# ANALYSIS 2
# RECITATION SESSION OF WEEK 4

JACOB SHAPIRO

## 1. Continuity of Product Spaces

1.1. *Claim.* Let $f:\mathbb{R}^n\to\mathbb{R}$. Assume that for all $i\in\{0,\dots,n-1\}$ and for all $z\in\mathbb{R}^{n-1}$, the map

$$\begin{array}{ccc}\mathbb{R} & \overset{g_i(z)}{\to} & \mathbb{R}\\ \alpha & \mapsto & f(z_1,z_2,\dots,z_i,\alpha,z_{i+1},\dots,z_{n-1})\end{array}$$

is *uniformly* continuous (also in $z$). Then $f$ is continuous.

*Proof.* First note that as $\mathbb{R}^n$ is a finite dimensional vector space, all norms on it are "equivalent" and we have proven previously that equivalent norms generate the same topology. Thus, we could just as well work with the norm $\|x\|_\infty\equiv\max(\{|x_i|\mid i\in\{1,\dots,n\}\})$.

Next, for $f$ to be continuous at $x$, we need that $\forall\varepsilon>0$, $\exists\delta(\varepsilon)>0$ such that if $\|x-y\|_\infty<\delta(\varepsilon)$ then $|f(x)-f(y)|<\varepsilon$. We know that the maps $g_i(z):\mathbb{R}\to\mathbb{R}$ are uniformly continuous (also in $z$), namely, $\forall\varepsilon>0$ $\exists\delta_i(\varepsilon)>0$ such that $\forall(\alpha,\beta)\in\mathbb{R}^2$, if $|\alpha-\beta|<\delta_i(\varepsilon)$ then $|f(z_1,\dots,z_i,\alpha,z_{i+1},\dots,z_{n-1})-f(z_1,z_2,\dots,z_i,\beta,z_{i+1},\dots,z_{n-1})|<\varepsilon$ (and this condition holds $z\in\mathbb{R}^{n-1}$, in such a way that $\delta_i(\varepsilon)$ does not depend on $z$).

Now for the actual proof:

Let $\varepsilon>0$ be given. Define $\delta(\varepsilon):=\min\left(\left\{\delta_i\left(\frac{\varepsilon}{n}\right)\mid i\in\{1,\dots,n\}\right\}\right)$. Then if $\|x-y\|_\infty<\delta(\varepsilon)$, then the uniform continuity condition on the $g_i$ functions are all fulfilled:

$$\begin{aligned}|x_i-y_i| &\leqslant \max\left(\{|x_{i'}-y_{i'}|\mid i'\in\{1,\dots,n\}\}\right)\\ &\equiv \|x-y\|_\infty\\ &\overset{!}{<} \delta(\varepsilon)\\ &\equiv \min\left(\left\{\delta_{i'}\left(\frac{\varepsilon}{n}\right)\middle| i'\in\{1,\dots,n\}\right\}\right)\\ &\leqslant \delta_i\left(\frac{\varepsilon}{n}\right)\end{aligned}$$

and so we may conclude that

$$|f(z_1,\dots,z_i,y_i,z_{i+1},\dots,z_{n-1})-f(z_1,\dots,z_i,x_i,z_{i+1},\dots,z_{n-1})|<\frac{\varepsilon}{n}$$

for all $i\in\{1,\dots,n\}$ and for all $z\in\mathbb{R}^n$. But we have

$$\begin{aligned}|f(y)-f(x)| &\leqslant |f(y)-f(x_1,y_2,y_3,y_4,\dots,y_n)|+\\ &\quad+|f(x_1,y_2,y_3,y_4\dots,y_n)-f(x_1,x_2,y_3,y_4,\dots,y_n)|\\ &\quad+|f(x_1,x_2,y_3,y_4,\dots,y_n)-f(x_1,x_2,x_3,y_4,\dots,y_n)|\\ &\quad\dots\\ &\quad+|f(x_1,\dots,x_{n-1},y_n)-f(x)|\\ &\leqslant n\frac{\varepsilon}{n}\\ &= \varepsilon\end{aligned}$$

