

MATH 20410 (Analysis in \mathbb{R}^n II – Accelerated) Notes

Steven Labalme

February 23, 2022

Contents

6	The Riemann-Stieltjes Integral	1
6.1	Notes	1
6.2	Exam 1 Additional Topics	3
7	Sequences and Series of Functions	4
7.1	Notes	4
9	Functions of Several Variables	8
9.1	Notes	8
9.2	Chapter 9: Functions of Several Variables	11

Chapter 6

The Riemann-Stieltjes Integral

6.1 Notes

1/28:

- Plan:

1. Finish up Fundamental Theorem of Calculus proof.
2. Basic consequences.
3. Rectifiable curves.

- Recall that we're given $f : [a, b] \rightarrow \mathbb{R}$ continuous, $f : [a, b] \rightarrow \mathbb{R}$, and $x \mapsto \int_a^x f(t) dt$.

- Goal: Show $F'(x_0) = f(x_0)$.

- WTS: Find δ such that $|x - x_0| < \delta$ implies

$$\begin{aligned} \left| \frac{1}{x - x_0} \int_{x_0}^x f(t) dt - f(x_0) \right| &= \left| \frac{1}{x - x_0} \int_{x_0}^x f(t) dt - \frac{1}{x - x_0} \int_{x_0}^x f(x_0) dt \right| \\ &= \frac{1}{|x - x_0|} \left| \int_{x_0}^x (f(t) - f(x_0)) dt \right| \\ &\leq \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| dt \\ &< \epsilon \end{aligned}$$

- Since f is continuous, there exists δ such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$.

- Now

$$\begin{aligned} \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| dt &< \frac{1}{|x - x_0|} \int_{x_0}^x \epsilon dt \\ &= \epsilon \end{aligned}$$

- Applications:

1. Theorem (MVT for integration): $f : [a, b] \rightarrow \mathbb{R}$ continuous, then there exists $x_0 \in [a, b]$ such that

$$f(x_0) = \frac{1}{b - a} \int_a^b f(x) dx$$

- Apply MVT to $F(x) = \int_a^x f(t) dt$. Then

$$F'(x_0) = f(x_0) = \frac{F(b) - F(a)}{b - a}$$

as desired.

2. Theorem (Integration by parts): Let $F, G : [a, b] \rightarrow \mathbb{R}$ be differentiable with $F' = f$, $G' = g$ and with f and g both integrable. Then

$$\int_a^b Fg = F(b)G(b) - F(a)G(a) - \int_a^b fG$$

- Just use the product rule plus the FTC to prove.
- We have

$$\begin{aligned} \int_a^b (FG)' &= \int_a^b fG + \int_a^b Fg \\ F(b)G(b) - F(a)G(a) &= \int_a^b fG + \int_a^b Fg \\ \int_a^b Fg &= F(b)G(b) - F(a)G(a) - \int_a^b fG \end{aligned}$$

3. Theorem (u -substitution).

- Follows similarly from the chain rule and FTC.

- Integration of vector-valued functions.

- If $f : [a, b] \rightarrow \mathbb{R}^k$, we define $\int_a^b f$ by

$$\int_a^b f = \left(\int_a^b f_1, \dots, \int_a^b f_k \right)$$

- Alternatively, you can define $\int_a^b f$ using P , $U(f, P)$, $L(f, P)$, etc. and then prove that the integral exists iff all f_i are integrable and in this case the above definition holds.
- Rectifiable curves: Let $\gamma : [a, b] \rightarrow \mathbb{R}^k$ be a continuous function.
- Plan: Define the length of γ and show that we can compute it with an integral.
 - Idea: For polygonal paths, we know how to define length. So let's approximate γ by polygons and take a limit.
 - Ref: Given a partition P , then define the length of γ with respect to P as $\Lambda(\gamma, P)$. Let the length of γ be $\Lambda(\gamma) = \sup_P \Lambda(\gamma, P)$ if this limit exists in this case, we call γ **rectifiable**.
- Fractals are not rectifiable — their length diverges.
- Theorem: Suppose γ is continuously differentiable (i.e., γ is differentiable and γ' is continuous). Then γ is rectifiable and

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)| dt$$

- Notice: If $P \leq P'$, then $\Lambda(\gamma, P) \leq \Lambda(\gamma, P')$. (Prove with triangle inequality.)
- WTS: For all partitions P , $\Lambda(\gamma, P) \leq \int_a^b |\gamma'(t)| dt$ and thus $\Lambda(\gamma) \leq \int_a^b |\gamma'(t)| dt$.
- We have that

$$\begin{aligned} \Lambda(\gamma, P) &= \sum_{i=1}^n |\gamma(x_i) - \gamma(x_{i-1})| \\ &= \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} \gamma'(t) dt \right| \\ &\leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |\gamma'(t)| dt \\ &= \int_a^b |\gamma'(t)| dt \end{aligned}$$

- Catch up.
 - I should make up PSets 1-2.
 - Exams have less than Rudin-strength problems.
 - Exams are mostly true/false (and of that, mostly false, provide a counterexample).

6.2 Exam 1 Additional Topics

- A continuous function that is not always differentiable.

$$f(x) = |x|$$

- A differentiable function with a discontinuous derivative.

$$f(x) = x^2 \sin \frac{1}{x}$$

- A vector-valued function that doesn't satisfy the MVT.

$$\mathbf{f}(x) = e^{ix}$$

- Between 0 and 2π .

