(x, y, z)\,dx \wedge dy$.

If $k = 3$, $w = f(x, y, z)\,dx \wedge dy \wedge dz$ is the 3-form.

Theorem [total derivative]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. Then $df = D_1 f\,dx^1 + D_2 f\,dx^2 + \dots + D_n f\,dx^n$ i.e. $\frac{\partial f}{\partial x^1} dx^1 + \frac{\partial f}{\partial x^2} dx^2 + \dots + \frac{\partial f}{\partial x^n} dx^n$.

Proof – At $p \in \mathbb{R}^n$, $(p, v) = v_p$ s.t. $v \in \mathbb{R}^n$. $df(p)(v_p) \overset{\text{def}}{=} Df(p)(v) = \overset{\text{Jacobian}}{f'(p)}(v) = (D_1 f(p), D_2 f(p), \dots, D_n f(p)) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = \sum_{j=1}^n D_j f(p) v^j$. Then, we have for the RHS,

$(D_1 f\,dx^1 + \dots + D_n f\,dx^n)(p)(v_p) = [D_1 f(p)\,dx^1(p) + D_2 f(p)\,dx^2(p) + \dots + D_n f(p)\,dx^n(p)](v_p) = D_1 f(p)\,dx^1(p)(v_p) + \dots + D_n f(p)\,dx^n(p)(v_p) = D_1 f(p) v^1 + \dots + D_n f(p) v^n$, q.e.d.

Let w be a 0-form on $\mathbb{R}^n$. $w = f(x^1, \dots, x^n)$. We define $dw = \frac{\partial f}{\partial x^1} dx^1 + \frac{\partial f}{\partial x^2} dx^2 + \dots + \frac{\partial f}{\partial x^n} dx^n$. Then now let w be a k-form with $k \geq 1$. Then $w = \sum_{i_1 < \dots < i_k} w_{i_1 \dots i_k}\, dx^{i_1} \wedge \dots \wedge dx^{i_k}$.

Definition $dw = \sum_{i_1 < \dots < i_k} \sum_{\alpha=1}^n D_\alpha w_{i_1 i_2 \dots i_k}\, dx^\alpha \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}$. This is now a $k+1$ form, and is called the <u>exterior derivative</u>.
Note – Hence we can consider an environment of p.

For $\mathbb{R}^1$, $k = 0$, $w = f(x)$, $dw = f'(x)\,dx$. If $k = 1$, $w = f(x)\,dx$. So $dw = f'(x)\,dx \wedge dx = 0$.

For $\mathbb{R}^2$, $k = 0$, $w = f(x, y)$, $dw = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$. For $k = 1$, $w = P(x, y)\,dx + Q(x, y)\,dy$. Then $dw = \frac{\partial P}{\partial x} \overbrace{dx \wedge dx}^{0} + \frac{\partial P}{\partial y} dy \wedge dx + \frac{\partial Q}{\partial x} dx \wedge dy + \frac{\partial Q}{\partial y} \overbrace{dy \wedge dy}^{0} = \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) dx \wedge dy$.

If $k = 2$, $w = f(x, y)\,dx \wedge dy$. $dw = \frac{\partial f}{\partial x} dx \wedge dx \wedge dy + \frac{\partial f}{\partial y} dy \wedge dx \wedge dy = 0$ (clear, since there is no 3-dim alternating form on $\mathbb{R}^2$).

For $\mathbb{R}^3$, $k=0$. $w = f(x,y,z)$, $dw = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$. For $k=1$, $w = f(x,y,z)\,dx + g(x,y,z)\,dy + h(x,y,z)\,dz$. Then we expand terms in exterior derivative to get

$dw = \frac{\partial f}{\partial x}dx\wedge dx + \frac{\partial f}{\partial y}dy\wedge dx + \frac{\partial f}{\partial z}dz\wedge dx + \frac{\partial g}{\partial x}dx\wedge dy + \frac{\partial g}{\partial y}dy\wedge dy + \frac{\partial g}{\partial z}dz\wedge dy + \frac{\partial h}{\partial x}dx\wedge dz + \frac{\partial h}{\partial y}dy\wedge dz + \frac{\partial h}{\partial z}dz\wedge dz = \left(\frac{\partial g}{\partial z} - \frac{\partial h}{\partial y}\right)dz\wedge dy + \left(\frac{\partial f}{\partial y} - \frac{\partial g}{\partial x}\right)dy\wedge dx + \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right)dx\wedge dz$

This corresponds to $\mathrm{curl}(f,g,h) = \begin{vmatrix} i & j & k \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ f & g & h \end{vmatrix} = \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)i - \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right)j + \left(-\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x}\right)k$. Then for $k=2$, $w = f_1\,dy\wedge dz + f_2\,dz\wedge dx + f_3\,dx\wedge dy$. Then we have

$dw = \frac{\partial f_1}{\partial x}dx\wedge dy\wedge dz + \frac{\partial f_2}{\partial y}dy\wedge dz\wedge dx + \frac{\partial f_3}{\partial z}dz\wedge dx\wedge dy = \left(\frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z}\right)dx\wedge dy\wedge dz = (\mathrm{div}\,F)\,dx\wedge dy\wedge dz$ with $F = (f_1, f_2, f_3)$.

29 November 2013.
Dr Yiannis PETRIDIS
Maths 706.

Let w be a differential k-form on $\mathbb{R}^n$, $w = \sum_{i_1<\dots<i_k} w_{i_1\dots i_k}\,dx^{i_1}\wedge\dots\wedge dx^{i_k}$, $\eta = \sum_{j_1<\dots<j_k} \eta_{j_1\dots j_\ell}\,dx^{j_1}\wedge\dots\wedge dx^{j_\ell}$ $\Rightarrow$ $\overset{(k+\ell)\text{-form}}{w\wedge\eta} = \sum_{i_1<\dots<i_k}\sum_{j_1<\dots<j_\ell} w_{i_1\dots i_k}\eta_{j_1\dots j_\ell}\,dx^{i_1}\wedge\dots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\dots\wedge dx^{j_\ell}$

Theorem Exterior derivative has the following properties:

(1) $d(w+\eta) = dw + d\eta$ for w,η k-forms (2) $d(dw) = 0$ for any w k-form (3) $d(w\wedge\eta) = dw\wedge\eta + (-1)^k w\wedge d\eta$ (Product Rule)

Proof – (1) Easy, left as exercise. (2). $dw = \sum_{i_1<\dots<i_k}\sum_{\alpha=1}^n D_\alpha w_{i_1\dots i_k}\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}$. Then $d(dw) = \sum_{\beta=1}^n\sum_{i_1<\dots<i_k}\sum_{\alpha=1}^n D_\beta(D_\alpha w_{i_1\dots i_k})\,dx^\beta\wedge dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}$

If $\alpha=\beta$, $dx^\alpha\wedge dx^\alpha\wedge\dots = 0$. Thus, suppose $\alpha\neq\beta$. For each (α,β), each contributing term is $D_\beta(D_\alpha w_{i_1\dots i_k})\,dx^\beta\wedge dx^\alpha\wedge\dots$. Since (α,β) is off-diagonal, we also have (β,α) as well, giving $D_\alpha(D_\beta w_{i_1\dots i_k})\,dx^\alpha\wedge dx^\beta\wedge\dots$. Since w is infinitely differentiable, mixed partial derivatives are the same, then by anti-commutativity, (α,β) term cancels out (β,α) term // q.e.d. (3) Note first that $d(w\wedge\eta)$ is a $(k+\ell)$-form. Then $w\wedge\eta = \sum_{\substack{i_1<\dots<i_k\\ j_1<\dots<j_\ell}} w_{i_1\dots i_k}\eta_{j_1\dots j_k}\,dx^{i_1}\wedge\dots\wedge dx^{j_\ell}$

Then $d(w\wedge\eta) = \sum_{\substack{i_1<\dots<i_k\\ j_1<\dots<j_\ell}}\sum_{\alpha=1}^n D_\alpha(w_{i_1\dots i_k}\eta_{j_1\dots j_\ell})\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\wedge dx^{j_1}\wedge\dots\wedge dx^{j_\ell} = \sum_{\substack{i_1<\dots<i_k\\ j_1<\dots<j_\ell}}\sum_{\alpha=1}^n [D_\alpha(w_{i_1\dots i_k})\eta_{j_1\dots j_\ell} + w_{i_1\dots i_k}D(\eta_{j_1\dots j_\ell})]\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{j_\ell}$

