18.101 Analysis II

Notes for part 2

Lecture 1. ODE's

Let U be an open subset of $\mathbb{R}^n$ and $g_i : U \to \mathbb{R}$ $i = 1, \dots, n$, C^1 functions. The ODE's that we will be interested in in this section are $n \times n$ systems of first order differential equations

$$\frac{dx_i}{dt} = g_i(x_1(t), \dots, x_n(t)), \quad i = 1, \dots, n \tag{1.1}$$

where the $x_i(t)$'s are C^1 functions on an open interval, I, of the real line. We will call $x(t) = x_1(t), \dots, x_n(t)$ an *integral curve* of (1.1), and for the equation to make sense we'll require that this be a curve in U. We will also frequently rewrite (1.1) in the more compact vector form

$$\frac{dx}{dt}(t) = g(x(t)) \tag{1.2}$$

where $g = (g_1, \dots, g_n)$. The questions we'll be investigating below are:

Existence. Let t_0 be a point on the interval, I, and x_0 a point in U. Does there exist an integral curve $x(t)$ with prescribed initial data, $x(t_0) = x_0$?

Uniqueness. If $x(t)$ and $y(t)$ are integral curves and $x(t_0) = x_0 = y_0 = y(t_0)$, does $x(t) = y(t)$ for all $t \in I$?

Today we'll concentrate on the uniqueness issue and take up the more complicated existence issue in the next lecture. We'll begin by recalling a result which we proved earlier in the semester.

Theorem 1.1. *Suppose U is convex and the derivatives of g satisfy bounds*

$$\left| \frac{\partial g_i}{\partial x_j}(x) \right| \leq C . \tag{1.3}$$

Then for all x and y in U

$$|g(x) - g(y)| \leq nC|x - y| . \tag{1.4}$$

Proof. To prove this it suffices to prove that for all i

$$|g_i(x) - g_i(y)| \leq nC|x - y| .$$

By the mean value theorem there exists a point, c, on the line joining x to y such that

$$g_i(x) - g_i(y) = Dg_i(c)(x - y)$$

hence

$$|g_i(x) - g_i(y)| \leq n|Dg_i(c)(x-y)|$$

and hence by (1.3)

$$|g_i(x) - g_i(y)| \leq nC|x-y|\,.$$

$\square$

Remark. It is easy to rewrite this inequality with sup norms replaced by Euclidean norms. Namely

$$\|g(x) - g(y)\| \leq n|g(x) - g(y)|$$

and $|x-y| \leq \|x-y\|$. Hence by (1.4):

$$\|g(x) - g(y)\| \leq L\|y-x\| \tag{1.5}$$

where $L = n^2C$. We will use this estimate to prove the following.

Theorem 1.2. *Let $x(t)$ and $y(t)$ be two solutions of (1.2) and let t_0 be a point on the interval I. Then for all $t \in I$*

$$\|x(t) - y(t)\| \leq e^{L|t-t_0|}\|x(t_0) - y(t_0)\|\,. \tag{1.6}$$

Remarks.

1. This result says in particular that if $x(t_0) = y(t_0)$ then $x(t) = y(t)$ and hence proves the uniqueness assertion that we stated above.

2. Let I_1 be a bounded subinterval of I. Then from (1.6) one easily deduces

Theorem 1.3. *For every $\epsilon > 0$ there exists a $\delta > 0$ such that if $\|x(t_0) - y(t_0)\|$ is less than δ, then $\|x(t) - y(t)\|$ is less than ϵ on the interval, $t \in I_1$.*

In other words the solution, $x(t)$, depends continuously on the initial data, $x_0 = x(t_0)$.

To prove Theorem 1.2 we need the following $1-D$ calculus lemma.

Lemma 1.4. *Let $\sigma : I \to \mathbb{R}$ be a C^1 function. Suppose*

$$\left|\frac{d\sigma}{dt}\right| \leq 2L\sigma(t) \tag{1.7}$$

on the interval I. Let t_0 be a fixed point on this interval. Then for all $t \in I$

$$\sigma(t) \leq e^{2L|t-t_0|}\sigma(t_0)\,. \tag{1.8}$$

Proof. First assume that $t > t_0$. Differentiating $\sigma(t)e^{-2Lt}$ we get

$$\begin{aligned}\frac{d}{dt}\sigma(t)e^{-2Lt} =& \frac{d\sigma}{dt}e^{-2Lt} - 2L\sigma(t)e^{-2Lt}\\ =& \left(\frac{d\sigma}{dt} - 2L\sigma(t)\right)e^{-2Lt}\,.\end{aligned}$$

But the estimate (1.6) implies that the right hand is less than or equal to zero so $\sigma(t)e^{-2Lt}$ is decreasing, i.e.

$$\sigma(t_0)e^{-2Lt_0} \geq \sigma(t)e^{-2Lt}\,.$$

Hence

$$\sigma(t_0)e^{2L(t-t_0)} \geq \sigma(t)\,.$$

Suppose now that $t < t_0$. Let I_1 be the interval: $s \in I_1 \Leftrightarrow -s \in I$ and let $\sigma_1 : I_1 \to \mathbb{R}$ be the function $\sigma(s) = \sigma(-s)$. Then

$$\frac{d\sigma_1}{ds}(s) = -\frac{d\sigma}{ds}(-s) \leq 2L\sigma(-s) = 2L\sigma_1(s)\,.$$

Thus with $s_0 = -t_0$

$$\sigma_1(s) \leq e^{2L(s-s_0)}\sigma_1(s_0)$$

for $s > s_0$. Thus if we substitute $-t$ for s this inequality becomes:

$$\sigma(t) \leq e^{2L|t-t_0|}\sigma(t_0)$$

for $t < t_0$. Q.E.D.

We'll now prove Theorem 1.2.

Let

$$\begin{aligned}\sigma(t) &= \|x(t) - y(t)\|^2\\ &= (x(t) - y(t)) \cdot (x(t) - y(t))\,.\end{aligned}$$

Then

$$\begin{aligned}\frac{d\sigma}{dt} &= 2\left(\frac{dx}{dt} - \frac{dy}{dt}\right) \cdot (x(t) - y(t))\\ &= 2(g(x(t)) - g(y(t)) \cdot (x(t) - y(t))\end{aligned}$$

so

$$\begin{aligned}\left|\frac{d\sigma}{dt}\right| &\leq 2\|g(x(t)) - g(y(t))\|\,\|x(t) - y(t)\|\\ &\leq 2L\|x(t) - y(t)\|\,\|x(t) - y(t)\|\\ &\leq 2L\sigma(t)\end{aligned}$$

by Schwarz's inequality and the estimate, (1.5). Now apply Lemma 1.4.

$\square$

For every $p \in U$ and $t \in I_1$, let $x(p,t)$ be the unique solution of equation 1.2 with $x(p,t_0) = p$.

By Theorem 1.3 $x(p,t)$ is continuous in p; and, in fact, it is easy to see that it is continuous in both p and t. To see this let's assume $|g(x)| \le M$ on U and note that for t and t_1 on the interval, I_1,

$$x(p,t_1) - x(p,t) = \int_t^{t_1} g(x(p,s))\, ds$$

and hence (see (2.2) below)

$$|x(p,t_1) - x(p,t)| \le M|t_1 - t|\,.$$

We'll conclude this section by pointing out a couple obvious but useful facts about solutions of (1.2).

Fact 1. If $x(t)$ is a solution of (1.2) on the interval I and a is any point on the real line then $x(t-a)$ is a solution of (1.2) on the interval $I_a = \{t \in \mathbb{R}\,,\, t + a \in I\}$.

Proof. Substitute $x(t-a)$ for $x(t)$ in (1.2) and note that $\frac{d}{dt}(x(t-a)) = \frac{dx}{dt}(t-a)$.

Fact 2. Let I and J be open intervals on the real line and $x(t)$, $t \in I$, and $y(t)$, $t \in J$, integral curves of (1.2). Suppose that for some $t_0 \in I \cap J$, $x(t_0) = y(t_0)$. Then

$$x(t) = y(t) \quad \text{on} \quad I \cap J\,, \tag{1.9}$$

and the curve

$$z(t) = \begin{cases} x(t), & t \in I \\ y(t), & t \in J \end{cases} \tag{1.10}$$

is an integral curve of (1.2) on the interval $I \cup J$.

Proof. Our uniqueness result implies (1.7) and since $x(t)$ and $y(t)$ are C^1 curves which coincide on $I \cap J$ the curve (1.10) is C^1. Moreover since it satisfies (1.2) on I and J it satisfies (1.2) on $I \cup J$.

Exercises.

1. a. Describe all solutions of the system of first order ODEs.

$$\begin{aligned} \frac{dx_1}{dt}(t) &= x_2(t) \\ \frac{dx_2}{dt}(t) &= -x_1(t) \end{aligned} \tag{I}$$

on the interval $-\infty < t < \infty$.

Hint: Show that $x_1(t)$ and $x_2(t)$ have to satisfy the second order ODE

$$\frac{d^2u}{dt}(t) + u(t) = 0\,. \tag{II}$$

b. Conversely show that if $u(t)$ is a solution of this equation, then setting $x_1(t) = u(t)$ and $x_2(t) = \frac{du}{dt}$ we get a solution of (I).

2. Let $F : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function and let $u(t) \in C^k(\mathbb{R})$ be a solution of the k^{th} order ODE

$$\frac{d^k u}{dt^k} = F\left(u, \frac{du}{dt}, \cdots, \frac{d^{k-1}u}{dt^{k-1}}\right)\,. \tag{1.11}$$

Show that if one sets

$$\begin{aligned} x_1(t) &= u(t) \\ x_2(t) &= \frac{du}{dt} \\ &\cdots \\ x_k(t) &= \frac{d^{k-1}u}{dt}\,. \end{aligned} \tag{1.12}$$

These functions satisfy the $k \times k$ systems of first order ODEs

$$\begin{aligned} \frac{dx_1}{dt} &= x_2(t) \\ \frac{dx_2}{dt} &= x_1(t) \\ &\cdots \\ \frac{dx_k}{dt} &= F(x_1(t), \ldots, x_k(t))\,. \end{aligned} \tag{1.13}$$

Conversely show that any solution of (1.13) can be converted into a solution of (1.11) with the properties (1.12).

3. Let $H(x, y)$ be a C^2 function on $\mathbb{R}^2$ and let $(x(t), y(t))$ be a solution of the Hamilton-Jacobi equations

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial}{\partial y} H(x(t), y(t)) \\ \frac{dy}{dt} &= -\frac{\partial}{\partial x} H(x(t), y(t))\,. \end{aligned} \tag{1.14}$$

Show that the function, H, is constant along the integral curves of (1.14).

4. Describe all solutions of (1.14) for the harmonic oscillator Hamiltonian:

$$H(x, y) = x^2 + y^2 .$$

5. Let $H(x, y) = \frac{y^2}{2m} + V(x)$ where V is in $\mathcal{C}^\infty(\mathbb{R})$. Show that every solution of (1.14) gives rise to a solution of Newton's equation

$$m\frac{d^2x}{dt^2} = -\frac{dV}{dx}(x(t))$$

where m is mass, $\frac{d^2x}{dt^2}$ is acceleration and V is potential energy.

6. Show that the functions

$$F(x, y, z) = x + y + z$$

and

$$G(x, y, z) = x^2 + y^2 + z^2$$

are *integrals* of the system of equations

$$\text{(III)} \qquad \begin{aligned} \frac{dx}{dt} &= y - z \\ \frac{dy}{dt} &= z - x \\ \frac{dz}{dt} &= x - y \end{aligned}$$

i.e., on any solution curve $(x(t), y(t), z(t))$ these functions are constant. Interpret this result geometrically

Lecture 2. The Picard theorem

Our goal in this lecture will be to prove a local existence theorem for the system of OED's (1.1). In the course of the proof we will need the following 1-D integral calculus result. Let $x(t) \in \mathbb{R}^n$, $a \le t \le b$, be a continuous curve. Then

$$\left| \int_a^b x(t)\, dt \right| \le \int_a^b |x(t)|\, dt . \tag{2.2}$$

Proof. Since $|x(t)| = \sup |x_i(t)|$, $1 \leq i \leq n$, it suffices to prove

$$\left| \int_a^b x_i(t)\, dt \right| \leq \int_a^b |x_i(t)|\, dt$$

and for the proof of this see §13 in Munkres.

$\square$

Now let U be an open subset of $\mathbb{R}^n$ and $g : U \to \mathbb{R}^n$ a C^1 map. Given a point $x_0 \in U$ we will let $B_\epsilon(x_0)$ be the closed rectangle: $|x - x_0| \leq \epsilon$. For ϵ sufficiently small this rectangle is contained in U. Let $M = \sup\{|g(x)|, x \in B_\epsilon(x_0)\}$ and let

$$0 < T < \frac{\epsilon}{M}\,. \tag{2.3}$$

We will prove

Theorem 2.1. *There exists an integral curve, $x(t)$, $-T \leq t \leq T$ of (1.2) with $x(t) \in B_\epsilon(x_0)$ and $x(0) = x_0$.*

The proof will be by a procedure known as *Picard iteration.* Given a C^1 curve $x(t)$, $-T \leq t \leq T$ on the rectangle $B_\epsilon(x_0)$ we will define its Picard iterate to be the curve

$$\tilde{x}(t) = x_0 + \int_0^t g(x(s))\, ds\,. \tag{2.4}$$

The estimates (2.2) and (2.3) imply that

$$|\tilde{x}(t) - x_0| \leq \int_0^t |g(x(s))|\, ds \leq M|t| \leq \epsilon \tag{2.5}$$

for $|t| \leq T$ so this curve is well-defined on the interval $-T \leq t \leq T$ and is also contained in $B_\epsilon(x_0)$. Let's now define by induction a sequence of curves $x_k(t)$, $-T \leq t \leq T$, by letting $x_0(t)$ be the constant curve $x(t) = x_0$ and letting

$$x_k(t) = x_0 + \int_0^t g(x_{k-1}(s))\, ds, -T \leq t \leq T \tag{2.6}$$

(i.e., by letting $x_k(t)$ be the Picard iterate of $x_{k-1}(t)$). We will prove below that these curves converge to a solution curve of (1.2). To prove this we'll need estimates for $|x_k(t) - x_{k-1}(t)|$. Let

$$C = \sup_{i,j} \left| \frac{\partial g_i}{\partial x_j} \right|, \quad x \in B_\epsilon(x_0)$$

then by (1.4) one has the estimate

$$|g(x) - g(y)| \leq L|x - y| \tag{2.7}$$

for x and y in $B_\epsilon(x_0)$ where $L = nC$. We will show by induction that

$$|x_k(t) - x_{k-1}(t)| \leq \frac{ML^{k-1}|t|^k}{k!}\,. \tag{2.6k}$$

Proof. To get the induction started observe that

$$x_1(t) = x_0 + \int_0^t g(x_0)\,ds = x_0 + tg(x_0)$$

so

$$|x_1(t) - x_0(t)| = |x_1(t) - x_0(t) \le |t|M\,.$$

Let's now prove that $(2.6)_{k-1}$ implies $(2.6)_k$. By (2.4)

$$x_k(t) - x_{k-1}(t) = \int_0^t (g(x_{k-1}(s)) - g(x_{k-2}(s)))\,ds\,.$$

Thus by (2.7) and (2.2) one has, for $t > 0$,

$$\begin{aligned} |x_k(t) - x_{k-1}(t)| &\le \int_0^t |g(x_{k-1}(s)) - g(x_{k-2}(s))|\,ds \\ &\le L \int_0^t |x_{k-1}(s) - x_{k-2}(s)|\,ds \end{aligned}$$

and hence by $(2.6)_{k-1}$

$$\begin{aligned} |x_k(t) - x_{k-1}(t)| &\le ML^{k-1} \int_0^t \frac{s^{k-1}}{(k-1)!}\,ds \\ &\le \frac{ML^{k-1}t^k}{k!} \end{aligned}$$

and with a change of sign the same argument works for $t < 0$.

□

We'll next show that as k tends to infinity $x_k(t)$ converges uniformly on the interval, $-T \le t \le T$. To see this, note that the series

$$\sum_1^\infty |x_i(t) - x_{i-1}(t)|$$

is majorized by the series

$$\frac{M}{L} \sum \frac{L^k |t|^k}{k!} = \frac{M}{L} e^{L|t|}\,.$$

Hence the sums

$$x_k(t) = \sum_{i=1}^k x_i(t) - x_{i-1}(t) + x_0$$

converge uniformly to a continuous limit

$$x(t) = \mathrm{Lim}_{\,k\to\infty}\, x_k(t)$$

as claimed. Now note that since g is a continuous function of x we can let k tend to infinity on both sides of (2.6), and this gives is in the limit the integral identity

$$x(t) = x_0 + \int_0^t g(x(s))\, ds\,. \tag{2.8}$$

Moreover since $x(t)$ is continuous the second term on the right is the anti-derivation of a continuous function and hence is C^1. Thus $x(t)$ is C^1, and we can differentiate both sides of (2.8) to get

$$\frac{dx}{dt} = g(x(t))\,. \tag{2.9}$$

Also, by (2.8), $x(0) = x_0$, so this proves the assertion above: that $x(t) = \text{Lim}\, x_k(t)$ satisfies (1.2) with initial data $x(0) = x_0$. Q.E.D.

We'll next show that we can, with very little effort, make a number of cosmetic improvements on this result.

Remark 1. For any $a \in \mathbb{R}$ there exists a solution $x_a(t)$, $-T + a \leq t \leq T + a$ of (1.2) with $x(a) = x_0$.

Proof. Replace the solution, $x(t)$, that we've just constructed by $x_a(t) = x(t-a)$. As was pointed out in Lecture 1, this is also a solution of (1.2).

□

Remark 2. If g is C^k the solution of (1.2) constructed above is C^{k+1},

Proof. (by induction on k)
We've already observed that $x(t)$ is C^1 hence if g is C^1 the second term on the right hand side of (2.8) is the anti-derivative of a C^1 function and hence is C^2. Continue.

□

Remark 3. We have proved that if $0 < T < \frac{\epsilon}{M}$ there exists a solution, $x(t)$, $-T \leq t \leq T$ of (1.2) with $x(0) = x_0$ and $x(t) \in B_\epsilon(x_0)$.
We claim

Theorem 2.2. *For every $p \in B_{\epsilon/2}(x_0)$ there exists a solution*

$$x_p(t)\,, \quad -T/2 \leq t \leq T/2$$

of (1.2) with $x_p(0) = p$ and $x(t) \in B_\epsilon(x_0)$.

Proof. In Theorem 2.1 replace x_0 by p and ϵ by $\epsilon/2$ to conclude that there exists a solution $x_p(t)$, $-T/2 \leq t \leq T/2$ of (1.2) with $x_p(0) = p$ and $x_p(t) \in B_{\epsilon/2}(p)$. Now note that if p is in $B_{\epsilon/2}(x_0)$ then $B_{\epsilon/2}(p)$ is contained in $B_\epsilon(x_0)$.

□

Remark 4. We can convert Remark 3 into a slightly more global result. We claim

Theorem 2.3. *Let W be a compact subset of U. Then for some $T > 0$ there exists, for* every *$q \in W$ a solution $x_q(t)$, $-T \leq t \leq T$ of (1.2) with $x_q(0) = q$.*

Proof. For each $p \in W$ we can, by Theorem 2.2, find a neighborhood, U_p, of p in U and a $T_p > 0$ such that for *every* $q \in U_p$, a solution, $x_q(t)$, $-T_p \leq t \leq T_p$, of (1.2) exists with $x_q(0) = q$. By compactness we can cover W by a finite number, U_{p_i}, $i = 1, \dots, N$ of these U_p's. Hence if we let $T = \min T_{p_i}$ there exists for every $q \in W$ a solution $x_q(t)$, $-T \leq t \leq T$ of (1.2) with $x_q(0) = q$. □

We will conclude this section by making a global application of this result which will play an important role later in this course when we study vector fields on manifolds and the flows they generate.

Definition 2.4. *We will say that a sequence $p_1, p_2, p_3, \dots$ of points in U* tends to infinity *in U if for every compact subset, W, of U there exists an i_0 such that $p_i \in U - W$ for $i > i_0$.*

Now let $x(t) \in U$, $0 \leq t < a$ be an integral curve of (1.2). We will say that $x(t)$ is a *maximal* integral curve if it can't be extended to an integral curve, $x(t)$, $0 \leq t < b$, on a larger interval, $b > a$. We'll prove

Theorem 2.5. *If $x(t)$, $0 \leq t < a$ is a maximal integral curve of (1.2) then either*

(a) $a = +\infty$,

or

(b) There exists a sequence, $t_i \in [0, a)$, $i = 1, 2, 3, \dots$ such that t_i tends to a and $x(t_i)$ tends to infinity in U as i tends to infinity.

Proof. We'll prove: "if not (b) then (a)". Suppose there exists a compact set W such that $x(t)$ is in W for all $t < a$. Let T be a positive number for which the hypotheses of Theorem 2.3 hold (i.e., with the property that for every $q \in W$ there exists an integral curve, $x_q(t)$, $-T \leq t \leq T$ for which $x_q(0) = q$). Let $q = x(a - T/2)$. Then the curve,

$$y(t) = x_q(t - (a - T/2)), \quad a - T \leq t \leq a + \frac{T}{2}$$

is an integral curve with the property,

$$y(a - T/2) = x(a - T/2) = q,$$

hence as we showed in Lecture 1

$$z(t) = \begin{cases} x(t), & t < a \\ y(t), & a - T < t < a + T/2 \end{cases}$$

is an integral curve of v contradicting the maximality of $x(t)$.

□

Remark 5. We will see later on that there are a lot of geometric criteria which prevent scenario (b) from occurring and in these situations we'll be able to conclude: For every point, $x_0 \in U$, there exists a solution, $x(t)$, of (1.2) for $0 \le t < \infty$ with $x(0) = x_0$.

Exercises.

1. Consider the ODE

$$\text{(I)} \qquad \frac{dx}{dt}(t) = x(t)\,, \qquad x(t) \in \mathbb{R}\,.$$

Construct a solution of (I) with $x(0) = x_0$ by Picard iteration. Show that the solution you get coincides with the solution of (I) obtained by elementary calculus techniques, i.e., by differentiating $e^{-t}x(t)$.

2. Solve the 2×2 system of ODE's

$$\text{(II)} \qquad \frac{dx_1}{dt} = x_2(t)$$

$$\frac{dx_2}{dt} = -x_1(t)$$

with $(x_1(0), x_2(0)) = (a_1, a_2)$ by Picard iteration and show that the answer coincides with the answer you obtained by more elementary means in Lecture 1, Exercise 1.

3*. Let A and $X(t)$ be $n \times n$ matrices. Solve the matricial ODE

$$\frac{dX}{dt}(t) = AX(t)\,,$$

with $X(0) =$ Identity, by Picard iteration.

4. Let $x(t) \in U$, $a < t < b$ be an integral curve of (1.2). We'll call $x(t)$ a maximal integral curve if it can't be extended to a larger interval.

(a) Show that if $x(t)$ is a maximal integral curve and $a > -\infty$ then there exists a sequence of points, s_i, $i = 1, 2, 3, \ldots$, on the interval (a, b) such that $\text{Lim}\ s_i = a$ and $x(s_i)$ tends to infinity in U as i tends to infinity.

