Miscellaneous Notes on Differentiable Manifolds

Simon Xiang

February 3, 2021

These notes cover a variety of topics related to or required for the study of smooth manifolds, taken over Winter break 2020-2021 in preparation for my geometry overload next semester. Source files: `https://git.simonxiang.xyz/math_notes/files.html`

Contents

Part I
Vector Calculus

Lecture 1

Vector Calculus Fundamentals

Here we review some basics, and cover other stuff that should be taught in a standard multivariable calculus course, but wasn't at UNT (div, grad, curl).

1.1 Tangent planes

Definition 1.1 (Tangent plane)**.** Let $z = f(x, y)$ represent some surface S in $\mathbb{R}^3$, $P = (a, b, c)$, $Q = (x, y, z)$ be points on S, and T a plane containing P. If the angle between the vector $\overrightarrow{PQ}$ and the plane T approaches zero as $Q \to P$ along S, then T is the **tangent plane** to S at P.

Since two lines determine a plane, the tangent lines from the partial derivatives will be in the tangent plane, if it exists: lines may determine the plane, but not the existence of it. However, if $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in nbd of (a, b) and are continuous at (a, b), then the tangent plane $z = f(x, y)$ at $(a, b, f(a, b))$ also exists. An equation for T is given by

$$A(x-a) + B(y-b) + C(z - f(a, b)) = 0$$

where $\mathbf{n} = (A, B, C)$ is a normal vector to T. Since T contains the tangent lines L_x and L_y, we just need vectors $\mathbf{v}_x$ and $\mathbf{v}_y$ parallel to L_x and L_y respectively, and then set $\mathbf{n} = \mathbf{v}_x \times \mathbf{v}_y$. Now the slope of L_x is $\frac{\partial f}{\partial x}(a, b)$, so $\mathbf{v}_x = (1, 0, \frac{\partial f}{\partial x}(a, b))$ is parallel to L_x. Similarly, $\mathbf{v}_y = (0, 1, \frac{\partial f}{\partial y}(a, b))$ is parallel to L_y, so the normal vector to T is given by

$$\mathbf{n} = \mathbf{v}_x \times \mathbf{v}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x}(a, b) \\ 0 & 1 & \frac{\partial f}{\partial y}(a, b) \end{vmatrix} = -\frac{\partial f}{\partial x}(a, b)\mathbf{i} - \frac{\partial f}{\partial y}(a, b)\mathbf{j} + \mathbf{k}.$$

So T can be represented by the equation $-\frac{\partial f}{\partial x}(a, b)(x-a) = \frac{\partial f}{\partial y}(a, b)(y-b) + z - f(a, b) = 0$, which simplifies to

$$\frac{\partial f}{\partial x}(a, b)(x-a) + \frac{\partial f}{\partial y}(a, b)(y-b) - z + f(a, b) = 0.$$

In general, if the surface is defined by an equation of the form $F(x, y, z) = 0$, then the tangent plane at (a, b, c) is

$$F_x(a, b, c)(x-a) + F_y(a, b, c)(y-b) + F_z(a, b, c)(z-c) = 0.$$

Our previous formula was just this applied to the case where $F(x, y, z) = f(x, y) - z$.

Example 1.1. To find the tangent plane to the surface $z = x^2 + y^2$ at $(1, 2, 5)$, note that for $f(x, y) = x^2 + y^2$, we have $f_x = 2x$ and $f_y = 2y$, so the equation is just $2(1)(x-1) + 2(2)(y-2) - z + (1^2 + 2^2) = 0$, or $2x + 4y - z - 5 = 0$.

Example 1.2. For the surface $x^2 + y^2 + z^2 = 9$, we have $F(x, y, z) = x^2 + y^2 + z^2 - 9$, so $F_x = 2x$, $F_y = 2y$, and $F_z = 2z$. Therefore the equation for the tangent plane is $2x + 2y - z - 9 = 0$.

1.2 Directional derivatives and the gradient

Definition 1.2 (Directional derivative)**.** Let $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}^2$, and $(a, b) \in D$. Let $\mathbf{v} \in \mathbb{R}^2$ be a unit vector ($\|v\| = 1$). Then the **directional derivative** of f at (a, b) in the direction of $\mathbf{v}$, denoted $D_v f(a, b)$, is defined as

$$D_v f(a, b) = \lim_{h \to 0} \frac{f((a, b) + h\mathbf{v}) - f(a, b)}{h}.$$

Note that if we write $\mathbf{v} = (v_1, v_2)$, then this expression becomes $\lim_{h \to 0}(f(a + hv_1, b + hv_2) - f(a, b))/h$.

Note that the partial derivatives f_x and f_y are just the cases where $\mathbf{v} = \mathbf{i} = (1,0)$, etc. We can express this by saying $f_x = D_i f$ and $f_y = D_j f$.

Theorem 1.1. *Let $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^2$ such that f_x and f_y exists and are continuous on D. Let $(a,b) \in D$, and $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ be a unit vector. Then*

$$D_v f = v_1 f_x + v_2 f_y.$$

Proof. The proof is annoying so it has been skipped. ⊠

Definition 1.3 (Gradient)**.** For $f : \mathbb{R}^2 \to \mathbb{R}$, the **gradient** of f denoted $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$ is the vector

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

In general, if $f : \mathbb{R}^n \to \mathbb{R}$, we have $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)$$

for $(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$.

Note that $D_v f = \mathbf{v} \cdot \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$, which is the same as saying $D_v f = \mathbf{v} \cdot \nabla f$.

Example 1.3. To find $D_v f(1,2)$ where $f : (x,y) \mapsto xy^2 + x^3 y$ and $\mathbf{v} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$, note that $\nabla f = (y^2 + 3x^2 y, 2xy + x^3)$, so

$$D_v f(1,2) = \mathbf{v} \cdot \nabla f(1,2) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \cdot (10,5) = \frac{15}{\sqrt{2}}.$$

If $f : \mathbb{R}^2 \to \mathbb{R}$ has continuous partial derivatives f_x and f_y, then f is **continuously differentiable.** Say f is continuously differentiable with $\nabla f \neq 0$, $c \in \operatorname{im} f$, and $\mathbf{v} \in \mathbb{R}^2$ be a unit vector tangent to the contour $f(x,y) = c$. Since $\mathbf{v}$ is tangent to the constant contour, the rate of change in the direction of $\mathbf{v}$ is zero, or $D_v f = 0$. We also know $D_v f = \mathbf{v} \cdot \nabla f = \|\mathbf{v}\| \|\nabla f\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and ∇f. Since $\|\mathbf{v}\| = 1$, $D_v = \|\nabla f\| \cos\theta = 0$, and since ∇f is nonzero, $\cos\theta = 0$, and therefore $\theta = 90°$. We conclude that $\nabla f \perp \mathbf{v}$, which says that ∇f is *normal* to the contour.

In general, for a unit vector $\mathbf{v} \in \mathbb{R}^2$ we have $D_v f = \|\nabla f\| \cos\theta$. At a point (x,y) the length $\|\nabla f\|$ is fixed, and $D_v f$ varies with θ. The maximum value of $D_v f$ is when $\theta = 0$ such that $\cos\theta = 1$, and the smallest value is when $\theta = \pi$ such that $\cos\theta = -1$. So f increases the fastest in the direction of ∇f (this is the case $\theta = 0$) and slowest in the direction of $-\nabla f$. We can formulate our findings as a theorem.

Theorem 1.2. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable, and $\nabla f \neq 0$. Then*

(a) The gradient ∇f is normal to any level curve $f(x,y) = c$.

(b) The value of f increases the fastest in the direction of ∇f.

(c) The value of f decreases the fastest in the direction of $-\nabla f$.

Example 1.4. In which direction does $f : (x,y) \mapsto xy^2 + x^3 y$ increase the fastest from the point $(1,2)$? What about the fastest rate of decrease?

Solution. We have $\nabla f = (y^2 + 3x^2 y, 2xy + x^3)$, so $\nabla f(1,2) = (10,5) \neq 0$. So a unit vector in that direction is $\mathbf{v} = \frac{\nabla f}{\|\nabla f\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right)$, similarly $\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}} \right)$ in the direction of $-\nabla f$. You can fill in the rest. ■

1.3 Line integrals

Here we review the concept of a line integral.

Definition 1.4 (Line integral)**.** For a function $f : \mathbb{R}^2 \to \mathbb{R}$ and a curve $C \subseteq \mathbb{R}^2$ parametrized by $x = x(t), y = y(t), a \le t \le b$, the **line integral** of $f(x,y)$ along C is

$$\int_C f(x,y)\,ds = \int_a^b f(x(t),y(t))\sqrt{x'(t)^2 + y'(t)^2}\,dt,$$

where $s = s(t) = \int_a^t \sqrt{x'(u)^2 + y'(u)^2}\,du$ denotes the arc length of the curve. So $ds = s'(t)dt = \sqrt{x'(t)^2 + y'(t)^2}dt$ by the FTC.

Some basic things: traversing the curve in the opposite direction doesn't change anything. We can also define a line integral with respect to x as opposed to s, where $\int_C f(x,y)\,ds = \int_a^b f(x(t),y(t))x'(t)\,dt$. For the physically inclined, you can think of the line integral as work done by a force moving along a curve.

Some of these constructions seem similar: we can work toward generalizing this. Define a function $f : \mathbb{R}^2 \to \mathbb{R}^2$ by $f : (x,y) \mapsto P\mathbf{i} + Q\mathbf{j}$ where $P, Q : \mathbb{R}^2 \to \mathbb{R}$, then f is a **vector field** on $\mathbb{R}^2$. This function takes in points and outputs vectors. For a curve C with parametrization $x = x(t), y = y(t), a \le t \le b$, let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be the position vector for a point $(x(t), y(t))$ on C. Then $\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$ and so

$$\begin{aligned}\int_C P(x,y)\,dx + \int_C Q(x,y)\,dy &= \int_a^b P(x(t),y(t))x'(t)\,dt + \int_a^b Q(x(t),y(t))y'(t)\,dt \\ &= \int_a^b \big(P(x(t),y(t))x'(t) + Q(x(t),y(t))y'(t)\big)\,dt \\ &= \int_a^b \mathbf{f}(x(t),y(t))\cdot \mathbf{r}'(t)\,dt.\end{aligned}$$

This leads us to the following definition.

Definition 1.5. For a vector field $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$ and a curve C with parametrization $y = y(t), a \le t \le b$, the **line integral** of f along C is

$$\int_C \mathbf{f}\cdot d\mathbf{r} = \int_C P(x,y)\,dx + \int_C Q(x,y)\,dy = \int_a^b \mathbf{f}(x(t),y(t))\cdot \mathbf{r}'(t)\,dt,$$

where $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ is the position vector for points on C.

We distinguish Definition 1.4 and Definition 1.5 by calling one the line integral of a *scalar field* and the other the line integral of a *vector field*. Which one corresponds to which should be clear from context. We use the notation $d\mathbf{f} = \mathbf{r}'(t)dt = dx\mathbf{i} + dy\mathbf{j}$ to denote the **differential** of $\mathbf{r}$. Often we denote $\int_C P(x,y)\,dx + \int_C Q(x,y)\,dy$ by $\int_C P(x,y)\,dx + Q(x,y)\,dy$ for convenience. $P(x,y)dx + Q(x,y)dy$ is known as a **differential form**. For $F : \mathbb{R}^2 \to \mathbb{R}$, the **differential** of F is $dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy$. A form is **exact** if it's the form of some function F.

Recall that $\mathbf{r}'(t)$ is a tangent vector to points on C in the direction of C. C is smooth, therefore $\mathbf{r}'(t) \neq 0$ on $[a,b]$ and so the unit tangent vector to C at a point is given by $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$. This naturally leads us to the following theorem:

Theorem 1.3. *For a vector field* $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$ *and a smooth curve* C *parametrized on* $[a,b]$ *with position vector* $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$*, we have*

$$\int_C \mathbf{f}\cdot d\mathbf{r} = \int_C \mathbf{f}\cdot \mathbf{T}ds,$$

where $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$.

This also works for piecewise smooth curves. If $C = C_1 \cup C_2 \cup \cdots \cup C_n$, then

$$\int_C \mathbf{r} \cdot d\mathbf{r} = \int_{C_1} \mathbf{f} \cdot d\mathbf{r}_1 + \int_{C_2} \mathbf{f} \cdot d\mathbf{r}_2 + \cdots + \int_{C_n} \mathbf{f} \cdot d\mathbf{r}_n.$$

Example 1.5. Evaluate $\int_C (x^2 + y^2)\,dx + 2xy\,dy$ on the curves $x = t,\ y = 2t$ and $x = t,\ y = 2t^2$ for $t \in [0, 1]$.