$= \sum_{\substack{i_1<\dots<i_k\\ j_1<\dots<j_\ell}}\sum_\alpha \overbrace{D_\alpha(w_{i_1\dots i_k})}^{\text{at } p,\ \mathbb{R}}\overbrace{\eta_{j_1\dots j_\ell}}^{\text{at } p,\ \mathbb{R}}\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{j_\ell} + \sum_{\substack{i_1<\dots<i_k\\ j_1<\dots<j_\ell}}\sum_\alpha w_{i_1\dots i_k}D_\alpha(\eta_{j_1\dots j_\ell})\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{j_\ell}$. Then we can apply commutativity of real numbers to terms,

taken at p. $= \sum_{i_1<\dots<i_k}\sum_{\alpha=1}^n D_\alpha(w_{i_1\dots i_k})\,dx^{i_1}\wedge\dots\wedge dx^{i_k}\wedge\sum_{j_1<\dots<j_\ell}\eta_{j_1\dots j_\ell}\,dx^{j_1}\wedge\dots\wedge dx^{j_\ell} + \sum_{i_1<\dots<i_k} w_{i_1\dots i_k}\,dx^\alpha\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\wedge\sum_{j_1<\dots<j_\ell} D_\alpha(\eta_{j_1\dots j_\ell})\,dx^{j_1}\wedge\dots\wedge dx^{j_\ell} = dw\wedge\eta + (-1)^k w\wedge d\eta$ // q.e.d.

Closed and exact forms.

Definition A k-form w is called closed if $dw=0$. A k-form w is called exact if $\exists$ a $(k-1)$ form η s.t. $d\eta = w$.

Theorem An exact form is closed. [Contrapositive: A form that is not closed is not exact].

Proof – If w is exact then $w = d\eta$, so $dw = d(d\eta) = 0$ // q.e.d.

Let us consider $\mathbb{R}^2$. $n=2$. k-form for $k=1$ is $w = P(x,y)\,dx + Q(x,y)\,dy$. w is closed, so $dw=0 \Rightarrow dw = P_x\,dx\wedge dx + P_y\,dy\wedge dx + Q_x\,dx\wedge dy + Q_y\,dy\wedge dy = -\frac{\partial P}{\partial y}dx\wedge dy + \frac{\partial Q}{\partial x}dx\wedge dy$

$= \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)dx\wedge dy = 0 \Rightarrow -\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x} = 0 \Rightarrow \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. This gives an irrotational vector field (or conservative), which has a potential.

If the form is exact, $w = df = D_1 f\,dx + D_2 f\,dy \Rightarrow \begin{matrix} P = D_1 f = \partial f/\partial x \\ Q = D_2 f = \partial f/\partial y \end{matrix}$. Then f is called a potential. We know that an exact form is closed, but does the converse apply?

Examples –

1. Let $xy^2\,dx + y\,dy = w$. Then $dw = d(xy^2\,dx + y\,dy) = 2xy\,dy\wedge dx + 0\,dx\wedge dy \neq 0 \Rightarrow$ not closed $\Rightarrow$ not exact.

2. Let $w = xy^2\,dx + x^2y\,dy$, $dw = 2xy\,dy\wedge dx + 2xy\,dx\wedge dy = (-2xy + 2xy)\,dx\wedge dy = 0 \Rightarrow$ closed. Is it exact? We search for potential f with $\frac{\partial f}{\partial x} = xy^2$, $\frac{\partial f}{\partial y} = x^2y$. Then

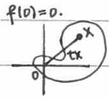
$f = \int xy^2\,dx = \frac{x^2}{2}y^2 + c(y) \Rightarrow f_y = x^2y + c'(y) = x^2y \Rightarrow c'(y) = 0 \Rightarrow c(y) = c$. So $f(x,y) = \frac{x^2y^2}{2} + c$, and form is exact.

If $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, then the function $f(x,y) = \int_0^x P(t,0)\,dt + \int_0^y Q(x,t)\,dt$ satisfies $\frac{\partial f}{\partial x} = P$, $\frac{\partial f}{\partial y} = Q$. (c.f. Homework 5, Q5). Take $w = \overbrace{xy^2}^{P}dx + \overbrace{x^2y}^{Q}dy$ as in Example 2. Then we have

$f(x,y) = \int_0^x t\cdot 0^2\,dt + \int_0^y x^2t\,dt = \frac{x^2t^2}{2}\Big|_{t=0}^{t=y} = \frac{x^2y^2}{2}$, as per before.

Now we explore higher dimensions: $n\geq 2$, $k=1$. $w = w_1\,dx^1 + w_2\,dx^2 + \dots + w_n\,dx^n = \sum_{j=1}^n w_j\,dx^j$ where w_j are functions on n variables. Then $dw = \sum_{j=1}^n\sum_{\alpha=1}^n D_\alpha w_j\,dx^\alpha\wedge dx^j$. If w is closed, $dw = 0 \Rightarrow D_\alpha w_j\,dx^\alpha\wedge dx^j + D_j w_\alpha\,dx^j\wedge dx^\alpha = 0 \Rightarrow D_\alpha w_j = D_j w_\alpha$. w is exact means $w = df = \sum_{j=1}^n D_j f\,dx^j \Leftrightarrow w_j = D_j f$ $j = 1,2,\dots,n$. Fix $f(0) = 0$.

Then $f(x) = f(x) - f(0) = f(1\cdot x) - f(0\cdot x) = \int_0^1 \frac{d}{dt}(f(tx))\,dt = \int_0^1\sum_{i=1}^n D_i f(tx)\,x^i\,dt$ (by chain rule) $= \sum_{i=1}^n\int_0^1 w_i(tx)\,x^i\,dt$ by exactness.



Cases where we can use the formula are restricted to regions containing the ray from 0 to x.

Definition A region $A\subseteq\mathbb{R}^n$ is called star-shaped w.r.t. 0 if $\forall x\in A$, the segment $[0,x]$ belongs to A i.e. $\forall t\in[0,1]$, $tx\in A$.

Theorem (Poincaré Lemma)

If A is a star-shaped region and w is a closed form on A, then w is exact.

Proof – Refer to Handout 5.

4 December 2013.
Dr. Yiannis PETRIDIS
Maths 706.

Geometric Preliminaries.

Definition The standard cube I^k in $\mathbb{R}^k$ is $I^k = [0,1]^k = \overbrace{[0,1]\times\dots\times[0,1]}^{k\text{ times.}}$. A singular cube is a continuous map in A $c: I^k\to A\subseteq\mathbb{R}^n$

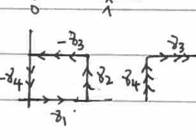
Examples – 1-cube $c:[0,1]\to A$ curve in A, 2-cube $c:[0,1]^2\to A$ surface in A, 3-cube $c:[0,1]^3\to A$ solid in A.

If $k=0$, $[0,1]^0 = \{0\}$ is zero cube, $c:\{0\}\to A$ is just a point in A.

We then examine the boundary of a k-cube. If $k=1$, boundary is given by Fundamental Theorem of Calculus – $\int_0^1 F' = F(1) - F(0)$, $\partial I^1 = +1 - 0$

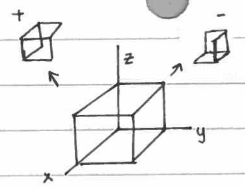
If $k=2$, we have the following cube: traversed anticlockwise. $[0,1]^2 = I^2$. $\gamma_1 = \{(x,0), 0\leq x\leq 1\}$, $\gamma_3 = \{(x,1), 0\leq x\leq 1\}$.

Likewise, $\gamma_2 = \{(1,y), 0\leq y\leq 1\}$, $\gamma_4 = \{(0,y), 0\leq y\leq 1\}$. Then $\partial I^2 = \gamma_1 + \gamma_2 - \gamma_3 - \gamma_4$. This notation is clumsy, and breaks down in higher dimensions.



Instead, we write $\gamma_1: \underset{\text{second var.}\;\;\text{set to }0}{I^2_{(2,0)}} = \{(x,0),\ 0\le x\le 1\}$, $\gamma_2: I^2_{(1,1)} = \{(1,y),\ 0\le y\le 1\}$, $\gamma_3: I^2_{(2,1)} = \{(x,1),\ 0\le x\le 1\}$, $\gamma_4: I^2_{(1,0)} = \{(0,y),\ 0\le y\le 1\}$.

Note that we can predict the sign of $\gamma_i = I^2_{(\alpha,\beta)}$, which is given by $(-1)^{\alpha+\beta}$. We investigate further for $k=3$.