- A pair of vector-valued functions that don't satisfy L'Hôpital's rule.

$$f(x) = x \qquad g(x) = x + x^2 e^{i/x^2}$$

Chapter 7

Sequences and Series of Functions

7.1 Notes

- 1/31:
- Midterm on differentiation and integration, and a bit of stuff from this week.
 - Plan:
 - Talk about sequences of functions, all with the same domain and range, converging.
 - Address what properties of f_n remain in the limit (e.g., continuity, differentiability, integrability).
 - The answer depends on what we mean by “convergence.”
 - $f_n \rightarrow f$ pointwise implies basically nothing.
 - $f_n \rightarrow f$ uniformly implies that basically everything works out nicely.
 - We’ll restrict ourselves to real functions because those have all the properties (integrability, differentiability, etc.) that we care about.
 - **Pointwise** (convergent sequence $\{f_n\}$ to f): A sequence of functions $\{f_n\}$ such that for all $x \in X$, the sequence $\{f_n(x)\}$ converges to $f(x)$, where $f_n : X \rightarrow \mathbb{R}$ for all $n \in \mathbb{N}$ and $f : X \rightarrow \mathbb{R}$. Denoted by $f_n \rightarrow f$.
 - Bad functions.
 - Consider $f_n : [0, 1] \rightarrow \mathbb{R}$ defined by $x \mapsto x^n$. Each f_n is continuous, but f is not (zero everywhere except $f(1) = 1$)^[1].
 - Consider $f_n : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f_n(x) = x^2/(1 + x^2)^n$, and $f(x) = \sum_{n=0}^{\infty} f_n(x)$. As a geometric series, $f(x) = 1 + x^2$ when $x \neq 0$ but $f(0) = 0$. Thus, the limit exists but is not continuous once again.
 - Consider $f_m : \mathbb{R} \rightarrow \mathbb{R}$ defined by $x \mapsto \lim_{n \rightarrow \infty} \cos^{2n}(m!\pi x)$. Each f_m is integrable, but the limit f is the function that’s 1 for rationals and zero for irrationals. In particular, f is not integrable.
 - We take even powers of the cosine to make it always positive.
 - We use $\cos^2(x)$ just because it’s always between $[0, 1]$, and we know when it is equal to 1.
 - In particular, $\cos^2(\pi x)$ is equal to 1 at every integer, $\cos^2(2\pi x)$ is equal to 1 at every half integer. $\cos^2(6\pi x)$ is equal to 1 at every one-sixth of an integer.
 - Then raising it to the n^{th} power just makes it spiky.
 - Aside: Interchanging limits.
 - If all f_n are continuous, then $\lim_{x \rightarrow x_0} f_n(x) = f_n(x_0)$.

¹Questions that require counterexamples like this could show up on the midterm!

- The question “is f continuous” is equivalent to being able to interchange limits:

$$\lim_{x \rightarrow x_0} \lim_{n \rightarrow \infty} f_n(x) = f(x_0) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow x_0} f_n(x)$$

- Sequence example showing we need to be careful interchanging limits: $s_{n,m} = m/(m+n)$.
- All of this pathology goes away with the right definition, though.
- **Uniformly** (convergent sequence $\{f_n\}$ to f): A sequence of functions $\{f_n\}$ such that for all $\epsilon > 0$, there exists an N such that if $n \geq N$, then $|f_n(x) - f(x)| < \epsilon$ for all $x \in X$, where $f_n : X \rightarrow \mathbb{R}$ for all $n \in \mathbb{N}$ and $f : X \rightarrow \mathbb{R}$.
- Proposition (Cauchy criterion for uniform convergence): $f_n \rightarrow f$ uniformly iff for all $\epsilon > 0$, there exists N such that for all $m, n \geq N$ and for all $x \in X$, $|f_n(x) - f_m(x)| < \epsilon$.
 - Forward direction: Let $\epsilon > 0$. Suppose $f_n \rightarrow f$ uniformly. Choose N such that the functions are within $\epsilon/2$. Then

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

2/2:

- Office hours tomorrow 4-5 PM.
- Plan:
 1. More on uniform convergence.
 - Limit of continuous functions is continuous.
 - Limit of the integral of functions is the integral of the limit.
 2. $\mathcal{C}(X)$ perspectives on uniform convergence.
- Corollary (Weierstraß M-test): If there exist constants $M_n \in \mathbb{R}$ such that $|f_n(x)| \leq M_n$ for all x and $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly.
- Theorem: $f_n : X \rightarrow \mathbb{R}$, f_n continuous at $x_0 \in X$ for all n , and $f_n \rightarrow f$ uniformly imply f continuous at x_0 .
 - Idea:
 - “ $\epsilon/3$ trick”: Find δ such that if $|x - x_0| < \delta$, then

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

- Proof:

- $f_n \rightarrow f$ uniformly implies there exists $N \in \mathbb{N}$ such that $|f_N(x) - f(x)| < \epsilon/3$ for all $x \in X$.
 - f_N continuous at x_0 : There exists δ such that if $d(x, x_0) < \delta$, then $|f_N(x) - f_N(x_0)| < \epsilon/3$.
 - Thus, by the $\epsilon/3$ trick, we have the continuity of f .
- Defining a norm on $\mathcal{C}(X)$.

$$\|f\| = \sup_{x \in X} |f(x)|$$

- This makes $\mathcal{C}(X)$ into a vector space.
- We can now define our metric $d(f, g)$ by $d(f, g) = \|f - g\|$.
- $f_n \rightarrow f \iff f$ is bounded.
 - $f_n \rightarrow f$ uniformly $\iff \lim_{n \rightarrow \infty} \sup |f_n(x) - f(x)| = 0 \iff f_n \rightarrow f$ is $\mathcal{C}(X)$.
- Corollary to the Weierstraß M-test: $\mathcal{C}(X)$ is complete (i.e., all uniformly Cauchy sequences converge).

- Assume $\{f_n\}$ is Cauchy. Then by the Cauchy criterion for uniform convergence, f_n converges uniformly to some f . But this f must be continuous, too, meaning $f \in \mathcal{C}(X)$.

2/4:

- Plan.

1. $\int \lim f_n = \lim \int f_n$.
2. $\mathrm{d}x \lim f_n = \lim \mathrm{d}x f_n$.
3. Definitions: Pointwise/uniform boundedness, equicontinuity.

- Theorem: $f_n : [a, b] \rightarrow \mathbb{R}$ integrable and $f_n \rightarrow f$ uniformly implies f is integrable and

$$\int_a^b f = \lim_{n \rightarrow \infty} \int_a^b f_n$$

- Plan:

1. Show f is integrable.
2. Show $\int f = \lim \int f_n$.

- Proof:

- Let $\epsilon_n = \sup_{x \in [a, b]} |f(x) - f_n(x)|$.
- Since $f_n \rightarrow f$ uniformly, $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$.
- By definition, $f_n - \epsilon_n \leq f \leq f_n + \epsilon_n$.
- Thus, by Theorems 6.4 and 6.5,

$$\int_a^b (f_n - \epsilon_n) = \int_a^b (f_n - \epsilon_n) \leq \int_a^b f \leq \int_a^b (f_n + \epsilon_n)$$

- It follows since

$$0 \leq \int_a^b f - \int_a^b f_n \leq \int_a^b (f_n + \epsilon_n) - \int_a^b (f_n - \epsilon_n) = (b - a) \dots$$

that f is integrable.