Solution. For the first curve, note that $x'(t) = 1$ and $y'(t) = 2$, so

$$\int_C (x^2 + y^2)\,dx + 2xy\,dy = \int_0^1 (t^2 + 4t^2)x'(t) + 2t(2t)y'(t)\,dt = \int_0^1 5t^2 + 8t^2\,dt = \frac{13}{3}.$$

Similarly, for the second curve $x'(t) = 1$ and $y'(t) = 4t$, and so

$$\int_C (x^2 + y^2)\,dx + 2xy\,dy = \int_0^1 (t^2 + 4t^4) + (2t \cdot 2t^2 \cdot 4t)\,dt = \int_0^1 t^2 + 20t^4\,dt = \frac{13}{3}.$$

■

1.4 Closed curves and conservative vector fields

Recall that $\int_C f(x, y)\,ds = \int_{-C} f(x, y)\,ds$ for line integrals of scalar fields, but for vector fields this does not hold, namely, $\int_{-C} \mathbf{f} \cdot dr = -\int_C \mathbf{f} \cdot dr$. Recall that our definition of line integrals depends on the parametrization of the curve: what if we parametrize C by some alternative parametrization? Then our definition would not be well defined, and this would be very bad. Thankfully, this is not the case, as long as the orientation of C is invariant under parametrization.

Theorem 1.4. *Let $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$ be a vector field and C be a smooth curve parametrized by $x = x(t), y = y(t)$ for $t \in [a, b]$. Say $t = \alpha(u)$ for $u \in [c, d]$ such that $a = \alpha(c), b = \alpha(d)$, and $\alpha'(u) > 0$ on the interval (c, d). Then $\int_C \mathbf{f} \cdot d\mathbf{r}$ has the same value for the alternate parametrization $x = \tilde{x}(u) = x(\alpha(u)), y = \tilde{y}(u) = y(\alpha(u)), u \in [c, d]$.*

Proof. Not too interesting, chain rule and u-sub. ⊠

A **closed curve** is a loop, and a **simple closed curve** has no self-intersections, and we denote line integrals along closed curves with $\oint$.

Theorem 1.5. *In a region R, the line integral $\int_C \mathbf{f} \cdot d\mathbf{r}$ is independent of path between two points of R iff $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for $C \subseteq R$ a closed curve.*

Proof. Split $C = C_1 \cup -C_2$ and go from there. ⊠

This doesn't completely determine path independence, but it does relate some things. We work toward a more practical condition for independence of path.

Theorem 1.6 (Chain Rule)**.** *If $z = f(x, y)$ is a continuously differentiable function of x and y, and both $x = x(t)$ and $y = y(t)$ are differentiable functions of t, then Z is a differentiable function of t, and*

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Theorem 1.7. *Let $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$ be a vector field on some region R, where $P, Q\colon \mathbb{R}^2 \to \mathbb{R}$ are continuously differentiable. Let $C \subseteq R$ be a smooth curve with parametrization $x = x(t), y = y(t), t \in [a, b]$. Suppose we have a function $F\colon \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla F = \mathbf{f}$ on R. Then*

$$\int_C \mathbf{f} \cdot d\mathbf{r} = F(B) - F(A),$$

where $A = (x(a), y(a))$ and $B = (x(b), y(b))$ are the endpoints of C. So the line integral depends only on the endpoints.

Proof. We have

$$\begin{aligned}\int_C \mathbf{f}\cdot d\mathbf{r} &= \int_a^b \big(P(x(t),y(t))x'(t)+Q(x(t),y(t))y'(t)\big)\,dt\\ &= \int_a^b \left(\frac{\partial F}{\partial x}\frac{dx}{dt}+\frac{\partial F}{\partial y}\frac{dy}{dt}\right)dt \text{ (since } \nabla F=\mathbf{f} \implies \frac{\partial F}{\partial x}=P \text{ and } \frac{\partial F}{\partial y}=Q)\\ &= \int_a^b F'(x(t),y(t))\,dt\\ &= F(x(t),y(t))\Big|_a^b = F(B)-F(A).\end{aligned}$$

⊠

We can think of Theorem 1.7 as the line integral version of the FTC. A function $F:\mathbb{R}^2\to\mathbb{R}$ such that $\nabla F=\mathbf{f}$ is a **potential** for $\mathbf{f}$. A **conservative** vector field is one that has a potential.

Corollary 1.1. *If a vector field* $\mathbf{f}$ *has a potential in a region* R*, then* $\oint_C \mathbf{f}\cdot d\mathbf{r}=0$ *for any closed curve* $C\subseteq R$*, in other words,* $\oint_C \nabla F\cdot d\mathbf{r}=0$ *for any* $F:\mathbb{R}^2\to\mathbb{R}$.

Example 1.6. Show that the line integral $\int_C (x^2+y^2)\,dx+2xy\,dy$ is path independent.

Solution. We want to show that $\mathbf{f}$ is conservative, that is, we want to find a potential F such that

$$\frac{\partial F}{\partial x}=x^2+y^2 \quad\text{and}\quad \frac{\partial F}{\partial y}=2xy.$$

If $\frac{\partial F}{\partial x}=x^2+y^2$, then $F=\frac{1}{3}x^3+xy^2+g(y)$ for some function $g(y)$. This satisfies $\frac{\partial F}{\partial y}=2xy$ if $g'(y)=0$, so g is a constant (say $g=0$). Then a potential F exists, where $F(x,y)=\frac{1}{3}x^3+xy^2$. So $\int_C(x^2+y^2)\,dx+2xy\,dy$ is path independent. By Theorem 1.7, we can also see that any value of $\int_C \mathbf{f}\cdot dr=\frac{13}{3}$ for C from $(0,0)$ to $(1,2)$ since

$$\int_C \mathbf{f}\cdot d\mathbf{f}=F(1,2)-F(0,0)=\frac{1}{3}+4=\frac{13}{3}.$$

■

1.5 Green's Theorem

We now examine a way of evaluating line integrals on smooth vector fields (a vector field whose components P and Q are smooth) on simple closed curves.

Green's Theorem. *Let* R *be a region in* $\mathbb{R}^2$ *whose boundary is a simply closed curve* C *which is piecewise smooth. Let* $\mathbf{f}=P\mathbf{i}+Q\mathbf{j}$ *be a smooth vector field defined on both* R *and* C*. Then*

$$\oint_C \mathbf{f}\cdot d\mathbf{r}=\iint_R \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dA.$$

Proof. We prove Green's theorem for a simple region R, where $C=C_1\cup C$. We can write C in two distinct ways, one where C_1 is the curve $y=y_1(x)$ from the farthest horizontal points X_1 and X_2 (C_2 is similarly defined) and the other where C_1 is the curve $x=x_1(y)$ from the farthest vertical points Y_2 to Y_1. Integrate P around C where

C_1 is $y = y_1(x)$, then

$$\begin{aligned}\oint_C P(x,y)\,dx &= \int_{C_1} P(x,y)\,dx + \int_{C_2} P(x,y)\,dx \\ &= \int_a^b P(x,y_1(x))\,dx + \int_b^a P(x,y_2(x))\,dx \\ &= \int_a^b (P(x,y_1(x)) - P(x,y_2(x)))\,dx \\ &= -\int_a^b \left(P(x,y)\Big|_{y=y_1(x)}^{y=y_2(x)} \right) dx \\ &= -\int_a^b \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x,y)}{\partial y}\,dy\,dx \\ &= -\iint\limits_R \frac{\partial P}{\partial y}\,dA\end{aligned}$$

A similar calculation shows that $\int_C Q(x,y)\,dy = \iint_R \frac{\partial Q}{\partial x}\,dA$ by integrating along C where $C_1 = x_1(y)$. This finishes the proof. Of course, we can generalize this if we wish. ⊠

Example 1.7. To evaluate $\int_C (x^2+y^2)\,dx + 2xy\,dy$ where C is the boundary enclosed by $y = 2x$ and $y = 2x^2$, by Green's Theorem we have

$$\oint_C (x^2+y^2)\,dx + 2xy\,dy = \iint\limits_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint\limits_R (2y-2y)dA = 0.$$

Of course we already knew this, since **f** has a potential function.

Example 1.8. To see where Green's Theorem does not hold, let **f** be defined by $P = -\frac{y}{x^2+y^2}$ and $Q = \frac{x}{x^2+y^2}$ on a punctured disk homeomorphic to the annulus given by $R = \{(x,y) \mid 0 < x^2+y^2 \le 1\}$. An exercise shows that $\oint_C \mathbf{f} \cdot d\mathbf{r} = 2\pi$, but since both $\frac{\partial Q}{\partial x}$ and $\frac{\partial P}{\partial y}$ are equal to $\frac{y^2-x^2}{(x^2+y^2)^2}$ we have $\iint\limits_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0$. The key thing is that the we integrated over a region not contained in R. If we outright define R as an annulus with boundary, then this works.

As seen in the example, we can extend Green's Theorem for multiply connected closed regions, just subdivide while preserving orientation and use the fact that it still works for curves $C_1 = C_2$. Say we have a smooth potential F in R of a vector field **f**, then $\frac{\partial F}{\partial x} = P$ and $\frac{\partial F}{\partial y} = Q$, so $\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y}$ implies that $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R. Conversely, if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R, then $\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint\limits_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0$. Then for a simply connected region R, the following are equivalent:

(a) $\mathbf{f} = P\mathbf{i} + Q\mathbf{j}$ has a smooth potential $F : \mathbb{R}^2 \to \mathbb{R}$,

(b) $\int_C \mathbf{f} \cdot d\mathbf{r}$ is independent of path for curves $C \subseteq R$,

(c) $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for every simple closed curve $C \subseteq R$,

(d) $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R (in this case, the differential form $P\,dx + Q\,dy$ is exact).

1.6 Surface Integrals and the Divergence Theorem

Similar to how curves are parametrized with a variable t, we can parametrize surfaces in $\mathbb{R}^3$ with two variables u, v, in essence a continuous map $f : I^2 \to \mathbb{R}^3$. In this case, a position vector is given by $\mathbf{r}(u,v) = x(u,v) + y(u,v)\mathbf{j} + z(u,v)\mathbf{k}$

for $(u,v)\in R$ (where R is a region in $\mathbb{R}^2$). Then define the partial derivatives as $\frac{\partial \mathbf{r}}{\partial u}(u,v)=\frac{\partial x}{\partial u}(u,v)\mathbf{i}+\frac{\partial y}{\partial u}(u,v)\mathbf{j}+\frac{\partial z}{\partial u}(u,v)\mathbf{k}$ and $\frac{\partial \mathbf{r}}{\partial v}(u,v)$ analogously. Tangent vectors to points on vertical gridlines are given by $\frac{\partial \mathbf{r}}{\partial v}$, and similarly for horizontal gridlines.

Take a rectangle at a point (u,v) with width Δu and height Δv, so it has area $\Delta u\Delta v$. It gets mapped onto some surface Σ by a parametrization, with small enough Δ (say $\Delta\sigma$) is approximately the area of the rectangle. Recall that $\frac{\partial \mathbf{r}}{\partial u}\approx\frac{\mathbf{r}(u+\Delta u,v)-\mathbf{r}(u,v)}{\Delta u}$ and $\frac{\partial \mathbf{r}}{\partial v}\approx\frac{\mathbf{r}(u,v+\Delta v)-\mathbf{r}(u,v)}{\Delta v}$, then the surface area is approximately

$$\|(\mathbf{r}(u+\Delta u)-\mathbf{r}(u,v))\times(\mathbf{r}(u,v+\Delta v)-\mathbf{r}(u,v))\|\approx\left\|\left(\Delta u\frac{\partial \mathbf{r}}{\partial u}\right)\times\left(\Delta v\frac{\partial \mathbf{r}}{\partial v}\right)\right\|=\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|\Delta u\Delta v.$$

This is just taking the norm of the cross product of the two vectors that make up the rectangle, nothing special. So the surface area of a surface Σ is the sum of the $\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|\Delta u\Delta v$ over the rectangles in R, therefore

$$S=\iint\limits_R\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|du\,dv.$$

We'll notate this $\iint\limits_\Sigma d\sigma=\iint\limits_R\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|du\,dv$, which is a special case of a *surface integral* over a surface Σ.

Definition 1.6. Let Σ be a surface in $\mathbb{R}^3$ parametrized by $x=x(u,v),\ y=y(u,v),\ z=z(u,v)$ for (u,v) in some region $R\subseteq\mathbb{R}^2$. Let $\mathbf{r}(u,v)=x(u,v)\mathbf{i}+y(u,v)\mathbf{j}+z(u,v)\mathbf{k}$ be the position vector for Σ, and let f be a function on some subset of $\mathbb{R}^3$ containing Σ. Then the **surface integral** of f over Σ is

$$\iint\limits_\Sigma f\,d\sigma=\iint\limits_R f(x(u,v),y(u,v),z(u,v))\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|du\,dv.$$

In particular, the surface area S of Σ is equal to $\iint\limits_\Sigma 1\,d\sigma$.