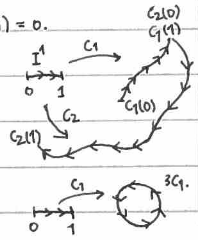


For 3-cube, boundary is the 6 faces (each found by fixing one variable and varying the other two. We get:

(bottom) $I^3_{(3,0)} = \{(x,y,0),\ 0\le x,y\le 1\}$ (top) $I^3_{(3,1)} = \{(x,y,1),\ 0\le x,y\le 1\}$ (front) $I^3_{(1,1)} = \{(1,y,z),\ 0\le y,z\le 1\}$ (back) $I^3_{(1,0)} = \{(0,y,z),\ 0\le y,z\le 1\}$

(left) $I^3_{(2,0)} = \{(x,0,z),\ 0\le x,z\le 1\}$ (right) $I^3_{(2,1)} = \{(x,1,z),\ 0\le x,z\le 1\}$. Then $\partial I^3 = -I^3_{(3,0)} + I^3_{(3,1)} + I^3_{(1,1)} - I^3_{(1,0)} + I^3_{(2,0)} - I^3_{(2,1)}$.

Earlier, we showed that $\partial I^2 = I^2_{(2,0)} + I^2_{(1,1)} - I^2_{(2,1)} - I^2_{(1,0)}$. Thus, $\partial(\partial I^2) = (+(1,0)-(0,0)) + (+(1,1)-(1,0)) - (+(1,1)-(0,1)) - (+(0,1)-(0,0)) = 0$.



We form formal sums of singular n-cubes with integer coefficients (this constructs an abelian group or $\mathbb{Z}$-module). Such formal sums are called n-chains.

Given the standard n-cube $I^n = [0,1]^n$, we define $I^n_{(i,0)} = \{(x^1,\dots,\overset{i}{0},x^{i+1},\dots,x^n),\ 0\le x^j\le 1\}$. $I^n_{(i,1)} = \{(x^1,\dots,\overset{i}{1},x^{i+1},\dots,x^n),\ 0\le x^j\le 1\}$.

Then define $\partial I^n = \sum_{j=1}^{n}\sum_{\alpha=0,1}(-1)^{j+\alpha} I^n_{(j,\alpha)}$. Let c be a singular n-cube, $c: [0,1]^n = I^n \to A \subseteq \mathbb{R}^m$, $\partial c = \sum_{j=1}^{n}\sum_{\alpha=0,1}(-1)^{\alpha+j} c(I^n_{(j,\alpha)})$.

Let c be a singular chain $c = \sum_{m=1}^{\ell} a_m c_m$, $a_m \in \mathbb{Z}$ and c_m singular cubes, then $\partial c = \sum_{m=1}^{\ell} a_m \partial(c_m)$.

In the 3-cube:  the signs along common edges cancel out.

For today's lecture, take $\mathbb{R}^k$ s.t. ω is a k-form, η is a $(k-1)$-form. Then ω will be integrated over I^k, η will be integrated over ∂I^k. Since $\binom{k}{k}=1$, let $\omega = f(x^1,\dots,x^k)\, dx^1\wedge\dots\wedge dx^k$.

Then we have the following definition:

$\boxed{\text{Definition}}$ $\int_{I^k}\omega = \int\!\!\int\!\cdots\!\int_{[0,1]^k} f(x^1,\dots,x^k)\, dx^1 dx^2\cdots dx^k$.

Remark – since f is continuous, it is Riemann integrable. Note also that we could exchange the order of integration, as a consequence of Fubini's theorem.

Since η is a k-1 form on $\mathbb{R}^k$, $\binom{k}{k-1} = k$, then $\eta = \sum_{i=1}^{k} f_i(x^1,\dots,x^k)\, dx^1\wedge dx^2\wedge\dots\wedge\widehat{dx^i}\wedge\dots\wedge dx^k$, where $\widehat{dx^i}$ denotes that the term is missing. Then we define integral of η on its individual pieces. i.e. the faces of I^k, which are $I^k_{(j,\alpha)}$. Then we have:

$\boxed{\text{Definition}}$ $\int_{I^k_{(j,\alpha)}}\eta = \begin{cases}\int_{[0,1]^{k-1}} f_i(x^1,\dots,\underset{i=j\text{ slot}}{\alpha},\dots,x^k)\, dx^1 dx^2\cdots\widehat{dx^i}\cdots dx^k & i=j\\ 0 & i\ne j.\end{cases}$

Note – Recall that in MATH1402, $\int_\gamma A\, dy = 0$ if A is a horizontal line. So if $i\ne j$, integral is 0.

$\int_c \omega_1+\omega_2 = \int_c\omega_1 + \int_c\omega_2$, $\int_c\lambda\omega = \lambda\int_c\omega$ by linearity. If $c = \sum_{m=1}^{\ell} a_m c_m$, $a_m\in\mathbb{Z}$. $\int_c\omega = \sum_{m=1}^{\ell} a_m\int_{c_m}\omega$. A 0-cube is a point A, $\int_c\omega = \omega(A)$.

$\boxed{\text{Theorem}}$ (Stokes' theorem)

For ω a $(k-1)$-form in $\mathbb{R}^k$, by linearity it suffices to prove when $\omega = f(x^1,\dots,x^k)\, dx^1\wedge\dots\wedge\widehat{dx^i}\wedge\dots\wedge dx^k$, $\int_{I^k} d\omega = \int_{\partial I^k}\omega$.

Proof – $\partial I^k = \sum_{j=1}^{k}\sum_{\alpha=0,1}(-1)^{j+\alpha} I^k_{(j,\alpha)}$. Then $\int_{\partial I^k}\omega \overset{\text{def}}{=} \sum_{j=1}^{k}\sum_{\alpha=0,1}(-1)^{j+\alpha}\int_{I^k_{(j,\alpha)}}\omega = \sum_{j=1}^{k}\sum_{\alpha=0,1}(-1)^{j+\alpha}\int_{I^k_{(j,\alpha)}} f(x^1,\dots,x^k)\, dx^1\wedge\dots\wedge\widehat{dx^i}\wedge\dots\wedge dx^k \overset{\text{def}}{=} \sum_{\alpha=0,1}(-1)^{i+\alpha}\int f(x^1,\dots,\underset{i^{\text{th}}\text{ slot}}{\alpha},\dots,x^k)\, dx^1 dx^2\dots\widehat{dx^i}\dots dx^k$

Since $\omega = f(x^1,\dots,x^k)\, dx^1\wedge\dots\wedge\widehat{dx^i}\wedge\dots\wedge dx^k$, $d\omega = D_i f(x^1,\dots,x^k)\, dx^i\wedge dx^1\wedge\dots\wedge\widehat{dx^i}\wedge\dots\wedge dx^k$ (only remaining term) $= D_i f(x^1,\dots,x^k)(-1)^{i-1} dx^1\wedge\dots\wedge dx^i\wedge\dots\wedge dx^k$.

We have $\int_{I^k} d\omega \overset{\text{def}}{=} (-1)^{i-1}\int_{[0,1]^k} D_i f(x^1,\dots,x^k)\, dx^1 dx^2\dots dx^k \overset{\text{Fubini}}{=} (-1)^{i-1}\int_{[0,1]^{k-1}}\left(\int_0^1 D_i f(x^1,\dots,x^k)\, dx^i\right) dx^1 dx^2\dots\widehat{dx^i}\dots dx^k = (-1)^{i-1}\int_{[0,1]^{k-1}}[f(x^1,\dots,1,\dots,x^k) - f(x^1,\dots,0,\dots,x^k)]\, dx^1\dots\widehat{dx^i}\dots dx^k$

$= (-1)^{i+1}\int_{[0,1]^{k-1}} f(x^1,\dots,1,x^{i+1},\dots,x^k)\, dx^1\dots\widehat{dx^i}\dots dx^k - (-1)^{i+1}\int_{[0,1]^{k-1}} f(x^1,x^2,\dots,0,\dots,x^k)\, dx^1\dots\widehat{dx^i}\dots dx^k$ $[\because (-1)^{i-1} = (-1)^{i+1}]$ $= \sum_{\alpha=0,1}(-1)^{i+\alpha} f(x^1,\dots,\alpha,\dots,x^k)\, dx^1\dots\widehat{dx^i}\dots dx^k$ // q.e.d.