(b) Show that if $x(t)$ is a maximal integral curve and $b < +\infty$ then there exists a sequence of points, t_i, $i = 1, 2, 3, \ldots$, on the interval (a, b) such that $\text{Lim}\ t_i = b$ and $x(t_i)$ tends to infinity in U as i tends to infinity.

5. Show that every integral curve, $x(t)$, $t \in I$, of (1.2) can be extended to a maximal integral curve of (1.2). *Hint:* Let J be the union of the set of open intervals, $I' \supset I$ to which $x(t)$ can be extended. Show that $x(t)$ can be extended to J.

6. Let $g : U \to \mathbb{R}^n$ be the function on the right hand side of (1.2). Show that if $g(x_0) = 0$ at some point $x_0 \in U$ the constant curve

$$x(t) = x_0 \,, \quad -\infty < t < \infty$$

is the (unique) solution of (1.2) with $x(0) = x_0$.

7. Suppose $g : U \to \mathbb{R}^n$ is compactly supported, i.e., suppose there exists a compact set, W, such that for $x_0 \notin W$, $g(x_0) = 0$. Prove that for every $x_0 \in U$ there exists an integral curve, $x(t)$, $-\infty < t < +\infty$, with $x(0) = x_0$. *Hint:* Exercises 5 and 6.

Lecture 3. Flows

Let U be an open subset of $\mathbb{R}^n$ and $g : U \to \mathbb{R}^n$ a C^1 function. We showed in lecture 2 that given a point, $p_0 \in U$, there exists an $\epsilon > 0$ and a neighborhood, V of p_0 in U such that for every $p \in V$ one has an integral curve, $x(p,t)$, $-\epsilon < t < \epsilon$ of the system of equation (1.2) with $x(p,0) = p$. We also showed that the map, $(p,t) \to x(p,t)$ is a continuous map of $V \times (-\epsilon, \epsilon)$ into U. What we will show in this lecture is that if g is C^{k+1}, $k > 1$, then this map is a C^k map of $V \times (-\epsilon, \epsilon)$ into U. The idea of the proof is to assume for the moment that this assertion is true and see what its implications are.

The details: Fix $1 \le j \le n$ and let

$$y(p,t) = \frac{\partial x}{\partial p_j}(p,t) = \left(\frac{\partial x_1}{\partial p_j}, \cdots, \frac{\partial x_n}{\partial p_j} \right) .$$

Then

$$\begin{aligned}
\frac{d}{dt} y(p,t) &= \frac{d}{dt} \frac{\partial}{\partial p_j} x(p,t) \\
&= \frac{\partial}{\partial p_j} \frac{d}{dt} x(p,t) \\
&= \frac{\partial}{\partial p_j} g(x(p,t)) \\
&= \sum_k \frac{\partial g}{\partial x_k}(x(p,t)) \frac{\partial x_k}{\partial p_j} \\
&= h(x(p,t), y(p,t))
\end{aligned}$$

where

$$h(x,y,t) = \sum \frac{\partial g}{\partial x_k}(x,t)y_k\,. \tag{3.1}$$

This gives us, for each j, a solution $(x(p,t), y(p,t))$ of the $(2n \times 2n)$ first order system of ODE's

$$\begin{aligned} \frac{dx}{dt}(p,t) &= g(x(p,t)) \\ \frac{dy}{dt}(p,t) &= h(x(p,t), y(p,t)) \end{aligned} \tag{3.2}$$

with initial data

$$x(p,0) = p\,, \quad y(p,0) = \frac{\partial x}{\partial p_j}(p,0) = e_j \tag{3.3}$$

the e_j's being the standard basis vectors of $\mathbb{R}^n$. Moreover, since g is C^{k+1}, h is, by (3.1), C^k. Shrinking V and ϵ if necessary, we can, as in lecture 2, solve these equations by Picard iteration. Starting with $x_0(p,t) = p$ and $y_0(p,t) = e_j$ this generates a sequence

$$(x_r(p,t), y_r(p,t))\,, \quad r = 1, 2, 3, \dots\,,$$

and we will prove

Lemma 3.1. *If at stage* $r-1$

$$y_{r-1}(p,t) = \frac{\partial}{\partial p_j} x_{r-1}(p,t)$$

then at stage r*,*

$$y_r(p,t) = \frac{\partial}{\partial p_j} x_r(p,t)\,.$$

Proof. By Picard iteration

$$\begin{aligned} x_r(p,t) &= p + \int_0^t g(x_{r-1}(p,s))\,ds \\ y_r(p,t) &= e_j + \int_0^t h(x_{r-1}(p,s), y_{r-1}(p,s))\,ds\,. \end{aligned} \tag{3.4}$$

Thus

$$\begin{aligned}
\frac{\partial}{\partial p_j} x_r(p,t) &= e_j + \int_0^t Dg(x_{r-1}(p,s)\frac{\partial}{\partial p_j} x_{r-1}(p,s))\,ds \\
&= e_j + \int_0^t Dg(x_{r-1}(p,s)\, y_{r-1}(p,s))\,ds \\
&= e_j + \int_0^t h(x_{r-1}(p,s)\,, y_{r-1}(p,s))\,ds \\
&= y_r(p,s)\,.
\end{aligned}$$

□

We showed in lecture 2 that the sequence $(x_r(p,t), y_r(p,t))$, $r = 0, 1, \ldots$, converges uniformly in $V \times (-\epsilon, \epsilon)$ to a solution $(x, (p,t), y(p,t))$ of (3.2). Moreover, we showed in lecture 1 that this solution is continuous in (p,t). Let's prove that $x(p,t)$ is C^1 in (p,t) by proving

$$\frac{\partial x}{\partial p_j}(p,t) = y(p,t)\,. \tag{3.5}$$

Since $\frac{\partial}{\partial p_j} x_r(p,t) = y_r(p,t)$ this follows from our next result:

Lemma 3.2. *Let U be an open subset of $\mathbb{R}^n$ and $f_r : U \to \mathbb{R}$, $r = 0, 1, \ldots$ a sequence of C^1 functions. Suppose that $f_r(p)$ converges uniformly to a function $f(p)$, and that it's j^{th} derivative, $\frac{\partial f_r}{\partial p_j}$, converges uniformly to a function, $h_j(p)$. Then f is C^1 and*

$$\frac{\partial f}{\partial p_j} = h_j(p)\,. \tag{3.6}$$

Proof. Since the convergence is uniform the functions, f and h_j, are continuous. Moreover for t small

$$\begin{aligned}
f_r(p + te_j) &= \int_0^t \frac{d}{ds} f_r(p + se_j)\,ds \\
&= \int_0^t \frac{\partial f_r}{\partial p_j}(p + se_j)\,ds
\end{aligned}$$

and since the integrand on the right converges uniformly to h_j and f_r converges to f

$$f(p + te_j) = \int_0^t h_j(p + se_j)\,ds$$

and differentiating with respect to t:

$$\frac{\partial f}{\partial p_j}(p) = h_j(p)\,.$$

□

The identity (3.5) shows that the partial derivatives of $x(p,t)$ with respect to the p_j's are continuous in (p,t) and since

$$\frac{\partial}{\partial t} x(p,t) = g(x(p,t))$$

the same is true for the partial with respect to t.

We'll now prove that if g is C^{k+1} the solution, $x(p,t)$, above of the systems of ODE:

$$\frac{dx}{dt}(p,t) = g(x(p,t))\,, \quad x(p,0) = p \tag{3.7}$$

depends C^k on p and t. The argument is bootstrapping argument: We've shown that if g is C^2, $x(p,t)$ is a C^1 function of p and t. This remark applies as well to the system (3.2), with initial data (3.3). If g is C^3 the function h, on the second line of (3.2) is C^2 by (3.1), and hence the result we've just proved shows that the solution, $(x(p,t), y(p,t))$ of this equation is C^1. However, since $y(p,t) = \frac{\partial}{\partial p_j} x(p,t)$, this implies that the solution, $x(p,t)$, of (1.2) is a C^2 function of (p,t). Repeating this argument k times we obtain

Theorem 3.3. *Assume $g : U \to \mathbb{R}^n$ is C^{k+1}. Then, given a point, $p_0 \in U$, there exists a neighborhood, V, of p_0 in U and an $\epsilon > 0$ such that*

> *(a) for every $p \in V$, there is an integral curve, $x(p,t)$, $-\epsilon < t < \epsilon$, of (1.2) with $x(p,0) = p$ and*
>
> *(b) the map*
>
> $$(p,t) \in V \times (-\epsilon, \epsilon) \to U\,, \quad (p,t) \to x(p,t)$$
>
> *is a C^k map.*

Let's denote this map by F, i.e., set $F(p,t) = x(p,t)$ and for each $t \in (-\epsilon, \epsilon)$ let

$$f_t : V \to U \tag{3.8}$$

be the map: $f_t(p) = F(p,t)$. We will call the family of maps, f_t, $-\epsilon < t < \epsilon$ *the flow* generated by the system of equation (1.2) and leave for you to prove the following two propositions.

Proposition 3.4. *Let t be a point on the interval $(-\epsilon, \epsilon)$ and let W be a neighborhood of p_0 in V with the property $f_t(W) \subset V$. Then for all $|s| < \epsilon - |t|$*

$$f_s \circ f_t = f_{s+t}\,. \tag{3.9}$$

Proposition 3.5. *In Proposition (3.4) let $|t|$ be less than $\frac{\epsilon}{2}$ and let $W_t = f_t(W)$. Then*

(a) *W_t is an open subset of V.*

(b) *$f_t : W \to W_t$ is a C^k diffeomorphism.*

(c) *The map $f_t^{-1} : W_t \to W$ is the restriction of f_{-t} to W_t.*

Theorem 3.3 can be slightly strengthened. Namely let A be a compact subset of U. Then we claim

Theorem 3.6. *There exists a neighborhood of W of A in U and an $\epsilon > 0$ such that*

(a) *for every $p \in W$ there is an integral curve, $x(p,t)$, $-\epsilon < t < \epsilon$, of (1.2) with $x(p,0) = p$*

(b) *the map*

$$(p,t) \in W \times (-\epsilon, \epsilon) \to U$$

is a C^k map.

The proof of this is essentially identical with the proof of theorem (2.3). Namely for every $q \in A$ we can, by Theorem 3.3, find a neighborhood, V_q, of q and an $\epsilon_q > 0$ such that for every $p \in V_q$ there exists an integral curve $x(p,t)$, $-\epsilon_q < t < \epsilon_q$ of (1.2) with $x(p,0) = p$. By compactness we can cover A by a finite number, U_{q_i}, $i = 1, \dots, N$, of these U_q's and we get a proof of (a)–(b) by taking W to be the union of the U_{q_i}'s and ϵ the minmum of the ϵ_{q_i}'s.

Exercises.

1. In the proof of Lemma 3.2 we quoted without proof the well-known theorem:

Theorem. *Let U be an open subset of $\mathbb{R}^n$ and $f_r : U \to \mathbb{R}$, $r = 1, 2, \dots,$ a sequence of continuous functions. If f_r converges uniformly on U the function*

$$f = \mathrm{Lim}\ f_r$$

is continuous.

Prove this.

2. Prove Proposition 3.4. *Hint:* For $p \in W$ and $q = f_t(p)$ show that the curve,

$$\gamma_1(s) = f_s(q)\,, \quad -(\epsilon - |t|) < s < \epsilon - |t|$$

is an integral curve of the system (1.2) and that it coincides with the integral curve

$$\gamma_2(s) = f_{s+t}(p)\,, \quad -(\epsilon - |t|) < s < e - |t|\,.$$

3. Prove Proposition 3.5.

Lecture 4. Vector fields

In this lecture we'll reformulate the theorems about ODEs that we've been discussing in the last few lectures in the language of vector fields.

First a few definitions. Given $p \in \mathbb{R}^n$ we define the tangent space to $\mathbb{R}^n$ at p to be the set of pairs

$$T_p\mathbb{R}^n = \{(p, \mathrm{v})\}\,; \quad \mathrm{v} \in \mathbb{R}^n\,. \tag{4.1}$$

The identification

$$T_p\mathbb{R}^n \to \mathbb{R}^n\,, \quad (p, \mathrm{v}) \to \mathrm{v} \tag{4.2}$$

makes $T_p\mathbb{R}^n$ into a vector space. More explicitly, for v, v_1 and $\mathrm{v}_2 \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we define the addition and scalar multiplication operations on $T_p\mathbb{R}^n$ by the recipes

$$(p, \mathrm{v}_1) + (p, \mathrm{v}_2) \;\; = \;\; (p, \mathrm{v}_1 + \mathrm{v}_2)$$

and

$$\lambda(p, \mathrm{v}) \;\; = \;\; (p, \lambda\mathrm{v})\,.$$

Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ a C^1 map. We recall that the derivative

$$Df(p) : \mathbb{R}^n \to \mathbb{R}^m$$

of f at p is the linear map associated with the $m \times n$ matrix

$$\left[\frac{\partial f_i}{\partial x_j}(p)\right]\,.$$

It will be useful to have a "base-pointed" version of this definition as well. Namely, if $q = f(p)$ we will define

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^m$$

to be the map

$$df_p(p, \mathrm{v}) = (q, Df(p)\mathrm{v})\,. \tag{4.3}$$

It's clear from the way we've defined vector space structures on $T_p\mathbb{R}^n$ and $T_q\mathbb{R}^m$ that this map is linear.

Suppose that the image of f is contained in an open set, V, and suppose $g : V \to \mathbb{R}^k$ is a C^1 map. Then the "base-pointed"" version of the chain rule asserts that

$$dg_q \circ df_p = d(f \circ g)_p\,. \tag{4.4}$$

(This is just an alternative way of writing $Dg(q)Df(p) = D(g \circ f)(p)$.)

The basic objects of 3-dimensional vector calculus are vector fields, a *vector field* being a function which attaches to each point, p, of $\mathbb{R}^3$ a base-pointed arrow, $(p, \vec{\mathrm{v}})$. The n-dimensional generalization of this definition is straight-forward.

Definition 4.1. *Let U be an open subset of $\mathbb{R}^n$. A vector field on U is a function, v, which assigns to each point, p, of U a vector $v(p)$ in $T_p\mathbb{R}^n$.*

Thus a vector field is a vector-valued function, but its value at p is an element of a vector space, $T_p\mathbb{R}^n$ that itself depends on p.

Some examples.

1. Given a fixed vector, $\mathrm{v} \in \mathbb{R}^n$, the function

   $$p \in \mathbb{R}^n \to (p, \mathrm{v}) \tag{4.5}$$

   is a vector field. Vector fields of this type are *constant* vector fields.

2. In particular let $e_i, i = 1, \dots, n$, be the standard basis vectors of $\mathbb{R}^n$. If $\mathrm{v} = e_i$ we will denote the vector field (4.5) by $\partial/\partial x_i$. (The reason for this "derivation notation" will be explained below.)

3. Given a vector field on U and a function, $f : U \to \mathbb{R}$ we'll denote by fv the vector field

   $$p \in U \to f(p)v(p)\,.$$

4. Given vector fields v_1 and v_2 on U, we'll denote by $v_1 + v_2$ the vector field

   $$p \in U \to v_1(p) + v_2(p)\,.$$

5. The vectors, (p, e_i), $i = 1, \dots, n$, are a basis of $T_p\mathbb{R}^n$, so if v is a vector field on U, $v(p)$ can be written uniquely as a linear combination of these vectors with real numbers, $g_i(p)$, $i = 1, \dots, n$, as coefficients. In other words, using the notation in example 2 above, v can be written uniquely as a sum

   $$v = \sum_{i=1}^{n} g_i \frac{\partial}{\partial x_i} \tag{4.6}$$

   where $g_i : U \to \mathbb{R}$ is the function, $p \to g_i(p)$.

We'll say that v is a C^k vector field if the g_i's are in $C^k(U)$.

A basic vector field operation is *Lie differentiation.* If $f \in C^1(U)$ we define $L_v f$ to be the function on U whose value at p is given by

$$Df(p)\mathrm{v} = L_v f(p) \tag{4.7}$$

where $v(p) = (p, \mathrm{v})\,, \quad \mathrm{v} \in \mathbb{R}^n$. If v is the vector field (4.6) then

$$L_v f = \sum g_i \frac{\partial}{\partial x_i} f \tag{4.8}$$

(motivating our "derivation notation" for v).

Exercise.

Check that if $f_i \in C^1(U)$, $i = 1, 2$, then

$$L_v(f_1 f_2) = f_1 L_v f_2 + f_1 L_v f_2\,. \tag{4.9}$$

We now turn to the main object of this lecture: formulating the ODE results of Lectures 1–3 in the language of vector fields.

Definition 4.2. *A C^1 curve $\gamma : (a, b) \to U$ is an integral curve of v if for all $a < t < b$ and $p = \gamma(t)$*

$$\left(p, \frac{d\gamma}{dt}(t)\right) = v(p)$$

i.e.., if v is the vector field (4.6) and $g : U \to \mathbb{R}^n$ is the function $(g_1, \dots, g_n)$ the condition n for $\gamma(t)$ to be an integral curve of v is that it satisfy the system of ODEs

$$\frac{d\gamma}{dt}(t) = g(\gamma(t))\,. \tag{4.10}$$

Hence the ODE results of the previous three lectures, give us the following theorems about integral curves.

Theorem 4.3 (Existence)**.** *Given a point $p_0 \in U$ and $a \in \mathbb{R}$, there exists an interval $I = (a - T, a + T)$, a neighborhood, U_0, of p_0 in U and for every $p \in U_0$ an integral curve, $\gamma_p : I \to U$ with $\gamma_p(a) = p$.*

Theorem 4.4 (Uniqueness)**.** *Let $\gamma_i : I_i \to U$, $i = 1, 2$, be integral curves. If $a \in I_1 \cap I_2$ and $\gamma_1(a) = \gamma_2(a)$ then $\gamma_1 \equiv \gamma_2$ on $I_1 \cap I_2$ and the curve $\gamma : I_1 \cup I_2 \to U$ defined by*

$$\gamma(t) = \begin{cases} \gamma_1(t)\,, & t \in I_1 \\ \gamma_2(t)\,, & t \in I_2 \end{cases}$$

is an integral curve.

Theorem 4.5 (Smooth dependence on initial data)**.** *Let v be a C^{k+1}-vector field, on an open subset, V, of U, $I \subseteq \mathbb{R}$ an open interval, $a \in I$ a point on this interval and $h : V \times I \to U$ a mapping with the properties:*

(i) $h(p, a) = p$.

(ii) For all $p \in V$ the curve

$$\gamma_p : I \to U \qquad \gamma_p(t) = h(p, t)$$

is an integral curve of v. Then the mapping, h, is C^k.

Theorem 4.6. *Let $I = (a, b)$ and for $c \in \mathbb{R}$ let $I_c = (a - c, b - c)$. Then if $\gamma : I \to U$ is an integral curve, the reparameterized curve*

$$\gamma_c : I_c \to U\,, \quad \gamma_c(t) = \gamma(t + c) \tag{4.11}$$

is an integral curve.

Finally we recall that a C^1-function $\varphi : U \to \mathbb{R}$ is an *integral* of the system (4.10) if for every integral curve $\gamma(t)$, the function $t \to \varphi(\gamma(t))$ is constant. This is true if and only if for all t and $p = \gamma(t)$

$$0 = \frac{d}{dt}\varphi(\gamma(t)) = (D\varphi)_p\left(\frac{d\gamma}{dt}\right) = (D\varphi)_p(\mathrm{v})$$

where $(p, \mathrm{v}) = v(p)$. But by (4.6) the term on the right is $L_v\varphi(p)$.

Hence we conclude

Theorem 4.7. *$\varphi \in C^1(U)$ is an integral of the system (4.10) if and only if $L_v\varphi = 0$.*

I'll devote the second half of this lecture to discussing some properties of vector fields which we will need to extend the notion of "vector field" to manifolds. Let U and W be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $f : U \to W$ be a C^{k+1} map. If v is a C^k-vector field on U and w a C^k-vector field on W we will say that v and w are "f-related" if, for all $p \in U$ and $q = f(p)$

$$df_p(v_p) = \mathrm{w}_q\,. \tag{4.12}$$

Writing

$$v \;=\; \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i}\,, \quad v_i \in C^k(U)$$

and

$$\mathrm{w} \;=\; \sum_{j=1}^{m} \mathrm{w}_j \frac{\partial}{\partial y_j}\,, \quad \mathrm{w}_j \in C^k(V)$$

this equation reduces, in coordinates, to the equation

$$\mathrm{w}_i(q) = \sum \frac{\partial f_i}{\partial x_j}(p) v_j(p)\,. \tag{4.13}$$

In particular, if $m = n$ and f is a C^{k+1} diffeomorphism, the formula (4.13) defines a C^k-vector field on V, i.e.,

$$\mathrm{w} = \sum_{j=1}^{n} \mathrm{w}_i \frac{\partial}{\partial y_j}$$

is the vector field defined by the equation

$$\mathrm{w}_i(y) = \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j} v_j \right) \circ f^{-1}\,. \tag{4.14}$$

Hence we've proved

Theorem 4.8. *If $f : U \to V$ is a C^{k+1} diffeomorphism and v a C^k-vector field on U, there exists a unique C^k vector field, w, on W having the property that v and w are f-related.*

We'll denote this vector field by $f_* v$ and call it the *push-forward of v by f*.
I'll leave the following assertions as easy exercises.

Theorem 4.9. *Let U_i, $i = 1, 2$, be an open subset of $\mathbb{R}^{n_i}$, v_i a vector field on U_i and $f : U_1 \to U_2$ a C^1-map. If v_1 and v_2 are f-related, every integral curve*

$$\gamma : I \to U_1$$

of v_1 gets mapped by f onto an integral curve, $f \circ \gamma : I \to U_2$, of v_2.

Theorem 4.10. *Let U_i, $i = 1, 2, 3$, be an open subset of $\mathbb{R}^{n_i}$, v_i a vector field on U_i and $f_i : U_i \to U_{i+1}$, $i = 1, 2$ a C^1-map. Suppose that, for $i = 1, 2$, v_i and v_{i+1} are f_i-related. Then v_1 and v_3 are $f_2 \circ f_1$-related.*

In particular, if f_1 and f_2 are diffeomorphisms and $v = v_1$

$$(f_2)_*(f_1)_* v = (f_2 \circ f_1)_* v\,.$$

Exercises.

1. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$ and $f : U \to V$ a C^1 map. Show that if u and v are vector fields on U and V then they are f-related if and only if
$$L_u f^* \varphi = f^* L_v \varphi$$
for every φ in $C^1(V)$.

2. (a) Let v be the vector field, $\sum x_i \partial/\partial x_i$, $1 \le i \le n$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be the projection map, $(x_1, \ldots, x_n) \to x_1$. Show that v and w are f-related where $w = x_1 \partial/\partial x_1$.
   (b) Verify that f maps integral curves of v onto integral curves of w.

3. Let v be the vector field, $x_1 \partial/\partial x_2 - x_2 \partial/\partial x_1$ and $f : \mathbb{R}^2 \to \mathbb{R}^2$ the diffeomorphism, $(x_1, x_2) \to (x_2, x_1)$. What is $f_* v$?

4. Let v be a constant vector field, $\sum c_i \partial/\partial x_i$, $1 \le i \le n$, and let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a bijective linear map. What is the vector field $f_* v$?

5. Let U be an open subset of $\mathbb{R}^n$, v a vector field on U and $\gamma : \mathbb{R} \to U$, $t \to \gamma(t)$, an integral curve of v. Show that $\partial/\partial t$ and v are γ-related.