Since $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ are tangent to the surface Σ, $\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}$ is perpendicular to the tangent plane at each point of Σ, so the surface integral of a function f can be expressed as $\iint_R f\,\|\mathbf{n}\|\,d\sigma$, where $\mathbf{n}=\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}$ is a **normal vector** to Σ. An **outward unit normal vector** points away from the "top" part of the surface. Let's define surface integrals of three dimensional vector fields.

Definition 1.7. Let Σ be a surface in $\mathbb{R}^3$ and let $\mathbf{f}=f_1\mathbf{i}+f_2\mathbf{j}+f_3\mathbf{k}$ be a vector field defined on some subset of $\mathbb{R}^3$ containing Σ. The **surface integral** of $\mathbf{f}$ over Σ is

$$\iint\limits_\Sigma \mathbf{f}\cdot d\sigma=\iint\limits_\Sigma \mathbf{f}\cdot\mathbf{n}\,d\sigma,$$

where $\mathbf{n}$ is the outward unit normal vector to Σ.

Example 1.9. Let's find the surface area T of a torus, created by revolving a circle of radius a in the yz-plane around the z-axis, at a distance b from the z-axis. Say points on longitudinal circles make an angle of u and meridional circles an angle of v with their midpoints. Then we can parametrize the torus as

$$x=(b+a\cos u)\cos v,\quad y=(b+a\cos u)\sin v,\quad z=a\sin u,\quad 0\le u\le 2\pi,\quad 0\le v\le 2\pi.$$

So $\frac{\partial \mathbf{r}}{\partial u}=-a\sin u\cos v\,\mathbf{i}-a\sin u\sin v\,\mathbf{j}+a\cos u\,\mathbf{k}$ and $\frac{\partial \mathbf{r}}{\partial v}=-(b+a\cos u)\sin v\,\mathbf{i}+(b+a\cos u)\cos v\,\mathbf{j}+0\mathbf{k}$, therefore the cross product $\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial r}{\partial v}$ is $-a(b+a\cos u)\cos v\cos u\,\mathbf{i}-a(b+a\cos u)\sin v\cos u\,\mathbf{j}-a(b+a\cos u)\sin u\,\mathbf{k}$, which has magnitude $a(b+a\cos u)$. So the surface area S is equal to $\iint_\Sigma 1\,d\sigma$, which is equal to

$$\int_0^{2\pi}\int_0^{2\pi}\left\|\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right\|du\,dv=\int_0^{2\pi}\int_0^{2\pi}a(b+a\cos u)\,du\,dv=\int_0^{2\pi}\left(abu+a^2\sin u\Big|_{u=0}^{u=2\pi}\right)dv,$$

which simplifies to $4\pi^2ab$.

Example 1.10. Let's calculate a surface integral. Hopefully this shouldn't take too much time. Let $\mathbf{f} = yz\mathbf{i}+xz\mathbf{j}+xy\mathbf{k}$ and Σ be the plane $x+y+z = 1$ bounded by $x \geq 0,\ y \geq 0,\ z \geq 0$. The outward unit normal vector is $\mathbf{n} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$, and Σ is parametrized by $z = 1-(u+v)$ for $u \in [0,1]$, $v \in [0,1-u]$ (call this region R) by projecting Σ onto the xy-plane. Then $\mathbf{f}\cdot\mathbf{n} = (yz,xz,xy)\cdot\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \frac{1}{\sqrt{3}}(yz+xz+xy) = \frac{1}{\sqrt{3}}((x+y)z+xy) = \frac{1}{\sqrt{3}}((u+v)(1-(u+v))+uv) = \frac{1}{\sqrt{3}}((u+v)-(u+v)^2+uv)$ for $(u,v) \in R$. For $\mathbf{r} = u\mathbf{j}+v\mathbf{j}+(1-(u+v))\mathbf{k}$ we have $\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v} = (1,0,-1)\times(0,1,-1) = (1,1,1)$, so $\left\|\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right\| = \sqrt{3}$. Then

$$\iint\limits_{\Sigma} \mathbf{f}\cdot\mathbf{n}\,d\sigma = \int_0^1\int_0^{1-u} \frac{1}{\sqrt{3}}((u+v)-(u+v)^2+uv)\sqrt{3}\,dv\,du = \frac{1}{8}.$$

Computing surface integrals can be tedious. If Σ is a **closed surface** (ie bounds a solid in $\mathbb{R}^3$, or is a 2-manifold), the Divergence Theorem makes this easier for us.

Divergence Theorem. *Let Σ be a closed surface in $\mathbb{R}^3$ that bounds a solid S, and let $\mathbf{f} = f_1\mathbf{i}+f_2\mathbf{j}+f_3\mathbf{k}$ be a vector field defined on some subset of $\mathbb{R}^3$ containing Σ. Then*

$$\iint\limits_{\Sigma} \mathbf{f}\cdot d\sigma = \iiint\limits_{S} \operatorname{div}\mathbf{f}\,dV,$$

where $\operatorname{div}\mathbf{f} = \frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z}$ *is the **divergence** of* $\mathbf{f}$.

Proof. The proof is similar to Green's Theorem, first being proved for when S is bounded above and below by one surface, and laterally by a number of surfaces. Then extend the proof to a general solid. ⊠

Example 1.11. To evaluate the surface integral $\iint_\Sigma \mathbf{f}\cdot d\sigma$ where $\mathbf{f} = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, $\Sigma = S^2$ using the Divergence Theorem, note that $\operatorname{div}\mathbf{f} = 1+1+1 = 3$, so

$$\iint\limits_{\Sigma} \mathbf{f}\cdot d\sigma = \iiint\limits_{S} \operatorname{div}\mathbf{f}\,dV = 3\iiint\limits_{S} 1\,dV + 3\operatorname{vol}(S) = 4\pi.$$

Note. Warning, physics will follow. The surface integral $\iint_\Sigma \mathbf{f}\cdot d\Sigma$ is often called the **flux** of $\mathbf{f}$ through Σ. If $\mathbf{f}$ represents the velocity of a field of a fluid, a positive flux means a net flow out of the fluid (in the direction of n), and similarly negative flux means net inward flow in the direction of $-\mathbf{n}$.

Divergence can be interpreted as how much a vector field diverges from a point, which makes more sense with the following definition equivalent to the Divergence Theorem:

$$\operatorname{div}\mathbf{f}(x,y,z) = \lim_{V\to 0}\frac{1}{V}\iint\limits_{\Sigma} \mathbf{f}\cdot d\sigma,$$

where V is the volume enclosed by Σ around (x,y,z). Taking the limit as $V \to 0$ means taking smaller and smaller neighborhoods around (x,y,z). The limit is the ratio of flux through a surface to the volume enclosed by the surface, giving a rough measure of a flow "leaving" a point. Vector fields with zero divergence are called *solenoidal* fields.

Corollary 1.2. *If the flux of a vector field* $\mathbf{f}$ *is zero through every closed surface containing a given point, then* $\operatorname{div}\mathbf{f} = 0$ *at such point.*

Proof. At a point (x,y,z) we have $\operatorname{div}\mathbf{f}(x,y,z) = \lim_{V\to 0}\frac{1}{V}\iint_\Sigma \mathbf{f}\cdot d\sigma$, but the surface integral is zero by assumption, so the limit is also zero, and we are done. ⊠

Note. Sometimes

$$\oiint\limits_{\Sigma} f(x,y,z)\,d\sigma \quad\text{and}\quad \oiint\limits_{\Sigma} \mathbf{f}\cdot d\sigma$$

are used to denote surface integrals of scalar and vector fields, respectively, over closed surfaces. In physics, you often see $\oint\limits_{\Sigma}$ instead of $\oiint\limits_{\Sigma}$ (it's just `\oiint` as opposed to `\oint`??).

1.7 Stokes' Theorem

This is the good stuff. Let's generalize things.

Definition 1.8. For a function $f : \mathbb{R}^3 \to \mathbb{R}$ and curve $C \subseteq \mathbb{R}^3$ parametrized by $x = x(t),\ y = y(t), z = z(t),\ t \in [a, b]$, the **line integral** of f along C with respect to arc length s is

$$\int_C f\, ds = \int_a^b f(x(t), y(t), z(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt.$$

The line integral of f along C with respect to x is $\int_C f\, dx = \int_a^b f(x(t), y(t), z(t))x'(t)\, dt$, and the line integrals of f with respect to y and z are similarly defined.

Definition 1.9. For a vector field $\mathbf{f} = P\mathbf{i}+Q\mathbf{j}+R\mathbf{k}$ for $P,Q,K : \mathbb{R}^3 \to \mathbb{R}$ and a curve $C \subseteq \mathbb{R}^3$ with smooth parametrization $x = x(t),\ y = y(t),\ z = z(t),\ t \in [a, b]$, the **line integral** of $\mathbf{f}$ along C is

$$\int_C f \cdot d\mathbf{r} = \int_C P\, dx + \int_C Q\, dy + \int_C R\, dz = \int_a^b \mathbf{f}(x(t), y(t), z(t)) \cdot \mathbf{r}'(t) dt,$$

where $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ is the position vector for C.

Theorem 1.8. *For a vector field* $\mathbf{f}$*, we have*

$$\int_C \mathbf{f} \cdot d\mathbf{r} = \int_C \mathbf{f} \cdot \mathbf{T}\, ds$$

where $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$ *is the unit tangent vector for* C.

Theorem 1.9. *For a vector field* $\mathbf{f}$ *with* P,Q,R *continuously differentiable functions on a solid* S*, if there exists an* $F : \mathbb{R}^3 \to \mathbb{R}$ *such that* $\nabla F = \mathbf{f}$ *on* S*, then for* A,B *endpoints of a curve* C *we have*

$$\int_C \mathbf{f} \cdot d\mathbf{r} = F(B) - F(A).$$

Corollary 1.3. *For any* $F : \mathbb{R}^3 \to \mathbb{R}$*, we have* $\oint_C \nabla F \cdot d\mathbf{r} = 0$.

Now that we've finished the theorem spam, let's talk about generalizing Green's Theorem to **orientable** surfaces, which require the existence of a continuous nonzero vector field $\mathbf{N}$ in $\mathbb{R}^3$ normal to the surface (i.e. perpendicular to the the tangent plane for all points of the surface). We call $\mathbf{N}$ a *normal vector field*. For an orientable surface Σ with boundary C, choose a unit normal vector $\mathbf{n}$ with the surface on the left, we say $\mathbf{n}$ is a *positive unit normal vector* and that C is traversed $\mathbf{n}$-positively.

Stoke's Theorem. *Let* Σ *be an orientable surface in* $\mathbb{R}^3$ *whose boundary is a simple closed curve* C*, and for* $P,Q,R : \mathbb{R}^3 \to \mathbb{R}$ *let* $\mathbf{f} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ *be a smooth vector field defined on a subset of* $\mathbb{R}^3$ *containing* Σ*. Then*

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint_\Sigma (\operatorname{curl} \mathbf{f}) \cdot \mathbf{n}\, d\sigma,$$

where

$$\operatorname{curl} \mathbf{f} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k},$$

$\mathbf{n}$ *is a positive unit normal vector over* Σ*, and* C *is traversed* $\mathbf{n}$*-positively.*

Proof. Homology! Omitted because we'll come back to this during diff top. ⊠

You can think of $\oint_C \mathbf{f}\cdot d\mathbf{r}$ as the **circulation** of $\mathbf{f}$ around C, for example, if $\mathbf{E}$ represents the electrostatic field due to a point charge, then $\operatorname{curl}\mathbf{E} = 0$, so the circulation $\oint_C \mathbf{E}\cdot d\mathbf{r} = 0$ by Stokes' Theorem. Such vector fields are called irrotational fields. The term "curl" was made to talk about electromagnetism, which measures something called *circulation density*. We can see this by considering the following definition for curl:

$$\mathbf{n}\cdot(\operatorname{curl}\mathbf{f})(x,y,z) = \lim_{S\to 0}\frac{1}{S}\oint_C \mathbf{f}\cdot d\mathbf{r},$$

where S is the surface area of a surface Σ containing a point (x,y,z) with boundary curve C and positive unit normal vector $\mathbf{n}$. Imagine C shrinking to encapsulate (x,y,z), causing S to approach zero. The ratio of circulation to surface area is what makes the curl a rough measure of circulation density.