Return to complex Analysis. We parametrise γ by $z(t)$, $a\le t\le b$, then $\int_\gamma f(z)\,dz = \int_a^b f(z(t))\,z'(t)\,dt$. We have thus far defined, for ω a k-form in $\mathbb{R}^k$, $\int_{I^k}\omega$. Then for c a singular cube, $c: I^k\to\mathbb{R}^n$, we have $\int_c\omega \overset{\text{def}}{=} \int_{I^k} c^*(\omega)$ where c^* is the pullback (analogous to parametrisation of differential form ω to I^k.

If $f: W\to V$ is a linear transformation, W,V vector spaces and T is a linear functional on V, then $f^*(T) = T\circ f$ is a linear functional on W.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "T\circ f"'] & W \arrow[d, "T"] \\ & \mathbb{R} \end{tikzcd}$$

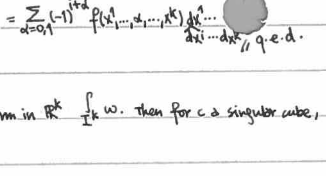
If S is a k-tensor on V, then $f^*(S)$ is a k-tensor on W defined by $f^*(S)(w_1,w_2,\dots,w_k) = S(f(w_1),\dots,f(w_k))$. $f(w) = \underset{\text{[pushforward]}}{f_*(w)} \in V$.

Let $\mathbb{R}^n \xrightarrow{g} \mathbb{R}^m$, g differentiable. Fix p, then $Dg(p): \mathbb{R}^n\to\mathbb{R}^m$ is a linear map with $\{(p,v),\ v\in\mathbb{R}^n\} = \mathbb{R}^n_p \longrightarrow \mathbb{R}^m_{g(p)} = \{(g(p),w),\ w\in\mathbb{R}^m\}$. Then we define $g_*: \mathbb{R}^n_p \longrightarrow \mathbb{R}^m_{g(p)}$, $g_*(p,v) = (g(p), Dg(p)(v))$. [or, $g_*(v_p) = (w_{g(p)})$ where $w = Dg(p)(v)$. Then g_* is "essentially" $Dg(p)$, which is linear. So $g_*: \mathbb{R}^n_p\to\mathbb{R}^m_{g(p)}$ is linear, and we can use g_* for f. If T is linear functional on $\mathbb{R}^m_{g(p)}$, $T\circ g_*$ is linear functional on $\mathbb{R}^n_p$.

$$\begin{tikzcd} \mathbb{R}^n_p \arrow[r, "g_*"] \arrow[dr] & \mathbb{R}^m_{g(p)} \arrow[d, "T"] \\ & \mathbb{R} \end{tikzcd}$$

Then $g^*(T) = T\circ g_*$ is the pullback of T from $\mathbb{R}^m_{g(p)}$ to $\mathbb{R}^n_p$.

In this case, if $S\in\Lambda^k(\mathbb{R}^m_{g(p)})$, then $g^*(S)\in\Lambda^k(\mathbb{R}^n_p)$ with $g^*(S)(v_1,\dots,v_k) = S(g_*(v_1),\dots,g_*(v_k))$ $= S(Dg(p)(v_1),\dots,Dg(p)(v_k))$ for $v_1,\dots,v_k\in\mathbb{R}^n_p$. $g^* = T\circ g_*$



Thus, we are pushing forward the vectors to the $\mathbb{R}^m_{g(p)}$ plane, and then we pull it back to define another differential form.

6 December 2013.
Dr. Yiannis PETRIDIS.
Maths 706

We stated that $\int_{I^k} d\omega = \int_{\partial I^k}\omega$, where ω is a k-1 form on $\mathbb{R}^k$. singular cubes $\to$ singular chains, $c: I^k\to A\subseteq\mathbb{R}^n$, $\int_c d\omega = \int_{\partial c}\omega$. Let $g: \mathbb{R}^n\to\mathbb{R}^m$, $p\mapsto g(p)$, $q\mapsto g(q)$.

For p define the linear map $Dg(p): \mathbb{R}^n\to\mathbb{R}^m$, or more specifically $\underset{\{v_p\}}{\mathbb{R}^n_p} \to \underset{\{w_{g(p)}\}}{\mathbb{R}^m_{g(p)}}$, tangent spaces at p, $g(p)$ respectively. We define this restricted map by $g_*: \mathbb{R}^n_p\to\mathbb{R}^m_{g(p)}$

Pullback of linear functionals on $\mathbb{R}^n_{g(p)}$:

$$\begin{tikzcd} \mathbb{R}^n_p \arrow[r, "g_*(p)"] \arrow[dr] & \mathbb{R}^m_{g(p)} \arrow[d, "T"] \\ & \mathbb{R} \end{tikzcd}$$

$g^*(T)(p) = T\circ g_*(p)$

$g^*(T)(p)(p,v) = T(g_*(p),(p,v)) = T(g_*(p)(v_p)) = T(Dg(p)(v))$, $T\in\mathcal{J}^1(\mathbb{R}^m_{g(p)})$, $g^*(T)\in\mathcal{J}^1(\mathbb{R}^n_p)$.

Pullback of tensors: if $S\in\mathcal{J}^k(\mathbb{R}^m_{g(p)})$, then $g^*(S)\in\mathcal{J}^k(\mathbb{R}^n_p)$. Given $\underset{(v_1)_p}{(p,v_1)},\dots,\underset{(v_k)_p}{(p,v_k)}\in\mathbb{R}^n_p$, then we have:

$g^*(S)((v_1)_p, (v_2)_p, \dots, (v_k)_p) = S(Dg(p)(v_1), Dg(p)(v_2), \dots, Dg(p)(v_k))$.

We can also find pullbacks of differential forms in particular — $g: \mathbb{R}^m \to \mathbb{R}^n$, and ω is a k-form on $\mathbb{R}^m$. If $q \in \mathbb{R}^m$, $\omega(q) \in \Lambda^k(\mathbb{R}^m_q)$. We define $g^*(\omega)$ as a k-form on $\mathbb{R}^n$.

Fix $p \in \mathbb{R}^n$, $g^*(\omega)(p) \in \Lambda^k(\mathbb{R}^n_p)$, $g^*(\omega)(p)((v_1)_p, \dots, (v_k)_p) = \omega(g(p))(g_*(p)(v_1), \dots, g_*(p)(v_k)) = \omega(g(p))(Dg(p)(v_1), \dots, Dg(p)(v_k))$.

$\boxed{\text{Theorem}}$ Let $f: \mathbb{R}^n \to \mathbb{R}^m$, $(y^1, \dots, y^n) \in \mathbb{R}^n$, $(x^1, \dots, x^m) \in \mathbb{R}^m$, $f = (f^1, \dots, f^m)$.

(a) $f^*(dx^i) = \sum_{j=1}^n D_j f^i dy^j$ (b) $f^*(c\omega_1 + \omega_2) = c f^*(\omega_1) + f^*(\omega_2)$ (c) $f^*(\omega_1 \wedge \omega_2) = f^*(\omega_1) \wedge f^*(\omega_2)$

(d) $f^*(g\omega) = (g \circ f) f^*(\omega)$ for $g: \mathbb{R}^m \to \mathbb{R}$ (e) $f^*(d\omega) = df^*(\omega)$.

Proof — (b), (c) easy, so left as exercise. (e) in Homework 9 Q1. Then for (a) — recall that $Df(p)(v) = \begin{pmatrix} \partial f^1/\partial y^1 & \cdots & \partial f^1/\partial y^n \\ \vdots & \ddots & \vdots \\ \partial f^m/\partial y^1 & \cdots & \partial f^m/\partial y^n \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}$. This is a differential form on $\mathbb{R}^n$, so we have

$f^*(dx^i)(p) \in \Lambda^1(\mathbb{R}^n_p)$, so $f^*(dx^i)(p)(v_p) \overset{\text{def}}{=} \overset{\text{1-form}}{dx^i}(f(p))(Df(p)(v)) \overset{\text{picks up } i^{\text{th}} \text{ coord of } Df(p)(v)}{=} \left(\frac{\partial f^i}{\partial y^1}(p) v^1 + \frac{\partial f^i}{\partial y^2}(p) v^2 + \dots + \frac{\partial f^i}{\partial y^n}(p) v^n\right) \in \mathbb{R}$. Then we perform calculation on RHS — this yields

$\left(\sum_{j=1}^n D_j f^i dy^j\right)(p) \in \Lambda^1(\mathbb{R}^n_p)$, then $\left(\sum_{j=1}^n D_j f^i dy^j\right)(p)(v_p) = \sum_{j=1}^n D_j f^i(p)\, dy^j(p)(v_p) = \sum_{j=1}^n D_j f^i(p)\, v_j$ = LHS // q.e.d. For (d), fix $p \in \mathbb{R}^n$, then where ω is a k-form on $\mathbb{R}^m$,

$f^*(g\omega)(p)((v_1)_p, (v_2)_p, \dots, (v_k)_p) \overset{\text{def}}{=} (g \cdot \omega)(f(p))(Df(p)(v_1), \dots, Df(p)(v_k)) = g(f(p))\, \omega(f(p))(Df(p)(v_1), \dots, Df(p)(v_k)) = (g \circ f)(p) \cdot f^*(\omega)(p)(v_1, \dots, v_k)$ // q.e.d.