6. Let U be an open subset of $\mathbb{R}^n$, v a vector field on U and $\varphi : U \to \mathbb{R}$ a C^1 function. Show that if $L_v \varphi = 1$, v and $\partial/\partial t$ are φ-related.

Lecture 5. Global properties of vector fields

Let U be an open subset of $\mathbb{R}^n$ and v a C^{k+1} vector field on U. We'll say that v is *complete* if, for every $p \in U$, there exists an integral curve, $\gamma : \mathbb{R} \to U$ with $\gamma(0) = p$, i.e., for every p there exists an integral curve that starts at p and *exists for all time.* To see what "completeness" involves, we recall that an integral curve
$$\gamma : [0, b) \to U \,,$$
with $\gamma(0) = p$, is called *maximal* if it can't be extended to an interval $[0, b')$, $b' > b$. For such curves we showed that either

i. $b = +\infty$
or

ii. $|\gamma(t)| \to +\infty$ as $t \to b$
or

iii. the limit set of

$$\{\gamma(t)\,, \quad 0 \leq t, b\}$$

contains points on BdU.

Hence if we can exclude ii. and iii. we'll have shown that an integral curve with $\gamma(0) = p$ exists for all positive time. A simple criterion for excluding ii. and iii. is the following.

Lemma 5.1. *The scenarios ii. and iii. can't happen if there exists a proper C^1-function, $\varphi : U \to \mathbb{R}$ with $L_v\varphi = 0$.*

Proof. $L_v\varphi = 0$ implies that φ is constant on $\gamma(t)$, but if $\varphi(p) = c$ this implies that the curve, $\gamma(t)$, lies on the compact subset, $\varphi^{-1}(c)$, of U; hence it can't "run off to infinity" as in scenario ii. or "run off the boundary" as in scenario iii.

□

Applying a similar argument to the interval $(-b, 0]$ we conclude:

Theorem 5.2. *Suppose there exists a proper C^1-function, $\varphi : U \to \mathbb{R}$ with the property $L_v\varphi = 0$. Then v is complete.*

Example.

Let $U = \mathbb{R}^2$ and let v be the vector field

$$v = x^3 \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\,.$$

Then $\varphi(x, y) = 2y^2 + x^4$ is a proper function with the property above.

If v is complete then for every p, one has an integral curve, $\gamma_p : \mathbb{R} \to U$ with $\gamma_p(0) = p$, so one can, as in Lecture 3, define a map

$$f_t : U \to U$$

by setting $f_t(p) = \gamma_p(t)$. We claim that the f_t's also have the property

$$f_t \circ f_a = f_{t+a}\,. \tag{5.1}$$

Indeed if $f_a(p) = q$, then by the reparameterization theorem, $\gamma_q(t)$ and $\gamma_p(t + a)$ are both integral curves of v, and since $q = \gamma_q(0) = \gamma_p(a) = f_a(p)$, they have the same initial point, so

$$\begin{aligned} \gamma_q(t) &= f_t(q) = (f_t \circ f_a)(p) \\ &= \gamma_p(t + a) = f_{t+a}(p) \end{aligned}$$

for all t. Since f_0 is the identity it follows from (5.1) that $f_t \circ f_{-t}$ is the identity, i.e.,

$$f_{-t} = f_t^{-1}\,.$$

Next we will prove

Theorem 5.3. *The map, f_t, is a C^k diffeomorphism.*

Proof. Fix $T > 0$ and $p \in U$ and note that since the map

$$[0, T] \to U\,, \quad t \to \gamma_p(t)$$

is continuous its image, A_T, is a compact subset of U. Hence by Theorem 3.6 there exists an open set W containing A_T and an $\epsilon > 0$ such that the map

$$(x, t) \in W \times (-\epsilon, \epsilon) \to \gamma_x(t)$$

is C^k. Now let N be a positive integer with the property $T/N < \epsilon$ and let $q = \gamma_p(T)$ and $q_r = \gamma_q\left(-\frac{rT}{N}\right)$ for $r = 1, \ldots, N$.

Thus $f_{T/N}(q_i) = q_{i-1}$, and we an inductively find open neighborhoods, U_i of q_i in W such that $f_{T/N}(U_i) \subset U_{i-1}$. This gives us a sequence of mappings

$$f_{T/N} : U_i \to U_{i-1} \qquad i = N, \ldots, 1$$

and since $T/N < \epsilon$ and every U_i is contained in W these mappings are all C^K. Hence since

$$f_T = f_{T/N} \circ \circ \circ f_{T/N}\,,$$

N times, $f_T : U_N \to U$ is C^k. However, p is an arbitrary point of U so we've shown that for every point, $p \in U$, there exists a neighborhood, U_N, of p in U on which f_T is C^k.

A nearly identical argument shows that f_{-T} is C^k and hence, since f_{-T} is the inverse of the mapping, f_T, that f_T is a C^k difffeomorphism. Q.E.D.

□

As a corollary of Theorem 5.3 we will prove

Theorem 5.4. *The mapping*

$$F : U \times \mathbb{R} \to U\,,\ (p, t) \to f_t(p)$$

is C^k.

Proof. For every p_0 in U there exists a neighborhood, V, of p_0 in U and an $\epsilon > 0$ such that the map

$$(p,t) \in V \times (-\epsilon, \epsilon) \to U\,,\ (p,t) \to f_t(p)$$

is C^k. Hence for all T the map

$$(p,t) \in V \times (-\epsilon, \epsilon) \to f_T \circ f_t(p)$$

is C^k. However, $f_T \circ f_t = f_{T+t}$ so the map

$$(p,t) \in V \times (-\epsilon, \epsilon) \to f_{T+t}(p)$$

is C^k and hence finally that the map

$$(p,t) \in V \times (-\infty, \infty) \to f_t(p) \tag{5.2}$$

is C^k on every interval, $T - \epsilon < t < T + \epsilon$. Hence the map (5.2) is C^k on all of $U \times (-\infty, \infty)$.

□

For v not complete there is an analogous result, but it's trickier to formulate precisely. Roughly speaking v generates a one-parameter group of diffeomorphisms, f_t, but these diffeomorphisms are not defined on all of U nor for all values of t. Moreover, the identity (5.1) only holds on the open subset of U where both sides are well-defined. (See Lecture 3.)

Exercises.

1. Let U_i, $i = 1, 2$, be open subsets of $\mathbb{R}^n$ and let v_i be a complete vector field on U_i. Suppose $f : U_1 \to U_2$ is a C^1 map. Show that if v_1 and v_2 are f-related and $(f_i)_t : U_i \to U_i$ is the one parameter group of diffeomorphisms generated by v_i then $f \circ (f_1)_t = (f_2)_t \circ f$.

2. Let v be a complete vector field on U and $f_t : U \to U$, the one parameter group of diffeomorphisms generated by v. Show that if $\varphi \in C^1(U)$

$$L_v \varphi = \left(\frac{d}{dt} f_t^* \varphi \right)_{t=0} .$$

3. (a) Let $U = \mathbb{R}^2$ and let $\mathfrak{v}$ be the vector field, $x_1 \partial/\partial x_2 - x_2 \partial/\partial x_1$. Show that the curve

$$t \in \mathbb{R} \to (r\cos(t+\theta)\,,\, r\sin(t+\theta))$$

is the unique integral curve of $\mathfrak{v}$ passing through the point, $(r\cos\theta, r\sin\theta)$, at $t = 0$.

(b) Let $U = \mathbb{R}^n$ and let $\mathfrak{v}$ be the constant vector field: $\sum c_i \partial/\partial x_i$. Show that the curve

$$t \in \mathbb{R} \to a + t(c_1, \ldots, c_n)$$

is the unique integral curve of $\mathfrak{v}$ passing through $a \in \mathbb{R}^n$ at $t = 0$.

(c) Let $U = \mathbb{R}^n$ and let $\mathfrak{v}$ be the vector field, $\sum x_i \partial/\partial x_i$. Show that the curve

$$t \in \mathbb{R} \to e^t(a_1, \ldots, a_n)$$

is the unique integral curve of $\mathfrak{v}$ passing through a at $t = 0$.

4. Let U be an open subset of $\mathbb{R}^n$ and $F : U \times \mathbb{R} \to U$ a $\mathcal{C}^\infty$ mapping. The family of mappings

$$f_t : U \to U\,, \quad f_t(x) = F(x, t)$$

is said to be a *one-parameter group of diffeomorphisms* of U if f_0 is the identity map and $f_s \circ f_t = f_{s+t}$ for all s and t. (Note that $f_{-t} = f_t^{-1}$, so each of the f_t's is a diffeomorphism.) Show that the following are one-parameter groups of diffeomorphisms:

(a) $f_t : \mathbb{R} \to \mathbb{R}\,, \quad f_t(x) = x + t$

(b) $f_t : \mathbb{R} \to \mathbb{R}\,, \quad f_t(x) = e^t x$

(c) $f_t : \mathbb{R}^2 \to \mathbb{R}^2\,, \quad f_t(x, y) = (\cos t\, x - \sin t\, y\,,\, \sin t\, x + \cos t\, y)$

5. Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping. Show that the series

$$\exp tA = I + tA + \frac{t^2}{2!}\, A^2 + \frac{t^3}{3!}\, A^3 + \cdots$$

converges and defines a one-parameter group of diffeomorphisms of $\mathbb{R}^n$.

6. (a) What are the generators of the one-parameter groups in exercise 4?

(b) Show that the generator of the one-parameter group in exercise 5 is the vector field

$$\sum a_{i,j} x_j \frac{\partial}{\partial x_i}$$

where $[a_{i,j}]$ is the defining matrix of A.

7. Let U be an open subset of $\mathbb{R}^n$ and v a vector field on U. We will say that v is compactly supported if the closure of the set

$$\{p \in U\,, \quad v(p) = 0\}$$

is compact. Show that if v is compactly supported it is complete.

Lecture 6. Generalizations of the inverse function theorem

In this lecture we will discuss two generalizations of the inverse function theorem. We'll begin by reviewing some linear algebra. Let

$$A : \mathbb{R}^m \to \mathbb{R}^n$$

be a linear mapping and $[a_{i,j}]$ the $n \times m$ matrix associated with A. Then

$$A^t : \mathbb{R}^n \to \mathbb{R}^m$$

is the linear mapping associated with the transpose matrix $[a_{j,i}]$. For $k < n$ we define the *canonical submersions*

$$\pi : \mathbb{R}^n \to \mathbb{R}^k$$

to be the map $\pi(x_1, \ldots, x_n) = (x_1, \ldots, x_k)$ and the canonical immersion

$$\iota : \mathbb{R}^k \to \mathbb{R}^n$$

to be the map, $\iota(x_1, \ldots, x_k) = (x_1, \ldots x_k, 0, \ldots 0)$. We leave for you to check that $\pi^t = \iota$.

Proposition 6.1. *If $A : \mathbb{R}^n \to \mathbb{R}^k$ is onto, there exists a bijective linear map $B : \mathbb{R}^n \to \mathbb{R}^n$ such that $AB = \pi$.*

We'll leave the proof of this as an exercise.

Hint: Show that one can choose a basis, $v_1, \ldots, v_n$ of $\mathbb{R}^n$ such that

$$Av_i = e_i\,, \quad i = 1, \ldots, k$$

is the standard basis of $\mathbb{R}^k$ and

$$Av_i = 0\,, \quad i > k\,.$$

Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$ and set $Be_i = v_i$.

Proposition 6.2. *If $A : \mathbb{R}^k \to \mathbb{R}^n$ is one–one, there exists a bijective linear map $C : \mathbb{R}^n \to \mathbb{R}^n$ such that $CA = \iota$.*

Proof. The rank of $[a_{i,j}]$ is equal to the rank of $[a_{j,i}]$, so if if A is one–one, there exists a bijective linear map $B : \mathbb{R}^n \to \mathbb{R}^n$ such that $A^t B = \pi$.

Letting $C = B^t$ and taking transposes we get $\iota = \pi^t = CB$

$\square$

Immersions and submersions

Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^k$ a $\mathcal{C}^\infty$ map. f is a *submersion* at $p \in U$ if

$$Df(p) : \mathbb{R}^n \to \mathbb{R}^k$$

is onto. Our first main result in this lecture is a non-linear version of Proposition 1.

Theorem 6.3 (Canonical submersion theorem)**.** *If f is a submersion at p and $f(p) = 0$, there exists a neighborhood, U_0 of p in U, a neighborhood, V, of 0 in $\mathbb{R}^n$ and a $\mathcal{C}^\infty$ diffeomorphism, $g : (V, 0) \to (U_0, p)$ such that $f \circ g = \pi$.*

Proof. Let $\tau_p : \mathbb{R}^n \to \mathbb{R}^n$ be the map, $x \to x+p$. Replacing f by $f \circ \tau_p$ we can assume $p = 0$. Let A be the linear map

$$Df(0) : \mathbb{R}^n \to \mathbb{R}^k \,.$$

By assumption this map is onto, so there exists a bijective linear map

$$B : \mathbb{R}^n \to \mathbb{R}^n$$

such that $AB = \pi$. Replacing f by $f \circ B$ we can assume that

$$Df(0) = \pi \,.$$

Let $h : U \to \mathbb{R}^n$ be the map

$$h(x_1, \ldots, x_n) = (f_1(x), \ldots, f_k(x)\,,\, x_{k+1}, \ldots, x_n)$$

where the f_i's are the coordinate functions of f. I'll leave for you to check that

$$Dh(0) = I \tag{6.1}$$

and

$$\pi \circ h = f \,. \tag{6.2}$$

By (6.1) $Dh(0)$ is bijective, so by the inverse function theorem h maps a neighborhood, U_0 of 0 in U diffeomorphically onto a neighborhood, V, of 0 in $\mathbb{R}^n$. Letting $g = f^{-1}$ we get from (6.2) $\pi = f \circ g$.

□

Our second main result is a non-linear version of Proposition 2. Let U be an open neighborhood of 0 in $\mathbb{R}^k$ and $f : U \to \mathbb{R}^n$ a $\mathcal{C}^\infty$-map.

Theorem 6.4 (Canonical immersion theorem)**.** *If f is an immersion at 0, there exists a neighborhood, V, of $f(0)$ in $\mathbb{R}^n$, a neighborhood, W, of 0 in $\mathbb{R}^n$ and a $\mathcal{C}^\infty$-diffeomorphism $g : V \to W$ such that $\iota^{-1}(W) \subseteq U$ and $g \circ f = \iota$ on $\iota^{-1}(W)$.*

Proof. Let $p = f(0)$. Replacing f by $\tau_{-p} \circ f$ we can assume that $f(0) = 0$. Since $Df(0) : \mathbb{R}^k \to \mathbb{R}^n$ is injective there exists a bijective linear map, $B : \mathbb{R}^n \to \mathbb{R}^n$ such that $BDf(0) = \iota$, so if we replace f by $B \circ f$ we can assume that $Df(0) = \iota$. Let $\ell = n - k$ and let

$$h : U \times \mathbb{R}^\ell \to \mathbb{R}^n$$

be the map

$$h(x_1, \ldots, x_n) = f(x_1, \ldots, x_k) + (0, \ldots, 0\ x_{k+1}, \ldots, x_n)\,.$$

I'll leave for you to check that

$$Dh(0) \;=\; I \tag{6.3}$$

and

$$h \circ \iota \;=\; f\,. \tag{6.4}$$

By (6.3) $Dh(0)$ is bijective, so by the inverse function theorem, h maps a neighborhood, W, of 0 in $U \times \mathbb{R}^\ell$ diffeomorphically onto a neighborhood, V, of 0 in $\mathbb{R}^n$. Let $g : V \to W$ be the inverse map. Then by (6.4), $\iota = g \circ f$. □

Problem set

1. Prove Proposition 1.

2. Prove Proposition 2.

3. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be the map

   $$(x_1, x_2, x_3) \to (x_1^2 - x_2^2\,,\, x_2^2 - x_3^2)\,.$$

   At what points $p \in \mathbb{R}^3$ is f a submersion?

4. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the map

   $$(x_1, x_2) \to (x_1, x_2\,,\, x_1^2, x_2^2)\,.$$

   At what points, $p \in \mathbb{R}^2$, is f an immersion?

5. Let U and V be open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, and let $f : U \to V$ and $g : V \to \mathbb{R}^k$ be C^1-maps. Prove that if f is a submersion at $p \in U$ and g a submersion at $q = f(p)$ then $g \circ f$ is a submersion at p.

6. Let f and g be as in exercise 5. Suppose that g is a submersion at q. Show that $g \circ f$ is a submersion at p if and only if

   $$T_q\mathbb{R}^n = \operatorname{Image} df_p + \operatorname{Kernel} dg_q\,,$$

   i.e., if and only if every vector, $\mathrm{v} \in T_q\mathbb{R}^n$ can be written as a sum, $\mathrm{v} = \mathrm{v}_1 + \mathrm{v}_2$, where v_1 is in the image of df_p and $dg_q(\mathrm{v}_2) = 0$.

Lecture 7. Manifolds

Let X be a subset of $\mathbb{R}^N$, Y a subset of $\mathbb{R}^n$ and $f : X \to Y$ a continuous map. We recall

Definition 7.1. *f is a $\mathcal{C}^\infty$ map if for every $p \in X$, there exists a neighborhood, U_p, of p in $\mathbb{R}^N$ and a $\mathcal{C}^\infty$ map, $g_p : U_p \to \mathbb{R}^n$, which coincides with f on $U_p \cap X$.*

We also recall:

Theorem 7.2 (Munkres, §16,#3)**.** *If $f : X \to Y$ is a $\mathcal{C}^\infty$ map, there exists a neighborhood, U, of X in $\mathbb{R}^N$ and a $\mathcal{C}^\infty$ map, $g : U \to \mathbb{R}^n$ such that g coincides with f on X.*

We will say that f is a *diffeomorphism* if it is one–one and onto and f and f^{-1} are both diffeomorphisms. In particular if Y is an open subset of $\mathbb{R}^n$, X is a simple example of what we will call a *manifold*. More generally,

Definition 7.3. *A subset, X, of $\mathbb{R}^N$ is an n-dimensional manifold if, for every $p \in X$, there exists a neighborhood, V, of p in $\mathbb{R}^N$, an open subset, U, in $\mathbb{R}^m$, and a diffeomorphism $\varphi : U \to X \cap V$.*

Thus X is an n-dimensional manifold if, locally near every point p, X "looks like" an open subset of $\mathbb{R}^n$.

We'll now describe how manifolds come up in concrete applications. Let U be an open subset of $\mathbb{R}^N$ and $f : U \to \mathbb{R}^k$ a $\mathcal{C}^\infty$ map.

Definition 7.4. *A point, $a \in \mathbb{R}^k$, is a regular value of f if for every point, $p \in f^{-1}(a)$, f is a submersion at p.*

Note that for f to be a submersion at p, $Df(p) : \mathbb{R}^N \to \mathbb{R}^k$ has to be onto, and hence k has to be less than or equal to N. Therefore this notion of "regular value" is interesting only if $N \geq k$.

Theorem 7.5. *Let $N - k = n$. If a is a regular value of f, the set, $X = f^{-1}(a)$, is an n-dimensional manifold.*

Proof. Replacing f by $\tau_{-a} \circ f$ we can assume without loss of generality that $a = 0$. Let $p \in f^{-1}(0)$. Since f is a submersion at p, the canonical submersion theorem tells us that there exists a neighborhood, $\mathcal{O}$, of 0 in $\mathbb{R}^N$, a neighborhood, U_0, of p in U and a diffeomorphism, $g : \mathcal{O} \to U_0$ such that

$$f \circ g = \pi \tag{7.1}$$

where π is the projection map

$$\mathbb{R}^N = \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}^k \,, \quad (x, y) \to x \,.$$

Hence $\pi^{-1}(0) = \{0\} \times \mathbb{R}^n = \mathbb{R}^n$ and by (7.1), g maps $\mathcal{O} \cap \pi^{-1}(0)$ diffeomorphically onto $U_0 \cap f^{-1}(0)$. However, $\mathcal{O} \cap \pi^{-1}(0)$ is a neighborhood, V, of 0 in $\mathbb{R}^n$ and $U_0 \cap f^{-1}(0)$ is a neighborhood of p in X, and, as remarked, these two neighborhoods are diffeomorphic. □

Some examples:

1. *The n-sphere.* Let

$$f : \mathbb{R}^{n+1} \to \mathbb{R}$$

be the map,

$$(x_1, \ldots, x_{n+1}) \to x_1^2 + \cdots + x_{n+1}^2 - 1\,.$$

Then

$$Df(x) = 2(x_1, \ldots, x_{n+1})$$

so, if $x \neq 0$ f is a submersion at x. In particular f is a submersion at all points, x, on the n-sphere

$$S^n = f^{-1}(0)$$

so the n-sphere is an n-dimensional submanifold of $\mathbb{R}^{n+1}$.

2. *Graphs.* Let $g : \mathbb{R}^n \to \mathbb{R}^k$ be a $\mathcal{C}^\infty$ map and let

$$X = \text{ graph } g = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^k\,, \quad y = g(x)\}\,.$$

We claim that X is an n-dimensional submanifold of $\mathbb{R}^{n+k} = \mathbb{R}^n \times \mathbb{R}^k$.

Proof. Let

$$f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$$

be the map, $f(x, y) = y - g(x)$. Then

$$Df(x, y) = [-Dg(x)\,,\, I_k]$$

where I_k is the identity map of $\mathbb{R}^k$ onto itself. This map is always of rank k. Hence graph $g = f^{-1}(0)$ is an n-dimensional submanifold of R^{n+k}.

□

3. Munkres, §24, #6. Let $\mathcal{M}_n$ be the set of all $n \times n$ matrices and let $\mathcal{S}_n$ be the set of all symmetric $n \times n$ matrices, i.e., the set

$$\mathcal{S}_n = \{A \in \mathcal{M}_n\,,\, A = A^t\}\,.$$

The map

$$[a_{i,j}] \to (a_{11}, a_{12}, \ldots, a_{1n}, a_{2,1}, \ldots, a_{2n}, \ldots)$$

gives us an identification

$$\mathcal{M}_n \cong \mathbb{R}^{n^2}$$

and the map

$$[a_{i,j}] \to (a_{11}, \dots a_{1n}, a_{22}, \dots a_{2n}, a_{33}, \dots a_{3n}, \dots)$$

gives us an identification

$$\mathcal{S}_n \cong \mathbb{R}^{\frac{n(n+1)}{2}}\,.$$

(Note that if A is a symmetric matrix,

$$a_{12} = a_{21}\ ,\ a_{13} = a_{13} = a_{31}\ ,\ a_{32} = a_{23},\ \text{etc.}$$

so this map avoids redundancies.) Let

$$O(n) = \{A \in \mathcal{M}_n\ ,\ A^t A = I\}\,.$$

This is the set of *orthogonal* $n \times n$ matrices, and the exercise in Munkres requires you to show that it's an $n(n-1)/2$-dimensional manifold.

Hint: Let $f : \mathcal{M}_n \to \mathcal{S}_n$ be the map $f(A) = A^t A - I$. Then

$$O(n) = f^{-1}(0)\,.$$

These examples show that lots of interesting manifolds arise as zero sets of submersions, $f : U \to \mathbb{R}^k$. We'll conclude this lecture by showing that locally *every* manifold arises this way. More explicitly let $X \subseteq \mathbb{R}^N$ be an n-dimensional manifold, p a point of X, U a neighborhood of 0 in $\mathbb{R}^n$, V a neighborhood of p in $\mathbb{R}^N$ and $\varphi : (U, 0) \to (V \cap X, p)$ a diffeomorphism. We will for the moment think of φ as a $\mathcal{C}^\infty$ map $\varphi : U \to \mathbb{R}^N$ whose image happens to lie in X.