Theorem 1.10. *Let S be a simply connected solid region in $\mathbb{R}^3$. Then the following are equivalent:*

(a) $\mathbf{f} = P\mathbf{i}+Q\mathbf{j}+R\mathbf{k}$ *has a smooth potential F in S,*

(b) $\int_C \mathbf{f}\cdot d\mathbf{r}$ *is path independent of curves C in S,*

(c) $\oint_C \mathbf{f}\cdot d\mathbf{r} = 0$ *for every simple closed curve C in S,*

(d) $\frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z}$, $\frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}$, *and* $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ *in S (i.e.* $\operatorname{curl}\mathbf{f} = 0$ *in S, or the differential form $P\,dx + Q\,dy + R\,dz$ is exact).*

1.8 Div, grad, curl

A wrap up section. For $f:\mathbb{R}^n\to\mathbb{R}$, ∇f is a vector-falued function $\mathbb{R}^n\to\mathbb{R}^n$, so we can "apply" the del operator to f to get a new function. Think of $\nabla = \frac{\partial}{\partial x}\mathbf{i}+\frac{\partial}{\partial y}\mathbf{j}+\frac{\partial}{\partial z}\mathbf{k}$ as a "vector" in $\mathbb{R}^3$, this doesn't make sense on its own but we apply these to functions (to get $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$ for example). We denote the divergence $\operatorname{div}\mathbf{f}$ as $\nabla\cdot\mathbf{f}$, the dot product of $\mathbf{f}$ with ∇, the reasoning can be seen below:

$$\nabla\cdot\mathbf{f} = \left(\frac{\partial}{\partial x}\mathbf{i}+\frac{\partial}{\partial y}\mathbf{j}+\frac{\partial}{\partial z}\mathbf{k}\right)\cdot(f_1\mathbf{i}+f_2\mathbf{j}+f_3\mathbf{k}) = \frac{\partial f_1}{\partial x}+\frac{\partial f_2}{\partial y}+\frac{\partial f_3}{\partial z} = \operatorname{div}\mathbf{f}.$$

Similarly, we write $\operatorname{curl}\mathbf{f}$ as the cross product $\nabla\times\mathbf{f}$, this is a fairly straightforward calculation. The divergence of the gradient $\nabla\cdot\nabla f$ has a special name, the **Laplacian** of f, where $\nabla^2 f = \frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}$.

Theorem 1.11. *The curl of the gradient is zero, or* $\nabla\times(\nabla f) = 0$ *for any smooth* $f:\mathbb{R}^3\to\mathbb{R}$.

Proof. We have

$$\nabla\times(\nabla f) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z}\end{vmatrix} = \left(\frac{\partial^2 f}{\partial y\partial z}-\frac{\partial^2 f}{\partial z\partial y}\right)\mathbf{i} - \left(\frac{\partial^2 f}{\partial x\partial z}-\frac{\partial^2 f}{\partial z\partial x}\right)\mathbf{j} + \left(\frac{\partial^2 f}{\partial x\partial y}-\frac{\partial^2 f}{\partial y\partial x}\right)\mathbf{k}.$$

Since f is smooth, the mixed partial derivatives cancel, and this expression is equal to zero. ⊠

Corollary 1.4. *If a vector field $\mathbf{f}(x,y,z)$ has a potential, then* $\operatorname{curl}\mathbf{f} = 0$.

Theorem 1.12. *The divergence of the curl is zero, or* $\nabla\cdot(\nabla\times\mathbf{f}) = 0$ *for smooth vector fields $\mathbf{f}(x,y,z)$.*

Proof. We have

$$\begin{aligned}\nabla\cdot(\nabla\times\mathbf{f}) &= \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)+\frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)+\frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\\ &= \frac{\partial^2 Q}{\partial z\partial x}+\left(\frac{\partial^2 P}{\partial y\partial z}-\frac{\partial^2 P}{\partial z\partial y}\right)+\left(\frac{\partial^2 R}{\partial x\partial y}-\frac{\partial^2 R}{\partial y\partial x}\right)-\frac{\partial^2 Q}{\partial x\partial z}\\ &= 0 \quad \text{since } \mathbf{f} \text{ is smooth.}\end{aligned}$$

⊠

Corollary 1.5. *The flux of the curl of a smooth vector field $\mathbf{f}(x,y,z)$ through any closed surface is zero.*

Lecture 2

The Inverse Function Theorem

I feel like this needs its own section. The inverse function thoerem is a useful theorem from analysis, which talks about the local behavior of a C^∞ map from $\mathbb{R}^n$ to $\mathbb{R}^n$. (This part assumes that you've read at least the part on Euclidian spaces).

2.1 The Inverse Function Theorem

A C^∞ map $f : U \to \mathbb{R}^n$ is **locally invertible** or a **local diffeomorphism** at a point $p \in U$ if f has a C^∞ inverse in some neighborhood of p. The inverse function theorem tells us when maps are locally invertible. We say the matrix $Jf = [\partial f^i / \partial x^j]$ of partial derivatives of f is the **Jacobian matrix** of f and its determinant $\det[\partial f^i / \partial x^j]$ is the **Jacobian determinant** of f (or just the *Jacobian*).

Inverse Function Theorem. *Let $f : U \to \mathbb{R}^n$ be a C^∞ map defined on some open $U \subseteq \mathbb{R}^n$. At any $p \in U$, f is invertible in some neighborhood of p iff the Jacobian is nonzero.*

Proof. Let's accept this without proof. ⊠

Although this apparently reduces invertibility to a number at p, the Jacobian is continuous, so it's actually about the Jacobian nonvanishing in a neighborhood at p. The Jacobian $Jf(p)$ is the best linear approximation to f at p, so it makes sense that f is invertible iff $Jf(p)$ is also invertible.

2.2 The Implicit Function Theorem

The implicit function theorem gives a sufficient condition for a system of equations $f^i(x^1, \cdots, x^n) = 0$ under which *locally* a set of variables can be solved implicity as C^∞ functions of the other variables.

Example 2.1. Consider the circle, given by $f(x, y) = x^2 + y^2 - 1 = 0$. You can see that y is a function of x in any of the points besides $(\pm 1, 0)$, given by $y = \pm\sqrt{1 - x^2}$. If $x \neq \pm 1$, the functions are C^∞, which is consistent.

Implicit Function Theorem for $\mathbb{R}^2$. *Let $f : U \to \mathbb{R}$ be C^∞ for $U \subseteq \mathbb{R}^2$ open. At a point $(a, b) \in U$ where $f(a, b) = 0$ and $\frac{\partial f}{\partial y}(a, b) \neq 0$, there is a neighborhood $A \times B$ of (a, b) in U and a unique function $h \colon A \to B$ such that in $A \times B$,*

$$f(x, y) = 0 \text{ if and only if } y = h(x).$$

Moreover, h is C^∞.

Implicit Function Theorem. *Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open and $f : U \to \mathbb{R}^m$ a C^∞ function. Suppose $[\partial f^i / \partial y^j(a, b)]$ is nonsingular at a point $(a, b) \in \ker f$. Then a neighborhood $A \times B \subseteq U$ and a unique smooth function $h \colon A \to B$ both exist such that in $A \times B$, we have*

$$f(x, y) = 0 \text{ if and only if } y = h(x).$$

2.3 The Constant Rank Theorem

Note that the rank of a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ at a point p is the rank of its Jacobian $[\partial f^i / \partial x^j(p)]$.

Constant Rank Theorem. *If $f : U \to \mathbb{R}^m$ has constant rank k in a neighborhood of a point $p \in U$, then after a change of coordinates near p in U and $f(p)$ in $\mathbb{R}^m$, the map f assumes the form $(x^1, \cdots, x^n) \mapsto (x^1, \cdots, x^k, 0, \cdots, 0)$. More precisely, there are diffeomorphisms G of a neighborhood of p in U and F of a neighborhood $f(p)$ in $\mathbb{R}^m$ such that*

$$F \circ f \circ G^{-1}(x^1, \cdots, x^n) = (x^1, \cdots, x^k, 0, \cdots .0).$$

Part II
Euclidian Spaces

Lecture 3

Smooth Functions on a Euclidian Space

INTRODUCTION

Calculus talks about differentiation and integration on $\mathbb{R}$, while real analysis extends this to $\mathbb{R}^n$. Vector calculus talks about integrals on curves and surfaces, and now we extend these concepts to higher dimensions, the structures which with we work with are called manifolds. Things become simple: gradient, curl, and divergence are cases of the exterior derivative, and the FTC for line integrals, Green's theorem, Stokes' theorem, and the divergence theorem are manifestations of the generalized Stokes' theorem.

Manifolds arise even when dealing with the space we live in, for example the set of affine motions in $\mathbb{R}^3$ is a 6-manifold. This is our plan: recast calculus on $\mathbb{R}^n$ so we can generalize it to manifolds by differential forms. Working in $\mathbb{R}^n$ first isn't necessary, but much easier, since the examples are simple. Then, we define a manifold and talk about tangent spaces, working with the idea of approximating nonlinear things with linear things, with Lie groups and Lie algebras as examples. Finally, we do calculus on manifolds, generalizing the theorems of vector calculus, with the de Rham cohomology groups as C^∞ and topological invariants.

3.1 C^∞ Versus Analytic Functions

Let's talk about C^∞ functions on $\mathbb{R}^n$. Write a base for $\mathbb{R}^n$ as $x^1, \cdots, x^n$ and let $p = (p^1, \cdots, p^n)$ be a point in an open set U in $\mathbb{R}^n$. Differential geometry uses *superscripts*, not subscripts, more on this later.

Definition 3.1. Let k be a nonnegative integer. A function $f : U \to \mathbb{R}$ is C^k at p if its partial derivatives $\frac{\partial^j f}{\partial x^{i_1} \cdots \partial x^{i_j}}$ of all orders $j \leq k$ exist and are continuous at p. The function $f : U \to \mathbb{R}$ is C^∞ at p if it is C^k for all $k \geq 0$, that is, its partial derivatives of all orders

$$\frac{\partial^k f}{\partial x^{i_1} \cdots \partial x^{i_k}}$$

exist and are continuous at p. We say f is C^k on U if it is C^k for all points in U, and the concept of C^∞ on a set U is defined similarly. When we say "smooth", we mean C^∞.

Example 3.1.

(i) We call C^0 functions on U continuous on U.

(ii) Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^{1/3}$. Then $f'(x)$ is $\frac{1}{3}x^{-2/3}$ for $x \neq 0$ and undefined at zero, so f is C^0 but not C^1 at $x = 0$.

(iii) Let $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = \int_0^x f(t)\,dt = \int_0^x t^{1/3}\,dt = \frac{3}{4}x^{4/3}.$$

Then $g'(x) = f(x) = \frac{1}{3}$, so $g(x)$ is C^1 but not C^2 at $x = 0$. In general, we can construct functions that are C^k but not C^{k+1} at a point.

(iv) Polynomials, the sine and cosine functions, and the exponential functions on $\mathbb{R}$ are all C^∞.

A function f is **real-analytic** at p if in some neighborhood of p it is equal to its Taylor series at p, that is,

$$f(x)=f(p)+\sum_i \frac{\partial f}{\partial x^i}(p)(x^i-p^i)+\frac{1}{2!}\sum_{i,j}\frac{\partial^2 f}{\partial x^i \partial x^j}(p)(x^i-p^i)(x^j-p^j)+\cdots$$

Real-analytic functions are C^∞ because you can differentiate them termwise in their region of convergence. The converse does not hold: define

$$f(x)=\begin{cases} e^{-1/x} & \text{for } x>0; \\ 0 & \text{for } x\leq 0. \end{cases}$$

We can show f is C^∞ on $\mathbb{R}$ and the derivatives $f^{(k)}(0)=0$ for all $k\geq 0$ by induction, then the Taylor series must be zero in any neighborhood of the origin, but f is not. Then f isn't equal to its Taylor series, and we have a smooth non-analytic function.

3.2 Taylor's Theorem with Remainder

However, we have a Taylor's theorem with remainder for C^∞ functions that's good enough. Say a subset S of $\mathbb{R}^n$ is **star-shaped** with respect to a point p in S if for every $x\in S$, the line segment from p to x lies in S.

Lemma 3.1 (Taylor's theorem with remainder). *Let f be a C^∞ function on an open subset U of $\mathbb{R}^n$ star-shaped with respect to a point $p=(p^1,\cdots,p^n)$ in U. Then there are C^∞ functions $g_1(x),\cdots,g_n(x)$ on U such that*

$$f(x)=f(p)+\sum_{i=1}^n (x^i-p^i)g_i(x), \quad g_i(p)=\frac{\partial f}{\partial x^i}(p).$$

Proof. For any $x\in U$ the line segment $p+t(x-p)$, $0\leq t\leq 1$ lies in U. So $f(p+t(x-p))$ is defined, and by the chain rule we have

$$\frac{d}{dt}f(p+t(x-p))=\sum (x^i-p^i)\frac{\partial f}{\partial x^i}(p+t(x-p)).$$

Integrating both sides with respect to $t\in[0,1]$ we have

$$f(p+t(x-p))\Big|_0^1=\sum (x^i-p^i)\int_0^1 \frac{\partial f}{\partial x^i}(p+t(x-p))\,dt.$$

Let $g_i(x)=\int_0^1 \frac{\partial f}{\partial x^i}(p+t(x-p))\,dt$. Then $g_i(x)$ is C^∞ and the above expression simplifies to

$$f(x)-f(p)=\sum (x^i-p^i)g_i(x).$$

Furthermore, $g_i(p)=\int_0^1 \frac{\partial f}{\partial x^i}(p)\,dt=\frac{\partial f}{\partial x^i}(p)$. ⊠

If $n=1$ and $p=0$, this lemma says that $f(x)=f(0)+xf_1(x)$ for a C^∞ function $f_1(x)$. Applying repeatedly gives $f_i(x)=f_i(0)+xf_{i+1}(x)$, where f_i, f_{i+1} are C^∞ functions. So

$$\begin{aligned} f(x) &= f(0)+x(f_1(0)+xf_2(x)) \\ &= f(x)+xf_1(x)+x^2(f_2(0)+xf_3(x)) \\ &\vdots \\ &= f(0)+f_1(0)x+f_2(0)x^2+\cdots+f_i(0)x^i+f_{i+1}(x)x^{i+1}. \end{aligned}$$

If we differentiate this expression k times, we get $f^{(k)}(0)=k!f_k(x)$, which simplifies to $f_k(0)=\frac{1}{k!}f^{(k)}(0)$ for $k=1,2,\cdots,i$. Note that balls are star-shaped, and since U is open there exists an $\varepsilon>0$ such that $p\in B(p,\varepsilon)\subseteq U$. So when a function's domain is restricted to $B(p,\varepsilon)$, f is defined on a star-shaped neighborhood of p and Taylor's theorem with remainder applies.