Let ω be a 1-form on $\mathbb{R}^3$, with $\omega = P(x,y,z)\,dx + Q(x,y,z)\,dy + R(x,y,z)\,dz$. $f = (f^1, f^2, f^3): [0,1] \to \mathbb{R}^3$, $f(t) = (f^1(t), f^2(t), f^3(t)) = (x(t), y(t), z(t))$. $f(0)$ $f(t)$ $f(1)$

Then $f^*(\omega)$ will be a 1-form on $[0,1]$ (not $\mathbb{R}^3$!), of the form $f^*(\omega) = h(t)\,dt$. $f^*(\omega) = f^*(P\,dx + Q\,dy + R\,dz) \overset{\text{linearity}}{=} f^*(P\,dx) + f^*(Q\,dy) + f^*(R\,dz)$. We compute these pullbacks:

$f^*(P(x,y,z)\,dx) \overset{(d)}{=} (P \circ f)(t)\, f^*(dx) \overset{(a)}{=} (P \circ f)(t) \frac{d}{dt} f^1\, dt = P(x(t), y(t), z(t)) \frac{dx(t)}{dt}\, dt$. So overall, $f^*(P\,dx + Q\,dy + R\,dz) = \left(P(x(t), y(t), z(t)) \frac{dx}{dt} + Q(x(t), y(t), z(t)) \frac{dy}{dt} + R(x(t), y(t), z(t)) \frac{dz}{dt}\right) dt$.

$\boxed{\text{Definition}}$ If c is a singular k-cube i.e. $c: I^k \to A \subseteq \mathbb{R}^n$, then $\int_c \omega \overset{\text{def}}{=} \int_{I^k} c^*(\omega)$. If c is a singular k-chain, $c = \sum_{m=1}^{\ell} a_m c_m$, $a_m \in \mathbb{Z}$, c_m are k-singular cubes, then $\int_c \omega = \sum_{m=1}^{\ell} a_m \int_{c_m} \omega$

Consider $\int_c x\,dy$, where c is an ellipse — we parametrise $c: [0,1] \to (a\cos(2\pi t), b\sin(2\pi t))$. $c^1(t) = a\cos(2\pi t)$, $c^2(t) = b\sin(2\pi t)$. By definition, $\int_c x\,dy = \int_{I^1} c^*(x\,dy)$ 

$c^*(x\,dy) \overset{(d)}{=} (x \circ c(t))\, c^*(dy) = a\cos(2\pi t)\, c^*(dy) \overset{(a)}{=} a\cos(2\pi t) \frac{dc^2}{dt}\, dt = a\cos(2\pi t) \cdot b\cos(2\pi t) \cdot 2\pi\, dt = 2\pi ab\cos^2(2\pi t)\, dt$; so $\int_c x\,dy = \int_{I^1} 2\pi ab \cos^2(2\pi t)\, dt$

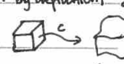
$= \int_0^1 2\pi ab\cos^2(2\pi t)\, dt = \int_0^1 2\pi ab \frac{1 + \cos(4\pi t)}{2}\, dt = 2\pi ab\left(\frac{1}{2}\right) \cdot (1 - 0) = \pi ab$. This is actually the area of the ellipse. — this is the principle used to calculate areas using a planimeter.

This is justified by Stokes theorem:

$\boxed{\text{Theorem}}$ (Stokes Theorem for singular chains).

Let c be a singular k-chain, ω a k-1 form. Then $\int_c d\omega = \int_{\partial c} \omega$.

Proof — We have proven it when c is the standard k-cube, and ω is k-1 form on $\mathbb{R}^k$. (1) If c is singular k-cube, $c: I^k \to A \subseteq \mathbb{R}^n$, ω is k-1 form on A. By definition, $\int_c d\omega = \int_{I^k} c^*(d\omega)$

$\overset{(e)}{=} \int_{I^k} d(c^*\omega)$. $c^*(\omega)$ is a k-1 form on I^k. Then by Fundamental Theorem of Calculus on $\mathbb{R}^k$, $\int_{I^k} d(c^*\omega) = \int_{\partial I^k} c^*(\omega) \overset{\text{def}}{=} \int_{c(\partial I^k)} \omega$. 

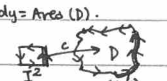
By definition of boundary of c, $\partial c = c(\partial I^k)$, so $\int_c d\omega = \int_{\partial c} \omega$. If c is a singular k-chain, $c = \sum_{m=1}^{\ell} a_m c_m$, $a_m \in \mathbb{Z}$, c_m singular k-cubes, $\partial c = \sum_{m=1}^{\ell} a_m \partial(c_m)$

$\int_c d\omega = \sum_{m=1}^{\ell} a_m \int_{c_m} d\omega = \sum_{m=1}^{\ell} a_m \int_{\partial c_m} \omega = \int_{\sum_{m=1}^{\ell} a_m \partial c_m} \omega = \int_{\partial c} \omega$ by linearity // q.e.d.

$\boxed{\text{Theorem}}$ (Green's Formula)

Given a region D with boundary γ traversed anticlockwise, $\int_\gamma P\,dx + Q\,dy = \iint_D \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) dx\,dy$.

Remark — Our example of the ellipse area has $\int_c x\,dy$, s.t. $P(x,y) = 0$ and $Q(x,y) = x$. Then by Green's formula, $\int_c x\,dy = \iint_D \left(0 + \frac{\partial x}{\partial x}\right) dx\,dy = \iint_D 1\, dx\,dy = \text{Area}(D)$.

Proof — We define function c from singular 2-cube to region D. $c(s^1, s^2) = (c^1(s^1, s^2), c^2(s^1, s^2))$. c preserves the orientation of I^2 in D. Moreover, 

the corners in I^2 are mapped to points in D (not necessarily corners). Then for instance, $c(I^2_{(1,1)}) = (c^1(1, s^2), c^2(1, s^2))$. $(s^1, s^2) \in [0,1]^2$

Assume ∂D is parametrised by $\gamma(t) = (\gamma^1(t), \gamma^2(t))$. Then $\int_\gamma P\,dx + Q\,dy = \int_{I^1} (P \circ \gamma)\, \gamma^*(dx) + (Q \circ \gamma)\, \gamma^*(dy) = \int_{I^1} P(\gamma^1(t), \gamma^2(t)) \frac{d\gamma^1}{dt}\, dt + Q(\gamma^1(t), \gamma^2(t)) \frac{d\gamma^2}{dt}\, dt$

Then $\int_{\gamma = \partial c} P\,dx + Q\,dy \overset{\text{Stokes}}{=} \int_c d(P\,dx + Q\,dy) = \int_c \frac{\partial P}{\partial x} dx \wedge dx + \frac{\partial P}{\partial y} dy \wedge dx + \frac{\partial Q}{\partial x} dx \wedge dy + \frac{\partial Q}{\partial y} dy \wedge dy = \int_c \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) dx \wedge dy \overset{\text{defn}}{=} \int_{I^2} c^*\left(\left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) dx \wedge dy\right)$

$= \int_{I^2} \left[-\frac{\partial P}{\partial y}(c^1(s^1, s^2), c^2(s^1, s^2)) + \frac{\partial Q}{\partial x}(c^1(s^1, s^2), c^2(s^1, s^2))\right] c^*(dx) \wedge c^*(dy) = \int_{I^2} \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)\left(\frac{\partial c^1}{\partial s^1} ds^1 + \frac{\partial c^1}{\partial s^2} ds^2\right) \wedge \left(\frac{\partial c^2}{\partial s^1} ds^1 + \frac{\partial c^2}{\partial s^2} ds^2\right)$

$= \int_{I^2} \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)(c^1(s^1, s^2), c^2(s^1, s^2))\left[\left(\frac{\partial c^1}{\partial s^1} \cdot \frac{\partial c^2}{\partial s^2}\right) ds^1 \wedge ds^2 + \left(\frac{\partial c^1}{\partial s^2} \frac{\partial c^2}{\partial s^1}\right) ds^2 \wedge ds^1\right] = \int_{I^2} \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)(c(s^1, s^2))\left(\frac{\partial c^1}{\partial s^1} \frac{\partial c^2}{\partial s^2} - \frac{\partial c^1}{\partial s^2} \frac{\partial c^2}{\partial s^1}\right) ds^1 \wedge ds^2 = \int_{I^2} \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)(c(s^1, s^2)) \det c'(s^1, s^2)\, ds^1 \wedge ds^2$

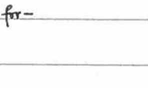
Recall the change of variables formula for double integrals: $\iint_{f(D)} g = \iint_D g \circ f\, |\det f'|$. Then our expression above gives us:

$= \int_0^1 \int_0^1 \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)(c(s^1, s^2)) \det c'(s^1, s^2)\, ds^1\, ds^2 \overset{\text{change of variables}}{=} \iint_{c(I^2)} \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)(x, y)\, dx\, dy$, which is acceptable as absolute value doesn't matter: $\det c' \geq 0$ by preservation of orientation // q.e.d.

