M611 Fall 2019 Assignment 3, due Friday Sept. 20

1. [10 pts] The following is taken from *Finite-dimensional vector spaces*, by P. Halmos.

Definition 1. A linear space X over $\mathbb{R}$ (or $\mathbb{C}$) (sometimes referred to as a vector space) is a collection of elements $\{x_1, x_2, ...\}$ that satisfy the following ten properties for all $\alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$) and $x, y, z \in X$:

1. $x + y \in X$
2. $x + y = y + x$
3. $x + (y + z) = (x + y) + z$
4. There exists a unique element $0 \in X$ so that $x + 0 = x$
5. There exists a unique element $-x \in X$ so that $x + (-x) = 0$.
6. $\alpha x \in X$
7. $1x = x$
8. $\alpha(\beta x) = (\alpha\beta)x$
9. $\alpha(x + y) = \alpha x + \alpha y$
10. $(\alpha + \beta)x = \alpha x + \beta x$

Note. As Halmos points out, these properties are not logically independent, and we could choose a reduced collections and obtain the others as lemmas.

Definition 2. A linear space X is referred to as a normed linear space if to each $x \in X$ there corresponds a real number $\|x\|$, called the norm of x, so that the following are satisfied:

(i) $\|x\| > 0$ for $x \neq 0$, and $\|0\| = 0$.

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$ (or $\mathbb{C}$), and all $x \in X$.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Definition 3. We say that a sequence $\{x_n\}_{n=1}^{\infty} \subset X$ converges in the normed linear space $(X, \|\cdot\|)$ to $x \in X$ provided

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

Definition 4. A complete normed linear space is called a Banach space.

Okay, finally some questions.

a. Show that any normed linear space is a metric space. (Note that we've already seen in class that some metric spaces aren't normed linear spaces.)

b. Verify that for any normed linear space $(X, \|\cdot\|)$ we have the inequality

$$|\|x\| - \|y\|| \leq \|x - y\|$$

for all $x, y \in X$.

c. Show that if $\{x_n\}_{n=1}^{\infty} \subset X$ converges to x in $(X, \|\cdot\|)$ then

$$\lim_{n\to\infty} \|x_n\| = \|x\|.$$

2. [10 pts] Determine the maximum interval of time for which the Picard-Lindelöf Theorem guarantees the existence/uniqueness of a solution to the ODE

$$\frac{dy}{dt} = y^4; \quad y(0) = 1.$$

Solve this ODE exactly and compare your result with the actual interval of existence.

3. [10 pts] In this problem, we'll work through quite a bit of material associated with fixed point theorems that won't be covered in lecture. If you would like to skip down to the actual assignment, find the bold **Assignment** below.

Theorem (Brower Fixed Point Theorem). *If $B \subset \mathbb{R}^n$ is a closed ball and $f : B \to B$ is continuous, then f has a fixed point $\vec{x} \in B$.*

(Note: This theorem is only given FYI; we won't use it.)

Definitions. *Let X denote a metric space with metric ρ.*

(i) We say that $A \subset X$ is open if for every element $x_0 \in A$ there exists some $r > 0$ so that if $x \in X$ and $\rho(x, x_0) < r$ then $x \in A$.

(ii) By an open cover of a set $A \subset X$ we mean a (possibly uncountable) collection $\{U_k\}$ of open subsets of X so that

$$A \subset \bigcup_k U_k.$$

*(iii) A subset $K \subset X$ is said to be **compact** provided every open cover of K contains a finite subcover. (This is the general, topological, definition of a compact set.)*

Lemma. *A metric space X is compact (as in (iii) above) if and only if every sequence $\{x_k\}_{k=1}^{\infty} \subset X$ has a subsequence that converges in X.*

Definition. *We say that a mapping $T : X \to X$ on a normed linear space $(X, \|\cdot\|)$ is continuous provided that for any $x_0 \in X$ and any given $\epsilon > 0$ there corresponds a $\delta > 0$ so that*

$$\|x - x_0\| < \delta \implies \|Tx - Tx_0\| < \epsilon.$$

Definition. *We say that a subset Ω of a Banach space X is **convex** if for all points $x, y \in \Omega$ and all $s \in [0, 1]$ we have that $(1 - s)x + sy \in \Omega$.*

Theorem (Schauder Fixed Point Theorem I). *Suppose Ω is a compact, convex set in a Banach space X and $T : \Omega \to \Omega$ is continuous. Then T has a fixed point $x \in \Omega$.*

In general it's difficult to establish that a set is compact. No worries, we turn to a slight modification of the Schauder Theorem.