Lecture 4

Tangent Vectors in $\mathbb{R}^n$ as Derivations

Vectors at a point p are usually represented by columns of points or arrows stemming from p. A vector at p is tangent to a surface if it lies in the tangent plane at p, the limiting position of the secant planes through p. This kind of definition assumes we live in $\mathbb{R}^n$ and would not work for a large class of manifolds. We will find a generalization that works for manifolds.

4.1 The Directional Derivative

We usually visualize the tangent space $T_p(\mathbb{R}^n)$ at $p \in \mathbb{R}^n$ as the vector space of all arrows emanating from p. This can be identified with the vector space $\mathbb{R}^n$. We write points as $p = (p^1, \cdots, p^n)$ and vectors v in $T_p(\mathbb{R}^n)$ as

$$v = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix} \quad \text{or} \quad \langle v^1, \cdots, v^n \rangle.$$

We denote the standard basis for $\mathbb{R}^n$ or $T_p(\mathbb{R}^n)$ by $\{e_1, \cdots, e_n\}$. Then $v = \sum v^i e_i$. Sometimes we denote $T_p(\mathbb{R}^n)$ by $T_p\mathbb{R}^n$. Elements of $T_p(\mathbb{R}^n)$ are called **tangent vectors** (or simply **vectors**) at p in $\mathbb{R}^n$.

The line through a point $p = (p^1, \cdots, p^n)$ with direction $\langle v_1, \cdots, v_n \rangle$ in $\mathbb{R}^n$ has parametrization $c(t) = (p^1 + tv^1, \cdots, p^n + tv^n)$, with ith component $c^i(t) = p^i + tv^i$. If f is C^∞ in a neighborhood of p in $\mathbb{R}^n$ and v is a tangent vector at p, the **directional derivative** of f in the direction v at p is defined to be

$$D_v f = \lim_{t \to 0} \frac{f(c(t)) - f(p)}{t} = \left.\frac{d}{dt}\right|_{t=0} f(c(t)).$$

By the chain rule,

$$D_v f = \sum_{i=1}^{n} \frac{dc^i}{dt}(0) \frac{\partial f}{\partial x^i}(p) = \sum_{i=1}^{n} v^i \frac{\partial f}{\partial x^i}(p).$$

Note that $D_v f$ is a number, since we evaluate partial derivatives at a point p. We write $D_v = \sum v^i \left.\frac{\partial}{\partial x^i}\right|_p$ for the operator that sends a function f to the number $D_v f$. Often times we omit the subscript p.

4.2 Germs of Functions

If two functions agree on some neighborhood of a point p, they will have the same directional derivatives at p. This suggests introducing an equivalence relation on the C^∞ functions defined in some neighborhood of p. Consider the set of pairs (f, U), where U is a neighborhood of p and $f : U \to \mathbb{R}$ is a C^∞ function. We say (f, U) is equivalent to (g, V) if there is an open set $W \subseteq U \cap V$ containing p such that $f = g$ when restricted to W. The equivalence class of (f, U) is called the **germ** of f at p. We write $C_p^\infty(\mathbb{R}^n)$ or simply C_p^∞ if there is no possibility of confusion, for the set of all germs of C^∞ functions on $\mathbb{R}^n$ at p.

Example 4.1. The functions $f(x) = \frac{1}{1-x}$ with domain $\mathbb{R} \setminus \{1\}$ and $g(x) = \sum_{n=0}^{\infty} x^n$ with domain $(-1, 1)$ have the same germ at any point p in the open inverval $(-1, 1)$.

An **algebra** over a field K is a vector space A over K with a multiplication map $\mu\colon A \times A \to A$, usually written $\mu(a, b) = a \times b$, such that for all $a, b, c \in A$ and $r \in K$,

(i) $(a \times b) \times c = a \times (b \times c)$ (associativity),

(ii) $(a + b) \times c = a \times c + b \times c$ and $a \times (b + c) = a \times b + a \times c$ (distributivity),

(iii) $r(a \times b) = (ra) \times b = a \times (rb)$ (homogeneity).

Equivalently, an algebra over a field K is a ring A also a K-vector space such that ring multiplication satisfies the homogeneity condition. So an algebra has three operations: the addition and multiplication of a ring, and the scalar multiplication of a vector space. Usually we write ab in place of $a \times b$.

Addition and multiplication of functions induce operations on C_p^∞, making it into an algebra over $\mathbb{R}$.

4.3 Derivations at a Point

A map $L: V \to W$ between vector spaces over a field K is called a **linear map** or a **linear operator** if for any $r \in K$ and $u, v \in V$,

(i) $L(u+v) = L(u) + L(v)$;

(ii) $L(rv) = rL(v)$.

Such a map can also be called K-*linear*.

For every tangent vector v at a point p in $\mathbb{R}^n$, the directional derivative at p gives a map of real vector spaces $D_v: C_p^\infty \to \mathbb{R}$. Since $D_v f = \sum_{i=1}^n v^i \frac{\partial f}{\partial x^i}(p)$, we have that D_v is $\mathbb{R}$-linear and satisfies the Leibniz rule

$$D_v(fg) = (D_v f)g(p) + f(p)D_v g,$$

since the partial derivatives $\partial/\partial x^i\big|_p$ have these properties. In general, a linear map $D: C_p^\infty \to \mathbb{R}$ satisfying the Leibniz rule is called a **derivation at p** or a **point-derivation** of C_p^∞. Denote the set of all derivations at p by $\mathscr{D}_p(\mathbb{R}^n)$. This set is a real vector space, since the sum of two derivations at p and a scalar multiple of a derivation at p are again derivations at p. We know that directional derivatives at p are all derivations at p, so we have a map

$$\phi: T_p(\mathbb{R}^n) \to \mathscr{D}_p(\mathbb{R}^n), \quad v \mapsto D_v = \sum v^i \left.\frac{\partial}{\partial x^i}\right|_p.$$

Since D_v is linear in v, the map ϕ is a linear operator of vector spaces.

Lemma 4.1. *If D is a point-derivation of C_p^∞, then $D(c) = 0$ for any constant function c.*

Proof. By $\mathbb{R}$-linearity, $D(c) = cD(1)$. To show $D(1) = 0$, we have

$$D(1) = D(1 \times 1) = D(1) \times 1 + 1 \times D(1) = 2D(1)$$

by the Leibniz rule. Subtracting $D(1)$ from both sides gives $D(1) = 0$. ⊠

Theorem 4.1. *The linear map $\phi: T_p(\mathbb{R}^n) \to \mathscr{D}_p(\mathbb{R}^n)$ is an isomorphism of vector spaces.*

Proof. To show injectivity, assume $D_v = 0$ for $v \in T_p(\mathbb{R}^n)$. Applying D_v to the coordinate function x^j (that sends $x \mapsto x^j$) gives

$$0 = D_v(x^j) = \sum_i v^i \left.\frac{\partial}{\partial x^i}\right|_p x^j = \sum_i v^i \delta_i^j = v^j.$$

[1] To prove surjectivity, let D be a derivation at p and let (f, V) be a representative of a germ in C_p^∞. Making V smaller if necessary, assume that V is an open (star-shaped) ball. By Taylor's theorem with remainder there are C^∞ functions $g_i(x)$ in a neighborhood of p such that

$$f(x) = f(p) + \sum (x^i - p^i) g_i(x), \quad g_i(p) = \frac{\partial f}{\partial x^i}(p).$$

Note that $D(f(p))$ and $D(p^i)$ equal zero since $f(p)$ and p^i are constant. Applying D to both sides, we get by the Leibniz rule

$$Df(x) = \sum (Dx^i) g_i(p) + \sum (p^i - p^i) Dg_i(x) = \sum (Dx^i) \frac{\partial f}{\partial x^i}(p).$$

So $D = D_v$ for $v = \langle Dx^1, \cdots, Dx^n \rangle$. ⊠

This theorem shows that we can identify tangent vectors at p with derivations at p. Under the identification $T_p(\mathbb{R}^n) \simeq \mathscr{D}_p(\mathbb{R}^n)$, the standard basis $\{e_1, \cdots, e_n\}$ for $T_p(\mathbb{R}^n)$ corresponds to the set $\{\partial/\partial x^1|_p, \cdots, \partial/\partial x^n|_p\}$ of partial derivatives. From now, we write a tangent vector $\langle v^1, \cdots, v^n \rangle = \sum v^i e_i$ as $v = \sum v^i \left.\frac{\partial}{\partial x^i}\right|_p$. Although the vector space $\mathscr{D}_p(\mathbb{R}^n)$ of derivations is not as intuitive, they turn out to be more suitable for generalization to manifolds.

[1] I think δ_i^j refers to the function that is one if $i = j$ and zero otherwise.

4.4 Vector Fields

A **vector field** X on an open subset U of $\mathbb{R}^n$ is a function that assigns to each point p in U a tangent vector X_p in $T_p(\mathbb{R}^n)$. Since $T_p(\mathbb{R}^n)$ has basis $\{\partial/\partial x^i|_p\}$, the vector X_p is a linear combination $X_p = \sum a^i(p) \frac{\partial}{\partial x^i}\Big|_p$, $p \in U$. We say that the vector field X is C^∞ on U if the coefficient functions a^i are all C^∞ on U. We can identify vector fields on U with column vectors of C^∞ functions on U:

$$X = \sum a^i \frac{\partial}{\partial x^i} \quad \longleftrightarrow \quad \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix}.$$

The ring of C^∞ functions on U is commonly denoted $C^\infty(U)$ or $\mathscr{F}(U)$. Since one can multiply a C^∞ vector field by a C^∞ function and still get a C^∞ vector field, the set of all C^∞ vector fields on U, denoted $\mathfrak{X}(U)$, is not only a vector space over $\mathbb{R}$ but also a $C^\infty(U)$-module.

4.5 Vector Fields as Derivations

If X is a C^∞ vector field on an open subset U of $\mathbb{R}^n$ and f is a C^∞ function on U, we define a new function Xf on U by $(Xf)(p) = X_p f$ for any $p \in U$. Writing $X = \sum a^i \partial/\partial x^i$, we get

$$(Xf)(p) = \sum a^i(p) \frac{\partial f}{\partial x^i}(p) \quad \text{or} \quad Xf = \sum a^i \frac{\partial f}{\partial x^i},$$

which shows Xf is a C^∞ function on U. So a C^∞ vector field gives rise to an $\mathbb{R}$-linear map $C^\infty(U) \to C^\infty(U)$, $f \mapsto Xf$.

Proposition 4.1. *If X is a C^∞ vector field and f and g are C^∞ functions on an open subset U of $\mathbb{R}^n$, then $X(fg)$ satisfies the product rule (Leibniz rule) $X(fg) = (Xf)g + fXg$.*

Proof. At each point $p \in U$, the vector X_p satisfies the Leibniz rule $X_p(fg) = (X_p f)g(p) + f(p)X_p g$. As p varies over U, this becomes an equality of functions $X(fg) = (Xf)g + fXg$. ⊠

If A is an algebra over a field K, a **derivation** of A is a K-linear map $D: A \to A$ such that

$$D(ab) = (Da)b + aDb \quad \text{for all } a, b \in A.$$

The set of all derivations of A is closed under addition and scalar multiplication and forms a vector space, denoted $\mathrm{Der}(A)$. As noted above, a C^∞ vector field on an open set U gives rise to a derivation of the algebra $C^\infty(U)$. We therefore have a map $\varphi: \mathfrak{X}(U) \to \mathrm{Der}(C^\infty(U))$, $X \mapsto (f \mapsto Xf)$. Similar to how tangent vectors at a point p can be identified with the point-derivations of C_p^∞, so the vector fields on an open set U can be identified with the derivations of the algebra $C^\infty(U)$, i.e., the map φ is an isomorphism of vector spaces. Note that a derivation at p is not a derivation of the algebra C_p^∞. A derivation at p is a map from C_p^∞ to $\mathbb{R}$, while a derivation of the algebra C_p^∞ is a map from $C_p^\infty \to C_p^\infty$.