Note that in each term in the above expression, only *one* coordinate is different between $x$ and $y$ and all the rest are the same (albeit mixed, but that's fine since our uniform continuity condition holds for any $z\in\mathbb{R}^{n-1}$. For example, in the first row, $z=(y_2,\dots,y_n)$, in the second row, $z=(x_1,y_3,\dots,y_n)$, in the third row, $z=(x_1,x_2,y_4,\dots,y_n)$ and in the last row, $z=(x_1,\dots,x_{n-1})$). □

1.2. *Remark.* How can 1.1 claim be improved?

## 2. Continuity of Derivatives in Banach Spaces

Note that if $X$ and $Y$ are Banach spaces, a map $f:E\to Y$ where $E\in\mathrm{Open}(X)$ is called differentiable at $x_0\in X$ (see [2] pp. 6) iff there exists a *continuous* linear map $f'(x_0):X\to Y$ such that

$$\lim_{h\to 0_X}\frac{\|f(x_0+h)-f(x_0)-(f'(x_0))(h)\|_Y}{\|h\|_X}=0 \quad (1)$$

2.1. *Claim.* If $\dim(X)<\infty$ then any linear map $A:X\to Y$ is continuous.

*Date*: March 6th, 2015.

*Proof.* Let $x_0 \in X$ be given. Because $\dim(X) < \infty$, $X \cong \mathbb{R}^n$ for some $n \in \mathbb{N}$. Thus, WLOG $X = \mathbb{R}^n$ with its standard Euclidean norm (because all norms on $\mathbb{R}^n$ are equivalent and generate the same topology, as we have seen), and let $\{e_i\}_{i=1}^n$ be the standard basis on $\mathbb{R}^n$. Thus for each $x \in X$ we may write $x_x = \sum_{i=1}^n x^i e_i$.

Let $\varepsilon > 0$ be given. Then we need to find some $\delta(\varepsilon) > 0$ such that for all $x \in \mathbb{R}^n$ such that $\|x_0 - x\|_{\mathbb{R}^n} < \delta(\varepsilon)$ we have $\|A(x) - A(x_0)\|_Y < \varepsilon$. But $A$ is linear and $(\alpha x + \beta y)^i = \alpha x^i + \beta y^i$ for any scalars $(\alpha, \beta) \in \mathbb{R}^2$, so that

$$\begin{aligned}
\|A(x) - A(x_0)\|_Y &= \|A(x - x_0)\|_Y \\
&= \left\| A\left(\sum_{i=1}^n (x - x_0)^i e_i\right)\right\|_Y \\
&= \left\| \sum_{i=1}^n (x - x_0)^i A(e_i)\right\|_Y \\
&\leqslant \sum_{i=1}^n \left\|(x - x_0)^i A(e_i)\right\|_Y \\
&= \sum_{i=1}^n \left|(x - x_0)^i\right| \|A(e_i)\|_Y \\
&\leqslant \max\left(\left\{ \|A(e_j)\|_Y \mid j \in \{1, \dots, n\}\right\}\right) \sum_{i=1}^n \left|(x - x_0)^i\right| \\
&\leqslant \max\left(\left\{ \|A(e_j)\|_Y \mid j \in \{1, \dots, n\}\right\}\right) \sum_{i=1}^n \|x - x_0\|_{\mathbb{R}^n} \\
&= \max\left(\left\{ \|A(e_j)\|_Y \mid j \in \{1, \dots, n\}\right\}\right) \|x - x_0\|_{\mathbb{R}^n} \cdot n
\end{aligned}$$

so that we see if we define $\delta(\varepsilon) := \frac{\varepsilon}{n \cdot \max(\{ \|A(e_j)\|_Y \mid j \in \{1, \dots, n\}\})}$ we get the desired condition. Observe that it is impossible for $\|A(e_j)\|_Y = \infty$ because $\|\|_Y$ is a map $Y \to \mathbb{R}$ and $\infty \notin \mathbb{R}$. □

Thus we see that if $\dim(X) < \infty$ we don't need to require the continuity of $A$, it comes automatically by linearity. However, otherwise, it is part of the definition of the differentiability of $f$.