- Hence,

$$\begin{aligned} \int_a^b (f_n - \epsilon_n) &\leq \int_a^b f \leq \int_a^b (f_n + \epsilon_n) \\ \left| \int_a^b f_n - \int_a^b f \right| &\leq \epsilon_n \\ \lim_{n \rightarrow \infty} \int_a^b f_n &= \int_a^b f \end{aligned}$$

- Theorem: $f_n : [a, b] \rightarrow \mathbb{R}$, each f_n differentiable, $f_n \rightarrow f$ pointwise, and $(f_n)' \rightarrow g$ uniformly implies that f is differentiable and $f' = g$.

- Note that you can do better: Substituting $f_n(x_0)$ converging for some $x_0 \in [a, b]$ for $f_n \rightarrow f$ pointwise still implies the desired result.
- Idea: We use the $\epsilon/3$ trick; $2/3$ will be easy and $1/3$ will be tricky.
- Goal: We want

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| < \epsilon$$

for some δ with $0 < |x - x_0| < \delta$. We will show that

$$\begin{aligned} &\left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_N(x) - f_N(x_0)}{x - x_0} + \frac{f_N(x) - f_N(x_0)}{x - x_0} - f'_N(x_0) + f'_N(x_0) - g(x_0) \right| \\ &\leq \left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_N(x) - f_N(x_0)}{x - x_0} \right| + \left| \frac{f_N(x) - f_N(x_0)}{x - x_0} - f'_N(x_0) \right| + |f'_N(x_0) - g(x_0)| \end{aligned}$$

- For the middle inequality, use Chapter 5, Exercise 8.
- For the right inequality, use the uniform convergence condition.
- For the left inequality, it will suffice to show the Cauchy condition

$$\left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - \frac{f_m(x) - f_m(x_0)}{x - x_0} \right| < \frac{\epsilon}{3}$$

so, noting that the left term above is equal to

$$\left| \frac{[f_n(x) - f_m(x)] - [f_n(x_0) - f_m(x_0)]}{x - x_0} \right|$$

which is equal to $|f'_n(c) - f'_m(c)|$ by the MVT, from which we can apply the Cauchy form of the uniform convergence of $(f_n)'$ condition.

- **Pointwise bounded** ($\{f_n\}$): A sequence of real functions $\{f_n\}$ such that for all $x \in X$, there exists $M_x \in \mathbb{R}$ such that $|f_n(x)| \leq M_x$ for all $n \in \mathbb{N}$.
- **Uniformly bounded** ($\{f_n\}$): A sequence of real functions $\{f_n\}$ for which there exists $M \in \mathbb{R}$ such that for all $x \in X$ and $n \in \mathbb{N}$, $|f_n(x)| \leq M$.
- Proposition: $f_n : E \rightarrow \mathbb{R}$, $\{f_n\}$ is pointwise bounded, and E is countable implies there is a subsequence $\{f_{n_k}\}$ that converges pointwise.
 - Enumerate $E = \{x_1, x_2, \dots\}$.
 - Then since $\{f_n(x_m)\}$ is bounded for all m by hypothesis, it always has a convergent subsequence.
 - The claim is if you look at the sequence of diagonal functions, it is such a subsequence, i.e., if $f_1(x_1)$ is the first term for x_1 , $f_3(x_2)$ is the second term for x_2 , $f_{11}(x_3)$ is the third term for x_3 , and so on, f_1, f_3, f_{11}, \dots is such a subsequence.

2/9: • Build up to the Arzelà-Ascoli theorem.

2/11: • The Arzelà-Ascoli theorem.

Chapter 9

Functions of Several Variables

9.1 Notes

2/14:

- Plan:
 1. Warm-up with matrices.
 2. The total derivatives of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($n = m = 2$, i.e., $f : \mathbb{C} \rightarrow \mathbb{C}$).
 3. Basic properties: Chain rule, relation with partial derivatives, implicit function theorem.
- Let V, W be finite-dimensional vector spaces over \mathbb{R} . We let $L(V, W)$ be the vector space of all linear transformations $\phi : V \rightarrow W$.
- If we pick bases $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \dots, \mathbf{w}_m$ of W , then $V \cong \mathbb{R}^n$ and $W \cong \mathbb{R}^m$. It follows that $L(V, W) \cong \mathbb{R}^{mn}$.
- $L(V, W) \times L(W, U) \xrightarrow{\text{compose}} L(V, U)$, i.e., $\mathbb{R}^{mn} \times \mathbb{R}^{nl} \xrightarrow[\text{mult.}]{\text{matrix}} \mathbb{R}^{ml}$.
- Sup norm: If A is an $m \times n$ real matrix, then $\|A\| = \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ |\mathbf{x}|=1}} |A\mathbf{x}|$.
 - Basic properties:
 1. $|A\mathbf{x}| \leq \|A\| |\mathbf{x}|$.
 2. $\|A\| < \infty$ and all $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are uniformly continuous.
 3. $\|A\| = 0 \iff A = 0$.
 4. $\|cA\| = |c| \|A\|$.
 5. $\|A + B\| \leq \|A\| + \|B\|$.
 6. $\|AB\| \leq \|A\| \|B\|$.
 - Note that we get a metric space structure on $L(V, W)$ by defining $d(A, B) = \|A - B\|$.
- Proves that 1 and 2 imply the uniform continuity of all A (via Lipschitz continuity).
- **Differentiable** (function \mathbf{f} at \mathbf{x}_0): A function $\mathbf{f} : U \rightarrow \mathbb{R}^m$ ($U \subset \mathbb{R}^n$) such that to $\mathbf{x}_0 \in U$ there corresponds some linear transformation $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that
$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{|\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - A\mathbf{h}|}{|\mathbf{h}|} = 0$$
- **Total derivative** (of \mathbf{f} at \mathbf{x}_0): The linear transformation A in the above definition. Denoted by $\mathbf{f}'(\mathbf{x}_0)$, $D\mathbf{f}(\mathbf{x}_0)$, $d\mathbf{f}(\mathbf{x}_0)$.
- “An proof and progress in mathematics” - Thurston.