Lemma 7.6. *The linear map*

$$D\varphi(0) : \mathbb{R}^n \to \mathbb{R}^N$$

is injective.

Proof. $\varphi^{-1} : V \cap X \to U$ is a diffeomorphism, so, shrinking V if necessary, we can assume that there exists a $\mathcal{C}^\infty$ map $\psi : V \to U$ which coincides with φ^{-1}on $V \cap X$ Since φ maps U onto $V \cap X$, $\psi \circ \varphi = \varphi^{-1} \circ \varphi$ is the identity map on U. Therefore,

$$D(\psi \circ \varphi)(0) = (D\psi)(p) D\varphi(0) = I$$

by the change rule, and hence if $D\varphi(0)v = 0$, it follows from this identity that $v = 0$. $\square$

Lemma 6 says that φ is an immersion at 0, so by the canonical immersion theorem there exists a neighborhood, U_0, of 0 in U a neighborhood, V_p, of p in V, a neighborhood, $\mathcal{O}$, of 0 in $\mathbb{R}^N$ and a diffeomorphism

$$g : (V_p, p) \to (\mathcal{O}, 0)$$

such that

$$\iota^{-1}(\mathcal{O}) = U_0 \tag{7.2}$$

and

$$g \circ \varphi = \iota \ \text{ on } U_0 \,, \tag{7.3}$$

ι being, as in lecture 1, the canonical immersion

$$\iota(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots 0) \,. \tag{7.4}$$

By (7.3) g maps $\varphi(U_0)$ diffeomorphically onto $\iota(U_0)$. However, by (7.2) and (7.3) $\iota(U_0)$ is the subset of $\mathcal{O}$ defined by the equations, $x_i = 0$, $i = n+1, \dots, N$. Hence if $g = (g_1, \dots, g_N)$ the set, $\varphi(U_0) = V_p \cap X$ is defined by the equations

$$g_i = 0 \,, \quad i = n+1, \dots, N \,. \tag{7.5}$$

Let $k = N - n$, let

$$\pi : \mathbb{R}^N = \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^k$$

be the canonical submersion,

$$\pi(x_1, \dots, x_N) = (x_{n+1}, \dots x_N)$$

and let $f = \pi \circ g$. Since g is a diffeomorphism, f is a submersion and (7.5) can be interpreted as saying that

$$V_p \cap X = f^{-1}(0) \,. \tag{7.6}$$

Thus we've proved

Theorem 7.7. *For every $p \in X$ there exists a neighborhood V_p of p in $\mathbb{R}^N$ and a submersion, $f : V_p \to \mathbb{R}^k$, with $f(p) = 0$, such that X is defined locally near p by the equation (7.6).*

A nice way of thinking about Theorem 7.7 is in terms of the coordinates of the mapping, f. More specifically if $f = (f_1, \dots, f_k)$ we can think of $f^{-1}(0)$ as being the set of solutions of the system of equations

$$f_i(x) = 0_i \,, \quad i = 1, \dots, k \tag{7.7}$$

and the condition that a be a regular value of f can be interpreted as saying that for every solution, p, of this system of equations the derivatives

$$(df_i)_p = \sum \frac{\partial f_i}{\partial x_j}(p)\,(dx_j)_p \tag{7.8}$$

are linearly independent, i.e., the system (7.7) is an "independent system of defining equations" for X.

Problem set

1. Show that the set of solutions of the system of equations

$$x_1^2 + \cdots + x_n^2 \;=\; 1$$

and

$$x_1 + \cdots + x_n \;=\; 0$$

is an $n-2$-dimensional submanifold of $\mathbb{R}^n$.

2. Let S^{n-1} be the n-sphere in $\mathbb{R}^n$ and let

$$X_a = \{x \in S^{n-1}\,, \quad x_1 + \cdots + x_n = a\}\,.$$

For what values of a is X_a an $(n-2)$-dimensional submanifold of S^{n-1}?

3. Show that if X_i, $i = 1, 2$, is an n_i-dimensional submanifold of $\mathbb{R}^{N_i}$ then

$$X_1 \times X_2 \subseteq \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$$

is an $(n_1 + n_2)$-dimensional submanifold of $\mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$.

4. Show that the set

$$X = \{(x, \mathrm{v}) \in S^{n-1} \times \mathbb{R}^n\,, \quad x \cdot \mathrm{v} = 0\}$$

is a $2n-2$-dimensional submanifold of $\mathbb{R}^n \times \mathbb{R}^n$. (Here "$x \cdot \mathrm{v}$" is the dot product, $\sum x_i v_i$.)

5. Let $g : \mathbb{R}^n \to \mathbb{R}^k$ be a $\mathcal{C}^\infty$ map and let $X = \operatorname{graph} g$. Prove directly that X is an n-dimensional manifold by proving that the map

$$\gamma : \mathbb{R}^n \to X\,, \qquad x \to (x, g(x))$$

is a diffeomorphism.

8. Tangent spaces

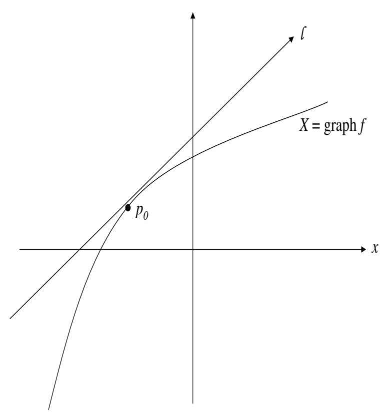
We recall that a subset, X, of $\mathbb{R}^N$ is an n-dimensional manifold, if, for every $p \in X$, there exists an open set, $U \subseteq \mathbb{R}^n$, a neighborhood, V, of p in $\mathbb{R}^N$ and a $\mathcal{C}^\infty$-diffeomorphism, $\varphi : U \to X \cap V$.

Definition 8.1. *We will call φ a* parameterization *of X at p.*

Our goal in this lecture is to define the notion of the *tangent space*, T_pX, to X at p and describe some of its properties. Before giving our official definition we'll discuss some simple examples.

Example 1.

Let $f : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function and let $X = \operatorname{graph} f$.



Then in this figure above the tangent line, ℓ, to X at $p_0 = (x_0, y_0)$ is defined by the equation

$$y - y_0 = a(x - x_0)$$

where $a = f'(x_0)$ In other words if p is a point on ℓ then $p = p_0 + \lambda \mathrm{v}_0$ where $\mathrm{v}_0 = (1, a)$ and $\lambda \in \mathbb{R}^n$. We would, however, like the tangent space to X at p_0 to be a subspace of the tangent space to $\mathbb{R}^2$ at p_0, i.e., to be the subspace of the space: $T_{p_0}\mathbb{R}^2 = \{p_0\} \times \mathbb{R}^2$, and this we'll achieve by defining

$$T_{p_0}X = \{(p_0, \lambda \mathrm{v}_0), \quad \lambda \in \mathbb{R}\}.$$

Example 2.

Let S^2 be the unit 2-sphere in $\mathbb{R}^3$. The tangent plane to S^2 at p_0 is usually defined to be the plane

$$\{p_0 + \mathrm{v}\,;\, v \in \mathbb{R}^3\,, \quad \mathrm{v} \perp p_0\}\,.$$

However, this tangent plane is easily converted into a subspace of $T_p\mathbb{R}^3$ via the map, $p_0 + \mathrm{v} \to (p_0, \mathrm{v})$ and the image of this map

$$\{(p_0, \mathrm{v})\,;\, \mathrm{v} \in \mathbb{R}^3\,, \quad \mathrm{v} \perp p_0\}$$

will be our definition of $T_{p_0}S^2$.

Let's now turn to the general definition. As above let X be an n-dimensional submanifold of $\mathbb{R}^N$, p a point of X, V a neighborhood of p in $\mathbb{R}^N$, U an open set in $\mathbb{R}^n$ and

$$\varphi : (U, q) \to (X \cap V, p)$$

a parameterization of X. We can think of φ as a $\mathcal{C}^\infty$ map

$$\varphi : (U, q) \to (V, p)$$

whose image happens to lie in $X \cap V$ and we proved last time that its derivative at q

$$(d\varphi)_q : T_q\mathbb{R}^n \to T_p\mathbb{R}^N \tag{8.1}$$

is injective.

Definition 8.2. *The tangent space, T_pX, to X at p is the image of the linear map (8.1). In other words, $w \in T_p\mathbb{R}^N$ is in T_pX if and only if $w = d\varphi_q(\mathrm{v})$ for some $\mathrm{v} \in T_q\mathbb{R}^n$. More succinctly,*

$$T_pX = (d\varphi_q)(T_q\mathbb{R}^n)\,. \tag{8.2}$$

(Since $d\varphi_q$ is injective this space is an n-dimensional vector subspace of $T_p\mathbb{R}^N$.)

One problem with this definition is that it appears to depend on the choice of φ. To get around this problem, we'll give an alternative definition of T_pX. Last time we showed that there exists a neighborhood, V, of p in $\mathbb{R}^N$ (which we can without loss of generality take to be the same as V above) and a $\mathcal{C}^\infty$ map

$$f : (V, p) \to (\mathbb{R}^k, 0)\,, \quad k = N - n\,, \tag{8.3}$$

such that $X \cap V = f^{-1}(0)$ and such that f is a submersion at all points of $X \cap V$, and in particular at p. Thus

$$df_p : T_p\mathbb{R}^N \to T_0\mathbb{R}^k$$

is surjective, and hence the kernel of df_p has dimension n. Our alternative definition of T_pX is

$$T_pX = \text{kernel } \ df_p\,. \tag{8.4}$$

The spaces (8.2) and (8.4) are both n-dimensional subspaces of $T_p\mathbb{R}^N$, and we claim that these spaces are the same. (Notice that the definition (8.4) of T_pX doesn't depend on φ, so if we can show that these spaces are the same, the definitions (8.2) and (8.4) will depend *neither* on φ *nor* on f.)

Proof. Since $\varphi(U)$ is contained in $X \cap V$ and $X \cap V$ is contained in $f^{-1}(0)$, $f \circ \varphi = 0$, so by the chain rule

$$df_p \circ d\varphi_q = d(f \circ \varphi)_q = 0\,. \tag{8.5}$$

Hence if $\mathrm{v} \in T_q\mathbb{R}^n$ and $w = d\varphi_q(\mathrm{v})$, $df_p(w) = 0$. This shows that the space (8.2) is contained in the space (8.4). However, these two spaces are n-dimensional so the coincide.

$\square$

From the proof above one can extract a slightly stronger result:

Theorem 8.3. *Let W be an open subset of $\mathbb{R}^\ell$ and $h : (W, q) \to (\mathbb{R}^N, p)$ a $\mathcal{C}^\infty$ map. Suppose $h(W)$ is contained in X. Then the image of the map*

$$dh_q : T_q\mathbb{R}^\ell \to T_p\mathbb{R}^N$$

is contained in T_pX.

Proof. Let f be the map (8.3). We can assume without loss of generality that $h(W)$ is contained in V, and so, by assumption, $h(W) \subseteq X \cap V$. Therefore, as above, $f \circ h = 0$, and hence $dh_q(T_q\mathbb{R}^\ell)$ is contained in the kernel of df_p.

$\square$

This result will enable us to define the *derivative* of a mapping between manifolds. Explicitly: Let X be a submanifold of $\mathbb{R}^N$, Y a submanifold of $\mathbb{R}^m$ and $g : (X, p) \to (Y, y_0)$ a $\mathcal{C}^\infty$ map. By Theorem 7.2 there exists a neighborhood, $\mathcal{O}$, of X in $\mathbb{R}^N$ and a $\mathcal{C}^\infty$ map, $\tilde{g} : \mathcal{O} \to \mathbb{R}^m$ extending g. We will define

$$(dg_p) : T_pX \to T_{y_0}Y \tag{8.6}$$

to be the restriction of the map

$$(d\tilde{g})_p : T_p\mathbb{R}^N \to T_{y_0}\mathbb{R}^m \tag{8.7}$$

to T_pX. There are two obvious problems with this definition:

1. Is the space

$$(d\tilde{g}_p)(T_pX)$$

contained in $T_{y_0}Y$?

2. Does the definition depend on $\tilde{g}$?

To show that the answer to 1. is yes and the answer to 2. is no, let

$$\varphi : (U, x_0) \to (X \cap V, p)$$

be a parameterization of X, and let $h = \tilde{g} \circ \varphi$. Since $\varphi(U) \subseteq X$, $h(U) \subseteq Y$ and hence by Theorem 2

$$dh_{x_0}(T_{x_0}\mathbb{R}^n) \subseteq T_{y_0}Y\,.$$

But by the chain rule

$$dh_{x_0} = d\tilde{g}_p \circ d\varphi_{x_0}\,, \tag{8.8}$$

so by (8.2)

$$(d\tilde{g}_p)(T_pX) \quad\subseteq\quad T_{y_0}Y \tag{8.9}$$

and

$$(d\tilde{g}_p)(T_pX) \quad=\quad (dh)_{x_0}(T_{x_0}\mathbb{R}^n) \tag{8.10}$$

Thus the answer to 1. is yes, and since $h = \tilde{g} \circ \varphi = g \circ \varphi$, the answer to 2. is no.

From (8.5) and (8.6) one easily deduces

Theorem 8.4 (Chain rule for mappings between manifolds)**.** *Let Z be a submanifold of $\mathbb{R}^\ell$ and $\psi : (Y, y_0) \to (Z, z_0)$ a $\mathcal{C}^\infty$ map. Then $d\psi_{y_0} \circ dg_p = d(\psi \circ g)_p$.*

Problem set

1. What is the tangent space to the quadric, $x_n^2 = x_1^2 + \cdots + x_{n-1}^2$, at the point, $(1, 0, \ldots, 0, 1)$?

2. Show that the tangent space to the $(n-1)$-sphere, S^{n-1}, at p, is the space of vectors, $(p, \mathrm{v}) \in T_p\mathbb{R}^n$ satisfying $p \cdot \mathrm{v} = 0$.

3. Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a $\mathcal{C}^\infty$ map and let $X = \mathrm{graph} f$. What is the tangent space to X at $(a, f(a))$?

4. Let $\sigma : S^{n-1} \to S^{n-1}$ be the anti-podal map, $\sigma(x) = -x$. What is the derivative of σ at $p \in S^{n-1}$?

5. Let $X_i \subseteq \mathbb{R}^{N_i}$, $i = 1, 2$, be an n_i-dimensional manifold and let $p_i \in X_i$. Define X to be the Cartesian product

$$X_1 \times X_2 \subseteq \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$$

and let $p = (p_1, p_2)$. Describe T_pX in terms of $T_{p_1}X_1$ and $T_{p_2}X_2$.

6. Let $X \subseteq \mathbb{R}^N$ be an n-dimensional manifold and $\varphi_i : U_i \to X \cap V_i$, $i = 1, 2$. From these two parameterizations one gets an overlap diagram

$$\begin{tikzcd} & X \cap V & \\ W_1 \arrow[ur, "\varphi_1"] \arrow[rr, "\psi"] & & W_2 \arrow[ul, "\varphi_2"'] \end{tikzcd} \tag{8.11}$$

where $V = V_1 \cap V_2$, $W_i = \varphi_i^{-1}(X \cap V)$ and $\psi = \varphi_2^{-1} \circ \varphi_1$.

(a) Let $p \in X \cap V$ and let $q_i = \varphi_i^{-1}(p)$. Derive from the overlap diagram (8.10) an overlap diagram of linear maps

$$\begin{tikzcd} & T_p\mathbb{R}^N & \\ T_{q_1}\mathbb{R}^n \arrow[ur, "(d\varphi_1)_{q_1}"] \arrow[rr, "(d\psi)_{q_1}"] & & T_{q_2}\mathbb{R}^n \arrow[ul, "(d\varphi_2)_{q_2}"'] \end{tikzcd} \tag{8.12}$$

(b) Use overlap diagrams to give another proof that T_pX is intrinsically defined.

9. Vector fields on manifolds

We defined a vector field on an open subset, U, of $\mathbb{R}^n$ to be a function v which assigns to each point, $p \in U$, a vector, $v(p)$, in T_pU. This definition makes perfectly good sense for manifolds as well:

Definition 9.1. *Let X be an n-dimensional manifold. A vector field on X is a function, v, which assigns to each point, $p \in X$, a vector $v(p)$ in T_pX.*

We'll begin our discussion of vector fields on manifolds by describing the manifold analogues of Theorems 4.8 and 4.10. Let X and Y be manifolds and $f : X \to Y$ a $\mathcal{C}^\infty$ map. Then for $p \in X$ and $q = f(p)$ we get, as explained in yesterday's lecture, a linear map, $df_p : T_pX \to T_qY$. Now let v be a vector field on X and w a vector field on Y. We will say that v and w are f-related if for all $p \in X$

$$df_p(v(p)) = w(q)\,. \tag{9.1}$$

In particular if f is a diffeomorphism the identity (9.1) enables one to associate to every vector field, v, on X a unique f-related vector field, w on Y, i.e. given a point, q, in Y one can define w at q by requiring that (9.1) hold for $p = f^{-1}(q)$. The vector field, w, defined by this recipe will be called the push forward of v by f and denoted by f_*v.

Now let Z be another manifold and $g : Y \to Z$ a $\mathcal{C}^\infty$ mapping. The manifold analogue of Theorem 4.10 asserts

Theorem 9.2. *Let v be a vector field on X, w a vector field on Y and u a vector field on Z. Suppose v and w are f-related and w and u are g-related. Then v and u are $g \circ f$-related.*

Proof. For $p \in X$ and $q = f(p)$ this follows from the chain rule

$$d(g \circ f)_p v(p) = (dg)_q \, df_p(v(p)) \, .$$

(See theorem 8.4.)

□

In particular if f and g are diffeomorphisms $(g \circ f)_* v = g_*(f_* v)$.

To prove the analogue of Theorem 4.8 we must first of all define what we mean by a "C^k vector field" on X.

Definition 9.3. *Let v be a vector field on X. We will say that v is C^k at a point p in X if there exists a neighborhood V of p in X and a parametrization, $\varphi : U \to V$ mapping $q \in U$ onto p such that $(\varphi^{-1})_* v$ is C^k in a neighborhood of q.*

To show that this is a legitimate definition we have to show that it is independent of the choice of parametrization. To do so let $\varphi_i : U_i \to V_i$, $i = 1, 2$ be a parametrization mapping $q_i \in U_i$ onto p, and let $V = V_1 \cap V_2$, $W_i = \varphi_i^{-1}(V)$ and $\psi = \varphi_2^{-1} \circ \varphi_1$. Then by the overlap diagram (8.11) and the chain rule

$$\psi_*(\varphi_1^{-1})_* v = (\varphi_2^{-1})_* v$$

and hence Theorem 4.8 tells us that $(\varphi_2^{-1})_* v$ is C^k in a neighborhood of q_2 if $(\varphi_1^{-1})_* v$ is C^k in a neighborhood of q_1.

Remarks.

1. We will say that v is a $\mathcal{C}^\infty$ vector field on X if it satisfies the conditions of Definition 9.3 for all k and all $p \in X$.

2. Henceforth we'll assume for the most part without explicitly saying so that the vector fields we're dealing with are $\mathcal{C}^\infty$.

We'll next prove

Theorem 9.4. *Let X and Y be n-dimensional manifolds and $f : X \to Y$ a diffeomorphism. Then if v is a $\mathcal{C}^\infty$ vector field on X, $f_* v$ is a $\mathcal{C}^\infty$ vector field on Y.*

Proof. If $\varphi_1 : U \to V_1$ is a parametrization of an open set in X, $\varphi_2 = f \circ \varphi_1$ is a parametrization, $\varphi_2 : U \to V_2$, of an open set $V_2 = f(V_1)$ in Y and by the chain rule:

$$(\varphi_2^{-1})_* f_* v = (\varphi_1^{-1})_* (f^{-1})_* f^* v = (\varphi_1^{-1})_* v\,.$$

Thus if v is $\mathcal{C}^\infty$ on V_1, $f_* v$ is $\mathcal{C}^\infty$ on V_2. □

There is an alternative way of defining the notion of $\mathcal{C}^\infty$ which is sometimes useful in practice: As above let $\varphi : U \to V$ be a parametrization of an open set V in X and let $w = (\varphi^{-1})_*(v\big|V)$. By definition w is a $\mathcal{C}^\infty$ vector field on an open set U in $\mathbb{R}^n$, i.e. a vector field of the form, $\sum_{i=1}^n w_i \frac{\partial}{\partial x_i}$, with $w_i \in \mathcal{C}^\infty(U)$. Suppose now that X sits in an ambient Euclidean space, $\mathbb{R}^N$. Then we can think of φ as a map from U into $\mathbb{R}^N$, and for $q \in U$ and $p = \varphi(q)$

$$v(p) = (d\varphi)_q w(q) = (p, \mathrm{v}(p)) \in T_p\mathbb{R}^N$$

where $\mathrm{v}(p) = (\mathrm{v}_1(p), \ldots, \mathrm{v}_N(p))$, and

$$\mathrm{v}_i(p) = \left(\sum \frac{\partial \varphi_i}{\partial x_j} \right)(q)\,. \tag{9.2}$$

In other words the functions, $\mathrm{v}_i : X \to \mathbb{R}$, $p \to \mathrm{v}_i(p)$, are the functions, $(\varphi^{-1})^* \left(\sum \frac{\partial \varphi_i}{\partial x_j} w_j \right)$ and hence are $\mathcal{C}^\infty$ functions on X. In other words an alternative definition of $\mathcal{C}^\infty$ for vector fields is that the functions, v_i, be $\mathcal{C}^\infty$ functions on X.

A third way of formulating the notion of $\mathcal{C}^\infty$ for vector fields is to note that by Munkres §16, #3, the functions, v_i, extend to $\mathcal{C}^\infty$ functions, $\tilde{\mathrm{v}}_i$ on an open set W in $\mathbb{R}^N$ containing X. Hence the vector field v on X extends to a $\mathcal{C}^\infty$ vector field

$$\tilde{v} = \sum_{i=1}^N \tilde{\mathrm{v}}_i \frac{\partial}{\partial x_i}$$

on W. Thus we've proved

Theorem 9.5. *v is a $\mathcal{C}^\infty$ vector field on X if and only if there exists an open neighborhood W of X in $\mathbb{R}^N$ and a $\mathcal{C}^\infty$ vector field $\tilde{v}$ on W such that for all $p \in X$, $v(p) = \tilde{v}(p)$.*

We will next discusss integral curves for vector fields on manifolds and prove the manifold version of Theorem 4.9. Let $\frac{\partial}{\partial t}$ be the coordinate vector field on $\mathbb{R}$, i.e. at $c \in \mathbb{R}$ let

$$\left(\frac{\partial}{\partial t} \right)_c = (c, 1) \in T_c\mathbb{R}\,.$$

Definition 9.6. *A $\mathcal{C}^\infty$ map, $\gamma : (a,b) \to X$ is an integral curve of v if the vector fields, $\frac{\partial}{\partial t}$ and v are γ-related. In other words for $a < c < b$ and $p = \gamma(c)$*

$$(d\gamma)_c \left(\frac{\partial}{\partial_t} \right)_c = \left(p, \frac{d\gamma}{dt}(c) \right) = v(p) \,. \tag{9.3}$$

The manifold version of Theorem 4.9 asserts:

Theorem 9.7. *Let Y be a manifold, $f : X \to Y$ a $\mathcal{C}^\infty$ map and w a vector field on Y. Then if v and w are f-related and $\gamma : (a,b) \to X$ is an integral curve of v, $f \circ \gamma : (a,b) \to Y$ is an integral curve of w.*

Proof. $\frac{\partial}{\partial t}$ and v are γ-related and v and w are f-related and hence by Theorem 9.2 $\frac{\partial}{\partial t}$ and w are $f \circ \gamma$-related. □

We will now turn to some applications of these "functorial" properties of vector fields.