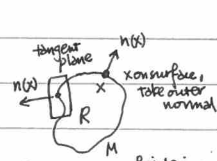
11 December 2013
Dr. Yiannis PETRIDIS
Maths 706.

Due to a lack of time, we have not been able to establish several results rigourously:

- change of basis formula
- $\int_C f = \int_{[a,b] \times [c,d]} f \cdot \mathbb{1}_C$



There will be several other aspects from here on where we will have to sacrifice rigour for —

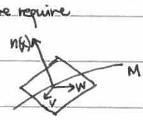


Consider an orientable smooth surface M, which is a manifold. For any $v, w \in \mathbb{R}^3_x$ (both vectors start from x). Then define $\omega(v, w) = \det\begin{pmatrix} v \\ w \\ n(x) \end{pmatrix}$.

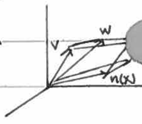
Then $\omega(w, v) = -\omega(v, w)$ and $\omega(cv_1 + v_2, w) = c\,\omega(v_1, w) + \omega(v_2, w)$. Then this is an alternating 2-tensor at x on $\mathbb{R}^3_x$: To ensure that ω is non-zero, we require

$\{v, w, n(x)\}$ to form a LI set: pick v, w on the tangent plane to be non-collinear, then $\det\begin{pmatrix} v \\ w \\ n(x) \end{pmatrix} \neq 0$, as $\{v, w, n(x)\}$ is LI.

We call $\omega(v, w) = dA$ the volume element at x.



Moreover, if we take $v=\begin{pmatrix}v^1\\v^2\\v^3\end{pmatrix}$, $w=\begin{pmatrix}w^1\\w^2\\w^3\end{pmatrix}$ and $n(x)=\begin{pmatrix}n^1\\n^2\\n^3\end{pmatrix}$. Then $\omega(v,w)=\begin{vmatrix}v^1 & v^2 & v^3\\ w^1 & w^2 & w^3\\ n^1 & n^2 & n^3\end{vmatrix} = \langle v\times w, n(x)\rangle$, the scalar cross product. To see this, we can expand each term to see that both equal $n^1(v^2w^3-v^3w^2)-n^2(v^1w^3-v^3w^1)+n^3(v^1w^2-v^2w^1)$. Recall that this gives the area of a parallelepiped subtended by edges v, w and $n(x)$, with the appropriate change of sign.



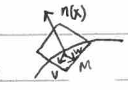
$\boxed{\text{Theorem}}$ $\omega(v,w) = dA(v,w) = n^1\, dy\wedge dz + n^2\, dz\wedge dx + n^3\, dx\wedge dy$.

Proof – $(n^1\, dy\wedge dz)(v,w) = n^1\,(dy\otimes dz - dz\otimes dy)(v,w) = n^1\,(dy(v)\cdot dz(w) - dz(v)\,dy(w)) = n^1(v^2w^3-v^3w^2)$, which is the matching first term in expansion of $\omega(v,w)$.

Likewise, $n^2(dz\wedge dx)(v,w) = n^2\,(dz\otimes dx\,(v,w) - dx\wedge dz(v,w)) = n^2(dz(v)\,dx(w) - dx(v)\,dz(w)) = n^2(v^3w^1 - v^1w^3)$. Repeat for third term, q.e.d.

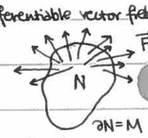
$\boxed{\text{Theorem}}$ Let v, w be on the tangent plane at x for M: then ① $n^1 dA = dy\wedge dz$, $n^1\, dA(v,w) = (dy\wedge dz)(v,w)$. ② $n^2 dA = dz\wedge dx$ and ③ $n^3 dA = dx\wedge dy$ with similar implications.

Proof – We just prove ①, then ②,③ are analogous. Previous calculations showed that $(dy\wedge dz)(v,w) = v^2w^3 - v^3w^2$. Since v, w are perpendicular to $v\times w$, $v\times w$ is parallel to $n(x) = (n^1, n^2, n^3)$. So $v\times w = a\,n(x)$ for $a\in\mathbb{R}$, $\langle v\times w, n(x)\rangle = \langle v\times w, n^1\hat{i} + n^2\hat{j} + n^3\hat{k}\rangle$. Then also, we have that $\langle v\times w, n(x)\rangle = \langle a\,n(x), n(x)\rangle = a\underbrace{\langle n(x), n(x)\rangle}_{\text{unit length}} = a$. Thus $n^1 dA\,(v,w) = n^1\cdot a$. So $\langle v\times w, \hat{i}\rangle = \begin{vmatrix}v^1 & v^2 & v^3\\ w^1 & w^2 & w^3\\ 1 & 0 & 0\end{vmatrix} = v^2w^3 - w^2v^3 = dy\wedge dz\,(v,w)$. Also, $\langle v\times w, \hat{i}\rangle = \langle a\,n(x), \hat{i}\rangle = a\cdot n^1 = n^1\, dA(v,w)$, q.e.d.



$\boxed{\text{Theorem}}$ (Gauss / Divergence Theorem).

Let N be an orientable, compact 3-dimensional manifold in $\mathbb{R}^3$ (solid), which is bounded with boundary ∂N, the 2-dimensional surface $M = \partial N$. Let $\vec{F}$ be a differentiable vector field on an open set containing N and ∂N. Then $\iiint_N (\operatorname{div}\vec{F})\, dx\, dy\, dz = \oint_{\partial N} \langle \vec{F}, \vec{n}(x)\rangle\, dA$.



Note – This equates a triple integral with a surface integral.

Proof – Define $F = (F^1, F^2, F^3)$, $\omega = F^1\, dy\wedge dz + F^2\, dz\wedge dx + F^3\, dx\wedge dy$ a 2-form: $d\omega = \frac{\partial F^1}{\partial x} dx\wedge dy\wedge dz + \frac{\partial F^2}{\partial y} dy\wedge dz\wedge dx + \frac{\partial F^3}{\partial z} dz\wedge dx\wedge dy \Rightarrow$

$d\omega = \left(\frac{\partial F^1}{\partial x} + \frac{\partial F^2}{\partial y} + \frac{\partial F^3}{\partial z}\right) dx\wedge dy\wedge dz = (\operatorname{div}\vec{F})\, dx\wedge dy\wedge dz$. Then $\int_N d\omega = \iiint_N (\operatorname{div}\vec{F})\, dx\, dy\, dz$. [Provided N is treated as a standard k-cube: more to follow].

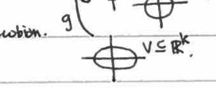
For RHS, $\int_N d\omega \overset{\text{stokes}}{=} \int_{\partial N} \omega = \int_{\partial N} F^1\, dy\wedge dx + F^2\, dz\wedge dx + F^3\, dx\wedge dy = \int_{\partial N} F^1 n^1\, dA + F^2 n^2\, dA + F^3 n^3\, dA = \int_{\partial N} \langle \vec{F}, \vec{n}(x)\rangle\, dA$, q.e.d.

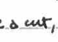
Remark – As seen, there are still a few gaps we need to fix to justify our arguments.

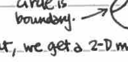
We return now to discussing manifolds. Use condition (C): using a chart/parametrisation. Recall this stated that M is a k-dimensional manifold in $\mathbb{R}^n$ if $\forall x\in M$, (C) holds, where (C) implies that $\exists\, U$ open in $\mathbb{R}^n$, W open in $\mathbb{R}^k$, $x\in U$, then $\exists\, f: W\to U$ differentiable s.t. (a) $f(W) = M\cap U$, (b) $f'(y)$ has rank k for all $y\in W$, and (c) $f^{-1}: f(W)\to W$ is continuous.