Theorem (Schauder's Fixed Point Theorem II). *Let Ω be a closed convex set in a Banach space X, and suppose $T : \Omega \to \Omega$ is continuous and such that $\overline{T(\Omega)}$ is compact in X. Then T has a fixed point in Ω.*

Notice in particular that we have removed the requirement that Ω be compact, though to be fair we have added the requirement that $\overline{T(\Omega)}$ be compact. But in many cases, this latter condition is a bit easier to verify. Let's verify it in the case of Peano's Theorem (see below). We'll set $J = [t_0 - \tau, t_0 + \tau]$,

$$\Omega = \{\vec{y} \in C(J) : \vec{y}(t_0) = \vec{y}_0, \max_{t \in J} |\vec{y}(t) - \vec{y}_0| \le \beta\},$$

$$\rho(\vec{y}_1, \vec{y}_2) := \max_{t \in J} |\vec{y}_1(t) - \vec{y}_2(t)|,$$

and

$$T\vec{y} = \vec{y}_0 + \int_{t_0}^{t} \vec{f}(\vec{y}_k(s), s) ds.$$

Claim. *$\overline{T(\Omega)}$ is a compact subset of the Banach space $X = C(J)$.*

In order to show that $\overline{T(\Omega)}$ is compact in X we need to show that given any sequence $\{\phi_k\}_{k=1}^{\infty} \subset \overline{T(\Omega)}$ there exists a subsequence $\{\phi_{k_j}\}_{j=1}^{\infty}$ so that

$$\phi_{k_j} \to \phi$$

for some $\phi \in \overline{T(\Omega)}$.

In order to accomplish this, we're going to use the Arzela–Ascoli Theorem, and before I state that I need to specify what an equicontinuous sequence of functions is, so here goes.

Definition. *Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence of real-valued functions defined on $U \subset \mathbb{R}^n$, and fix $\vec{x}_0 \in U$. We say that $\{f_k\}_{k=1}^{\infty}$ is equicontinuous at $\vec{x}_0$ if for every $\epsilon > 0$ there exists a $\delta > 0$, independent of k, so that for all $\vec{x} \in U$ such that $|\vec{x} - \vec{x}_0| < \delta$ we have*

$$|f_k(\vec{x}) - f_k(\vec{x}_0)| < \epsilon,$$

for all $k \in \mathbb{N}$.

Arzela–Ascoli Theorem. *Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence of real-valued functions defined on a compact subset $K \subset \mathbb{R}^n$, that there exists a value $M > 0$ so that $|f_k(\vec{x})| \le M$ for all $k = 1, 2, ...$ and all $\vec{x} \in U$ (we say the sequence is uniformly bounded), and that the sequence is equicontinuous at every $\vec{x} \in U$. Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ that converges uniformly on K.*

Now let's apply this to our problem. Consider any sequence $\{\vec{\phi}_k\}_{k=1}^{\infty} \subset T(\Omega)$ (dropping the closure for a moment), and note that $\vec{\phi}_k \in T(\Omega)$ implies there exists $\vec{y}_k \in \Omega$ so that $\vec{\phi}_k = T\vec{y}_k$:

$$\vec{\phi}_k(t) = \vec{y}_0 + \int_{t_0}^{t} \vec{f}(\vec{y}_k(s), s) ds.$$

We have, then, for (w.l.g.) $t_2 > t_1$

$$|\vec{\phi}_k(t_2) - \vec{\phi}_k(t_1)| \le \int_{t_1}^{t_2} |\vec{f}(\vec{y}_k(s), s)| ds \le M|t_2 - t_1|,$$

for all $t_1, t_2 \in J$. Since this is independent of k we conclude that $\{\vec{\phi}_k\}_{k=1}^{\infty}$ is equicontinuous on J. By a similar calculation we see that $\{\vec{\phi}_k\}_{k=1}^{\infty}$ is uniformly bounded, and so we can conclude from the Arzela–Ascoli Theorem that there exists a uniformly convergent subsequence, which must by definition converge to some $\vec{\phi} \in \overline{T(\Omega)}$. Finally, for the case $\{\vec{\phi}_k\}_{k=1}^\infty \subset \overline{T(\Omega)}$ (i.e., we bring back the closure) we note that each $\vec{\phi}_k$ is the (uniform, by passing to a subsequence) limit of a sequence in $T(\Omega)$, and so boundedness and equicontinuity can be recovered by taking a limit in the above inequalities. This gives that $\overline{T(\Omega)}$ is indeed compact.

Finally, while we're at it:

Definition. Let X denote a Banach space. A mapping $K : X \to X$ is called compact if it maps bounded sequences $\{u_k\}_{k=1}^\infty \subset X$ into precompact sequences $\{Ku_k\}_{k=1}^\infty$; that is, there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ so that $\{Ku_{k_j}\}_{j=1}^\infty$ converges to an element of X.

Theorem (Schaefer's Fixed Point Theorem). Suppose $K : X \to X$ is a continuous compact mapping, and the set

$$\{u \in X : u = \lambda K u \, \text{for some } \lambda \in [0,1]\}$$

is bounded. Then K has a fixed point.

Note. Schaefer's theorem appears as Theorem 9.2.4 in Evans, though we'll need it in M612 (for those of you who hang around) to prove Theorem 6.5.3. (A rare instance in which Evans does things out of order.)

Notice in particular that none of the fixed point theorems listed above asserts uniqueness.