Theorem 9.8 (Existence). *Given a point, $p_0 \in X$ and $a \in \mathbb{R}$ there exists an interval $I = (a - T, a + T)$, a neighborhood, U_0, of p_0 in X and, for every $p \in U_0$ an integral curve, $\gamma_p : I \to X$, of v with $\gamma_p(a) = p$.*

Theorem 9.9 (Uniqueness). *Let $\gamma_i : I_i \to X$, $i = 1, 2$ be integral curves. If $a \in I_1 \cap I_2$ and $\gamma_1(a) = \gamma_2(q)$ then $\gamma_1 \equiv \gamma_2$ on $I_1 \cap I_2$ and the curve, $\gamma : I_1 \cup I_2 \to X$ defined by*

$$\gamma(t) = \begin{cases} \gamma_1(t)\,, & t \in I \\ \gamma_2(t)\,, & t \in I_2 \end{cases}$$

is an integral curve.

Theorem 9.10 (Smooth dependence on initial data). *Let U_0 be as in Theorem 9.8. Then the map*

$$F : U_0 \times I \to X\,, \quad F(p,t) = \gamma_p(t)$$

is $\mathcal{C}^\infty$.

Proof. Let's call an open subset, V, of X parametrizable if there exists an open set U in $\mathbb{R}^n$ and a parametrization $\varphi : U \to V$. Since Theorems 9.8 and 9.10 are local results it suffices to show that they are true for parametrizable open subsets of X. However, if $\varphi : U \to V$ is the parametrization of such a subset and $w = (\varphi^{-1})_*(v \big| V)$ then, by Theorem 9.7, Theorems 9.8 and 9.10 for V follow from the analogous assertions for the vector field, w, on U, i.e. from Theorems 4.5 and 4.7. As for Theorem 9.9 we can't immediately deduce this from the uniqueness result in Theorem 4.6, but we do get a local uniqueness result which tells us that if $\gamma_1(t) = \gamma_2(t)$ at a point t_0 of $I_1 \cap I_2$ then $\gamma_1(t) = \gamma_2(t)$ on a neighborhood, $(-\epsilon + t_0\,,\, \epsilon + t_0)$ and a connectivity argument then shows that γ_1 and γ_2 have to be equal on all of $I_1 \cap I_2$. Thus Theorem 9.9 also follows from the local version of this result. □

Finally, we will leave for you to check

Theorem 9.11 (Reparametrization)**.** *Let $I = (a, b)$ and for $c \in \mathbb{R}$ let $I_c = (a-c, b-c)$. Then if $\gamma : I \to X$ is an integral curve of v the reparametrized curve*

$$\gamma_c : I_c \to X\,, \quad \gamma_c(t) = \gamma(t+c) \tag{9.4}$$

is an integral curve of v.

We now turn to global results. (However we will give a rather cursory treatment of these global results since for the most part the proofs are, verbatim, identical with the proofs of the anlogous global results for vector fields on open subsets of $\mathbb{R}^n$.)

Definition 9.12. *A sequence of points $p_i \in X$, $i = 1, 2, \ldots$, tends to infinity in X if, for every compact subset, W, of X there exists an i_0 such that $p_i \notin W$ for $i > i_0$.*

Now let $\gamma : [0, a) \to X$ be an integral curve of v. As in Lecture 2 we will say that γ is a maximal integral curve if it can't be extended to an integral curve on an interval $[0, b)$, $b > a$.

We claim

Theorem 9.13. *If γ is a maximal integral curve than either*

(a) $a = +\infty$

or

(b) there exists a sequence, $t_i \in [0, a)$, such that t_i tends to a and $\gamma(t_i)$, tends to infinity in X as i tends to infinity.

For the proof of this assertion we refer to Lecture 2 since the proof is, verbatim, identical to the proof of Theorem 2.5 (and we also refer to exercise 4 in Lecture 2 for the proof of the analogous assertion for intervals $(-a, 0]$).

In particular, if X is compact this rules out alternative (b) so every vector field on X is complete and one gets a map

$$F : X \times \mathbb{R} \to X \tag{9.5}$$

by setting $F(p, t) = \gamma_p(t)$. The following assertions about this map extend to manifolds, the results of Lecture 5 (and can be proved by repeating, verbatim, the proofs of these results).

Theorem 9.14. *The map, F, is $\mathcal{C}^\infty$ and hence for all t so are the maps*

$$f_t : X \to X\,, \quad f_t(p) = \gamma_p(t)\,. \tag{9.6}$$

Moreover, these maps satisfy

$$f_{t+a} = f_t \circ f_a \tag{9.7}$$

for all $a \in \mathbb{R}$.

In particular, since $f_0(p) = \gamma_p(0) = p$, f_0 is the identity map and hence for $a = -t$, $f_t \circ f_{-t}$ is the identity map i.e. $f_{-t} = f_t^{-1}$. Thus as in Lecture 5 the vector field, v, generates a "one parameter group of diffeomorphisms" of X.

Remark. Similar results are true if we replace the hypothesis, X compact, by the hypothesis, v compactly supported.

We'll conclude this lecture by showing that the Lie differentiation operation for vector fields that we discussed in Lecture 4 makes sense for vector fields on manifolds. If $\varphi : X \to \mathbb{R}$ is a $\mathcal{C}^\infty$ function we will define its Lie derivative, $(L_v\varphi)$, by the recipe

$$(L_v\varphi)(p) = d\varphi_p(v(p)) . \tag{9.8}$$

(To make sense of the right hand side we'll delete the "c" from the base-pointed map, $d\varphi_p : T_pX \to T_c\mathbb{R}$, $c = \varphi(p)$, and think of $d\varphi_p$ as being a linear map from T_pX to $\mathbb{R}$.) Thus by (9.8) $L_v\varphi$ is a well-defined real-valued function on X. We will show that this Lie differentiation operation has the following functorial property.

Theorem 9.15. *Let Y be a manifold, $f : X \to Y$ a $\mathcal{C}^\infty$ map and $f^* : \mathcal{C}^\infty(Y) \to \mathcal{C}^\infty(X)$ the pull-back operation $f^*\psi(p) = \psi(f(p))$. Then if w is a vector field on Y which is f-related to X*

$$L_v f^*\psi = f^* L_w\psi . \tag{9.9}$$

Proof. For $p \in X$ and $q = f(p)$

$$\begin{aligned}
(L_v f^*\psi(p)) &= d(\psi \circ f)_p(v_p) \\
&= d\psi_q \circ df_p(v_p) \\
&= d\psi_q(w_q) \\
&= (L_w\psi)(q) = (f^* L_w\psi)(p) .
\end{aligned}$$

□

We will leave the following as easy exercises.

1. Show that L_v maps $\mathcal{C}^\infty(X)$ into $\mathcal{C}^\infty(X)$. *Hint:* By applying Theorem 9.15 to parametrizable open subsets of X show that this assertion follows from analogous assertion for open subsets of $\mathbb{R}^n$.

2. Show that for φ_1 and $\varphi_2 \in \mathcal{C}^\infty(X)$

$$L_v(\varphi_1\varphi_2) = L_v\varphi_1\varphi_2 + \varphi_1 L_v\varphi_2 . \tag{9.10}$$

3. Show that if v is complete and $f_t : X \to X$, $-\infty < t < \infty$, is the one parameter group of diffeomorphisms generated by V

$$L_v\varphi = \frac{d}{dt} f_t^*\varphi\big|_{t=0}. \tag{9.11}$$

Problem Set

1. Let $X \subseteq \mathbb{R}^3$ be the paraboloid, $x_3 = x_1^2 + x_2^2$ and let w be the vector field

$$w = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + \partial x_3 \frac{\partial}{\partial x_3} \,.$$

   (a) Show that w is tangent to X and hence defines by restriction a vector field, v, on X.

   (b) What are the integral curves of v?

2. Let S^2 be the unit 2-sphere, $x_1^2 + x_2^2 + x_3^2 = 1$, in $\mathbb{R}^3$ and let w be the vector field

$$w = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \,.$$

   (a) Show that w is tangent to S^2, and hence by restriction defines a vector field, v, on S^2.

   (b) What are the integral curves of v?

3. As in problem 2 let S^2 be the unit 2-sphere in $\mathbb{R}^3$ and let w be the vector field

$$w = \frac{\partial}{\partial x_3} - x_3 \left(x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + x_3 \frac{\partial}{\partial x_3} \right)$$

   (a) Show that w is tangent to S^2 and hence by restriction defines a vector field, v, on S^2.

   (b) What do its integral curves look like?

4. Let S^1 be the unit sphere, $x_1^2 + x_2^2 = 1$, in $\mathbb{R}^2$ and let $X = S^1 \times S^1$ in $\mathbb{R}^4$ with defining equations

$$\begin{aligned} f_1 &= x_1^2 + x_2^2 - 1 = 0 \\ f_2 &= x_3^2 + x_4^2 - 1 = 0 \,. \end{aligned}$$

   (a) Show that the vector field

$$w = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} + \lambda \left(x_4 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_4} \right) ,$$

   $\lambda \in \mathbb{R}$, is tangent to X and hence defines by restriction a vector field, v, on X.

(b) What are the integral curves of v?

(c) Show that $L_w f_i = 0$.

5. For the vector field, v, in problem 4, describe the one-parameter group of diffeomorphisms it generates.

6. Let X and v be as in problem 1 and let $f : \mathbb{R}^2 \to X$ be the map, $f(x_1, x_2) = (x_1, x_2, x_1^2 + x_2^2)$. Show that if u is the vector field,

$$u = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2},$$

then $f_* u = v$.

7. Verify Theorem 9.11.

8. Let X be a submanifold of X in $\mathbb{R}^N$ and let v and w be the vector fields on X and U. Denoting by ι the inclusion map of X into U, show that v and w are ι-related if and only if w is tangent to X and its restriction to X is v.

9. Verify (9.10) and (9.11).

10.* An elementary result in number theory asserts

Theorem. *A number, $\lambda \in \mathbb{R}$, is irrational if and only if the set*

$$\{m + \lambda n, \quad m \text{ and } n \text{ intgers}\}$$

is a dense *subset of* $\mathbb{R}$.

Let v be the vector field in problem 4. Using the theorem above prove that if λ is irrational then for every integral curve, $\gamma(t)$, $-\infty < t < \infty$, of v the set of points on this curve is a dense subset of X.

Lecture 10. Integration on manifolds: algebraic tools

To extend the theory of the Riemann integral from open subsets of $\mathbb{R}^n$ to manifolds one needs some algebraic tools, and these will be the topic of this section. We'll begin by reformulating a few standard facts about determinants of $n \times n$ matrices in the language of linear mappings.[1]

[1] A good reference for these standard facts is Munkres, §2.

Let V be an n-dimensional vector space and $u_1, \dots, u_n$ a basis of V. Given a linear mapping, $A : V \to V$ one gets a matrix description of A by setting

$$Au_i = \sum a_{j,i} u_j \qquad i = 1, \dots, n\,. \tag{10.1}$$

Since A is linear, it is defined by these equalities and hence by the the $n \times n$ matrix, $\mathcal{A} = [a_{i,j}]$. Moreover if A and B are linear mappings and $\mathcal{A}$ and $\mathcal{B}$ the matrices defining them, the linear mapping, AB, is defined by the product matrix, $\mathcal{C} = \mathcal{A}\mathcal{B}$, where the entries of $\mathcal{C}$ are

$$c_{i,j} = \sum_{k=1}^{n} a_{ik} b_{kj}\,. \tag{10.2}$$

We will define the *determinant* of the linear mapping, A, to be the determinant of its matrix, $\mathcal{A}$. (This definition would appear, at first glance, to depend on our choice of $u_1, \dots, u_n$, but we'll show in a minute that it doesn't.) First, however, let's note that from the product law for determinants of matrices: $\det \mathcal{A}\mathcal{B} = \det \mathcal{A} \det \mathcal{B}$, one gets an analogous product law for determinants of linear mappings

$$\det AB = \det A \det B\,. \tag{10.3}$$

Let's now choose another basis, $\mathrm{v}_1, \dots, \mathrm{v}_n$, of V and show that the definition of $\det A$ for linear maps is the same regardless of whether the u_i's or v_i's are used as a basis for V. Let $L : V \to V$ be the unique linear map with the property

$$Lu_i = \mathrm{v}_i\,, \qquad i = 1, \dots, n \tag{10.4}$$

and L^{-1} the inverse map

$$L^{-1}\mathrm{v}_i = u_i\,, \qquad i = 1, \dots, n\,. \tag{10.5}$$

Setting $B = LAL^{-1}$ we get from (10.1), (10.5) and (10.4):

$$\begin{aligned} B\mathrm{v}_i &= LAu_i = L\left(\sum a_{j,i} u_j\right) \\ &= \sum a_{j,i} Lu_j = \sum a_{j,i} \mathrm{v}_j\,. \end{aligned}$$

Thus $\det \mathcal{A}$ is the determinant of B computed using the v_i's as a basis for V. However, using the v_i's as a basis for V we get from the multiplicative law

$$\begin{aligned} \det B &= \det(LAL^{-1}) \\ &= \det L \det A \det L^{-1} \\ &= \det L (\det L)^{-1} \det A \\ &= \det A\,. \end{aligned}$$

Thus $\det \mathcal{A}$ is also the determinant of A computed using the v_i's as a basis for V.

The proof above not only shows that $\det A$ is intrinsically defined but that it is an invariant of "isomorphisms" of vector spaces.

Proposition 10.1. *Let W be an n-dimensional vector space and $L : V \to W$ a bijective linear map. Then* $\det A = \det(LAL^{-1})$.

One important property of determinant which we'll need below is the following.

Proposition 10.2. *A linear map $A : V \to V$ is onto if and only if* $\det A \neq 0$.

Proof. A is onto if and only if the vectors Au_i in (10.1) are linearly independent and hence if and only if the columns of $\mathcal{A}$ are linearly independent. However, a standard fact about determinants says that this is case if and only if $\det A \neq 0$.

$\square$

Now let

$$V^n = V \times \cdots \times V \qquad (n \text{ copies}) .$$

Definition 10.3. *A map $\sigma : V^n \to \mathbb{R}$ is a* density on V *if for all n-tuples of vectors,* $\mathrm{v}_1, \ldots, \mathrm{v}_n$*, and all linear mappings, $A : V \to V$*

$$\sigma(A\mathrm{v}_1, \ldots A\mathrm{v}_n) = |\det A|\sigma(\mathrm{v}_1, \ldots, \mathrm{v}_n) . \tag{10.6}$$

Check:

1. If $\sigma_i : V^n \to \mathbb{R}$ $i = 1, 2$ is a density on V, $\sigma_1 + \sigma_2$ is a density on V.

2. If $\sigma : V^n \to \mathbb{R}$ is a density on V and $c \in \mathbb{R}$, $c\sigma$ is a density on V.

Thus the set of density on V form a vector space. We'll denote this vector space by $|V|$.

Claim:

$|V|$ is a one-dimensional vector space.

Proof. Let $u = (u_1, \ldots, u_n)$ be a basis of V. Then for every $(\mathrm{v}_1 \ldots, \mathrm{v}_n) \in V^n$ there exists a unique linear mapping, $A : V \to V$, with

$$Au_i = \mathrm{v}_i \qquad i = 1, \ldots, n . \tag{10.7}$$

Hence if $\sigma : V^n \to \mathbb{R}$ is a density on V

$$\sigma(\mathrm{v}_1, \ldots, \mathrm{v}_n) = \sigma(Au_1, \ldots, Au_n)$$

and hence

$$\sigma(\mathrm{v}_1, \ldots, \mathrm{v}_n) = |\det A|\sigma(u_1, \ldots, u_n) \tag{10.8}$$

i.e., σ is completely determined by its value at u.

Conversely, note that if we let $\sigma(u_1, \ldots, u_n)$ be an arbitrary constant, $c \in \mathbb{R}$, and define σ by the formula (10.8), then for any linear mapping, B, one has, by definition:

$$\begin{aligned}\sigma(B\mathrm{v}_1, \ldots, B\mathrm{v}_n) &= \sigma(BAu_1, \ldots, BAu_n) \\ &= |\det BA|\sigma(u_1, \ldots, u_n)\end{aligned}$$

and by (10.8) and (10.3), the term on the right is $|\det B|\sigma(\mathrm{v}_1, \ldots, \mathrm{v}_n)$, so the map defined by (10.8) *is* a density on V.

□

Exercise 1. Show that if $\mathrm{v}_1, \ldots, \mathrm{v}_n$ are linearly dependent, the mapping, A, defined by (10.7) is not onto and conclude that $\sigma(\mathrm{v}_1, \ldots, \mathrm{v}_n) = 0$. *Hint:* Proposition 10.2.

We'll discuss now some examples.

1. In formula 2 set $\sigma(u_1, \ldots, u_n) = 1$. Then the density on V defined by this formula will be denoted by σ_u.

2. In particular let $V = \mathbb{R}^n$ and let $(e_1, \ldots, e_n) = e$ be the standard basis of $\mathbb{R}^n$. Then $\sigma_e \in |\mathbb{R}^n|$ is the unique density on V which is 1 on $(e_1, \ldots, e_n)$.

3. More generally if $p \in \mathbb{R}^n$ and

$$V = T_p\mathbb{R}^n = \{(p, \mathrm{v}), \mathrm{v} \in \mathbb{R}^n\},$$

$\sigma_{p,e} \in |T_p\mathbb{R}^n|$ is the unique density on V which is 1 on the basis of vectors: $(p, e_1), \ldots, (p, e_n)$.

4. To describe the next example, we recall that an *inner product* on a vector space, V, is a map, $B : V \times V \to \mathbb{R}$ with the properties

$$B(\mathrm{v}, w) = B(w, \mathrm{v}) \tag{10.9a}$$

$$B(\mathrm{v}_1 + \mathrm{v}_2, w) = B(\mathrm{v}_1, w) + B(\mathrm{v}_2, w) \tag{10.9b}$$

$$B(c\mathrm{v}, w) = cB(\mathrm{v}, w) \tag{10.9c}$$

$$B(\mathrm{v}, \mathrm{v}) > 0 \text{ if } \mathrm{v} \neq 0\,. \tag{10.9d}$$

For example if $V = \mathbb{R}^n$ and v and w are the vectors, $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ the usual dot product:

$$B(\mathrm{v}, w) = \mathrm{v} \cdot w = \sum_{i=1}^{n} a_i b_i$$

is an inner product on V. (For more about inner products see Munkres §1.) Given an inner product let $\sigma_B : V^n \to \mathbb{R}$ be the map

$$\sigma_B(\mathrm{v}_1, \ldots, \mathrm{v}_n) = (\det[b_{i,j}])^{\frac{1}{2}} \tag{10.10}$$

where

$$b_{i,j} = B(\mathrm{v}_i, \mathrm{v}_j)\,. \tag{10.11}$$

We claim

Proposition 10.4. *The map, σ_B, is a density on V.*

Proof. Let $A : V \to V$ be a linear map and let

$$\mathrm{v}_i' = A\mathrm{v}_i = \sum a_{k,i}\mathrm{v}_k\,.$$

Then by (10.8)

$$b_{i,j}' = B(\mathrm{v}_i', \mathrm{v}_j') = \sum_{k,\ell} a_{k,i}a_{\ell,j}B(\mathrm{v}_k, \mathrm{v}_\ell)\,. \tag{10.12}$$

Using matrix notation we can rewrite this identity in the more succinct form

$$\mathcal{B}' = \mathcal{A}^t\mathcal{B}\mathcal{A} \tag{10.13}$$

where $\mathcal{B}' = [b_{i,j}']$ and $\mathcal{A}^t$ is the transpose of $\mathcal{A}$. Therefore, since the determinant of an $n \times n$ matrix is equal to the determinant of its transpose:

$$\begin{aligned} \det\mathcal{B}' &= \det\mathcal{A}^t\mathcal{B}\mathcal{A} \\ &= \det\mathcal{A}^t\det\mathcal{B}\det\mathcal{A} \\ &= (\det\mathcal{A})^2\det\mathcal{B} \end{aligned}$$

and

$$\begin{aligned} \sigma_B(A\mathrm{v}_1, \ldots, A\mathrm{v}_n) &= (\det\mathcal{B}')^{\frac{1}{2}} \\ &= |\det\mathcal{A}|(\det\mathcal{B})^{\frac{1}{2}} \\ &= |\det A|\sigma_B(\mathrm{v}_1, \ldots, \mathrm{v}_n) \end{aligned}$$

verifying that σ_B is a density on V.

□

5. In example 4 take V to be $\mathbb{R}^n$ and B the dot product.

   Exercise 2. Show that $\sigma_B = \sigma_e$.

   Hint: for the standard basis, $e_1, \dots, e_n$, of $\mathbb{R}^n$, $e_i \cdot e_j = 1$ if $i = j$ and 0 if $i \neq j$.

   More generally, let $V = T_p\mathbb{R}^n$ and let B be the inner product defined by setting

$$B(\mathrm{v}, w) = x \cdot y$$

   for vectors, $\mathrm{v} = (p, x)$ and $w = (p, y)$ in $T_p\mathbb{R}^n$. Show that $\sigma_B = \sigma_{p,e}$.

6. Let W be an $(n-1)$-dimensional subspace of V. Given $\mathrm{v} \in V$, we define $\iota(\mathrm{v})\sigma$ to be the density on W defined by

$$\iota_{\mathrm{v}}\sigma(\mathrm{v}_1, \dots, \mathrm{v}_{n-1}) = \sigma(\mathrm{v}, \mathrm{v}_1, \dots, \mathrm{v}_{n-1})\,. \tag{10.14}$$

   Exercise 3. Check that this is a density on W.

   Hint: If $\mathrm{v} \in W$ the vectors, v, $\mathrm{v}_1, \dots, \mathrm{v}_{n-1}$ are linearly dependent, so by exercise 1 there is nothing to prove. Hence we can assume that $\mathrm{v} \notin W$. Let $A : W \to W$ be a linear map. Show that A can be extended to a linear map, $A^\sharp : V \to V$, by setting $A^\sharp \mathrm{v} = \mathrm{v}$. Check that

$$\begin{aligned} \iota(\mathrm{v})\sigma(A\mathrm{v}_1, \dots, A\mathrm{v}_{n-1}) &= \sigma(A^\sharp \mathrm{v}, A^\sharp \mathrm{v}_1, \dots, A^\sharp \mathrm{v}_{n-1}) \\ &= |\det A^\sharp| \iota(\mathrm{v})\sigma(\mathrm{v}_1, \dots, \mathrm{v}_{n-1}) \end{aligned}$$

   and check that $\det A = \det A^\sharp$.