Lecture 5

Alternating k-Linear Functions

5.1 Dual Space

If V and W are real vector spaces, we denote the vector space of all linear maps $f: V \to W$ by $\mathrm{Hom}(V, W)$. Define the **dual space** V^* to be the set of all real valued linear functions on V, denoted $V^* := \mathrm{Hom}(V, \mathbb{R})$. Elements of V^* are called **covectors** or 1-**covectors** on V. Assume V is finite-dimensional, and let $\{e_1, \cdots, e_n\}$ be a basis for V. Then every $v \in V$ is uniquely a linear combination $v = \sum v^i e_i$ with $v^i \in \mathbb{R}$. Let $\alpha^i: V \to \mathbb{R}$ be the linear function that picks out the ith coordinate, given by $\alpha^i(v) = v^i$. Note that $\alpha^i(e_j) = \delta^i_j$.

Proposition 5.1. *The functions $\alpha^1, \cdots, \alpha^n$ form a basis for V^*.*

Proof. Let $f \in V^*$ and $v = \sum v^i e_i \in V$. Then $f(v) = \sum v^i f(e_i) = \sum f(e_i)\alpha^i(v)$. So $f = \sum f(e_i)\alpha^i$, and so the α^i span V^*. Now suppose $\sum c_i \alpha^i = 0$ for some $c_i \in \mathbb{R}$. Then $0 = \sum c_i \alpha^i(e_j) = \sum c_i \delta^i_j = c_j$ for $j = 1, \cdots, n$. So the α^i are LI. ⊠

This basis $\{\alpha^1, \cdots, \alpha^n\}$ is said to be *dual* to the basis $\{e_1, \cdots, e_n\}$ for V.

Corollary 5.1. *The dual space V^* of a finite-dimensional vector space V has the same dimension as V.*

Example 5.1. If $e_1, \cdots, e_n$ is a basis for a vector space V, every $v \in V$ can be uniquely written as a linear combo $v = \sum b^i(v) e_i$, where $b_i(v) \in \mathbb{R}$. Let $\alpha^1, \cdots, \alpha^n$ be the basis of V^* dual to $e_1, \cdots, e_n$. Then

$$\alpha^i(v) = \alpha^i\left(\sum_j b^j(v) e_j\right) = \sum_j b^j(v)\alpha^i(e_j) = \sum_j b^j(v)\delta^i_j = b^i(v).$$

So the set of coordinate functions $b^1, \cdots, b^n$ WRT the basis $e_1, \cdots, e_n$ is precisely the dual basis.

5.2 Permutations

Quick review since you know what permutations are. They're self-bijections, or elements of the symmetric group on such set. You use cycle notation to denote them, and a transposition is a 2-cycle. Recall that the sign of a permutation denoted $\operatorname{sgn}(\sigma)$ is ± 1 depending on whether the permutation is even or odd. Since $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$, one way to compute signs is to decompose and count (eg if its a product of odd and even cycles it must be odd).

An **inversion** in a permutation σ is an ordered pair $(\sigma(i), \sigma(j))$ such that $i < j$ but $\sigma(i) > \sigma(j)$. For example, the permutation $(124)(35)$ has the inversions $(21), (41), (51), (43), (53)$. Another way to compute the sign of a permutation is to count the number of inversions:

Proposition 5.2. *A permutation is even iff it has an even number of inversions.*

Proof. We can multiply by transpositions corresponding to inversions and recover our original list.

(1) Find 1 in the list $\sigma(1), \sigma(2), \cdots, \sigma(k)$, then every number before 1 gives rise to an inversion. Say $\sigma(i) = 1$, then $(\sigma(1), 1), \cdots, (\sigma(i-1), 1)$ are all inversions. Apply the $i-1$ transpositions to move 1 to the front, the number of inversions ending in 1.

(2) Now find 2 in the list $1, \sigma(1), \cdots, \sigma(\hat{i}), \cdots, \sigma(k)$ (using deletion notation, note that we moved 1 to the front). Every number (besides 1) preceding 2 gives rise to an inversion $(\sigma(m), 2)$, suppose we have i_2 such inversions. Applying i_2 transformations, we move 2 to the front.

If you continue to the sort, you'll see the number of transpositions required to order the list is the same as the number of inversions. So the transposition decomposition is the number of inverstions, and $\operatorname{sgn}(\sigma) = (-1)^{\#\text{ of inversions in }\sigma}$. ⊠

5.3 Multilinear Functions

Let $V^k = \overbrace{V \times \cdots \times V}^{\text{k times}}$ for V a real vector space. A function $f : V^k \to \mathbb{R}$ is **k-linear** if it is linear in each of its k arguments, that is, $f(\cdots, av + bw, \cdots) = af(\cdots, v, \cdots) + bf(\cdots, w, \cdots)$ for $a, b \in \mathbb{R}$, $v, w \in V$. Usually we say "bilinear" and "trilinear" instead of 2-linear and 3-linear. A k-linear function on V is also called a **k-tensor** on V. We denote the vector space of k-tensors on V by $L_k(V)$. If f is a k-tensor on V, we call k the **degree** of f.

Example 5.2. The dot product $f(v, w) = v \cdot w$ on $\mathbb{R}^n$ is bilinear.

Example 5.3. If we view the determinant $f(v_1, \cdots, v_n) = \det[v_1 \cdots v_n]$ as a function on the n column vectors in $\mathbb{R}^n$, then the determinant is n-linear.

Definition 5.1. A k-linear function $f : V^k \to \mathbb{R}$ is **symmetric** if $f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) = f(v_1, \cdots, v_k)$ for all permutations $\sigma \in S_k$; it is **alternating** if $f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) = (\operatorname{sgn}\sigma) f(v_1, \cdots, v_k)$ for all $\sigma \in S_k$.

Example 5.4.

(i) The dot product $f(v,w) = v \cdot w$ on $\mathbb{R}^n$ is symmetric.

(ii) The determinant $f(v_1, \cdots, v_n) = \det[v_1 \cdots v_n]$ is alternating.

Intuitively, symmetric k-linear functions don't care which order you input the variables in, while alternating multilinear maps preserve even orientation, but get flipped with an odd number of shuffles.

We are especially interested in the space $A_k(V)$ of all alternating k-linear functions on a vector space V for $k > 0$. These are called **alternating k-tensors**, **k-covectors**, or **multicovectors** on V. For $k = 0$, we define a 0-covector to be a constant so that $A_0(V)$ is the vector space $\mathbb{R}$. A 1-covector is just a covector.

5.4 Permutation Action on k-Linear Functions

If f is a k-linear function on a vector space V and $\sigma \in S_k$, we define a new k-linear function by $(\sigma f)(v_1, \cdots, v_k) = f(v_{\sigma(1)}, \cdots, v_{\sigma(k)})$. So f is symmetric iff $\sigma f = f$ and alternating iff $\sigma f = (\operatorname{sgn}\sigma) f$ for all $\sigma \in S_k$. In the trivial case, S_1 is the identity group, and a 1-linear function is both symmetric and alternating, in particular, $A_1(V) = L_1(V) = V^*$.

Lemma 5.1. *If $\sigma, \tau \in S_k$ and f is a k-linear function on V, then $\tau(\sigma f) = (\tau\sigma)f$. In other words, S_k acts (from the left) on $L_k(V)$, the space of k-linear functions on V.*

Proof. Exercise. ⊠

5.5 The Symmetrizing and Alternating Operators

We can turn a k-linear function f on a vector space V into a symmetric k-linear function Sf by defining

$$(Sf)(v_1, \cdots, v_k) = \sum_{\sigma \in S_k} f(v_{\sigma(1)}, \cdots, v_{\sigma(k)})$$

or with our new notation, $Sf = \sum_{\sigma \in S_k} \sigma f$. Similarly, we the define the alternization $Af = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma)\sigma f$.

Proposition 5.3.

(i) The k-linear function Sf is symmetric.

(ii) The k-linear function Af is alternating.

Proof.

(i) If $\tau \in S_k$, we have $\tau(Sf) = \sum_{\sigma \in S_k} \tau(\sigma f) = \sum_{\sigma \in S_k} (\tau\sigma) f = Sf$, since by applying τ^{-1} we can obtain every permutation in S_k.

(ii) If $\tau \in S_k$, we have $\tau(Af) = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma)\tau(\sigma f) = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma)(\tau\sigma) f$, since S_k acts on $L_k(V)$ from the left. Since $(\operatorname{sgn}\tau)(\operatorname{sgn}\tau) = 1$, this expression is equal to $(\operatorname{sgn}\tau)\sum_{\sigma \in S_k} (\operatorname{sgn}\tau\sigma)(\tau\sigma) f = (\operatorname{sgn}\tau) Af$. ⊠

Lemma 5.2. *If f is an alternating k-linear function on a vector space V, then $Af = (k!)f$.*

Proof.

$$Af = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma)\sigma f = \sum_{\sigma \in S_k} (\operatorname{sgn}\sigma)(\operatorname{sgn}\sigma) f = \sum_{\sigma \in S_k} f = k! f.$$ ⊠

5.6 The Tensor Product

Let f be a k-linear function and g be an ℓ-linear function on a vector space V. Their **tensor product** is the $(k+\ell)$-linear function $f \otimes g$ defined by

$$(f \otimes g)(v_1, \cdots, v_{k+\ell}) = f(v_1, \cdots, v_k)g(v_{k+1}, \cdots, v_{k+\ell}).$$

Example 5.5. Let $e_1, \cdots, e_n$ be the standard basis for $\mathbb{R}^n$ and let $\alpha^1, \cdots, \alpha^n$ be its dual basis. The Euclidian inner product on $\mathbb{R}^n$ is the bilinear function $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by $\langle v, w \rangle = \sum v^i w^i$ for $v = \sum v^i e_j$ and $w = \sum w^i e_i$. We can express $\langle , \rangle$ in terms of the tensor product:

$$\begin{aligned}\langle v, w \rangle &= \sum_i v^i w^i = \sum_i \alpha^i(v)\alpha^i(w) \\ &= (\alpha^i \otimes \alpha^i)(v, w).\end{aligned}$$

So $\langle , \rangle = \sum_i \alpha^i \otimes \alpha^i$. This notation is often used in differential geometry to describe an inner product on a vector space.

5.7 The Wedge Product

We would like for the product of two alternating multilinear functions to also be alternating. This motivates the **wedge product**: for $f \in A_k(V)$ and $g \in A_\ell(V)$, $f \wedge g = \frac{1}{k!\ell!}A(f \otimes g)$; or explicitly,

$$(f \wedge g)(v_1, \cdots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \sigma) f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+\ell)}).$$

Since the definition involves the alternization, $f \wedge g$ is alternating.

When $k = 0$, the element $f \in A_0(V)$ is a constant c: then the wedge product $c \wedge g$ is scalar multiplication, since $\frac{1}{\ell!} \sum_{\sigma \in S_\ell} (\operatorname{sgn} \sigma) c g(v_{\sigma(1)}, \cdots, c_{\sigma(\ell)}) = c g(v_1, \cdots, v_\ell)$. So $c \wedge g = cg$ for $c \in \mathbb{R}$ and $g \in A_\ell(V)$.

The coefficient $1/(k!\ell!)$ compensates for the repetitions in the sum: there are $k!$ permutations that permute the arguments of f and similarly $\ell!$ permutations of the arguments of g. So we divide to get rid of repeating terms.

Example 5.6. For $f \in A_2(V)$ and $g \in A_1(V)$, $A(f \otimes g)(v_1, v_2, v_3) = f(v_1, v_2)g(v_3) - f(v_1, v_3)g(v_2) + f(v_2, v_3)g(v_1) - f(v_2, v_1)g(v_3) + f(v_3, v_1)g(v_2) - f(v_3, v_2)g(v_1)$. There are three pairs of equal terms, for example $f(v_1, v_2)g(v_3) = -f(v_2, v_1)g(v_3)$ (since f is alternating) and so on. So after dividing by two, $(f \wedge g)(v_1, v_2, v_3) = f(v_1, v_2)g(v_3) - f(v_1, v_3)g(v_2) + f(v_2, v_3)g(v_1)$.

A way to avoid such redundancies in the definition of $f \wedge g$ is to stipulate that $\sigma(1), \cdots, \sigma(k)$ be in ascending order and $\sigma(k+1), \cdots, \sigma(k+\ell)$ also be in ascending order in the sum of the wedge product. We call a permutation $\sigma \in S_{k+\ell}$ a **$(\mathbf{k}, \ell)$-shuffle** if $\sigma(1) < \cdots < \sigma(k)$ and $\sigma(k+1) < \cdots < \sigma(k+\ell)$. Then we can rewrite the definition of the wedge product as

$$(f \wedge g)(v_1, \cdots, v_{k+\ell}) = \sum_{(k,\ell)\text{-shuffles } \sigma} (\operatorname{sgn} \sigma) f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \cdots, v_{\sigma(k+\ell)}).$$

Written this way, the wedge $f \wedge g$ is a sum of $\binom{k+\ell}{k}$ terms, rather than $(k+\ell)!$ terms.