However, this chart/map f is not unique. We could also have another chart g s.t. $f^{-1}\circ g: V\to W$, $g^{-1}\circ f: W\to V$ are differentiable with nonzero Jacobian.

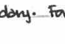
In this sense, f and g should be chosen to be "compatible".

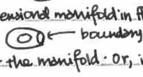






We also examine the concept of manifolds with boundary. For instance, consider the 2-sphere, which is a 2-dimensional manifold in $\mathbb{R}^3$. If we take a cut, we get a 2-D manifold with boundary.





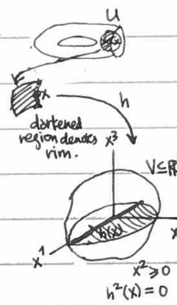
Or consider the hollow doughnut – like a torus. If we make a cut, then the cut becomes a boundary for the manifold. Or, if we consider the solid (rather than hollow) torus, it already has a boundary surface to start with. Formally, we have:

$\boxed{\text{Definition}}$ A set $M\subseteq\mathbb{R}^n$ is a k-dimensional manifold with boundary if, given $x\in M$ either condition (M) holds for x, or (exclusive) condition (M') below holds:

(M) $\exists\, U, V$ open in $\mathbb{R}^n$, $x\in U$, $\exists$ diffeomorphism $h: U\to V$ s.t. $h(M\cap U) = \{y\in V,\ y^{k+1} = y^{k+2} = \dots = y^n = 0,\ y^k \geq 0\}$, and $h^k(x) = 0$.

Example – Consider the hollow torus, with an open ball $U\subseteq\mathbb{R}^3$ near rim. Then for $U\cap M$, we get a curved surface containing the rim on a side. The diffeomorphism h flattens out the surface to a region in x^1x^2-plane. Here, the rim corresponds to the line $x^2=0$, with $h(x)$ lying along it. Then the part where $x^2\geq 0$ corresponds to the original surface, whereas $x<0$ is not included – if torus were whole, this would reflect the (now empty) space away from the rim that has been cut away.

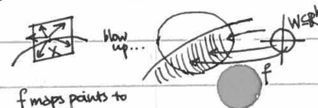
We also define $\partial M = \{x\in M \mid \text{condition (M') holds for } x\}$.



$\boxed{\text{Theorem}}$ (General version of Stoke's Theorem)

For a k-dimensional manifold with corresponding $k-1$ dimensional manifold ∂M in $\mathbb{R}^n$, if ω is a $k-1$ form on M, $\int_M d\omega = \int_{\partial M} \omega$. [i.e. for M, $y^{k+1} = \dots = y^n = 0$ and ∂M has extra condition $y^k = 0$]

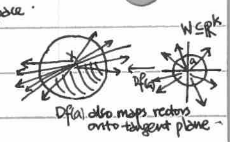
Consider a manifold M with a tangent plane. To map vectors from $W\subseteq\mathbb{R}^k$ onto tangent plane, we use f_* [i.e. $Df(a)$].



$\boxed{\text{Definition}}$ The tangent space at x of M is $f_*(\mathbb{R}^k_a)$ where $f(a) = x$, and f satisfies condition (C).

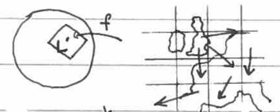
Notation – let $M_x = f_*(\mathbb{R}^k_a)$ be the tangent space of M at x. Then $(a, v) \xrightarrow{f_*} (x, Df(a)(v))\ \forall\, v\in\mathbb{R}^k$.

If M is a k-dimensional manifold in $\mathbb{R}^n$, we expect that M_x is a k-dimensional vector space. $f_*: \mathbb{R}^k_a \to M_x$ is a linear map, because we used the linear transformation $Df(a): \underset{(\dim k)}{\mathbb{R}^k} \to \underset{(\dim n)}{\mathbb{R}^n}$. $M_x = f_*(\mathbb{R}^k_a)$. $\dim f_*(\mathbb{R}^k_a) = \operatorname{rank} Df(a) \overset{(C)}{=} k$, so $M_x \subseteq \mathbb{R}^n_x$.

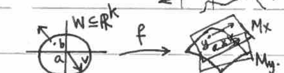


Imagine the earth, with a tangent plane at London (point L). L is on the surface of the sphere (with a set distance from centre of earth C). Now imagine Dr Petridis wants to go home, to the North Pole (point N). Then there is another tangent plane there. Suppose we want to measure wind at these points, these translate to an arrow on each of these tangent planes. Taken together, wind at each point of the earth gives a set of vectors defined on tangent planes, giving a vector field on manifold.

$\boxed{\text{Definition}}$ A vector field F on a manifold M is a function F on M s.t. $\forall x \in M$, $F(x) \in M_x$.



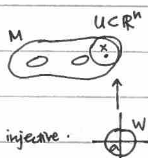
Just like in weather reports, we get a 2D plan of the world to approximate a 3D sphere, we need a chart/map f to interpret data on the earth (manifold) itself.



Let $f: W \to \mathbb{R}^n$ be a chart around $x \in M$ s.t. $f(W) = U \cap W$. Suppose we have a point $a \in W$. Now, we move to another point $b \in W$.

Let F be a vector field on $f(W)$, and let $G(b) = (f^{-1})_* F(f(b))$. Since f_* has trivial kernel, by rank-nullity theorem, $\mathbb{R}^k_a \longleftrightarrow M_x$ is a bijection, $\mathbb{R}^k_b \xleftrightarrow{\text{bijection}} M_y$ etc. This way we determine a vector field G on W s.t. $f_*(G) = F$ on $M \cap U$. If G is a continuous (resp differentiable) vector field on W, then we say that F is a continuous (resp differentiable) vector field on $M \cap U$.

$f(a) = x$, $f(b) = y$, $M_y = f_*(\mathbb{R}^k_b)$.



13 December 2013
Dr. Yiannis PETRIDIS.
Maths 706.

Let M be a k-dimensional manifold, $U \subset \mathbb{R}^n$ open set. $f(W) = U \cap M$. $f'(y)$ has rank k. $\forall y \in W$, $Df(y): \mathbb{R}^k \to \mathbb{R}^n$

By rank-nullity theorem, $\dim \mathbb{R}^k = k = \dim \operatorname{Ker} Df(y) + \operatorname{rank} Df(y)$ $\Rightarrow \dim \operatorname{Ker} Df(y) = 0$, $\operatorname{Ker} Df(y) = \{0\} \subseteq \mathbb{R}^k \Rightarrow Df(y)$ injective.

$Df(y): \mathbb{R}^k \xrightarrow{\text{bijection}} Df(y)(\mathbb{R}^k)$. Let f_* be the push forward, then if $(v, a) = v_a \in \mathbb{R}^k_a$, $f_*(v_a) = w_x = (x, Df(a)(v))$. $f_*: \mathbb{R}^k_a \to \mathbb{R}^n_x$, $f_*(\mathbb{R}^k_a) = M_x$, tangent space of M at x. If we choose $\forall x \in M$ a vector $F(x) \in M_x$, then we have a vector field on M.

Since $f_*: \mathbb{R}^k_y \to M_{f(y)}$ $(y \in W)$ is an isomorphism, then $F(f(y)) \in M_{f(y)}$ gives vector $G(y) \in \mathbb{R}^k_y$, $f_*(G(y)) = F(f(y))$, $G(y) = (f^{-1})_*(F(f(y)))$.

If G is continuous, then we say F is continuous. If G is differentiable, then we say F is differentiable.

$\boxed{\text{Definition}}$ A function ω on M such that $\forall x \in M$, $\omega(x) \in \Lambda^m(M_x)$ is called a <u>m-form on M</u>.

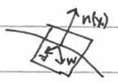
To understand continuous (respectively differentiable) m-forms on M, we pullback them to W. If $f^*(\omega)$ is a continuous (resp. differentiable) m-form on $W \subseteq \mathbb{R}^k$, then we say ω is a continuous (resp. differentiable) m-form on M. $\omega(x) = \sum_{i_1 < \dots < i_m} \omega_{i_1 \cdots i_m}(x)\, dx^{i_1} \wedge \dots \wedge dx^{i_m}$. We can no longer apply exterior derivatives here as they make no sense.

$\boxed{\text{Theorem}}$ Given ω a differential m form on M, there exists a unique differential $m+1$ form on M called $d\omega$ s.t. $f^*(d\omega) = d(f^*\omega)$.