7. Let V and W be n-dimensional vector spaces and $A : V \to W$ a bijective linear mapping. Given $\sigma \in |W|$, one defines $A^*\sigma \in |V|$ by the recipe

$$A^*\sigma(\mathrm{v}_1, \dots, \mathrm{v}_n) = \sigma(A\mathrm{v}_1, \dots, A\mathrm{v}_n)\,. \tag{10.15}$$

   We call $A^*\sigma$ the *pull-back* of σ to V by A.

   Exercise 4. Check that $A^*\sigma$ is a density on V.

   Hint: If $B : V \to V$ is a linear map then

$$\begin{aligned} A^*\sigma(B\mathrm{v}_1, \dots, B\mathrm{v}_n) &= \sigma(AB\mathrm{v}_1, \dots, AB\mathrm{v}_n) \\ &= \sigma(B'A\mathrm{v}_1, \dots, B'A\mathrm{v}_n) \end{aligned} \tag{10.16}$$

   where $B' = ABA^{-1}$. Now use Proposition 10.1.

Exercise 5.

Let $u = (u_1, \ldots, u_n)$ be a basis of V and $w = (w_1, \ldots, w_n)$ a basis of W. Show that if

$$Au_i = \sum a_{j,i} w_j$$

and $\mathcal{A} = [a_{i,j}]$

$$A^* \sigma_w = |\det \mathcal{A}| \sigma_u .$$

Hint: Observe that $A^* \sigma_w = c\sigma_\mu$ for some constant $c \in \mathbb{R}$. Now evaluate both sides of this equation on the n-tuple of vectors $(u_1, \ldots, u_n)$.

8. In particular let U and U' be open subsets of $\mathbb{R}^n$ and

$$f : (U, p) \to (U', q)$$

a diffeomorphism. Then if $u_i = (p, e_i)$ and $w_i = (q, e_i)$,

$$df_p(u_i) = \sum \frac{\partial f_i}{\partial x_j}(p) w_j$$

so

$$(df_p)^* \sigma_{q,e} = |\det \partial f_i / \partial x_j(p)| \sigma_{p,e} . \tag{10.17}$$

(This identity will play a major role in the theory of integration that we'll develop in lecture 12.)

Additional Exercises

Exercise 6.

Let B be an inner product on the vector space, V. Show that if $\mathrm{v}_1, \ldots, \mathrm{v}_n$ is an orthogonal basis of V, i.e., $B(\mathrm{v}_i, \mathrm{v}_j) = 0$ for $i \neq j$, then $\sigma_B(\mathrm{v}_1, \ldots, \mathrm{v}_n) = |\mathrm{v}_1| \cdots |\mathrm{v}_n|$, where $|\mathrm{v}_i|^2 = B(\mathrm{v}_i, \mathrm{v}_i)$.

Exercise 7.

Let $V = \mathbb{R}^2$ and let B be the dot product. Show that $\sigma_B(\mathrm{v}_1, \mathrm{v}_2)$ is the area of the parallelogram having v_1 and v_2 as adjacent edges.

Exercise 8.

Let $V = \mathbb{R}^n$ and let B be the dot product. Show that $\sigma_B(\mathrm{v}_1, \ldots, \mathrm{v}_2)$ is the volume of the parallelepiped $P(\mathrm{v}_1, \ldots, \mathrm{v}_n)$ having $\mathrm{v}_1, \ldots, \mathrm{v}_n$ as adjacent edges. (See Munkres §20, page 170.)

Exercise 9.

Let V_i, $i = 1, 2$, be an n-dimensional vector space and $W_i \subset V_i$ an $(n-1)$-dimensional subspace of V. Suppose that $A : V_1 \to V_2$ is a bijective linear mapping mapping W_1 onto W_2. Show that for $v_1 \in V$, and $\sigma \in |V_2|$

$$\iota(\mathrm{v}_1)A^*\sigma = B^*\iota(\mathrm{v}_2)\sigma \tag{10.18}$$

where $\mathrm{v}_2 = A\mathrm{v}_1$ and B is the restriction of A to W_1.

Exercise 10.

Let V be an n-dimensional vector space and W an $(n-1)$-dimensional subspace of V. Let v_1 and v_2 be elements of V, w an element of W and σ and element of $|V|$.

(a) Show that if $\mathrm{v}_1 = a\mathrm{v}_2$

$$\iota(\mathrm{v}_1)\sigma = |a|\,\iota(\mathrm{v}_2)\sigma\,.$$

(b) Show that $\iota(w)\sigma = 0$.

(c) Show that if $\mathrm{v}_1 = a\mathrm{v}_2 + w$

$$\iota(\mathrm{v}_1)\sigma = |a|\,\iota(\mathrm{v}_2)\sigma\,. \tag{10.19}$$

Exercise 11.

Let $V = \mathbb{R}^n$ and let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$. Let $W = \operatorname{span}\{e_2, \dots, e_n\}$ and let $\mathrm{v} = \sum a_i \mathrm{v}_i$. Show that for $\sigma \in |\mathbb{R}^n|$

$$\iota(\mathrm{v})\sigma(e_2, \dots, e_n) = |a_1|\sigma(e_1, \dots, e_n)\,. \tag{10.20}$$

Exercise 12.

Let V_i, $i = 1, 2, 3$, be n-dimensional vector spaces and $L_i : V_i \to V_{i+1}$, $i = 1, 2$, bijective linear maps. Show that for $\sigma \in |V_3|$

$$L_1^* L_2^* \sigma = (L_2 L_1)^* \sigma\,. \tag{10.21}$$

Lecture 11. Densities on manifolds

Let U be an open subset of $\mathbb{R}^n$. Given a function, $\varphi \in \mathcal{C}_0^\infty(U)$ we denoted the Riemann integral of φ over U by the expression

$$\int_U \varphi$$

following the conventions in Munkres and Spivak. However, most traditional text books in multivariable calculus denote this integral by

$$\int_U \varphi\, dx\,.$$

What is the meaning of the "dx" in this expression? The usual explanation is that it's a mnemonic device to remind one how the change of variables formula works for integrals. Namely, if V is an open subset in $\mathbb{R}^n$ and $f : V \to U$ a diffeomorphism, then for $\psi \in \mathcal{C}_0^\infty(U)$

$$\int_U \psi(y)\, dy = \int_V \psi(f(x)) \left|\frac{dy}{dx}\right| dx \tag{11.1}$$

where $|\frac{dy}{dx}|$ is shorthand for the expression $|\det(\partial f_i/\partial y_i)|$. The topic of this last segment of 18.101 is "integration over manifolds", and to extend the theory of the Riemann integral to manifolds we'll have to take the expressions "dx" and "dy" more seriously, that is, not just treat them as mnemonic devices. In fact the objects that one integrates when one does integration over manifolds are manifold versions of dx and dy, and to define these objects we'll first have to show how expressions like dx and $\varphi(x)\, dx$ can be converted from mnemonic devices to well-defined mathematical objects.

Definition 11.1. *Let U be an open subset of $\mathbb{R}^n$. A* density *on U is a function, σ, which assigns to each point, p, of U an element, $\sigma(p)$, of $|T_p\mathbb{R}^n|$.*

Some examples:

1. The Lebesgue density, σ_{Leb}. This is the density which assigns to each $p \in U$ the element, $\sigma_{p,e}$ of $|T_p\mathbb{R}^n|$. (We'll show in the next section that the "dx" in the paragraph above is essentially σ_{Leb}.)

2. Given any density, σ, on U and given a real-valued function, $\varphi : U \to \mathbb{R}$ one defined the density, $\varphi\sigma$, by defining

$$\varphi\sigma(p) = \varphi(p)\sigma(p)\,. \tag{11.2}$$

3. Since $\sigma_{p,e}$ is a basis of the one-dimensional vector space, $|T_p\mathbb{R}^n|$, it is clear by (11.2) that every density on U can be written as a product

$$\sigma = \varphi\sigma_{\text{Leb}}\;.$$

We'll say that σ is $\mathcal{C}^\infty$ if φ is in $\mathcal{C}^\infty(U)$ and is *compactly supported* if φ is in $\mathcal{C}_0^\infty(U)$ and we'll denote the spaces

$$\{\varphi\sigma_{\text{Leb}}\,, \quad \varphi \in \mathcal{C}^\infty(U)\}$$

and

$$\{\varphi\sigma_{\text{Leb}}\,, \quad \varphi \in \mathcal{C}_0^\infty(U)\}$$

by $\mathcal{D}^\infty(U)$ and $\mathcal{D}_0^\infty(U)$.

4. Given densities, σ_i, $i = 1, 2$, on U $\sigma_1 + \sigma_2$ is the density whose value at p is the sum, $\sigma_1(p) + \sigma_2(p) \in |T_p\mathbb{R}^n|$. It's clear that if σ_1 and σ_2 are both in $\mathcal{D}^\infty(U)$ or in $\mathcal{D}_0^\infty(U)$ then so is $\sigma_1 + \sigma_2$.

5. Let V be an open subset of $\mathbb{R}^n$ and $f : V \to U$ a diffeomorphism. Given $\sigma \in \mathcal{D}^\infty(U)$ one defines a density, $f^*\sigma$, on V by the following recipe. For each $p \in V$ and $q = f(p)$ the bijective linear map

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^n$$

gives rise (as in example 7 in the notes for Lecture 10) to a map

$$(df_p)^* : |T_q\mathbb{R}^n| \to |T_p\mathbb{R}^n|$$

and so $f^*\sigma$ is defined at p by

$$f^*\sigma = (df_p)^*\sigma(q)\,. \tag{11.3}$$

For example, if $\sigma = \sigma_{\text{Leb}}$ then by (10.17)

$$\begin{aligned}
(df_p)^* \sigma_{\text{Leb}}(q) &= (df_p)^* \sigma_{q,e} \\
&= \left| \det \left[\frac{\partial f_i}{\partial x_j}(p) \right] \right| \sigma_{p,e} \\
&= \left| \det \left[\frac{\partial f_i}{\partial x_j}(p) \right] \right| \sigma_{\text{Leb}}(p)
\end{aligned}$$

and hence

$$f^* \sigma_{\text{Leb}} = \left| \det \left[\frac{\partial f_i}{\partial x_j} \right] \right| \sigma_{\text{Leb}} \,. \tag{11.4}$$

More generally, if $\sigma = \psi \sigma_{\text{Leb}}$ with $\psi \in \mathcal{C}^\infty(V)$

$$f^* \sigma = \psi \circ f \left| \det \left[\frac{\partial f_i}{\partial x_j} \right] \right| \sigma_{\text{Leb}} \,. \tag{11.5}$$

Since the function on the right is $\mathcal{C}^\infty$ this proves

Proposition 11.2. *If σ is in $\mathcal{D}^\infty(U)$, $f^*\sigma$ is in $\mathcal{D}^\infty(V)$.*

We will next show how to extend the results above to manifolds.

We'll begin with the definition of density (which is exactly the same for manifolds as for open sets in $\mathbb{R}^n$).

Definition 11.3. *Let X be an n-dimensional submanifold of $\mathbb{R}^N$. A* density *on X is a function which assigns to each $p \in X$ an element $\sigma(p)$ of $|T_pX|$.*

Examples.

Example 1. The volume density, σ_{vol}. For each $p \in X$, T_pX is by definition a vector subspace of $T_p\mathbb{R}^N$. Moreover since

$$T_p\mathbb{R}^N = \{(p, v)\,, \quad v \in \mathbb{R}^N\}$$

one has a bijective linear map,

$$T_p\mathbb{R}^N \to \mathbb{R}^N \tag{11.6}$$

mapping (p, v) to v. Since $\mathbb{R}^N$ is equipped with a natural inner product (the "dot product") one can equip $T_p\mathbb{R}^N$ with this inner product via the identification (11.6), and since T_pX sits inside $T_p\mathbb{R}^N$ as a vector subspace, this gives us an inner product on T_pX. Let's call this inner product B_p.

Definition 11.4. *The volume density on X is the density defined by the formula*

$$\sigma_{\mathrm{vol}}(p) = \sigma_{B_p} \tag{11.7}$$

at each point $p \in X$.

N.B. The term on the right is by definition an element of $|T_pX|$, so this formula does assign to each $p \in X$ an element of $|T_pX|$ as required. (See example 4 in the notes for Lecture 10.)

Exercise 1. Check that for $X = \mathbb{R}^n$, $\sigma_{\mathrm{vol}} = \sigma_{\mathrm{Leb}}$.

Example 2. Given a density, σ, on X and a real-valued function, $\varphi : X \to \mathbb{R}$ one defines the density $\varphi\sigma$, as above, by

$$(\varphi\sigma)(p) = \varphi(p)\sigma(p)$$

at $p \in X$.

Example 3. Given densities, σ_i, $i = 1, 2$, one defines their sum, $\sigma_1 + \sigma_2$, as above, by

$$(\sigma_1 + \sigma_2)(p) = \sigma_1(p) + \sigma_2(p)\,.$$

Example 4. Let X and Y be n-dimensional manifolds and $f : X \to Y$ a diffeomorphism. Given a density, σ, on Y one defines as above, the pull-back density, $f^*\sigma$ on X by the recipe (11.3). Namely if $p \in X$ and $q = f(p)$ the derivative of f at p is a bijective linear map

$$df_p : T_pX \to T_pY$$

and from this map one gets a linear map

$$df_p^* : |T_qY| \to |T_pX|\,,$$

and one defines $f^*\sigma$ by:

$$(f^*\sigma)(p) = (df_p)^*\sigma(q)\,.$$

Exercise 2. Check that if σ is a density on Y and $\psi : Y \to \mathbb{R}$ a real valued function

$$f^*(\psi\sigma) = f^*\psi f^*\sigma \tag{11.8}$$

i.e., the pull-back operation on densities, is consistent with the pull-back operation on functions that we defined earlier in the course.

Exercise 3. Let Z be an n-dimensional manifold and $g : Y \to Z$ a diffeomorphism. Check that if σ is a density on Z

$$f^*(g^*\sigma) = (g \circ f)^*\sigma\,. \tag{11.9}$$

Hint: The manifold version of the chain rule for derivatives of mappings.

We will next show how to define analogues of the spaces $\mathcal{D}^\infty(U)$ and $\mathcal{D}_0^\infty(U)$ for manifolds. Recall that if U is an open subset of X, a *parametrization* of U is a pair, (U_0, φ), where U_0 is an open subset of $\mathbb{R}^n$ and $\varphi : U_0 \to U$ is a diffeomorphism.

Definition 11.5. *Let σ be a density on X. We will say that σ is $\mathcal{C}^\infty$ on U if $\varphi^*\sigma \in \mathcal{D}^\infty(U_0)$.*

Claim: This definition doesn't depend on (U_0, φ).

Proof. Let (U_0', φ') be another parametrization of U and let $f : U_0 \to U_0'$ be the diffeomorphism, $f = (\varphi')^{-1} \circ \varphi$. Letting $\sigma_0 = \varphi^*\sigma$ and $\sigma_0' = (\varphi_0')^*\sigma$ we get from (11.9)

$$\sigma_0 = f^*\sigma_0' \tag{11.10}$$

and hence by Proposition 11.2, σ_0' is in $\mathcal{D}^\infty(U_0')$ if and only if σ_0 is in $\mathcal{D}^\infty(U_0)$. □

Example 5. σ_{vol}. Let (U_0, φ) be a parameterization of U. Since X sits inside $\mathbb{R}^N$ we can regard φ as a $\mathcal{C}^\infty$ map

$$\varphi : U_0 \to \mathbb{R}^N, \tag{11.11}$$

and for $p \in U_0$ and $q = \varphi(p)$ we can regard $d\varphi_p$ as an injective linear map

$$d\varphi_p : T_p\mathbb{R}^n \to T_p\mathbb{R}^N \tag{11.12}$$

which maps $T_p\mathbb{R}^n$ bijectively onto the subspace, T_pX, of $T_p\mathbb{R}^N$. Let $\varphi_1, \ldots, \varphi_N$ be the coordinates of the map (11.11) and let $e_1, \ldots, e_n$ be the standard basis vectors of $\mathbb{R}^N$. Then if $u_i = (p, e_i)$

$$d\varphi_p(u_i) = w_i = (q, \mathrm{v}_i) \tag{11.13}$$

where

$$\mathrm{v}_i = \left(\frac{\partial \varphi_1}{\partial x_i}(p), \cdots, \frac{\partial \varphi_N}{\partial x_i}(p)\right) \tag{11.14}$$

and hence

$$\begin{aligned} ((d\varphi_p)^*\sigma_{B_q})(u_1, \ldots, u_n) &= \sigma_{B_q}(w_1, \ldots, w_n) \\ &= (\det[b_{i,j}(p)])^{\frac{1}{2}} \end{aligned} \tag{11.15}$$

where $b_{i,j}(p)$ is the matrix

$$b_{i,j}(p) = \mathrm{v}_i \cdot \mathrm{v}_j = \sum_{k=1}^N \frac{\partial \varphi_k}{\partial x_i}(p) \frac{\partial \varphi_k}{\partial x_j}(p). \tag{11.16}$$

Since $\sigma_{p,e}$ is a basis of the one-dimensional vector space, $|T_p\mathbb{R}^n|$,

$$(d\varphi_p)^*\sigma_{B_q} = c\sigma_{p,e} \tag{11.17}$$

for some constant, c, and since, by definition,

$$\sigma_{p,e}(u_1, \dots, u_n) = 1$$

we see from (11.15) that this constant has to be equal to the expression on the right hand side of (11.15) i.e.,

$$(d\varphi)_p^* \sigma_{B_q} = (\det[b_{i,j}(p)])^{\frac{1}{2}} \sigma_{p,e} \,. \tag{11.18}$$

Therefore since $\sigma_{B_q} = \sigma_{\text{vol}}(q)$ and $\sigma_{p,e} = \sigma_{\text{Leb}}(p)$ we get from (11.18)

$$\varphi^* \sigma_{\text{vol}} = (\det[\psi_{i,j}])^{\frac{1}{2}} \sigma_{\text{Leb}} \tag{11.19}$$

where $\psi_{i,j}$ is the function

$$\psi_{i,j} = \sum \frac{\partial \varphi_k}{\partial x_i} \frac{\partial \varphi_k}{\partial x_j} \,. \tag{11.20}$$

Thus σ_{vol} is $\mathcal{C}^\infty$ on U.

Definition 11.6. *We will say that a density, σ, on X is $\mathcal{C}^\infty$ if for every point, $p \in X$ it is $\mathcal{C}^\infty$ on a neighborhood, U of p.*

Notation: We will denote by $\mathcal{D}^\infty(X)$ the space of $\mathcal{C}^\infty$ densities on X and by $\mathcal{D}_0^\infty(X)$ the space of compactly supported $\mathcal{C}^\infty$ densities on X.

Exercise 4. Show that

$$\mathcal{D}^\infty(X) \;=\; \{\varphi \sigma_{\text{vol}} \,, \quad \varphi \in \mathcal{C}^\infty(X)\}$$

and

$$\mathcal{D}_0^\infty(X) \;=\; \{\varphi \sigma_{\text{vol}} \,, \quad \varphi \in \mathcal{C}_0^\infty(X)\}.$$

Hint: For every $p \in X$, σ_{B_p} is a basis vector for the one-dimensional vector space, $|T_p X|$.

Let Y be an n-dimensional manifold and $f : X \to Y$ a diffeomorphism. We will show that Proposition 11.2 is true for manifolds.

Proposition 11.7. *If $\sigma \in \mathcal{D}^\infty(Y)$, then $f^*\sigma \in \mathcal{D}^\infty(X)$.*

Proof. Let U be an open subset of X and $\varphi : U_0 \to U$ a parameterization of U. Let $V = f(U)$. Then $f \circ \varphi : U_0 \to V$ is a parameterization of V and by (11.9)

$$\varphi^* f^* \sigma = (f \circ \varphi)^* \sigma \,.$$

Since σ is a $\mathcal{C}^\infty$ density on Y the right side of this identity is in $\mathcal{D}^\infty(U_0)$ and hence so is the left side.

□

Some additional exercises.

Exercise 5. Given a $\mathcal{C}^\infty$ function $f : \mathbb{R} \to \mathbb{R}$ its graph

$$X = \{(x, f(x))\,, \quad x \in \mathbb{R}\}$$

is a submanifold of $\mathbb{R}^2$ and

$$\varphi : \mathbb{R}^2 \to X\,, \quad x \to (x, f(x))$$

is a diffeomorphism. Show that

$$\varphi^* \sigma_{\mathrm{vol}} = \left(1 + \left(\frac{df}{dx}\right)^2\right)^{\frac{1}{2}} \sigma_{\mathrm{Leb}}\,. \tag{11.21}$$

Exercise 6. Given a $\mathcal{C}^\infty$ function $f : \mathbb{R}^2 \to \mathbb{R}$ its graph

$$X = \{(x, f(x))\,, \quad x \in \mathbb{R}^2\}$$

is a submanifold of $\mathbb{R}^3$ and

$$\varphi : \mathbb{R}^2 \to X\,, \quad x \to (x, f(x))$$

is a diffeomorphism. Show that

$$\varphi^* \sigma_{\mathrm{vol}} = \left(1 + \left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2\right)^{\frac{1}{2}} \sigma_{\mathrm{Leb}}\,. \tag{11.22}$$

Exercise 7*. Given a $\mathcal{C}^\infty$ function, $f : \mathbb{R}^n \to \mathbb{R}$ its graph

$$X = \{(x, f(x))\,, \quad x \in \mathbb{R}^n\}$$

is a submanifold of $\mathbb{R}^{n+1}$ and

$$\varphi : \mathbb{R}^n \to X\,, \quad x \to (x, f(x)) \tag{11.23}$$

is a diffeomorphism. Show that

$$\varphi^* \sigma_{\mathrm{vol}} = \left(1 + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2\right)^{\frac{1}{2}} \sigma_{\mathrm{Leb}}\,. \tag{11.24}$$

Hints:

a. Let $\mathrm{v} = (c_1, \ldots, c_n) \in \mathbb{R}^n$. Show that if $C : \mathbb{R}^n \to \mathbb{R}^n$ is the linear mapping defined by the matrix $[c_i c_j]$ then $C\mathrm{v} = (\sum c_i^2)\mathrm{v}$ and $Cw = 0$ if $w \cdot \mathrm{v} = 0$.

b. Conclude that the eigenvalues of C are $\lambda_1 = \sum_{i=1}^n c_i^2$ and $\lambda_2 = \cdots = \lambda_n = 0$.

c. Show that the determinant of $I + C$ is $1 + \sum c_i^2$.

d. Use the preceding results to compute the determinant of the matrix (11.20) where φ is the mapping (11.23).

Exercise 8. Let X be the unit $(n-1)$-sphere in $\mathbb{R}^n$ and let v be the vector field,

$$v = \sum x_i \, \partial/\partial x_i \, .$$

For $p \in X$ let $V = T_p\mathbb{R}^n$ and let $W = T_pX$. Show that if σ_{vol} is the volume density on X and σ_{Leb} the Lebesgue density on $\mathbb{R}^n$ then at p these two densities are related by

$$\sigma_{\text{vol}}(p) = \iota(v_p)\sigma_{\text{Leb}}(p) \, . \tag{11.25}$$

(For the definition of the term on the right see example 6 in the notes for Lecture 10.)