Observe that there is another map defined by this process, namely, the map $E \to \mathcal{L}(X, Y)$ given by $x \mapsto f'(x)$, where $\mathcal{L}(X, Y)$ is the set of all continuous linear maps from $X$ to $Y$. Actually, $\mathcal{L}(X, Y)$ is a Banach space in and of itself, with the norm on it defined as

$$\|A\|_{\mathcal{L}(X,Y)} := \sup\left(\left\{ \frac{\|A(x)\|_Y}{\|x\|_X} \,\middle|\, x \in X \backslash \{0_X\}\right\}\right)$$

which gives $\mathcal{L}(X, Y)$ its own topological structure and thus we may speak about whether the map $X \ni x \mapsto f'(x) \in \mathcal{L}(X, Y)$ is continuous. So, even though $X \ni x \mapsto (f'(x_0))(x) \in Y$ is always continuous (by definition) for any $x_0 \in E$ where $f$ is differentiable, $X \ni x \mapsto f'(x) \in \mathcal{L}(X, Y)$ may fail to be continuous.

If $X \ni x \mapsto f'(x) \in \mathcal{L}(X, Y)$ is continuous, we say $f$ is *continuously differentiable* (see [3] definition 9.20) and is denoted by $f \in C^1(E, Y)$.

2.2. *Claim.* (Theorem 9.21 in [3]) $f \in C^1(E, Y)$ at $x_0 \in E$ iff $\partial_h f$ exists at $x_0 \in E$ for all $h \in X$.

Question 1 in exercise sheet number 4 shows you a counter example of this. Here's another example:

2.3. *Claim.* Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function at $x_0 \in \mathbb{R}$ such that $f' : \mathbb{R} \to \mathbb{R}$ is not continuous at $x_0$ (for example, $x \mapsto |x|$ with $x_0 = 0$). Then the function $g : \mathbb{R}^2 \to \mathbb{R}$ given by $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto f(x)$ is not continuously differentiable, but is differentiable, at $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ for any $y_0 \in \mathbb{R}$.

*Proof.* Using 2.2 we see that since

$$\begin{aligned}
(\partial_x g)\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) &\equiv \lim_{t \to 0} \frac{g\left(\begin{bmatrix} x + t \\ y \end{bmatrix}\right) - g\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)}{t} \\
&= \lim_{t \to 0} \frac{f(x + t) - f(x)}{t} \\
&\equiv f'(x)
\end{aligned}$$

then $\partial_x g$ is discontinuous at $\begin{bmatrix} x_0 \\ y \end{bmatrix}$ for any $y \in \mathbb{R}$ and in particular, that means $g \notin C^1(\mathbb{R}^2, \mathbb{R})$. Note that $\partial_y g = 0$.

To see that $g'$ exists at $(x_0, y_0)$ for all $y_0 \in \mathbb{R}$ and is equal

$$\begin{aligned}
g'\left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}\right) &= \begin{bmatrix} \partial_x g\left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}\right) & \partial_y g\left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}\right) \end{bmatrix} \\
&= \begin{bmatrix} f'(x_0) & 0 \end{bmatrix}
\end{aligned}$$

we compute the limit from 1:

$$
\begin{aligned}
0 &\overset{!}{=} \lim_{(h_x,h_y)\to(0,0)} \frac{\left\| g\left(\begin{bmatrix}x_0\\y_0\end{bmatrix}+\begin{bmatrix}h_x\\h_y\end{bmatrix}\right) - g\left(\begin{bmatrix}x_0\\y_0\end{bmatrix}\right) - g'\left(\begin{bmatrix}x_0\\y_0\end{bmatrix}\right)\begin{bmatrix}h_x\\h_y\end{bmatrix}\right\|_{\mathbb{R}}}{\left\|\begin{bmatrix}h_x\\h_y\end{bmatrix}\right\|_{\mathbb{R}^2}} \\
&= \lim_{(h_x,h_y)\to(0,0)} \frac{\left\| f(x_0+h_x) - f(x_0) - \begin{bmatrix}f'(x_0) & 0\end{bmatrix}\begin{bmatrix}h_x\\h_y\end{bmatrix}\right\|_{\mathbb{R}}}{\left\|\begin{bmatrix}h_x\\h_y\end{bmatrix}\right\|_{\mathbb{R}^2}} \\
&= \lim_{(h_x,h_y)\to(0,0)} \frac{\| f(x_0+h_x) - f(x_0) - f'(x_0)h_x\|_{\mathbb{R}}}{\sqrt{h_x^2+h_y^2}} \\
&= \lim_{h_x\to 0} \frac{| f(x_0+h_x) - f(x_0) - f'(x_0)h_x|}{|h_x|}
\end{aligned}
$$

which is indeed zero as

$$
f'(x_0) \equiv \lim_{h_x\to 0} \frac{f(x_0+h_x)-f(x_0)}{h_x}
$$

□

## 3. Derivatives in Banach Algebras

Let $\mathcal{A}$ be a Banach algebra.

3.1. *Claim.* Let $n\in\mathbb{N}$ be given. The derivative of the map $f:\mathcal{A}\to\mathcal{A}$ given by $A\mapsto A^n$ is

$$
\left(f'(A)\right)(B) = \sum_{k=0}^{n-1} A^k B A^{n-k-1}
$$

*Proof.* We proceed by computing the limit of 1:

$$
\begin{aligned}
0 &\overset{!}{=} \lim_{B\to 0} \frac{\|f(A+B)-f(A)-(f'(A))(B)\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&= \lim_{B\to 0} \frac{\left\|(A+B)^n - A^n - \sum_{k=0}^{n-1} A^k B A^{n-k-1}\right\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}}
\end{aligned}
$$

but

$$
(A+B)^n = A^n + \sum_{k=0}^{n-1} A^k B A^{n-k-1} + \sum_{k=2}^{n} g_k(A) B^k h_k(A)
$$

where $g_k$ and $h_k$ are the non-commutative powers of $A$ that will occur for every given pwoer of $B$ (you can show this by induction. The actual form of $g_k$ and $h_k$ is not important for the argument). Thus we have:

$$
\begin{aligned}
&= \lim_{B\to 0} \frac{\left\|(A+B)^n - A^n - \sum_{k=0}^{n-1} A^k B A^{n-k-1}\right\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&= \lim_{B\to 0} \frac{\left\|A^n + \sum_{k=0}^{n-1} A^k B A^{n-k-1} + \sum_{k=2}^{n} g_k(A) B^k h_k(A) - A^n - \sum_{k=0}^{n-1} A^k B A^{n-k-1}\right\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&= \lim_{B\to 0} \frac{\left\|\sum_{k=2}^{n} g_k(A) B^k h_k(A)\right\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&\leqslant \lim_{B\to 0} \sum_{k=2}^{n} \frac{\|g_k(A) B^k h_k(A)\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&\leqslant \lim_{B\to 0} \sum_{k=2}^{n} \frac{\|g_k(A)\|_{\mathcal{A}} \|B^k\|_{\mathcal{A}} \|h_k(A)\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&\leqslant \lim_{B\to 0} \sum_{k=2}^{n} \frac{\|g_k(A)\|_{\mathcal{A}} \left[\|B\|_{\mathcal{A}}\right]^k \|h_k(A)\|_{\mathcal{A}}}{\|B\|_{\mathcal{A}}} \\
&= \|g_k(A)\|_{\mathcal{A}} \|h_k(A)\|_{\mathcal{A}} \lim_{B\to 0} \sum_{k=2}^{n} \left[\|B\|_{\mathcal{A}}\right]^{k-1} \\
&= 0
\end{aligned}
$$