Example 5.7. If f and g are covectors on a vector space V and $v_1, v_2 \in V$, then $(f \wedge g)(v_1, v_2) = f(v_1)g(v_2) - f(v_2)g(v_1)$.

5.8 Anticommutativity of the Wedge Product

It follows from the definition of the wedge product that $f \wedge g$ is bilinear in f and g.

Proposition 5.4. *The wedge product is anticommutative: if $f \in A_k(V)$ and $g \in A_\ell(V)$, then $f \wedge g = (-1)^{k\ell} g \wedge f$.*

Proof. Define $\tau \in S_{k+\ell}$ to be the permutation

$$\tau = \begin{pmatrix} 1 & \cdots & \ell & \ell+1 & \cdots & \ell+k \\ k+1 & \cdots & k+\ell & 1 & \cdots & k \end{pmatrix}.$$

Then $\sigma(1) = \sigma\tau(\ell+1), \cdots, \sigma(k) = \sigma\tau(\ell+k),\ \sigma(k+1) = \sigma\tau(1), \cdots, \sigma(k+\ell) = \sigma\tau(\ell)$. For any $v_1, \cdots, v_{k+\ell} \in V$,

$$\begin{aligned} A(f \otimes g)(v_1, \cdots, v_{k+\ell}) &= \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn}\sigma) f(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) g(v_{\sigma(k+1)}, \cdots, v_{k\sigma(k+\ell)}) \\ &= \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn}\sigma) f(v_{\sigma\tau(\ell+1)}, \cdots, v_{\sigma\tau(\ell+k)}) g(v_{\sigma\tau(1)}, \cdots, v_{\sigma\tau(\ell)}) \\ &= (\operatorname{sgn}\tau) \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn}\sigma\tau) g(v_{\sigma\tau(1)}, \cdots, v_{\sigma\tau(\ell)}) f(v_{\sigma\tau(\ell+1)}, \cdots, v_{\sigma\tau(\ell+k)}) \\ &= (\operatorname{sgn}\tau) A(g \otimes f)(v_1, \cdots, v_{k+\ell}). \end{aligned}$$

So $A(f \otimes g) = (\operatorname{sgn}\tau)A(g \otimes f) = (-1)^{k\ell}A(g \otimes f)$. Dividing by $k!\ell!$ gives $f \wedge g = (-1)^{k\ell} g \wedge f$. ⊠

Corollary 5.2. *If f is a k-covector on V and k is odd, then $f \wedge f = 0$.*

Proof. We have $f \wedge f = (-1)^{k^2} f \wedge f = -f \wedge f$ by anticommutativity, so $f \wedge f = 0$. ⊠

5.9 Associativity of the Wedge Product

Lemma 5.3. *Suppose f is a k-linear function and g is an ℓ-linear function on a vector space V. Then*

(i) $A(A(f) \otimes g) = k!A(f \otimes g)$, and

(ii) $A(f \otimes A(g)) = \ell!A(f \otimes g)$.

Proof. We have

$$A(A(f) \otimes g) = \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn}\sigma)\sigma \left(\sum_{\tau \in S_k} (\operatorname{sgn}\tau)(\tau f) \otimes g \right).$$

We can view $\tau \in S_k$ as a permutation in $S_{k+\ell}$ such that $\tau(i) = i$ for $i = k+1, \cdots, k+\ell$. For such a τ, $(\tau f) \otimes g = \tau(f \otimes g)$. So

$$A(A(f) \otimes g) = \sum_{\sigma \in S_{k+\ell}} \sum_{\tau \in S_k} (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)(\sigma\tau)(f \otimes g).$$

Let $\mu = \sigma\tau \in S_{k+\ell}$, there are $k!$ ways to write $\mu = \sigma\tau$ with $\sigma \in S_{k+\ell}$ and $\tau \in S_k$, because each $\tau \in S_k$ determines a unique σ by the formula $\sigma = \mu\tau^{-1}$. So the double sum above can be rewritten as

$$A(A(f) \otimes g) = k! \sum_{\mu \in S_{k+\ell}} (\operatorname{sgn}\mu)\mu(f \otimes g).$$

The proof for (ii) is similar. ⊠

Proposition 5.5. *Let V be a real vector space and f, g, h alternating multilinear functions on V of degrees k, ℓ, m respectively. Then*

$$(f \wedge g) \wedge h = f \wedge (g \wedge h).$$

Proof. By the definition of the wedge product, we have

$$\begin{aligned} (f \wedge g) \wedge h &= \frac{1}{(k+\ell)!m!} A((f \wedge g) \otimes h) \\ &= \frac{1}{(k_\ell)!m!} \frac{1}{m!\ell!} A(A(f \otimes g) \otimes h) \\ &= \frac{(k+\ell)!}{(k+\ell)!m!k!\ell!} A((f \otimes g) \otimes h) \\ &= \frac{1}{k!\ell!m!} A((f \otimes g) \otimes h). \end{aligned}$$

Similarly, since the tensor product is associative, we conclude that $(f \wedge g) \wedge h = f \wedge (g \wedge h)$. ⊠

Corollary 5.3. *Let V be a real vector space and f, g, h be alternating multilinear functions on V of degrees k, ℓ, m respectively. Then*

$$f \wedge g \wedge h = \frac{1}{k!\ell!m!} A(f \otimes g \otimes h).$$

More generally, if $f_i \in A_{d_i}(V)$, then

$$f_1 \wedge \cdots \wedge f_r = \frac{1}{(d_1)! \cdots (d_r)!} A(f_1 \otimes \cdots \otimes f_r).$$

Proposition 5.6. *Let $[b^i_j]$ denote the matrix whose (i,j)-entry is b^i_j. If $\alpha^1, \cdots, \alpha^k$ are linear functions on a vector space V and $v_1, \cdots, v_k \in V$, then*

$$(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \cdots, v_k) = \det[\alpha^i(v_j)].$$

Proof.

$$\begin{aligned}(\alpha^1 \wedge \cdots \wedge \alpha^k)(v_1, \cdots, v_k) &= A(\alpha^1 \otimes \cdots \otimes \alpha^k)(v_1, \cdots, v_k) \\ &= \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \alpha^1(v_{\sigma(1)}) \cdots \alpha^k(v_{\sigma(k)}) \\ &= \det[\alpha^i(v_j)].\end{aligned}$$

⊠

5.10 A Basis for k-Covectors

Let $e_1, \cdots, e_n$ be a basis for a real vector space V, and let $\alpha^1, \cdots, \alpha^n$ be the dual basis for V^*. We introduce the multi-index notation $I = (i_i, \cdots, i_k)$ and write e_I for $(e_{i_1}, \cdots, e_{i_k})$ and α^I for $\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$. A k-linear function f on V is completely determined by its values on all k-tuples e_I. If f is alternating, then it is completely determined by its values on e_I with $1 \le i_1 < \cdots < i_k \le n$; that is, we can just talk about e_I with I in ascending order. Suppose I, J are ascending, then since the wedge of covectors forms the determinant, we have

$$\alpha^I(e_J) = \begin{cases} 1 & \text{if } I = J; \\ 0 & \text{if } I \neq J. \end{cases}$$

Proposition 5.7. *The alternating k-linear functions α^I where $I = (i_1 < \cdots < i_k)$ form a basis for the space $A_k(V)$ of alternating k-linear functions on V.*

Proof. First we show linear independence. Suppose $\sum c_I \alpha^I = 0$, where $c_I \in \mathbb{R}$ and I runs over ascending multi-indices of length k. Applying both sides to e_J, we have $0 = \sum c_I \alpha^I(e_J) = c_J$, since among all the ascending multi-indices I of length k there is only one equal to J. So the α^I are LI.

To show that the α^I span $A_k(V)$, let $f \in A_k(V)$. We claim that $f = \sum f(e_I)\alpha^I$, where I runs over all ascending multi-indices of length k. Let $g = \sum f(e_I)\alpha^I$. By k-linearity and the alternating property, if two covectors agree on all e_J, then they are equal. But

$$g(e_J) = \sum f(e_I)\alpha^I(e_J) = \sum f(e_I)\delta^I_J = f(e_J).$$

So $f = g = \sum f(e_I)\alpha^I$. ⊠

Corollary 5.4. *If a vector space V is of dimension n, then $A_k(V)$ has dimension $\binom{n}{k}$.*

Proof. We obtain an ascending multi-index $I = (i_1 < \cdots < i_k)$ by choosing a subset of k numbers from $1, \cdots, n$: This can be done in $\binom{n}{k}$ ways. ⊠

Corollary 5.5. *If $k > \dim V$, then $A_k(V) = 0$.*

Proof. In $\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k}$, at least two of the factors must be the same, say α. Because α is a 1-covector, $\alpha \wedge \alpha = 0$, so $\alpha^{i_1} \wedge \cdots \wedge \alpha^{i_k} = 0$. ⊠

Lecture 6

Differential Forms on $\mathbb{R}^n$

6.1 Differential 1-Forms and the Differential of a Function

The **cotangent space** to $\mathbb{R}^n$ at p, denoted by $T_p^*(\mathbb{R}^n)$ or $T_p^*\mathbb{R}^n$, is defined as the dual space of $T_p(\mathbb{R}^n)$. So elements are covectors or linear functionals on $T_p(\mathbb{R}^n)$. Similarly, a **covector field** or **differential 1-form** (1-form for short) ω on an open subset U of $\mathbb{R}^n$ is a function that assigns to each point U a covector $\omega_p \in T_p^*(\mathbb{R}^n)$.

From any C^∞ function $f : U \to \mathbb{R}$, we can construct a 1-form df, called the **differential** of f as follows. For $p \in U$ and $X_p \in T_pU$, define $(df)_p(X_p) = X_pf$. Let $x^1, \cdots, x^n$ be the standard coordinates on $\mathbb{R}^n$. Recall that $\{\partial/\partial x^1|_p, \cdots, \partial/\partial x^n|_p\}$ is a basis for the tangent space $T_p(\mathbb{R}^n)$.

Proposition 6.1. *If $x^1, \cdots, x^n$ are the standard coordinates on $\mathbb{R}^n$, then at each point $p \in \mathbb{R}^n$, $\{(dx^1)_p, \cdots, (dx^n)_p\}$ is the basis for the cotangent space $T_p^*(\mathbb{R}^n)$ dual to the basis $\{\partial/\partial x^1|_p, \cdots, \partial/\partial x^n|_p\}$.*

Proof.

$$(dx^i)_p\left(\left.\frac{\partial}{\partial x^j}\right|_p\right) = \left.\frac{\partial}{\partial x^j}\right|_p x^i = \delta^i_j.$$

⊠

If ω is a 1-form on an open $U \subseteq \mathbb{R}^n$, then at each point $p \in U$ we have a linear combination $\omega_p = \sum a_i(p)(dx^i)_p$ for some $a_i(p) \in \mathbb{R}$, since the $(dx^i)_p$ form a basis for the cotangent space $T_p^*(\mathbb{R}^n)$. As p varies over U, the coefficients a_i become functions on U, and we write $\omega = \sum a_i dx^i$. The covector field ω is said to be C^∞ on U if the coefficient functions a_i are all C^∞ on U. If x, y, and z are coordinates on $\mathbb{R}^3$, then dx, dy, and dz are 1-forms on $\mathbb{R}^3$.

Proposition 6.2. *If $f : U \to \mathbb{R}$ is a C^∞ function on an open $U \subseteq \mathbb{R}^n$, then $df = \sum \frac{\partial f}{\partial x^i} dx^i$.*

Proof. At each point $p \in U$, we have $(df)_p = \sum a_i(p)(dx^i)_p$ for some constants $a_i(p)$ depending on p. So $df = \sum a_i dx^i$ for some functions a_i on U. To evaluate a_j, apply both sides of the first expression to the coordinate vector field $\partial/\partial x^j$; that is,

$$df\left(\frac{\partial f}{\partial x^j}\right) = \sum_i a_i dx^i \left(\frac{\partial}{\partial x^j}\right) = \sum_i a_i \delta^i_j = a_j.$$

On the other hand, $df\left(\frac{\partial}{\partial x^j}\right) = \frac{\partial f}{\partial x^i}$ by the definition of the differential. ⊠

So if f is C^∞, df is also C^∞.

Example 6.1. Every tangent vector $X_p \in T_p(\mathbb{R}^n)$is a linear combination of the standard basis vectors: $X_p = \sum_i b^i(X_p) \left.\frac{\partial}{\partial x^i}\right|_p$.