- Relating to the old one dimensional derivative.
- A paper we'd find rather impressionistic right now.

- Propositions ahead of us.

- Proposition: Suppose that \mathbf{f} is differentiable at $\mathbf{x}_0 \in U$ and A, B are both derivatives of \mathbf{f} at \mathbf{x}_0 . Then $A = B$.
- Proposition: Differentiable implies continuous.
- Proposition: Sum rule, product rule, quotient rule.

2/16:

- Plan: Derivatives of functions $\mathbf{f} : U \rightarrow \mathbb{R}^m$ where $U \subset \mathbb{R}^n$.

- Basic properties: Differentiability implies continuity, $(\mathbf{f} + \mathbf{g})' = \mathbf{f}' + \mathbf{g}'$, $(c\mathbf{f})' = c\mathbf{f}'$, chain rule, $\mathbf{f}' = 0$ iff \mathbf{f} is constant.
- Relationship with partial derivatives (how we compute everything and anything).
- When is \mathbf{f} differentiable?
- Inverse function theorem.
- Implicit function theorem.

- **Continuously differentiable** (function \mathbf{f}): A function $\mathbf{f} : U \rightarrow \mathbb{R}^m$ that is differentiable for all $\mathbf{x}_0 \in U$ and such that $\mathbf{f}' : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. *Also known as \mathcal{C}^1 .*

- Proposition: Let $\mathbf{f} : U \rightarrow \mathbb{R}^m$ be differentiable at $\mathbf{x}_0 \in U$. Then \mathbf{f} is continuous at \mathbf{x}_0 .

- The proof makes use of the fact that $\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0)\mathbf{h} + \mathbf{r}(\mathbf{h})$.

- Proposition: Given $\mathbf{f}, \mathbf{g} : U \rightarrow \mathbb{R}^m$ both differentiable at $\mathbf{x}_0 \in U$, then $\mathbf{f} + \mathbf{g}$ is also differentiable at \mathbf{x}_0 with

$$(\mathbf{f} + \mathbf{g})'(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0) + \mathbf{g}'(\mathbf{x}_0)$$

- The proof is immediate via the triangle inequality.

- Theorem (Chain Rule): Given $\mathbf{f} : U \rightarrow \mathbb{R}^m$ and $\mathbf{g} : V \rightarrow \mathbb{R}^k$, where $U \subset \mathbb{R}^n$ and $\mathbf{f}(U) \subset V \subset \mathbb{R}^m$, with \mathbf{f} differentiable at $\mathbf{x}_0 \in U$ and \mathbf{g} differentiable at $\mathbf{f}(\mathbf{x}_0)$, the composition $\mathbf{g} \circ \mathbf{f}$ is differentiable at \mathbf{x}_0 with

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0)) \cdot \mathbf{f}'(\mathbf{x}_0)$$

- The proof is rather subtle.

- **Partial derivative** (of f_i wrt. x_j at \mathbf{x}_0): The following limit, if it exists, where $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $1 \leq i \leq m$, and $1 \leq j \leq n$. Denoted by $(\partial \mathbf{f}_i / \partial x_j)(\mathbf{x}_0)$, $(D_j \mathbf{f}_i)(\mathbf{x}_0)$. Given by

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{e}_j) - f_i(\mathbf{x}_0)}{t}$$

- **Directional derivative** (of f_i toward $\mathbf{u} \in \mathbb{R}^n$): The following limit, if it exists, where $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ and $1 \leq i \leq m$. Denoted by $D_{\mathbf{u}} \mathbf{f}_i$. Given by

$$D_{\mathbf{u}} \mathbf{f}_i = \lim_{t \rightarrow 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{u}) - f_i(\mathbf{x}_0)}{t}$$

- **Jacobian**: The following matrix. Given by

$$\left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right]$$

- Theorem: Let $\mathbf{f} = (f_1, \dots, f_m) : U \rightarrow \mathbb{R}^m$, where $U \subset \mathbb{R}^n$, be differentiable at some $\mathbf{x}_0 \in U$. Then the partial derivatives $\partial f_i / \partial x_j$ ($1 \leq i \leq m$; $1 \leq j \leq n$) exist at \mathbf{x}_0 and, with respect to the usual choice of bases,

$$\mathbf{f}'(\mathbf{x}_0) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right]$$

2/18: – We have that

$$\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) + \mathbf{r}(t\mathbf{e}_j)$$

- Since \mathbf{f} is differentiable at \mathbf{x}_0 , $\mathbf{f}(t\mathbf{e}_j)/t \rightarrow 0$ as $t \rightarrow 0$.
- Additionally, $\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j)/t = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$.
- Therefore,

$$\lim_{t \rightarrow 0} \frac{\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0)}{t} = \lim_{t \rightarrow 0} \frac{\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) - \mathbf{r}(t\mathbf{e}_j)}{t} = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) - \lim_{t \rightarrow 0} \frac{\mathbf{r}(t\mathbf{e}_j)}{t} = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$$

as desired.

- Unpacking the definition of the linear transformation as a matrix gives the rest of the proof.

- Today:

- More on differentiation (recall the Jacobian).
- Sufficient condition for differentiability.
- $\mathbf{f}' = 0$ iff \mathbf{f} is constant.
- State the inverse function theorem.

- It is not true that having all partials exist implies that \mathbf{f} is differentiable at \mathbf{x}_0 .

- Theorem: \mathbf{f} continuously differentiable at \mathbf{x}_0 iff all partials exist and are continuous at \mathbf{x}_0 .

2/21: • Contraction mapping theorem.

2/23: • Plan.

1. Proof of the inverse function theorem.
2. Commuting partials.

- Theorem (Inverse function theorem): If $E \subset \mathbb{R}^n$ open, $\mathbf{f} : E \rightarrow \mathbb{R}^n$ is differentiable at $\mathbf{x}_0 \in E$, and $\mathbf{f}'(\mathbf{x}_0)$ is invertible, then there exist $U \subset E$ open with $\mathbf{x}_0 \in U$ and $V \subset \mathbb{R}^n$ open with $\mathbf{f}(\mathbf{x}_0) \in V$ such that $\mathbf{f}|_U : U \rightarrow V$ is a bijection and $(\mathbf{f}|_U)^{-1}$ is continuously differentiable.

- Idea.