Lecture 12. Integrating densities

In this section we will show how to integrate densities over manifolds. First, however, we will have to explain how to integrate densities over open subsets, U, of $\mathbb{R}^n$. Recall that if σ is in $\mathcal{D}^\infty(U)$ it can be written as a product, $\sigma = \psi\sigma_{\text{Leb}}$, where ψ is in $\mathcal{C}^\infty(U)$. We will say that σ is integrable over U if ψ is integrable over U, and will define the integral of σ over U to be the usual Riemann integral

$$\int_U \sigma = \int_U \psi \, dx \, . \tag{12.1}$$

The advantage of using "density" notation for this integral is that it makes the change of variables formula more transparent. Namely if U_1 is an open subset of $\mathbb{R}^n$ and $f : U_1 \to U$ a diffeomorphism, then by (11.5) $f^*\sigma = \psi_1\sigma_{\text{Leb}}$ where

$$\psi_1 = \psi \circ f \left| \det \left[\frac{\partial f_i}{\partial x_j} \right] \right| \tag{12.2}$$

and hence by the change of variables formula[2] ψ_1 is integrable over U_1 and

$$\int_{U_1} \psi_1 \, dx = \int_U \psi \, dx \, .$$

Thus using density notation the change of variables formula takes the much simpler form

$$\int_{U_1} f^*\sigma = \int_U \sigma \, . \tag{12.3}$$

[2]See Theorem 17.2 in Munkres.

Now let $X \subseteq \mathbb{R}^N$ be an n-dimensional manifold. Our goal below will be to define the integral

$$\int_W \sigma \tag{12.4}$$

where W is an open subset of X and σ is a compactly supported $\mathcal{C}^\infty$ density. We'll first show how to define this integral when the support of σ is contained in a "parametrizable" open subset of X and then, using a partition of unity argument, define it in general.

Definition 12.1. *An open subset, U, of X is* parametrizable *if there exists an open set, U_0, in $\mathbb{R}^n$ and a diffeomorphism, $\varphi_0 : U_0 \to U$.*

In other words "U is parametrizable" means that there exists a parameterization, (U_0, φ_0), of U. It's clear that if U is parametrizable every open subset of U is parametrizable, and, in particular, if U_1 and U_2 are parametrizable, so is $U_1 \cap U_2$. Moreover the definition of manifold says that every point, $p \in X$, is contained in a parametrizable open set.

Let σ be an element of $\mathcal{D}_0^\infty(X)$ whose support is contained in a parametrizable open set U. Picking a parameterization, $\varphi_0 : U_0 \to U$ we will define the integral of σ over W by defining it to be

$$\int_W \sigma = \int_{W_0} \varphi_0^* \sigma \tag{12.5}$$

where $W_0 = \varphi_0^{-1}(W)$. Note that since σ is compactly supported on U, $\varphi_0^*\sigma$ is a product, $\varphi_0^*\sigma = \psi\sigma_{\mathrm{Leb}}$, with ψ in $\mathcal{C}_0^\infty(U_0)$. Hence by Munkres, Theorem 15.2, ψ is integrable over W_0 and hence so is $\varphi^*\sigma$. We will prove

Lemma 12.2. *The definition (12.5) doesn't depend on the choice of the parameterization, (U_0, φ_0).*

Proof. Let (U_1, φ_1) be another parameterization of U and let $f = \varphi_1^{-1} \circ \varphi_0$. Since φ_0 and φ_1 are diffeomorphisms of U_0 and U_1 onto U f is a diffeomorphism of U_0 onto U_1 with the property

$$\varphi_1 \circ f = \varphi_0 \,. \tag{12.6}$$

In particular if $W_i = \varphi_i^{-1}(W)$, $i = 0, 1$ it follows from (12.6) that f maps W_0 diffeomorphically onto W_1 and from the chain rule it follows that $f^*\varphi_1^*\sigma = \varphi_0^*\sigma$. Hence by (12.3)

$$\int_{W_0} \varphi_0^*\sigma = \int_{W_1} \varphi_1^*\sigma \,. \tag{12.7}$$

In other words (12.5) is unchanged if we substitute (U_1, φ_1) for (U_0, φ_0).

□

From the additivity of the Riemann integral for integrable functions on open subsets of $\mathbb{R}^n$ we also conclude

Lemma 12.3. *If $\sigma_i \in \mathcal{D}_0^\infty(X)$, $i = 1, 2$, is supported on U*

$$\int_W \sigma_1 + \sigma_2 = \int_W \sigma_1 + \int_W \sigma_2$$

and if $\sigma \in \mathcal{D}_0^\infty(X)$ is supported on U and $c \in \mathbb{R}$

$$\int_W c\sigma = c \int_W \sigma\,.$$

To define the integral (12.4) for arbitrary elements of $\mathcal{D}_0^\infty(X)$ we will resort to the same partition of unity arguments that we used earlier in the course to define improper integrals of functions over open subsets of $\mathbb{R}^n$. To do so we'll need the following manifold version of Munkres' Theorem 16.3.

Theorem 12.4. *Let*

$$\mathbb{U} = \{U_\alpha\,,\ \alpha \in \mathcal{I}\} \tag{12.8}$$

be a covering of X be open subsets. Then there exists a family of functions, $\rho_i \in \mathcal{C}_0^\infty(X)$, $i = 1, 2, 3, \ldots$, with the properties

> *(a)* $\rho_i \geq 0$.
>
> *(b) For every compact set, $C \subseteq X$ there exists a positive integer N such that if $i > N$,* supp $\rho_i \cap C = \emptyset$.
>
> *(c)* $\sum \rho_i = 1$.
>
> *(d) For every i there exists an $\alpha \in \mathcal{I}$ such that* supp $\rho_i \subseteq U_\alpha$.

Remark. *Conditions (a)–(c) say that the ρ_i's are a partition of unity and (d) says that this partition of unity is subordinate to the covering (12.8).*

Proof. For each $p \in X$ choose an open set O_p in $\mathbb{R}^N$ containing p such that the closure of $O_p \cap X$ in X is compact and such that

$$O_p \cap X \subseteq U_\alpha \tag{12.9}$$

for some α. Let O be the union of the O_p's.

By the theorem in Munkres that we cited above there exists a partition of unity, $\tilde{\rho}_i \in \mathcal{C}_0^\infty(\mathcal{O})$, $i = 1, 2, \ldots$, subordinate to the covering of $\mathcal{O}$ by the $\mathcal{O}_p$'s. Let ρ_i be the restriction of $\tilde{\rho}_i$ to X. Since the support of $\tilde{\rho}_i$ is compact and contained in some O_p, the support of ρ_i is compact, so $\rho_i \in \mathcal{C}_0^\infty(X)$ and it's clear that the ρ_i's inherit from the $\tilde{\rho}_i$'s the properties (a)–(d).

∎

Now let the covering (12.8) be any covering of X by parametrizable open sets and let $\rho_i \in \mathcal{C}_0^\infty(X)$, $i = 1, 2, \dots$, be a partition of unity subordinate to this covering. Given $\sigma \in \mathcal{D}_0^\infty(X)$ we will define the integral of σ over W by the sum

$$\sum_{i=1}^{\infty} \int_W \rho_i \sigma \, . \tag{12.10}$$

Note that since each ρ_i is supported in some U_α the individual summands in this sum are well-defined and since the support of σ is compact all but finitely many of these summands are zero by part (b) of Theorem 12.4. Hence the sum itself is well-defined. Let's show that this sum doesn't depend on the choice of $\mathbb{U}$ and the ρ_i's. Let $\mathbb{U}'$ be another covering of X by parametrizable open sets and ρ_j', $j = 1, 2, \dots$, a partition of unity subordinate to $\mathbb{U}'$. Then

$$\begin{aligned} \sum_j \int_W \rho_j' \sigma &= \sum_j \int_W \sum_i \rho_j' \rho_i \sigma \\ &= \sum_j \left(\sum_i \int_W \rho_j' \rho_i \sigma \right) \end{aligned} \tag{12.11}$$

by Lemma 12.3. Interchanging the orders of summation and resumming with respect to the j's this sum becomes

$$\sum_i \int_W \sum_j \rho_j' \rho_i \sigma$$

or

$$\sum_i \int_W \rho_i \sigma \, .$$

Hence

$$\sum_i \int_W \rho_j' \sigma = \sum_i \int_W \rho_i \sigma \, ,$$

so the two sums are the same. Q.E.D.

From (12.10) and Lemma 12.3 one easily deduces

Proposition 12.5. *For $\sigma_i \in \mathcal{D}_0^\infty(X)$, $i = 1, 2$*

$$\int_W \sigma_1 + \sigma_2 = \int_W \sigma_1 + \int_W \sigma_2 \tag{12.12}$$

and for $\sigma \in \mathcal{D}_0^\infty(X)$ and $c \in \mathbb{R}$

$$\int_W c\sigma = c \int_W \sigma \, . \tag{12.13}$$

In the definition of the integral (12.4) we've allowed W to be an arbitrary open subset of X but required $\sigma \in \mathcal{D}^\infty(X)$ to be compactly supported. This integral is also well-defined if we allow σ to be an arbitrary element of $\mathcal{D}^\infty(X)$ but require the closure of W in X to be compact. To see this, note that under this assumption the sum (12.10) is still a finite sum, so the definition of the integral still makes sense, and the double sum on the right side of (12.11) is still a finite sum so it's still true that the definition of the integral doesn't depend on the choice of partitions of unity. In particular if the closure of W in X is compact we will define the volume of W to be the integral,

$$\mathrm{vol}(W) = \int_W \sigma_{\mathrm{vol}}\,, \tag{12.14}$$

and if X itself is compact we'll define its volume to be the integral

$$\mathrm{vol}(X) = \int_X \sigma_{\mathrm{vol}}\,. \tag{12.15}$$

(For an alternative way of defining the volume of a manifold see Munkres, §22.)

We'll conclude this discussion of integration by proving a manifold version of the change of variables formula (12.3).

Theorem 12.6. *Let X' and X be n-dimensional manifolds and $f : X' \to X$ a diffeomorphism. If W is an open subset of X and $W' = f^{-1}(W)$*

$$\int_{W'} f^*\sigma = \int_W \sigma \tag{12.16}$$

for all $\sigma \in \mathcal{D}_0^\infty(X)$.

Proof. By (12.10) the integrand of the integral above is a finite sum of $\mathcal{C}^\infty$ densities, each of which is supported on a parametrizable open subset, so we can assume that σ itself as this property. Let V be a parametrizable open set containing the support of σ and let $\varphi_0 : U \to V$ be a parameterization of V. Since f is a diffeomorphism its inverse exists and is a diffeomorphism of X onto X_1. Let $V' = f^{-1}(V)$ and $\varphi_0' = f^{-1} \circ \varphi_0$. Then $\varphi_0' : U \to V'$ is a parameterization of V'. Moreover, $f \circ \varphi_0' = \varphi$ so if $W_0 = \varphi_0^{-1}(W)$ we have

$$W_0 = (\varphi_0)^{-1}(f^{-1}(W)) = (\varphi_0')^{-1}(W')$$

and by the chain rule we have

$$\varphi_0^*\sigma \;=\; (f \circ \varphi_0')^*\sigma = (\varphi_0')^* f^*\sigma$$

hence

$$\int_W \sigma \;=\; \int_{W_0} \varphi_0^*\sigma = \int_{W_0} (\varphi_0')^*(f^*\sigma) = \int_{W'} f^*\sigma\,.$$

□

Lecture 13. Lie derivatives of densities

In the next three lectures of this course we will prove a manifold version of one of the fundamental theorems in multi-variable calculus: the divergence theorem. The calculus version of this theorem says that if $S \subseteq \mathbb{R}^3$ is a closed surface and v a vector field, the flux of v through S is equal to integral of the divergence of v over the region bounded by S. To extend this theorem to manifolds we will need manifold versions of the notion of divergence and flux. The notion of divergence is closely related to another important manifold notion: *the Lie derivative of a density by a vector field*. We'll discuss both of these concepts below, and discuss the concept of flux in Lecture 14.

Let X be an n-dimensional manifold and v a vector field on X. To simplify slightly the exposition in what follows we'll assume that v is complete, and hence that it generates a one-parameter group of diffeomorphisms

$$f_t : X \to X\,, \quad -\infty < t < \infty\,. \tag{13.1}$$

Let's recall that if φ is a $\mathcal{C}^\infty$ function on X its Lie derivative with respect to v can be defined by the formula

$$L_v\varphi = \left(\frac{d}{dt} f_t^*\varphi\right)(t=0)\,. \tag{13.2}$$

This formula makes sense for densities as well. Namely if σ is an element of $\mathcal{D}^\infty(X)$ we can define its Lie derivative by the recipe:

$$L_v\sigma = \left(\frac{d}{dt} f_t^*\sigma\right)(t=0) \quad (*)\,. \tag{13.3}$$

Moreover, the operations (13.2) and (13.3) are compatible: if φ is a $\mathcal{C}^\infty$ function and σ a $\mathcal{C}^\infty$ density then[3]

$$f_t^*\varphi\sigma \;=\; (f_t^*\varphi)f_t^*\sigma\,,$$

hence

[3]This definition makes sense without the assumption that v be complete, however it is slightly more complicated. In the vector field segment of this course we pointed out that for every point, p, in X there exists a neighborhood, U, of p, an interval, $-\epsilon < t < \epsilon$, and a family of local diffeomorphisms

$$f_t : U \to X$$

such that for $q \in U$ the curve, $\gamma_q(t) = f_t(q)$, $-\epsilon < t < \epsilon$, is an integral curve of v with initial point, $\gamma_q(0) = q$. In other words, v generates a "local" one-parameter group of diffeomorphisms on U, so the Lie derivative of σ with respect to v can still be defined by the recipe (13.3) in the vicinity of p for every $p \in X$.

$$\frac{d}{dt} f_t^*(\varphi\sigma) \;=\; \frac{d}{d^t}(f_t^*\varphi f_t^*\sigma) = \left(\frac{d}{dt} f_t^*\varphi\right) f_t^*\sigma + f_t^*\varphi\left(\frac{d}{dt} f_t^*\sigma\right)$$

which for $t = 0$ reduces to:

$$L_v(\varphi\sigma) = (L_v\varphi)\sigma + \varphi L_v\sigma\,. \tag{13.4}$$

To see what this Lie differentiation operation looks like "locally" let's compute (13.3) for the special case of open subsets of $\mathbb{R}^n$. We'll begin by proving a linear algebra lemma which we'll need for this computation.

Lemma 13.1. *Let $A(t) = [a_{i,j}(t)]$, $-\epsilon < t < \epsilon$, be an $n \times n$ matrix whose entries, $a_{i,j}(t)$, are $\mathcal{C}^\infty$ functions of t. Then if $A(0)$ is the identity matrix*

$$\frac{d}{dt}(\det A)(0) \;=\; \operatorname{trace}\frac{d}{dt}A(0) \tag{13.5}$$

where

$$\operatorname{trace}\frac{d}{dt}A(0) \;=\; \sum_{i=1}^{n}\frac{d}{dt}a_{i,i}(0)\,. \tag{13.6}$$

Proof. By Theorem 2.15 in Munkres, §2

$$\det A(t) = \sum_{i=1}^{n}(-1)^{1+i}a_{1,i}(t)\det A_{1,i}(t)$$

where $A_{1,i}(t)$ is the $(n-1)\times(n-1)$ matrix obtained by deleting from $A(t)$ its first row and i^{th} column. Thus $\frac{d}{dt}\det A(t)$ is equal at $t = 0$ to the sum of

$$\sum(-1)^{1+i}\frac{d}{dt}a_{1,i}(0)\det A_{1,i}(0) \tag{13.7}$$

and

$$\sum(-1)^{1+i}a_{1,i}(0)\frac{d}{dt}\det A_{1,i}(0)\,. \tag{13.8}$$

However, $A_{1,1}(0)$ is the identity $(n-1)\times(n-1)$ matrix and for $i \neq 1$, the first column of the matrix $A_{1,i}(0)$ consists entirely of zeros. Thus $\det A_{1,1}(0) = 1$ and, for $i \neq 1$, $\det A_{1,i}(0) = 0$, so (13.7) is just $\frac{d}{dt}a_{1,1}(0)$. Moreover, $a_{1,1}(0) = 1$ and $a_{1,i}(0) = 0$ for $i \neq 1$, so (13.8) is just $\frac{d}{dt}A_{1,1}(0)$. Arguing by induction we can assume that the theorem is true for $n-1$ and hence that (13.8) is equal to

$$\sum_{i=2}^{n}\frac{d}{dt}a_{i,i}(0)\,.$$

Adding to this the term (13.7), which we've just observed to be $\frac{d}{dt}a_{1,1}(0)$, we get the formula (13.5).

□

Now let U be an open subset of $\mathbb{R}^n$ and let v be the vector field

$$v = \sum v_i \frac{\partial}{\partial x_i}\,, \tag{13.9}$$

A density, $\sigma \in \mathcal{D}^\infty(U)$ can be written as a product $\varphi\sigma_{\mathrm{Leb}}$ with φ in $\mathcal{C}^\infty(U)$, so by (13.4)

$$L_v\sigma = (L_v\varphi)\sigma_{\mathrm{Leb}} + \varphi L_v\sigma_{\mathrm{Leb}}\,, \tag{13.10}$$

so to compute $L_v\sigma$ it suffices to compute $L_v\sigma_{\mathrm{Leb}}$. Let

$$f_t(x) = (f_1(x,t),\cdots,f_n(x,t))$$

then by (11.4):

$$f_t^*\sigma_{\mathrm{Leb}} = |\det J(t)|\sigma_{\mathrm{Leb}} \tag{13.11}$$

where

$$J(t) = \left[\frac{\partial f_i}{\partial x_j}(x,t)\right]\,.$$

Note that since f_0 is the identity map, $J(0)$ is the identity matrix, so its determinant is 1. Moreover, $J(t)$ is an invertible matrix, so its determinant is non-zero. Hence since $\det J(t)$ depends continuously on t it has to be positive for all t, so we can drop the absolute value sign from (13.11) and write (13.11) in the form

$$f_t^*\sigma_{\mathrm{Leb}} = \det J(t)\sigma_{\mathrm{Leb}}\,.$$

Thus

$$L_v\sigma_{\mathrm{Leb}} = \left(\frac{d}{dt}f_t^*\sigma_{\mathrm{Leb}}\right)(t=0) = \frac{d}{dt}(\det J)(0)\sigma_{\mathrm{Leb}}\,,$$

and hence by the lemma:

$$L_v\sigma_{\mathrm{Leb}} = \left(\sum_{i=1}^n \frac{d}{dt}\frac{\partial f_i}{\partial x_i}(x,0)\right)\sigma_{\mathrm{Leb}}\,. \tag{13.12}$$

Now recall that $f_t(x) = \gamma_x(t)$ where $\gamma_x(t)$ is the integral curve of v with initial point $\gamma_x(0) = x$, i.e,

$$\frac{d}{dt}\gamma_x(0) = v(x)$$

and hence

$$\frac{d}{dt} f_i(x,0) = v_i(x)\,.$$

Differentiating this identity with respect to x_i we get

$$\frac{d}{dt}\frac{\partial f_i}{\partial x_i}(x,0) = \sum \frac{\partial v_i(x)}{\partial x_i} \tag{13.13}$$

and hence, finally, by (13.12)

$$L_v \sigma_{\text{Leb}} \;=\; \operatorname{div}(v)\sigma_{\text{Leb}} \tag{13.14}$$

where

$$\operatorname{div}(v) \;=\; \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i} \tag{13.15}$$

is the *divergence* of v. Coming back to (13.10) we get for $L_v\sigma$ the formula

$$\left(\sum v_i \frac{\partial \varphi}{\partial x_i} + \varphi \sum \frac{\partial v_i}{\partial x_i}\right)\sigma_{\text{Leb}}$$

so we can write (13.10) in the more compact form:

$$L_v\sigma = \left(\sum \frac{\partial}{\partial x_i}(v_i\varphi)\right)\sigma_{\text{Leb}}\;. \tag{13.16}$$

We will next show how the Lie differentiation operation behaves under global change of variables. Let X and Y be n-dimensional manifolds and $\gamma : X \to Y$ a diffeomorphism. In the "theory of manifolds" segment of this course we showed that if v is a vector field on X and $w = \gamma_* v$ then for $\varphi \in \mathcal{C}^\infty(Y)$

$$L_v\gamma^*\varphi = \gamma^* L_w\varphi\,. \tag{13.17}$$

We will prove that the same identity holds for densities, i.e., for σ in $\mathcal{D}^\infty(Y)$

$$L_v\gamma^*\sigma = \gamma^* L_w\sigma\,. \tag{13.18}$$

Proof. Let $f_t : X \to X$ be the one-parameter group of diffeomorphisms generated by v. Then the one-parameter group of diffeomorphisms generated by w is the group

$$g_t \;=\; \gamma \circ f_t \circ \gamma^{-1}$$

so

$$\begin{aligned} g_t \circ \gamma \;&=\; \gamma \circ f_t \qquad \text{and hence} \\ \gamma^* g_t^* \sigma \;&=\; f_t^* \gamma^* \sigma\,. \end{aligned}$$

Differentiating this identity with respect to t and setting $t = 0$ we get (13.15). $\square$

One application of this change of variables formula is the following result (which we've implicitly been assuming to be true, but nonetheless requires a proof).

Theorem 13.2. *If σ is in $\mathcal{D}^\infty(X)$, so is $L_v\sigma$.*

Proof. One has to prove that $L_v\sigma$ is $\mathcal{C}^\infty$ on parametrizable open subset of X and hence, by (13.18), that $L_v\sigma$ is $\mathcal{C}^\infty$ when X is an open subset of $\mathbb{R}^n$. This, however, is obvious by the formula (13.16). □

The statement and proof of the divergence theorem requires some further machinery (which we'll develop in the next lecture) but we can already prove an important special case of this theorem.

Theorem 13.3. *If σ is in $\mathcal{D}_0^\infty(X)$*

$$\int_X L_v\sigma = 0\,. \tag{13.19}$$

Proof. Let $f_t : X \to X$ $-\infty < t < \infty$ be the one-parameter group of diffeomorphisms generated by v. Then by the global change of variables formula for integration which we proved in Lecture 12

$$\int_X f_t^*\sigma = \int_X \sigma$$

and hence

$$0 = \frac{d}{dt}\int_X f_t^*\sigma = \int_X \frac{d}{dt} f_t^*\sigma$$

and at $t = 0$ the term on the right is $\int_X L_v\sigma$. □

We still have to show how to extend the notion of divergence to manifolds. This we will do as follows: If v is a vector field on X then, as we observed in Lecture 11, we can write the $\mathcal{C}^\infty$ density, $L_v\sigma_{\mathrm{vol}}$, as the product of a $\mathcal{C}^\infty$ function with σ_{vol}, and we'll call this $\mathcal{C}^\infty$ function the *divergence* of v, i.e., we will define the divergence of v by the identity

$$L_v\sigma_{\mathrm{vol}} = \mathrm{div}\,(v)\sigma_{\mathrm{vol}}\,. \tag{13.20}$$

For $\mathbb{R}^n$ $\sigma_{\mathrm{vol}} = \sigma_{\mathrm{Leb}}$, so for vector fields on $\mathbb{R}^n$ this definition coincides with the calculus definition (13.15) of divergence.

Exercises.

1. For $X = \mathbb{R}^n$ derive Theorem 13.3 directly from the formula (13.15).

2. A vector field, v, on $\mathbb{R}^n$ is divergence-free if $\operatorname{div}(v) = 0$. Show that the vector fields below are divergence free.

(a) The coordinate vector fields, $\partial/\partial x_i$.