We have seen that at each $p \in \mathbb{R}^n$, we have $b^i(X_p) = (dx^i)_p(X_p)$. So the coefficient b^i of a vector WRT the standard basis $\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^n}$ is none other than the dual form dx^i on $\mathbb{R}^n$.

6.2 Differential k-Forms

In general, a **differential form ω of degree k** or a **k-form** on an open $U \subseteq \mathbb{R}^n$ is a function that assigns to each $p \in U$ an alternating k-linear function on $T_p(\mathbb{R}^n)$, i.e., $\omega_p \in A_k(T_p\mathbb{R}^n)$. Since $A_1(T_p\mathbb{R}^n) = T_p^*(\mathbb{R}^n)$, this generalizes the idea of a 1-form.

Part III
Manifolds

Lecture 7

Manifolds

7.1 Topological Manifolds

Definition 7.1. A topological space M is **locally Euclidian of dimension n** if every point p in M has a neighborhood U such that there is a homeomorphism ϕ from U onto an open subset of $\mathbb{R}^n$. We call the pair $(U, \phi: U \to \mathbb{R}^n)$ a **chart**, U a **coordinate neighborhood** or a **coordinate open set**, and ϕ a **coordinate map** or a **coordinate system** on U. We say a chart (U, ϕ) is **centered** at $p \in U$ if $\phi(p) = 0$. A **chart** (U, ϕ) **about p** simply means that (U, ϕ) is a chart and $p \in U$.

Definition 7.2. A **topological manifold of dimension n** is a Hausdorff, second countable, locally Euclidian space of dimension n.

For this concept to be well defined, we need to show that $\mathbb{R}^n$ and $\mathbb{R}^m$ are not homeomorphic for $n \neq m$. This is difficult in general (uses homology) but easier for *smooth* manifolds, what we're interested in. Usually this refers to a connected manifold, since manifolds with multiple connected components can have different dimensions for each.

Example 7.1. Euclidian space $\mathbb{R}^n$ is covered by a single chart $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$. Every open subset of $\mathbb{R}^n$ is also a topological manifold, with chart $(U, 1_U)$.

Example 7.2. The graph of $y = x^{2/3}$ in $\mathbb{R}^2$ is a topological manifold. Since it's a subspace of $\mathbb{R}^2$, it's T_2 and second countable. It's also locally Euclidian since it's homeomorphic to $\mathbb{R}$ via $(x, x^{2/3}) \mapsto x$.

7.2 Compatible Charts

Definition 7.3. Two charts $(U, \phi: U \to \mathbb{R}^n), (V, \psi: V \to \mathbb{R}^n)$ of a topological manifold are **$\mathbf{C}^\infty$-compatible** if the two maps

$$\phi \circ \psi^{-1}: \psi(U \cap V) \to \phi(U \cap V), \quad \psi \circ \phi^{-1}: \phi(U \cap V) \to \psi(U \cap V)$$

are C^∞. These two maps are called the **transition functions** between the charts. If $U \cap V$ is empty, then two charts are automatically C^∞-compatible. To simplify notation, we often write $U_{\alpha\beta}$ for $U_\alpha \cap U_\beta$ and $U_{\alpha\beta\gamma}$ for $U_\alpha \cap U_\beta \cap U_\gamma$. Non C^∞-compatible charts are not interesting, so we omit the C^∞ and only speak of compatible charts.

Definition 7.4. A C^∞ atlas or simply an **atlas** on a locally Euclidean space M is a collection $\{(U_\alpha, \phi_\alpha)\}$ of C^∞ compatible charts that cover M, i.e., such that $M = \bigcup_\alpha U_\alpha$.

Although C^∞ compatibility of charts is reflexive and symmetric, it's not transitive. This intuitively is the case since a triple intersection can be small, and two charts being compatible could miss an area of the third. We say a chart (V, ψ) is **compatible with an atlas** $\{(U_\alpha, \phi_\alpha)\}$ if it is compatible with all the charts (U_α, ϕ_α) of the atlas.

Lemma 7.1. *Let $\{(U_\alpha, \phi_\alpha)\}$ be an atlas on a locally Euclidian space. If two charts (V, ψ) and (W, σ) are both compatible with the atlas $\{(U_\alpha, \phi_\alpha)\}$ then they are compatible with each other.*

Proof. Let $p \in V \cap W$. We want to show that $\sigma \circ \psi^{-1}$ is C^∞ at $\psi(p)$. Since $\{(U_\alpha, \phi_\alpha)\}$ is an atlas for M, $p \in U_\alpha$ for some α. Then p is in the triple intersection $V \cap W \cap U_\alpha$. We have that $\sigma \circ \psi^{-1} = (\sigma \circ \phi_\alpha^{-1}) \circ (\phi_\alpha \circ \psi^{-1})$ is C^∞ on $\psi(V \cap W \cap U_\alpha)$, hence at $\psi(p)$. So $\sigma \circ \psi^{-1}$ is C^∞ on $\psi(V \cap W)$, and similarly $\psi \circ \sigma^{-1}$ is C^∞ on $\sigma(V \cap W)$. ⊠

7.3 Smooth Manifolds

An atlas $\mathfrak{A}$ on a locally Euclidian spsace is said to be **maximal** if it is not contained in a larger atlas; in other words, if $\mathfrak{M}$ is any other atlas containing $\mathfrak{A}$, then $\mathfrak{M} = \mathfrak{A}$.

Definition 7.5. A **smooth** or C^∞ manifold is a topological manifold M together with a maximal atlas. The maximal atlas is also called a **differentiable structure** on M. A manifold is said to have dimension n if all of its connected components have dimension n.

We will eventually prove that if an open set $U \subseteq \mathbb{R}^n$ is diffeomorphic to an open set $V \subseteq \mathbb{R}^m$, then $m = n$. So the dimension of a manifold is well-defined.

Usually we don't have to exhibit a maximal atlas to put a smooth structure on a manifold. The existence of *any* atlas will do, actually.

Proposition 7.1. *Any atlas $\mathfrak{A} = \{(U_\alpha, \phi_\alpha)\}$ on a locally Euclidian space is contained in a unique maximal atlas.*

Proof. Adjoin to $\mathfrak{A}$ all charts (V_i, ψ_i) compatible with $\mathfrak{A}$. By Lemma 7.1, the charts (V_i, ψ_i) are all compatible with each other, so the enlarged collection is an atlas. Any chart compatible with this new atlas must be compatible with $\mathfrak{A}$, and so by construction belongs to the new atlas. So the new atlas is maximal.

Let $\mathfrak{M}$ be the maximal atlas constructed above. If $\mathfrak{M}'$ is another maximal atlas containing $\mathfrak{A}$, then all charts in $\mathfrak{M}'$ are compatible with $\mathfrak{A}$ and so belong to $\mathfrak{M}$. So $\mathfrak{M}' \subset \mathfrak{M}$, and since both are maximal we have $\mathfrak{M}' = \mathfrak{M}$. ⊠

As a summary, to show a topological space M is a smooth manifold, we must check:

(i) M is Hausdorff and second countable,

(ii) M has a C^∞ atlas (not necessarily maximal).

7.4 Examples of Smooth Manifolds

Example 7.3. Euclidian space $\mathbb{R}^n$ is a smooth manifold with single chart $(\mathbb{R}^n, r^1, \cdots, r^n)$, where the r^i are the standard coordinates on $\mathbb{R}^n$.

Example 7.4. Any open subset V of a manifold M is also a manifold. If $\{(U_\alpha, \phi_\alpha)\}$ is an atlas for M, then $\{(U_\alpha \cap V, \phi_\alpha|_{U_\alpha \cap V})\}$ is an atlas for V, where $\phi_\alpha|_{U_\alpha \cap V} : U_\alpha \cap V \to \mathbb{R}^n$ denotes the restriction of ϕ_α to the subset $U_\alpha \cap V$.

Example 7.5. For $U \subseteq \mathbb{R}^n$ open and $f : U \to \mathbb{R}^m$ a C^∞ function, the *graph* of f is defined as the subspace $\Gamma(f) = \{(x, f(x)) \in U \times \mathbb{R}^m\}$. The maps $\phi : \Gamma(f) \to U,\ (x, f(x)) \mapsto x$ and $1 \times f : U \to \Gamma(f),\ x \mapsto (x, f(x))$ are continuous and homeomorphisms. The graph $\Gamma(f)$ of a C^∞ function $f : U \to \mathbb{R}^m$ has an atlas with a single chart $(\Gamma(f), \phi)$ and is therefore a C^∞ manifold.

Example 7.6. For any two positive integers m and n let $\mathbb{R}^{m\times n}$ be the vector space of all $m \times n$ matrices. Since $\mathbb{R}^{m\times n} \cong \mathbb{R}^{mn}$, we give it the topology of $\mathbb{R}^{mn}$. The **general linear group** $\mathrm{GL}(n, \mathbb{R})$ is defined as

$$\mathrm{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n\times n} \mid \det A \neq 0\} = \det^{-1}(\mathbb{R} \setminus \{0\}).$$

Since $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous, $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}$ and is therefore a manifold.

Lecture 8

Smooth Maps on a Manifold

Definition 8.1. Let M be a smooth n-manifold. A function $f : M \to \mathbb{R}$ is C^∞ or **smooth at a point** $p \in M$ if there is a chart (U, ϕ) containing p in the atlas of M such that $f \circ \phi^{-1}$, which is defined on the open subset $\phi(U)$ of $\mathbb{R}^n$, is C^∞ at $\phi(p)$. This definition is independent of the specific chart (U,ϕ), for if (V,ψ) is another chart in the atlas containing p, then on $\psi(U\cap V)$ we have

$$f\circ\psi^{-1}=(f\circ\phi^{-1})\circ(\phi\circ\psi^{-1}),$$

which is C^∞ at $\psi(p)$. We say f is C^∞ on M if it is smooth at every point of M.

Definition 8.2. Let $F\colon N\to M$ be a map and h a function on M. The **pullback** of h by F, denoted F^*h, is the composite function $h\circ F$.

So a function f on M is C^∞ on a chart (U,ϕ) if its pulllback by ϕ^{-1} is C^∞ on the subset $\phi(U)$ of a Euclidian space.

Definition 8.3. Let N and M be manifolds of dimension n and m. A map $F\colon N\to M$ is C^∞ at a point p in N if there is a chart (V,ψ) in M containing $F(p)$ and a chart (U,ϕ) in N containing p such that the composition $\psi\circ F\circ\phi^{-1}$, a map from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^m$, is C^∞ at $\phi(p)$. Since F is continuous, we can always choose U small enough such that $F(U)\subseteq V$.

Definition 8.4. The map $F\colon N\to M$ is said to be smooth if it is smooth at every point of N. It is a **diffeomorphism** if it is bijective and both F and its inverse F^{-1} are smooth.

Part IV
The Tangent Space

Lecture 9

Tangent Space

9.1 The Tangent Space at a Point

We have two ways to define tangent vectors, either by column vectors, or point-derivations of C_p^∞, the algebra of germs of C^∞ functions at p. These both generalize to manifolds: for the first approach, define a tangent vector at p by first choosing a chart (U,ϕ) around p, then defining a tangent vector to be an arrow at $\phi(p)$ in $\phi(U)$. While this may be intuitive, different charts give rise to different arrows, and reconciling these differences gets messy. The clean definition is by point derivations. $C_p^\infty(M)$ is defined the same way as $C_p^\infty(\mathbb{R}^n)$, and a **point-derivation** of $C_p^\infty(M)$ is a linear map $D\colon C_p^\infty(M)\to\mathbb{R}$ that satisfies

$$D(fg)=(Df)g(p)+f(p)Dg.$$

Definition 9.1. A **tangent vector** at a point $p\in M$ is a derivation at p. The set of tangent vectors at p form the **tangent space** of M at p, denoted $T_p(M)$ or T_pM.

Given a coordinate neighborhood $(U,\phi)=(U,x^1,\cdots,x^n)$ about a point p in a manifold M, recall the new definition of a partial derivative, where

$$\left.\frac{\partial}{\partial x^i}\right|_p f:=\frac{\partial f}{\partial x^i}(p)=\frac{\partial(f\circ\phi^{-1})}{\partial r^i}(\phi(p)).$$

This is if $x^i=r^i\circ\phi\colon U\to\mathbb{R}$. Furthermore, $\partial_{x^i}|_pf=\partial_{r^i}|_{\phi(p)}f\circ\phi^{-1}\in\mathbb{R}$.

9.2 The Differential of a Map

Let $F\colon N\to M$ be a smooth map between manifolds. At each $p\in N$, F induces a linear map of tangent spaces $dF\colon T_pN\to T_{F(p)}M$, called its **differential** at p. The differential is defined as follows: If $X_p\in T_pN$, then $dF(X_p)$ is the tangent vector in $T_{F(p)}M$ defined by

$$(dF(X_p))f=X_p(f\circ F)\in\mathbb{R}$$