$d\omega(x) \in \Lambda^{m+1}(M_x)$, $d\omega(x) \cdot (v_1, v_2, \dots, v_m, v_{m+1})$, $v_1, \dots, v_{m+1} \in M_x$. $\exists$ unique vectors $w_1, \dots, w_{m+1}$ s.t. $f_*(w_j) = v_j$

Then $d\omega(x)(v_1, \dots, v_{m+1}) = d(f^*\omega)(a)(w_1, w_2, \dots, w_{m+1})$.

Remark – Need to prove uniqueness.



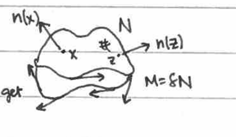
Stokes' Theorem for manifolds becomes $\int_M d\omega = \int_{\partial M} \omega$ where ω is a $k-1$ form on M (which is a k-dimensional manifold).

Remember for instance, if M is a 2-dimensional surface on $\mathbb{R}^3$, for any $x \in M$, $\dim M_x = 2$, $\dim \Lambda^2(M_x) = \binom{2}{2} = 1$. $dA(v, w) = \det\begin{pmatrix} v \\ w \\ n(x) \end{pmatrix} \in \Lambda^2(M_x)$, which is non-zero.

dA is called a <u>surface element</u>.

On a curve, let M be a 1-dimensional manifold on $\mathbb{R}^3$. Then $\dim \Lambda^1(M_x) = \binom{1}{1} = 1$. This is the <u>arclength element</u> $ds(x) \in \Lambda^1(M_x) = J^1(M_x)$.

$\boxed{\text{Theorem}}$ (Stokes's Theorem, classical version) — cont'd.



Let N be a 2-dimensional orientable manifold in $\mathbb{R}^3$ with boundary $M = \partial N$ (1-dimensional manifold) with an induced orientation [i.e. compatible with surface]. Find a vector $T \in M_x$ with $ds(T) = 1$ i.e. $ds(x)(T) = 1$. Let F be a vector field on an open set containing N. Then we get

$$\underbrace{\iint_N (\operatorname{curl} F) \cdot \vec{n}\, dA}_{\text{[surface integral]}} = \underbrace{\int_{\partial N = M} \vec{F} \cdot T\, ds}_{\text{[line integral]}}.$$

Proof – Since RHS is a line integral, we need a 1-form from $\vec{F} = (F^1, F^2, F^3)$. Then let $\omega = F^1 dx + F^2 dy + F^3 dy$. Then taking exterior derivative, $d\omega = G^1 dy \wedge dz + G^2 dz \wedge dx + G^3 dx \wedge dy$ with $(G^1, G^2, G^3) = \operatorname{curl} \vec{F}$. Moreover, $d\omega = G^1 dy \wedge dz + G^2 dz \wedge dx + G^3 dx \wedge dy = G^1 n^1 dA + G^2 n^2 dA + G^3 n^3 dA = (G^1 n^1 + G^2 n^2 + G^3 n^3)\, dA = (\vec{G} \cdot \vec{n})\, dA$. $n(x) = (n^1, n^2, n^3)$ is external normal. Then we have arclength elements ds s.t. $ds(T) = 1$. $\vec{F} \cdot \vec{T}\, ds = (F^1, F^2, F^3)(T^1, T^2, T^3)\, ds = (F^1 T^1 + F^2 T^2 + F^3 T^3)\, ds$

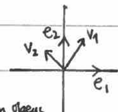
$= F^1 T^1 ds + F^2 T^2 ds + F^3 T^3 ds$, $\omega = F^1 dx + F^2 dy + F^3 dz$. We guess that $dx = T^1 ds$, $dy = T^2 ds$, $dz = T^3 ds$. We check, and this is true (exercise) // q.e.d.

Finally, we examine the concept of orientability. To do this, we return to some basics from linear algebra.

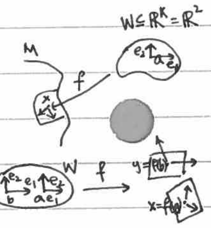
Let $\mathcal{B}$ be an ordered basis of a vector space V, e.g. ξ standard basis of $\mathbb{R}^n$, $\xi = \langle e_1, e_2, \dots, e_n \rangle$. Let $\mathcal{F}$ be an ordered basis of V. $V \xrightarrow{\text{Id}} V$, $M = [\text{Id}]_{\mathcal{B}}^{\mathcal{F}}$

If $\det M > 0$, we say $\mathcal{B}, \mathcal{F}$ define the same orientation. We denote $\mathcal{F} \sim \mathcal{B}$ to mean they have the same orientation.

Here, $\sim$ is an equivalence relation.



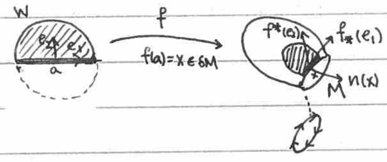
(translation obeys right hand rule)
Note that $\mathcal{F}' = \langle v_2, v_1 \rangle$ has opposite orientation.

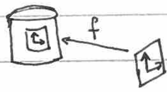


For a manifold M, vectors in tangent space M_x are in 1-to-1 correspondence with vectors in W. let $W \subseteq \mathbb{R}^2$ with ordered standard basis $\langle e_1, e_2 \rangle$, then $\langle f_*(e_1), f_*(e_2) \rangle$ is an ordered basis for M_x. We can generalise this to higher dimensions. This works for a point $a \in W$, but there is no reason to stick to a — we can consider also $b \in W$.

We then push forward again accordingly. Then $\langle f_*((e_1)_b), f_*((e_2)_b) \rangle$ is an ordered basis at $f(b)$, i.e. ordered basis for $M_{f(b)=y}$.

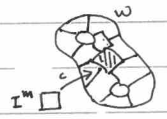
If we can do this is a continuous and consistent way around the manifold, it is considered to be orientable.





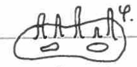
For a 2-dim manifold with boundary, we plot the chart accordingly:

Once we have an ordered basis on W, this "induces" a natural orientation for traversing the boundary of M by calculating $f_*(e_1)$ as the tangent.



How can we integrate on a manifold? let ω be a m-form on M, c be a singular m-cube on M, $c: I^m \to M$. ω vanishes $= 0$ on $M \setminus c(I^m)$.

Then by definition, $\int_c \omega = \int_{I^m} c^*(\omega)$. If the form cannot be contained, we can slice up the manifold into smaller units and sum results.



We could also use a more advanced intuition — by the partitions of unity. Take a collection of functions $\Phi = \{\varphi : 0 \le \varphi(x) \le 1, \sum_{\varphi \in \Phi} \varphi(x) = 1 \ \forall x \in M\}$.

Then $\int_M \omega = \int_M \sum_{\varphi \in \Phi} \varphi(x)\, \omega(x) \overset{\text{def}}{=} \sum_{\varphi \in \Phi} \int_M \varphi \omega$. This is a bit more advanced and will not be examined, but leads us to a better, more rigourous understanding of stokes' theorem.

END OF SYLLABUS.

Consider singular k-chains, $c = \sum_{l=1}^{m} a_l c_l$, $a_l \in \mathbb{Z}$, c_k is a singular k-cube. $\partial(c_k) = c(\partial I^k)$, then $\partial c = \sum_{l=1}^{m} a_l \partial(c_l)$ is a singular k-1 chain. let C_{k-1} be the $\mathbb{Z}$-module of k-chains: $C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}} C_{k-2} \to \cdots \xrightarrow{\partial_1} C_0$, then $\partial_k \circ \partial_{k-1} \circ \cdots \circ \partial_1 = 0$, then we can define the k-homology group $H_k(e, \mathbb{Z}) = \dfrac{\operatorname{Ker} \partial_k}{\operatorname{Im} \partial_{k+1}}$

let Ω^k be the vector space of differential k-forms on M. Then $\Omega^k \xrightarrow{d_k} \Omega^{k+1} \xrightarrow{d_{k+1}} \Omega^{k+2} \xrightarrow{d_{k+2}} \Omega^{k+3} \cdots$ by exterior derivatives. Then $d_{k+1} d_k = 0$. Then the k-cohomology is $H^k(M, \mathbb{R}) = \dfrac{\operatorname{Ker} d_k}{\operatorname{Im} d_{k-1}}$. Then we can have $H_k \times H^k \to \mathbb{R}$, $(c, \omega) \mapsto \int_c \omega$, which is the Poincaré duality.

END OF COURSE.

3109-18.