1. Find U and prove one-to-one restricted to U .
2. $\mathbf{f}(U)$ is open.
3. Prove the inverse is continuously differentiable (left as an exercise to us).

- There is a trick for 1-2: We introduce an auxiliary function $\varphi_{\mathbf{y}}$ and apply the contraction mapping theorem.

- Proof.

- Let $A = \mathbf{f}'(\mathbf{x}_0)$.
- Since \mathbf{f}' is continuous, there is an open ball $U \subset E$ with center \mathbf{x}_0 such that $\|\mathbf{f}'(\mathbf{x}) - A\| < \lambda$ for all $\mathbf{x} \in U$.
 - We'll pick $\lambda = 1/(2\|A^{-1}\|)$ without motivation for now.

- Note that if you need to pick a U (for an example function), this criterion gives you one (not necessarily the best one, but it gives you a one).
- Trick: For all $\mathbf{y} \in \mathbb{R}^n$, consider $\varphi_{\mathbf{y}} : U \rightarrow \mathbb{R}^n$ defined by

$$\varphi_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x}))$$

- Important property of this function: $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ iff $\varphi_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}$.
- Plan: Show that for all $\mathbf{y} \in \mathbf{f}(U)$ that $\varphi_{\mathbf{y}}$ is a contraction. Therefore, by the contraction mapping theorem, \mathbf{f} has exactly 1 fixed point, so $\mathbf{f}|_U$ is injective.
- Proving that $\varphi_{\mathbf{y}}$ is a contraction. Claim: $|\varphi_{\mathbf{y}}(\mathbf{x}_1) - \varphi_{\mathbf{y}}(\mathbf{x}_2)| \leq \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$. Use the Chain Rule, MVT, and the fact that $\|AB\| \leq \|A\|\|B\|$.
- Using the chain rule, we have that

$$\begin{aligned}\varphi'_{\mathbf{y}} &= I - A^{-1}\mathbf{f}'(\mathbf{x}) \\ &= A^{-1}(A - \mathbf{f}'(\mathbf{x}))\end{aligned}$$

- Thus,

$$\|\varphi_{\mathbf{y}}(\mathbf{x})\| \leq \|A^{-1}\|\|A - \mathbf{f}'(\mathbf{x})\| < \frac{1}{2}$$

for all \mathbf{x} .

- It follows by the MVT that

$$|\varphi_{\mathbf{y}}(\mathbf{x}_1) - \varphi_{\mathbf{y}}(\mathbf{x}_2)| \leq \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$$

- Therefore, $\varphi_{\mathbf{y}}$ is a contraction.
- We now prove that $\mathbf{f}(U)$ is open.
- Let $\mathbf{y}_0 \in \mathbf{f}(U)$ be such that $\mathbf{y}_0 = \mathbf{f}(\mathbf{p}_0)$.
- Pick $B_r(\mathbf{p}_0) \subset U$ such that $\overline{B} \subset U$.
- Claim: For all $\mathbf{y} \in \mathbb{R}^n$ with $|\mathbf{y} - \mathbf{y}_0| < \lambda r$, we have that $\mathbf{y} \in \mathbf{f}(U)$.
 - We are going to show that $\varphi_{\mathbf{y}}(\overline{B}) \subset \overline{B}$ and therefore $\varphi_{\mathbf{y}} : \overline{B} \rightarrow \overline{B}$ is a contraction and therefore by the contraction mapping theorem, there exists a fixed point $\mathbf{x}_{\mathbf{y}}$ of $\varphi_{\mathbf{y}}$ in \overline{B} . Therefore, $\mathbf{f}(\mathbf{x}_{\mathbf{y}}) = \mathbf{y}$ and so $\mathbf{f}(U)$ is open.
- $\varphi_{\mathbf{y}}$ is derived from Newton's method. The contraction mapping thing then is what substitutes for convergence. You have to start in the right area though, the chosen U !

9.2 Chapter 9: Functions of Several Variables

From Rudin (1976).

- 2/15:
- Defines a vector space by the closure of its elements under addition and scalar multiplication.
 - Defines a linear combination, span, independence and dependence, dimension, basis, coordinates, and the standard basis.
 - Theorem 9.2: If X is spanned by r vectors, $\dim X \leq r$.
 - Corollary: $\dim \mathbb{R}^n = n$.
 - Theorem 9.3: Let X a vector space with $\dim X = n$.
 - (a) $E \subset X$ containing n vectors spans X iff E is independent.
 - (b) X has a basis, and every basis contains n vectors.

- (c) If $1 \leq r \leq n$ and $\{\mathbf{y}_1, \dots, \mathbf{y}_r\}$ is independent in X , then X has a basis containing $\{\mathbf{y}_1, \dots, \mathbf{y}_r\}$.
- Defines linear transformation, linear operator.
- Notes that $A\mathbf{0} = \mathbf{0}$ if A is a linear transformation, and that A is completely determined by its action on any basis.
- **Invertible** (linear operator): A linear operator A that is one-to-one and onto.
- Theorem 9.5: A linear operator on X finite-dimensional is one-to-one iff it is onto.
- Defines $L(X, Y)$, $L(X)$, the product BA of two linear transformations, and the supremum norm of a linear transformation.
- Theorem 9.7:
 - (a) $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ implies $\|A\| < \infty$ and $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uniformly continuous.
 - (b) $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $c \in \mathbb{C}$ implies

$$\|A + B\| \leq \|A\| + \|B\| \qquad \|cA\| = |c|\|A\|$$

Defining $d(A, B) = \|A - B\|$ makes $L(\mathbb{R}^n, \mathbb{R}^m)$ a metric space.

- (c) $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^m, \mathbb{R}^k)$ implies

$$\|BA\| \leq \|B\|\|A\|$$

- Theorem 9.8: Let Ω be the set of all invertible linear operators on \mathbb{R}^n .

- (a) $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and $\|B - A\| \cdot \|A^{-1}\| < 1$ implies $B \in \Omega$.