(b) The vector field

$$v = \sum x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i}\,. \tag{13.21}$$

(c) The vector field

$$v = (x_1^2 + \cdots + x_n^2)^{-n/2} \sum x_i \frac{\partial}{\partial x_i}\,. \tag{13.22}$$

3. Let $[a_{i,j}(x)]$ be a skew-symmetric $n \times n$ matrix of functions, $a_{i,j} \in \mathcal{C}^\infty(\mathbb{R}^n)$, i.e.,

$$a_{i,j} = -a_{j,i}\,.$$

Show that the vector field

$$v = \sum \left(\frac{\partial}{\partial x_i} a_{i,j} \right) \frac{\partial}{\partial x_j} \tag{13.23}$$

is divergence-free.

4. Let $v = f_1 \frac{\partial}{\partial x_1} + f_2 \frac{\partial}{\partial x_2}$ be a divergence free vector field on $\mathbb{R}^2$. Show that v is of the form (13.23) with

$$a_{1,2} = -a_{2,1} = \int_0^{x_1} f_2(s, x_2)\, ds - \int_0^{x_2} f_1(0, t)\, dt\,.$$

5. Let v be a vector field on $\mathbb{R}^n$. Show that v can be written as a sum, $v = f_1 \frac{\partial}{\partial x_1} + w$ where w is a divergence free vector field.

6*. Prove by induction on n that every divergence-free vector field on $\mathbb{R}^n$, $n > 1$, is of the form (13.23).

7.

(a) Let $X \subseteq \mathbb{R}^n$ be the $(n-1)$-sphere. Show that the vector field

$$\sum x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i}$$

is tangent to X at all points, $p \in X$ and hence restricts to a vector field, v, on X.

(b) Prove that the divergence of this vector field is zero.

Lecture 14. Flux

Let X be an n-dimensional manifold and Z a closed, connected $(n-1)$-dimensional submanifold of X. We will say that Z is *two-sided* if there exists a neighborhood, U of Z in X and a vector field, v, on U which is nowhere tangent to Z, i.e., has the property

$$v(p) \notin T_pZ \text{ for all } p \in Z\,. \tag{14.1}$$

We'll say for short that a vector field with this property is a *non-tangential* vector field. (Hence two-sidedness means that a non-tangential vector field exists.)

The notion of two-sidedness is a nontrivial notion. For instance the Möbius band fails to have this property. (By staring at the figure on p. 285 of Munkres' book you can easily convince yourself that it's *one-sided*.)

Let v_1 and v_2 be non-tangential vector fields. Then for every $p \in Z$

$$v_1(p) = a(p)v_2(p) + w(p) \tag{14.2}$$

where $a(p)$ is nowhere zero constant and $w(p)$ is in T_pZ. We will say that v_1 and v_2 are *compatibly oriented* if $a(p) > 0$ for all p. We'll leave the following as an exercise.

Lemma 14.1. *Let v_i, $i = 1, 2, 3$, be non-tangential vector fields. Then*

i. If v_1 and v_2 are compatibly oriented, v_2 and v_1 are compatibly oriented.

ii. If v_1 and v_2 are compatibly oriented, and v_2 and v_3 are compatibly oriented, v_1 and v_3 are compatibly oriented.

iii. Either v_1 is compatibly oriented with v_2 or with $-v_2$.

In other words if v is a non-tangential vector field it determines a *co-orientation* of Z and there are just two ways in which Z can be co-oriented.

Let v be a non-tangential vector field and let σ be an element of $\mathcal{D}^\infty(X)$. Then we can define a $\mathcal{C}^\infty$ density on Z by defining it pointwise by the following procedure. For $p \in Z$ let $V = T_pX$ and $W = T_pZ$. Since W is an $(n-1)$-dimensional subspace of V one gets from $\sigma(p) \in |V|$ and $v(p) \in T_pX$ an element $\iota(v(p))\sigma(p)$ of $|W|$ by (10.14) and since $|W| = |T_pZ|$ the assignment

$$p \in Z \to \iota(v(p))\sigma(p) \tag{14.3}$$

defines a density on Z which we'll denote by $\iota(v)\sigma$.[4] Notice that if v_1 and v_2 are non-tangential vector fields then by (14.2)

$$v_1 = av_2 + w \tag{14.4}$$

[4] We'll postpone to the end of this lecture the proof that this is a $\mathcal{C}^\infty$ density.

along Z, where w is a vector field on Z itself and $a \in \mathcal{C}^\infty(Z)$ a non-vanishing function. Thus by (10.19)

$$\iota(v_1)\sigma = |a|\iota(v_2)\sigma \tag{14.5}$$

and in particular if v_1 and v_2 are compatibly oriented

$$\iota(v_1)\sigma = a\iota(v_2)\sigma\,. \tag{14.6}$$

Now let v be an arbitrary vector field on X. Then along Z

$$v = \varphi_1 v_1 + w_1 \tag{14.7}$$

where φ_1 is in $\mathcal{C}^\infty(Z)$ and w_1 is in a $\mathcal{C}^\infty$ vector field on Z itself. Let

$$\sigma_v = \varphi_1 \iota(v_1)\sigma_1\,. \tag{14.8}$$

Notice that by (14.4)

$$v = \varphi_2 v_2 + w_2 \tag{14.9}$$

along Z where $\varphi_2 = a\varphi_1$ and w_2 is a vector field on Z itself. Hence if v_1 and v_2 are compactly oriented

$$\varphi_1 \iota(v_1)\sigma_1 = \varphi_2 \iota(v_2)\sigma_2 \tag{14.10}$$

by (14.6). Thus if we fix a co-orientation of Z the form (14.8) is intrinsically defined.

Exercise 1. Show that if one changes the co-orientation the form (14.8) changes sign, i.e., $\sigma_{-v} = -\sigma_v$.

Now suppose that σ is compactly supported or alternatively that Z itself is compact. We define the *flux* of (v, σ) through Z to be the integral

$$\mathrm{Flux}(\mathrm{v}, \sigma) = \int_{\mathrm{Z}} \sigma_{\mathrm{v}}\,. \tag{14.11}$$

Remarks

1. Notice that this definition depends upon the co-orientation of Z. If we reverse the co-orientation the flux changes sign.

2. If v itself is a non-tangential vector field and its orientation is compatible with the co-orientation of Z then $\sigma_v = \iota(v)\sigma$ so we get for the flux of (v, σ) the simpler definition

$$\int_Z \iota(v)\sigma\,.$$

3. If Z is compact we can take σ to be σ_{vol}. For $\sigma = \sigma_{\text{vol}}$ the flux of (v, σ) through Z coincides with the standard definition of flux which one encounters in text books in physics and calculus courses.

We will next show that the definition of flux is invariant under "global changes of variables". Let X' be another n-dimensional manifold and $f : X' \to X$ a diffeomorphism. Then $Z' = f^{-1}(Z)$ is an $(n-1)$-dimensional submanifold of X' and it is easy to see that it, like Z, is two-sided. In fact if U is a neighborhood of Z in X and v_1 is a vector field on U having the non-tangency property (14.1) then $U' = f^{-1}(U)$ is a neighborhood of Z' in X' and the vector field

$$v_1' = (f^{-1})_* v_1$$

is a vector field on U' with the analogous non-tangency property:

$$v_1'(p) \notin T_p Z' \text{ for all } p \in Z' .$$

Hence from the co-orientation of Z associated with the vector field, v_1, we get an induced co-orientation of Z'.

Let $g : Z' \to Z$ be the restriction of f to Z'. We claim

Lemma 14.2. *For $\sigma \in \mathcal{D}^\infty(X)$*

$$\iota(v_1') f^* \sigma = g^* \iota(v_1) \sigma . \tag{14.12}$$

Proof. This amounts to showing that for every $p \in Z'$ and $q = f(p)$

$$\iota(v_1'(p))(df_p)^* \sigma(q) = (dg_p)^* \iota(v_1(q)) \sigma(q) .$$

But $v_1(q) = df_p(v_1'(p))$ and dg_p is the restriction of df_p to $T_p Z'$, so this follows from the linear algebra result (10.18).

□

We'll use this result to prove

Theorem 14.3. *Let v be any vector field on X. Then if $v' = (f^{-1})_* v$ and $\sigma' = f^* \sigma$*

$$\sigma'_{v'} = g^* \sigma_v . \tag{14.13}$$

Proof. By (14.7) $v = \varphi_1 v_1 + w_1$ along Z and hence along Z', $v' = g^* \varphi_1 v_1' + w_1'$ where $w_1' = (g^{-1})_* w_1$. Thus

$$\sigma'_{v'} = g^* \varphi_1 \iota(v_1') \sigma' = g^*(\varphi_1 \iota(v_1) \sigma)$$

by the lemma.

□

By the global change of variables formula that we proved in Lecture 12

$$\int_{Z'} \sigma'_{v'} = \int_{Z'} g^*(\sigma_v) = \int_Z \sigma_v$$

so as a corollary of Theorem 14.3 we get

Theorem 14.4. *The flux of (v, σ) through Z is equal to the flux of (v', σ') through Z'.*

We will conclude this lecture by proving a theorem about $(n-1)$-dimensional submanifolds of X which we'll need in the next lecture. (We'll also use this result to verify an assertion we made earlier in this lecture but whose proof we postponed: If v is non-tangential and σ is in $\mathcal{D}^\infty(X)$, $\iota(v)\sigma$ is in $\mathcal{D}^\infty(Z)$.)

Let's identify $\mathbb{R}^{n-1}$ with the submanifold

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n\,, \quad x_1 = 0\} \tag{14.14}$$

of $\mathbb{R}^n$ and also think of this submanifold as the boundary of the open set

$$\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n\,, \quad x_n < 0\}\,. \tag{14.15}$$

We will prove the following general result.

Theorem 14.5. *Let Z be an $(n-1)$-dimensional submanifold of X. Then for every $p \in Z$ there exists an open set, U, in X containing p and a parametrization*

$$\psi : U_0 \to U \tag{14.16}$$

of U with the property

$$\psi(U_0 \cap \mathbb{R}^{n-1}) = U \cap Z\,. \tag{14.17}$$

Proof. X is locally diffeomorphic at p to an open subset of $\mathbb{R}^n$ so it suffices to prove this assertion for submanifolds of $\mathbb{R}^n$. However, if Z is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ then as we showed in Lecture 7 there exists, for every $p \in Z$ a neighborhood, U, of p in $\mathbb{R}^n$ and a function, $\varphi \in \mathcal{C}^\infty(U)$ with the property

$$x \in U \cap Z \Leftrightarrow \varphi(x) = 0 \tag{14.18}$$

and

$$d\varphi_p \neq 0\,. \tag{14.19}$$

Without loss of generality we can assume by (14.19) that

$$\frac{\partial \varphi}{\partial x_1}(p) \neq 0\,. \tag{14.20}$$

Hence if $\rho : U \to \mathbb{R}^n$ is the map

$$\rho(x_1, \dots, x_n) = (\varphi(x), x_2, \dots, x_n) \tag{14.21}$$

$(d\rho)_p$ is bijective, and hence ρ is locally a diffeomorphism at p. Shrinking U we can assume that ρ is a diffeomorphism of U onto an open set, U_0. By (14.18) and (14.20) ρ maps $U \cap Z$ onto $U_0 \cap \mathbb{R}^{n-1}$ hence if we take ψ to be ρ^{-1}, it will have the property (14.17).

□

Let's now come back to the issue of proving that if σ is in $\mathcal{D}^\infty(X)$ and v is a non-tangential vector field, $\iota(v)\sigma$ is in $\mathcal{D}^\infty(Z)$. By the theorem we've just proved and Lemma 14.2 it suffices to prove this for $X = \mathbb{R}^n$ and $Z = \mathbb{R}^{n-1}$. Let

$$v = \sum_{i=1}^{n} a_i \frac{\partial}{\partial x_i} \tag{14.22}$$

and

$$\sigma = \varphi \sigma_{\mathrm{Leb}} \,. \tag{14.23}$$

Then by (14.5)

$$\iota(v)\sigma = \varphi |a_1| \iota\left(\frac{\partial}{\partial x_1}\right) \sigma_{\mathrm{Leb}} \,. \tag{14.24}$$

However, by (10.20) the density

$$\iota\left(\frac{\partial}{\partial x_1}\right) \sigma_{\mathrm{Leb}} \tag{14.25}$$

is the Lebesgue density on $\mathbb{R}^{n-1}$ and since v is non-tangential, $|a_1|$ is non-vanishing and hence is in $\mathcal{C}^\infty(\mathbb{R}^{n-1})$. Thus (14.23) is in $\mathcal{D}^\infty(\mathbb{R}^{n-1})$.

Exercises.

1. Let U be a bounded open subset of $\mathbb{R}^{n-1}$ and $f : U \to \mathbb{R}$ a $\mathcal{C}^\infty$ function. Let Z_f be the graph of f:

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n \,,\, x_n = f(x_1, \dots, x_{n-1})\} \,.$$

Compute the flux of (v, σ) through Z_f where $v = \partial/\partial x_n$ and $\sigma = \sigma_{\mathrm{Leb}}$ and show that it is equal to the volume of U.

2. Let $v = x_n \partial/\partial x_n$ and $\sigma = \sigma_{\mathrm{Leb}}$. Show that the flux of (v, σ) through Z_f is the integral of f over U.

3. Let $Z_1 \subseteq \mathbb{R}^n$ be the unit sphere, $x_1^2 + \cdots + x_n^2 = 1$, and let v be the vector field

$$v = x_1 \frac{\partial}{\partial x_1} + \cdots + x_n \frac{\partial}{\partial x1} .$$

Show that the flux of (v, σ_{Leb}) through Z_1 is the volume of Z_1. *Hint:* Exercise 11.8.

4. Let Z_a be the sphere

$$x_1^2 + \cdots + x_n^2 = a \, , \quad 0 < a \, .$$

What is the flux of (v, σ_{Leb}) through Z?

5. Prove Lemma 14.1.

6. Show that the density σ_v satisfies $\sigma_{-v} = -\sigma_v$ and conclude that the flux of (v, σ) through an $(n-1)$-dimensional submanifold, Z, of X changes sign if one changes the co-orientation of Z.

Lecture 15. The divergence theorem

To formulate the divergence theorem we need one final ingredient: Let D be an open subset of X and $\bar{D}$ its closure in X.

Definition 15.1. *D is a* smooth domain *if*

(a) its boundary is an $(n-1)$-dimensional submanifold of X and

(b) the boundary of D coincides with the boundary of $\bar{D}$.

Examples.

1. The n-ball, $x_1^2 + \cdots + x_n^2 < 1$, whose boundary is the sphere, $x_1^2 + \cdots + x_n^2 = 1$.
2. The n-dimensional annulus,

$$1 < x_1^2 + \cdots + x_n^2 < 2$$

whose boundary consists of the spheres,

$$x_1^2 + \cdots + x_n^2 = 1 \text{ and } x_1^2 + \cdots + x_n^2 = 2 \, .$$

3. Let S^{n-1} be the unit sphere, $x_1^2 + \cdots + x_2^2 = 1$ and let $D = \mathbb{R}^n - S^{n-1}$. Then the boundary of D is S^{n-1} but D is not a smooth domain since the boundary of $\bar{D}$ is empty.

The simplest example of a smooth domain is the half-space (14.15). We will show that every bounded domain looks locally like this example.

Theorem 15.2. *Let D be a smooth domain and p a boundary point of D. Then there exists a neighborhood, U, of p in X, an open set, U_0, in $\mathbb{R}^n$ and a diffeomorphism, $\varphi_0 : U_0 \to U$ such that φ_0 maps $U_0 \cap \mathbb{H}^n$ onto $U \cap D$.*

Proof. Let Z be the boundary of D. Then by Theorem 14.5 there exists a neighborhood, U of p in X, an open ball, U_0 in $\mathbb{R}^n$, with center at $q \in Bd\mathbb{H}^n$, and a diffeomorphism,

$$\varphi : (U_0, q) \to (U, p)$$

mapping $U_0 \cap Bd\mathbb{H}^n$ onto $U \cap Z$. Thus for $\varphi^{-1}(U \cap D)$ there are three possibilities.

ι $\varphi^{-1}(U \cap D) = (\mathbb{R}^n - Bd\mathbb{H}^n) \cap U_0$.

$\iota\iota$ $\varphi^{-1}(U \cap D) = \mathbb{H}^n \cap U_0$.

or

$\iota\iota\iota$ $\varphi^{-1}(U \cap D) = U_0 - \overline{\mathbb{H}}^n \cap U_0$.

However, scenario i. is excluded by the second hypothesis in Definition 15.1 and if scenario iii occurs we can rectify the situation by composing φ with the map, $(x_1, \ldots, x_n) \to (-x_1, \ldots, x_n)$.

□

Definition 15.3. *We will call an open set, U, with the properties above a D-adapted parametrizable open set.*

We will leave the following as an exercise:

Proposition 15.4. *The boundary of D is two-sided.*

Hint: At boundary points, p, of D there are two kinds of non-tangential vector fields: inward-pointing vector fields and outward-pointing vector fields. In the divergence theorem we will co-orient the boundary of D by giving it the *outward-pointing* co-orientation.

Theorem 15.5 (Divergence Theorem)**.** *If v is a vector field on X and σ a compactly supported $\mathcal{C}^\infty$ density the flux of (v, σ) through the boundary of D is equal to the integral over D of $L_v\sigma$.*

The key ingredient of the proof of this theorem is the following lemma.

Lemma 15.6. *Let X_i, $i = 1, 2$, be an n-dimensional manifold and $D_i \subseteq X_i$ a smooth domain. If (X_1, D_1) is diffeomorphic to (X_2, D_2) then the divergence theorem is true for (X_1, D_1) if and only if it is true for (X_2, D_2).*

Proof. This follows from Theorem 14.4, the identity (13.18) and the global change of variables formula for integrals of densities that we proved in Lecture 12 (see Theorem 12.6).

Let's now prove the theorem itself. By a partition of unity argument we can assume one of the following three alternatives holds.

1. σ is supported in the exterior of D.

2. σ is supported in D.

3. σ is supported in a D-adapted parametrizable open set of U.

In case 1 there is nothing to prove. The integral of $L_v\sigma$ over D and the integral of σ_v over the boundary are both zero. In case 2

$$\int_D L_v\sigma = \int_X L_v\sigma$$

and since σ is zero on the boundary the flux through the boundary is zero, so in this case Theorem 15.5 follows from Theorem 13.3. Let's prove the theorem in case 3. By Theorem 15.2 there exists an open ball, U_0, in $\mathbb{R}^n$ and a diffeomorphism of U_0 onto U mapping $U_0 \cap \mathbb{H}^n$ onto $U \cap D$, hence by Lemma 15.6 it suffices to prove the theorem for $\mathbb{H}^n$ and $\mathbb{R}^n$. Let's do so. Let v be the vector field

$$\sum a_i \frac{\partial}{\partial x_i}$$

and σ the density

$$\sigma = \varphi \sigma_{\text{Leb}}$$

with $\varphi \in \mathcal{C}_0^\infty(U_0)$. Let $v_1 = \partial/\partial x_1$. Then

$$\sigma_v = (a_1\varphi)(0, x_2, \ldots, x_n)\iota\left(\frac{\partial}{\partial x_i}\right)\sigma_{\text{Leb}}\ .$$

However, by (10.20) $\iota(\partial/\partial x_i)\sigma_{\text{Leb}}$ is the Lebesgue density on $\mathbb{R}^{n-1}$, so

$$\text{Flux}(v,\sigma) = \int_{\mathbb{R}^{n-1}} \varphi(0, x_2, \ldots, x_n) a_1(0, x_2, \ldots, x_n)\, dx_2 \cdots dx_n\ . \tag{15.1}$$

On the other hand by (13.16)

$$L_v\sigma = \sum \left(\frac{\partial}{\partial x_i} a_i\varphi\right)\sigma_{\text{Leb}}$$

so the right hand side of the divergence formula is the sum from 1 to n of the integrals

$$\int_{\mathbb{H}^n} \left(\frac{\partial}{\partial x_i} a_i \varphi \right) dx \,. \tag{15.2}$$

By Fubini's theorem we can write this integral as an iterated integral, integrating first with respect to the variable, x_i, then with respect to the other variables. For $i \neq 1$ the integration with respect to x_i is over the interval, $-\infty < x_i < \infty$, so we get

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial x_i} (a_i \varphi) \, dx_i = 0$$

since $a_i\varphi$ is compactly supported in the variable, x_i. On the other hand for $i = 1$, the integration is over the integral, $-\infty < x_1 < 0$ so we get

$$\int_{-\infty}^{0} \frac{\partial}{\partial x_i} (a_1 \varphi) \, dx_1 = a_1(0, x_2, \ldots, x_n) \varphi(0, x_1, \ldots, x_n) \,,$$

and integrating over the remaining variables we get (15.1).

□

Exercises.

1. Let B^n be the unit ball in $\mathbb{R}^n$ and S^{n-1} the unit $(n-1)$-sphere. Prove that

$$\operatorname{vol}(S^{n-1}) = n \operatorname{vol}(B^n) \,.$$

   Hint: Lecture 14, exercise 3.

2. Let D be the annulus,

$$a < x_1^2 + \cdots + x_2^2 < b \,, \quad 0 < a < b \,.$$

   From the divergence theorem conclude that if v is a divergence free vector field, the flux of (v, σ_{Leb}) through the sphere, $x_1^2 + \cdots + x_n^2 = a$ is equal to its flux through the sphere, $x_1^2 + \cdots + x_n^2 = b$.

3. Let U be a bounded open subset of $\mathbb{R}^{n-1}$ and let $X = U \times \mathbb{R}$. Given a positive $\mathcal{C}^\infty$ function $f : U \to \mathbb{R}$ let D be the open subset of X defined by

$$0 < x_n < f(x_1, \ldots, x_{n-1}) \,.$$

   For $v = x_n \partial / \partial x_n$ and $\sigma = \sigma_{\text{Leb}}$, verify the divergence theorem by computing the flux of (v, σ) through the boundary of D and the integral of the divergence of v over D and showing they're equal.

4. Let D be a bounded smooth domain in $\mathbb{R}^n$ and v a vector field. The classical divergence theorem of multivariable calculus asserts that

$$\int_D \operatorname{div}(v) = \int_{BdD} (n \cdot v)\sigma_{\text{Leb}}$$

where, at $p \in BdD$, n_p is the unit outward normal vector and $(n \cdot v)(p)$ is the dot product of $v(p)$ and n_p. Deduce this version of the divergence theorem from the divergence theorem that we proved above.

5. Let v be a vector field on $\mathbb{R}^n$ and for $a \in \mathbb{R}^n$ let Δ be the n-cube

$$-\epsilon + a_i < x_i < \epsilon + a_i\,, \quad i = 1, \ldots, n\,.$$

Prove that if $\mathrm{Flux(v, \Delta)}$ is the sums of the fluxes of (v, σ_{Leb}) over the $2n$ faces of Δ then for ϵ small,

$$\mathrm{Flux(v, \Delta)} \doteq \mathrm{div\,(v)(a)\,vol\,(\Delta)}\,.$$

6. Prove Proposition 15.4.

Hint: Let $\mathbb{U} = \{U_i\,, \, i = 1, 2, \ldots\}$ be a covering of the boundary of D by D-adapted parametrizable open sets. Let U be their union and $\rho_i \in \mathcal{C}_0^\infty(U)$, $i = 1, 2, \ldots,$ a partition of unity subordinate to $\mathbb{U}$. Show that on each U_i there exists and outward-pointing vector field, v_i, and let $v = \sum \rho_i v_i$.