The map $B\mapsto (f'(A))(B) = \sum_{k=0}^{n-1} A^k B A^{n-k-1}$ is clearly linear and continuous. □

As we have seen previously, any power series $\mathbb{R} \ni x \mapsto \sum_{n\in\mathbb{N}} \alpha_n x^n \in \mathbb{R}$ with $\{ \alpha_n \}_{n\in\mathbb{N}} \subset \mathbb{R}$ which converges absolutely defines a map on $\mathcal{A}$ by $A \mapsto \sum_{n\in\mathbb{N}} \alpha_n A^n$. This map defines a sequence of maps $\left\{ A \overset{f_N}{\mapsto} \sum_{n=0}^{N} \alpha_n A^n \right\}_{N\in\mathbb{N}}$ which converge *uniformly* to $A \overset{f}{\mapsto} \sum_{n\in\mathbb{N}} \alpha_n A^n$:

*Proof.* Let $\varepsilon > 0$ be given. Then we need to find some $m(\varepsilon) \in \mathbb{N}$ such that if $N \geqslant m(\varepsilon)$ then $\|f_N(A) - f(A)\| < \varepsilon$ for all $A \in \mathcal{A}$.

$$\begin{aligned}
\|f_N(A) - f(A)\| &= \left\| \sum_{n=0}^{N} \alpha_n A^n - \sum_{n\in\mathbb{N}} \alpha_n A^n \right\| \\
&= \left\| \sum_{n=N+1}^{\infty} \alpha_n A^n \right\| \\
&\equiv \left\| \lim_{M\to\infty} \sum_{n=N+1}^{M} \alpha_n A^n \right\| \\
&\overset{\text{norm cont.}}{=} \lim_{M\to\infty} \left\| \sum_{n=N+1}^{M} \alpha_n A^n \right\| \\
&\leqslant \lim_{M\to\infty} \sum_{n=N+1}^{M} |\alpha_n| \, \|A\|^n
\end{aligned}$$

which converges by the uniform convergence of $\sum_{n\in\mathbb{N}} \alpha_n \|A\|^n$, which follows from Rudin Theorems 7.10 and 8.1. □

and so, in particular, we may differentiate term by term of the power series (compare with theorem 7.17 in [3]). Thus we have shown:

3.2. *Claim.* The derivative of $A \mapsto \sum_{n\in\mathbb{N}} \alpha_n A^n$ is

$$B \mapsto \sum_{n\in\mathbb{N}} \alpha_n \sum_{k=0}^{n-1} A^k B A^{n-k-1}$$

3.3. **Exercise.** What is the derivative of $\mathcal{A}^* \ni A \mapsto A^{-1}$?

## 4. Taylor Polynomials

## 5. Homework Number 4 Preview

5.1. **Question 1.**

- An example of a function that is differentiable, but not continuously differentiable. Part (a) is straight forward, need to think separately about $(0, 0)$ and elsewhere.
- The partial derivatives are not continuous at the origin, and so the total derivative is not continuous.

5.2. **Question 3.**

- Derivative of power series

## 6. Homework Number 2 Overview

6.1. **Question 1.**

- Recall that it is not strictly necessary to have a multiplicative unit for an algebra (See [1] chapter 6 pp. 262).
- A matrix $A$ is not a sum $\sum_{i,j=1}^{n} a_{i,j}$! (basic linear algebra).

6.2. **Question 2.**

- Do (a), talk about exchanging limits.

## References

[1] I. N. Herstein. *Topics in Algebra, 2nd Edition*. John Wiley & Sons, 1975.
[2] Serge Lang. *Differential and Riemannian Manifolds (Graduate Texts in Mathematics)*. Springer, 1996.
[3] Walter Rudin. *Principles of Mathematical Analysis (International Series in Pure and Applied Mathematics)*. McGraw-Hill Science/Engineering/Math, 1976.