Proof. Let $\|A^{-1}\| = 1/\alpha$, and let $\|B - A\| = \beta$. Then

$$\begin{aligned} \|B - A\| \cdot \|A^{-1}\| &< 1 \\ \beta \cdot \frac{1}{\alpha} &< 1 \\ \beta &< \alpha \end{aligned}$$

To prove that $B \in \Omega$, the definition of invertibility and Theorem 9.5 tell us that it will suffice to show that B is 1-1. To do so, it will suffice to show that $B\mathbf{x} = \mathbf{0}$ iff $\mathbf{x} = \mathbf{0}$. Let's begin. Let $\mathbf{x} \in \mathbb{R}^n$ be arbitrary. Then

$$\begin{aligned} \alpha|\mathbf{x}| &= \alpha|A^{-1}A\mathbf{x}| \leq \alpha\|A^{-1}\| \cdot |A\mathbf{x}| = |A\mathbf{x}| \leq |(A - B)\mathbf{x}| + |B\mathbf{x}| \leq \beta|\mathbf{x}| + |B\mathbf{x}| \\ (\alpha - \beta)|\mathbf{x}| &\leq |B\mathbf{x}| \end{aligned}$$

It follows that if $\mathbf{x} \neq \mathbf{0}$, then $|B\mathbf{x}| > 0$. This combined with the fact that $B\mathbf{0} = \mathbf{0}$ implies the desired result. \square

- (b) Ω is open in $L(\mathbb{R}^n)$ and $A \mapsto A^{-1}$ is continuous on Ω .

Proof. To prove that Ω is open in $L(\mathbb{R}^n)$, it will suffice to show that for all $A \in \Omega$, there exists $N_r(A)$ such that if $\|B - A\| < r$, then $B \in \Omega$. Let's begin. Let $A \in \Omega$ be arbitrary. Choose $N_\alpha(A)$ to be our neighborhood, where α is defined as in part (a). Let $B \in L(\mathbb{R}^n)$ satisfy $\|B - A\| < \alpha$. Then $\|B - A\| \cdot \|A^{-1}\| < 1$, so $B \in \Omega$ by part (a), as desired.

To prove that $A \mapsto A^{-1}$ is continuous, it will suffice to show that $\|B^{-1} - A^{-1}\| \rightarrow 0$ as $B \rightarrow A$. First off, we have by part (a) and the substitution $\mathbf{x} = B^{-1}\mathbf{y}$ ($\mathbf{y} \in \mathbb{R}^n$) that

$$\begin{aligned} (\alpha - \beta)|B^{-1}\mathbf{y}| &\leq |BB^{-1}\mathbf{y}| = |\mathbf{y}| \\ \left| B^{-1} \left(\frac{\mathbf{y}}{|\mathbf{y}|} \right) \right| &\leq (\alpha - \beta)^{-1} \end{aligned}$$

Thus, since $|B^{-1}\mathbf{u}|$ is bounded by $(\alpha - \beta)^{-1}$ for every unit vector $\mathbf{u} \in \mathbb{R}^n$, $\|B^{-1}\|$ is bounded by $(\alpha - \beta)^{-1}$. This combined with the fact that

$$\begin{aligned} B^{-1} - A^{-1} &= B^{-1}I - IA^{-1} \\ &= B^{-1}AA^{-1} - B^{-1}BA^{-1} \\ &= B^{-1}(A - B)A^{-1} \end{aligned}$$

implies by Theorem 9.7c that

$$\|B^{-1} - A^{-1}\| \leq \|B^{-1}\| \|A - B\| \|A^{-1}\| \leq (\alpha - \beta)^{-1} \cdot \beta \cdot \frac{1}{\alpha} = \frac{\beta}{\alpha(\alpha - \beta)}$$

Therefore, since $\beta \rightarrow 0$ as $B \rightarrow A$, the above inequality establishes the desired result. \square

- Note that the mapping $A \mapsto A^{-1}$ defined in Theorem 9.8b is a 1-1 mapping of Ω onto Ω and its own inverse.
- Defines matrices, column vectors, and matrix multiplication.
- From the Schwarz inequality, we can show that

$$\|A\| \leq \left(\sum_{i,j} a_{i,j}^2 \right)^{1/2}$$

- “If S is a metric space, if a_{11}, \dots, a_{mn} are real continuous functions on S , and if for each $p \in S$, A_p is the linear transformation of \mathbb{R}^n into \mathbb{R}^m whose matrix has entries $a_{ij}(p)$, then the mapping $p \rightarrow A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, \mathbb{R}^m)$ ” (Rudin, 1976, p. 211).
- Rudin (1976) spends some time motivating the definition of the total derivative. He also discusses the natural 1-1 correspondence between \mathbb{R}^1 and $L(\mathbb{R}^1)$.
- Defines differentiability in \mathbb{R}^n .
- Theorem 9.12: A_1, A_2 the derivative of \mathbf{f} at \mathbf{x} implies $A_1 = A_2$.
- If $\mathbf{f} : E \rightarrow \mathbb{R}^m$ where $E \subset \mathbb{R}^n$, then $\mathbf{f}' : E \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$.
- \mathbf{f} differentiable implies \mathbf{f} continuous.
- Example (\mathbf{f} is linear):
 - If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $A'(\mathbf{x}) = A$ for all $\mathbf{x} \in \mathbb{R}^n$. Note that this means that $A' : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$, as expected.
- Theorem 9.15 (Chain Rule): E open in \mathbb{R}^n , $\mathbf{f} : E \rightarrow \mathbb{R}^m$ differentiable at $\mathbf{x}_0 \in E$, $I \supset \mathbf{f}(E)$ open in \mathbb{R}^m , and $\mathbf{g} : I \rightarrow \mathbb{R}^k$ differentiable at $\mathbf{f}(\mathbf{x}_0)$ implies $\mathbf{F} : E \rightarrow \mathbb{R}^k$ defined by

$$\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$$

is differentiable at \mathbf{x}_0 with

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0)^{[1]}$$

Proof. Largely symmetric to that of the one-dimensional chain rule in Chapter 5. \square

¹Note that the right-hand side of this equation contains the product of two linear transformations.

- **Components** (of $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$): The real functions f_1, \dots, f_m defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^m f_i(\mathbf{x}) \mathbf{u}_i$$

for all $\mathbf{x} \in E$ or, equivalently, by $f_i(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \cdot \mathbf{u}_i$ ($1 \leq i \leq m$), where $\mathbf{u}_1, \dots, \mathbf{u}_m$ is the standard basis of \mathbb{R}^m .