Assignment. Use the SFPT II to prove the following:

Theorem (Peano's Theorem). *Suppose $\vec{f}(\vec{y}, t)$ is continuous in*

$$R = \{(\vec{y}, t) : |\vec{y} - \vec{y}_0| \le \beta, |t - t_0| \le \alpha\}$$

for some constants $\alpha > 0$ and $\beta > 0$. Then the ODE

$$\frac{d\vec{y}}{dt} = \vec{f}(\vec{y}, t); \quad \vec{y}(t_0) = \vec{y}_0$$

has a solution for $|t - t_0| \le \tau$, where $\tau = \min\{\alpha, \frac{\beta}{M}\}$, where

$$M = \sup_{(\vec{y},t) \in R} |\vec{f}(\vec{y}, t)|.$$

Notes. (1) We do not assume Lipshitz continuity; (2) We do not conclude uniqueness.

4. [10 pts] In this problem I want to collect several important differentiation formulas. First, recall that for a function $\vec{f}(\vec{x})$, with $\vec{x} \in \mathbb{R}^n$ and $\vec{f} \in \mathbb{R}^m$, we're using the notation

$$D\vec{f}(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Two important special cases are $n = 1$ and $m = 1$. First, for $m = 1$ we obtain the gradient vector

$$Df(\vec{x}) = (f_{x_1}, f_{x_2}, \dots, f_{x_n}).$$

(The gradient vector is often regarded as a column vector, and in certain contexts this is important, so I may not typically refer to this as the gradient vector.) For $n = 1$ we obtain

$$D\vec{f}(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \\ \vdots \\ \frac{\partial f_m}{\partial x} \end{pmatrix} = \frac{d\vec{f}}{dx}.$$

Theorem (Chain Rule). Suppose $\vec{y} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{x}_0 \in \mathbb{R}^n$ and $\vec{f} : \mathbb{R}^m \to \mathbb{R}^k$ is differentiable at $\vec{y}(\vec{x}_0) \in \mathbb{R}^m$. Then $\vec{h} = \vec{f}(\vec{y}(\vec{x}))$ is differentiable at $\vec{x} = \vec{x}_0$ and

$$D_x\vec{h}(\vec{x}_0) = D_y\vec{f}(\vec{y}(\vec{x}_0))D_x\vec{y}(\vec{x}_0).$$

In particular, notice that this is the multiplication of a $k \times m$ matrix by a $m \times n$ matrix, resulting in a $k \times m$ matrix.

For the following problems assume all functions are as differentiable as necessary. In all cases you can apply the chain rule without proof.

a. Show that if A is any $n \times n$ matrix and $\vec{x} \in \mathbb{R}^n$ is a column vector, then

$$D_x(A\vec{x}) = A,$$

and likewise

$$D_x(\vec{x}^{tr}A^{tr}) = A$$

b. Show that if $f : \mathbb{R}^n \to \mathbb{R}$ and $\vec{x} : \mathbb{R} \to \mathbb{R}^n$ then

$$\frac{d}{dt}f(\vec{x}(t)) = D_x f \cdot \frac{d\vec{x}}{dt}.$$

c. Show that if $\vec{x} \in \mathbb{R}^n$ and $f : [0, \infty) \to \mathbb{R}$ then

$$D_x f(|\vec{x}|) = f'(|\vec{x}|)\frac{\vec{x}}{|\vec{x}|},$$

and in particular

$$D_x(|\vec{x}|^r) = r|\vec{x}|^{r-2}\vec{x}.$$

d. Show that if $\vec{x} \in \mathbb{R}^n$ and $f : [0, \infty) \to \mathbb{R}$ then

$$\Delta f(|\vec{x}|) = f''(|\vec{x}|) + \frac{n-1}{|\vec{x}|}f'(|\vec{x}|).$$

Here, $\Delta := \sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}$.

e. Show that if $u, v : \mathbb{R}^n \to \mathbb{R}$ then

$$\Delta(uv) = v\Delta u + 2Du \cdot Dv + u\Delta v.$$

f. Show that if $\vec{f}:\mathbb{R}^n \to \mathbb{R}^m$ and $\vec{g}:\mathbb{R}^n \to \mathbb{R}^m$ are regarded as row vectors then

$$D_x(\vec{f}(\vec{x}) \cdot \vec{g}(\vec{x})) = \vec{g}(\vec{x}) D_x \vec{f}(\vec{x}) + \vec{f}(\vec{x}) D_x \vec{g}(\vec{x}).$$

g. Show that if $f:\mathbb{R}^n \to \mathbb{R}$ and $g:\mathbb{R}^n \to \mathbb{R}^m$ then

$$D_x(f(\vec{x})\vec{g}(\vec{x})) = \vec{g}(\vec{x}) \otimes D_x f(\vec{x}) + f(\vec{x}) D_x \vec{g}(\vec{x}),$$

where $\otimes$ denotes the $m \times n$ tensor product matrix

$$\vec{g}(\vec{x}) \otimes D_x f(\vec{x}) = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix} (f_{x_1}, f_{x_2}, \ldots, f_{x_n}).$$