- Defines partial derivatives.
- Theorem 9.17: $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \rightarrow \mathbb{R}^m$ differentiable at $\mathbf{x} \in E$ imply the partial derivatives $(D_j f_i)(\mathbf{x})$ exist and

$$\mathbf{f}'(\mathbf{x}) \mathbf{e}_j = \sum_{i=1}^m (D_j f_i)(\mathbf{x}) \mathbf{u}_i$$

for $1 \leq j \leq n$.

- It follows that

$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} (D_1 f_1)(\mathbf{x}) & \cdots & (D_n f_1)(\mathbf{x}) \\ \vdots & & \vdots \\ (D_1 f_m)(\mathbf{x}) & \cdots & (D_n f_m)(\mathbf{x}) \end{bmatrix}$$

- Discusses the gradient and the directional derivative.
- Theorem 9.19: $E \subset \mathbb{R}^n$ convex and open, $\mathbf{f} : E \rightarrow \mathbb{R}^m$ differentiable in E , and there exists M such that

$$\|\mathbf{f}'(\mathbf{x})\| \leq M$$

for all $\mathbf{x} \in E$ implies

$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leq M|\mathbf{b} - \mathbf{a}|$$

for all $\mathbf{a}, \mathbf{b} \in E$.

- Corollary: If, in addition, $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in E$, then \mathbf{f} is constant.
- **Continuously differentiable** (mapping $\mathbf{f} : E \rightarrow \mathbb{R}^m$): A function $\mathbf{f} : E \rightarrow \mathbb{R}^m$ such that $\mathbf{f}' : E \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. *Also known as \mathcal{C}^1 -mapping. Denoted by $\mathbf{f} \in \mathcal{C}^1(E)$.*
- Theorem 9.21: Let $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \rightarrow \mathbb{R}^m$. Then $\mathbf{f} \in \mathcal{C}^1(E)$ iff the partial derivatives $D_j f_i$ ($1 \leq i \leq m; 1 \leq j \leq n$) exist and are continuous on E .

- 2/20: • **Contraction** (of X into X): A function $\varphi : X \rightarrow X$ for which there exists a number $c < 1$ such that

$$d(\varphi(x), \varphi(y)) \leq c \cdot d(x, y)$$

for all $x, y \in X$, where X is a metric space with metric d .

- Theorem 9.23: X a complete metric space and ϕ a contraction of X into X implies there exists a unique $x \in X$ such that $\varphi(x) = x$.

Proof. Let $x_0 \in X$ be arbitrary. Define $\{x_n\}$ recursively by

$$x_{n+1} = \phi(x_n)$$

for $n = 0, 1, 2, \dots$. Let $c < 1$ be the number corresponding to the contraction φ . Then for $n \geq 1$, we have

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq c \cdot d(x_n, x_{n-1})$$

or, for $n \geq 0$,

$$d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$$

by induction. Now to prove that $\{x_n\}$ is Cauchy, it will suffice to show that for all $\epsilon > 0$, there exists N such that $m \geq n \geq N$ implies $d(x_n, x_m) < \epsilon$. But since

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=n+1}^m d(x_i, x_{i-1}) \\ &\leq (c^n + c^{n+1} + \cdots + c^{m-1})d(x_1, x_0) \\ &\leq [(1-c)^{-1}d(x_1, x_0)]c^n \end{aligned}$$

we can simply choose N large enough that $[(1-c)^{-1}d(x_1, x_0)]c^N < \epsilon$. Thus, since $\{x_n\}$ is Cauchy and X is complete, there exists $x \in X$ such that $\lim_{n \rightarrow \infty} x_n = x$. Therefore, since φ is Lipschitz continuous, we have that

$$\varphi(x) = \lim_{n \rightarrow \infty} \varphi(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x$$

as desired.

Now suppose for the sake of contradiction that there exists $y \neq x$ such that $\varphi(y) = y$. Then since φ is a contraction,

$$d(y, x) = d(\varphi(y), \varphi(x)) \leq c \cdot d(y, x) < d(y, x)$$

a contradiction. □

- Theorem 9.24 (Inverse Function Theorem): $E \subset \mathbb{R}^n$ open, $\mathbf{f} : E \rightarrow \mathbb{R}^n$ a \mathcal{C}^1 -mapping, $\mathbf{f}'(\mathbf{a})$ invertible for some $\mathbf{a} \in E$, and $\mathbf{b} = \mathbf{f}(\mathbf{a})$ implies

- (a) There exist $U, V \subset \mathbb{R}^n$ open with $\mathbf{a} \in U$, $\mathbf{b} \in V$ such that \mathbf{f} is 1-1 on U and $\mathbf{f}(U) = V$.

Proof. Let $A = \mathbf{f}'(\mathbf{a})$. Choose λ such that

$$2\lambda \|A^{-1}\| = 1$$

Define^[2] for each $\mathbf{y} \in \mathbb{R}^n$ a function φ by

$$\varphi(\mathbf{x}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x}))$$

for all $\mathbf{x} \in E$. (Note that a key property of φ is that as defined, $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ iff \mathbf{x} is a fixed point of φ .) Now since $\mathbf{f} \in \mathcal{C}^1$ and hence \mathbf{f}' is continuous at \mathbf{a} , there exists an open ball $B_r(\mathbf{a}) \subset E$ such that

$$\|\mathbf{f}'(\mathbf{x}) - A\| < \lambda$$

for all $\mathbf{x} \in B_r(\mathbf{a})$. Let $U = B_r(\mathbf{a})$. Clearly it follows that U is open. Thus, since each $\varphi'(\mathbf{x}) = I - A^{-1}\mathbf{f}'(\mathbf{x}) = A^{-1}(A - \mathbf{f}'(\mathbf{x}))$, we have that

$$\|\varphi'(\mathbf{x})\| \leq \|A^{-1}\| \|A - \mathbf{f}'(\mathbf{x})\| < \frac{1}{2\lambda} \cdot \lambda = \frac{1}{2}$$

Consequently, we have by Theorem 9.19 that for all $\mathbf{x}_1, \mathbf{x}_2 \in U$,

$$|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leq \frac{1}{2} |\mathbf{x}_1 - \mathbf{x}_2|$$

Thus, by the uniqueness argument in the proof of Theorem 9.23, φ has at most one fixed point in U , so $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ for at most one $\mathbf{x} \in U$. Therefore, \mathbf{f} is 1-1 on U .

Let $V = \mathbf{f}(U)$. To prove that V is open, it will suffice to show that for all $\mathbf{y}_0 \in V$, there exists an open subset of V containing \mathbf{y}_0 such that. Let $\mathbf{y}_0 \in V$ be arbitrary. By the definition of V as the image of U under \mathbf{f} , there exists $\mathbf{x}_0 \in U$ such that $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$. As such, choose $B_r(\mathbf{x}_0)$ such that $\overline{B} \subset U$. Pick \mathbf{y} satisfying $|\mathbf{y} - \mathbf{y}_0| < \lambda r$. Then

$$|\varphi(\mathbf{x}_0) - \mathbf{x}_0| = |A^{-1}(\mathbf{y} - \mathbf{y}_0)| < \|A\| \lambda r = \frac{r}{2}$$

²How do we motivate this definition?

so for all $\mathbf{x} \in \overline{B}$,

$$\begin{aligned} |\varphi(\mathbf{x}) - \mathbf{x}_0| &\leq |\varphi(\mathbf{x}) - \varphi(\mathbf{x}_0)| + |\varphi(\mathbf{x}_0) - \mathbf{x}_0| \\ &< \frac{1}{2}|\mathbf{x} - \mathbf{x}_0| + \frac{r}{2} \\ &\leq \frac{1}{2} \cdot r + \frac{r}{2} \\ &= r \end{aligned}$$

Thus, $\varphi(\mathbf{x}_0) \in B$. Moreover, since $|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \leq \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$ naturally holds for all $\mathbf{x}_1, \mathbf{x}_2 \in \overline{B} \subset U$, we have that φ is a contraction of \overline{B} into \overline{B} . Additionally, since $\overline{B} \subset \mathbb{R}^n$ is closed, it is a complete metric space under the Euclidean metric. Thus, Theorem 9.23 implies that φ has a fixed point $\mathbf{x} \in \overline{B}$. In particular, $\mathbf{f}(\mathbf{x}) = \mathbf{y}$. Therefore, $\mathbf{y} \in f(\overline{B}) \subset f(U) = V$, as desired. \square

(b) If \mathbf{g} is the inverse of \mathbf{f} on V [which exists by (a)], i.e.,

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$$

for all $\mathbf{x} \in U$, then $\mathbf{g} \in \mathcal{C}^1(V)$.

Proof. We first show that for all $\mathbf{y} \in V$, $\mathbf{g}'(\mathbf{y}) = [\mathbf{f}'(\mathbf{g}(\mathbf{y}))]^{-1}$. Let $\mathbf{y} \in V$ be arbitrary, and choose \mathbf{k} such that $(\mathbf{y} + \mathbf{k}) \in V$. It follows by part (a) that there exist $\mathbf{x}, \mathbf{x} + \mathbf{h} \in U$ such that $\mathbf{y} = \mathbf{f}(\mathbf{x})$ and $\mathbf{y} + \mathbf{k} = \mathbf{f}(\mathbf{x} + \mathbf{h})$. Thus,

$$\varphi(\mathbf{x} + \mathbf{h}) - \varphi(\mathbf{x}) = \mathbf{h} + A^{-1}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x} + \mathbf{h})] = \mathbf{h} - A^{-1}\mathbf{k}$$

so

$$|\mathbf{h} - A^{-1}\mathbf{k}| = |\varphi(\mathbf{x} + \mathbf{h}) - \varphi(\mathbf{x})| \leq \frac{1}{2}|\mathbf{x} + \mathbf{h} - \mathbf{x}| = \frac{1}{2}|\mathbf{h}|$$

Consequently, $|A^{-1}\mathbf{k}| \geq \frac{1}{2}|\mathbf{h}|$, so

$$|\mathbf{h}| \leq 2\|A^{-1}\|\|\mathbf{k}\| = \frac{\|\mathbf{k}\|}{\lambda}$$

Additionally, we know that $\|\mathbf{f}'(\mathbf{x}) - A\|\|A^{-1}\| = 1/2 < 1$, so Theorem 9.8a implies that $\mathbf{f}'(\mathbf{x})$ is invertible with an inverse that we may call T . Thus, since

$$\begin{aligned} \mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k} &= \mathbf{h} - T\mathbf{k} \\ &= -T[(\mathbf{y} + \mathbf{k}) - \mathbf{y}] + T\mathbf{f}'(\mathbf{x})\mathbf{h} \\ &= -T[\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}] \end{aligned}$$

we have that

$$\frac{|\mathbf{g}(\mathbf{y} + \mathbf{k}) - \mathbf{g}(\mathbf{y}) - T\mathbf{k}|}{\|\mathbf{k}\|} \leq \frac{\|T\|\|\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\|}{\lambda|\mathbf{h}|}$$

Consequently, $\mathbf{k} \rightarrow \mathbf{0}$ implies that $\mathbf{h} \rightarrow \mathbf{0}$, which implies that the right side of the above inequality goes to zero, which implies that the left side of the above inequality goes to zero. Thus, $\mathbf{g}'(\mathbf{y}) = T$, so

$$\mathbf{g}'(\mathbf{y}) = [\mathbf{f}'(\mathbf{g}(\mathbf{y}))]^{-1}$$

for all $\mathbf{y} \in V$, as desired.

To prove that \mathbf{g}' is continuous on V , Theorem 4.7 and the above equation tell us that it will suffice to show that $\mathbf{g} : V \rightarrow U$ is continuous, $\mathbf{f}' : U \rightarrow L(\mathbb{R}^n)$ is continuous, and $M \mapsto M^{-1} : L(\mathbb{R}^n) \rightarrow L(\mathbb{R}^n)$ is continuous. But we have the first condition since differentiability implies continuity and \mathbf{g} is differentiable, we have the second condition since $\mathbf{f} \in \mathcal{C}^1$ by hypothesis, and we have the third condition by Theorem 9.8b, as desired. \square

- Theorem 9.25: $E \subset \mathbb{R}^n$ open, $\mathbf{f} : E \rightarrow \mathbb{R}^n$ a \mathcal{C}^1 -mapping, and $\mathbf{f}'(\mathbf{x})$ invertible for all $\mathbf{x} \in E$ implies $\mathbf{f}(W)$ open in \mathbb{R}^n for every open $W \subset E$.

– Note that the hypotheses of this theorem guarantee that \mathbf{f} is locally 1-1 at each $\mathbf{x} \in E$, but it may not be 1-1 in E under these conditions (see Exercise 9.17).