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

Steven Labalme

February 15, 2022

Contents

Chapter 6

The Riemann-Stieltjes Integral

6.1 Notes

- Plan:

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

- Recall that we're given $f : [a, b] \to \mathbb{R}$ continuous, $f : [a, b] \to \mathbb{R}$, 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

    $$\left| \frac{1}{x - x_0} \int_{x_0}^x f(t) \, dt - f(x_0) \right| = \left| \frac{1}{x - x_0} \int_{x_0}^x f(t) \, dt - \frac{1}{x - x_0} \int_{x_0}^x f(x_0) \, dt \right|$$

    $$= \frac{1}{|x - x_0|} \left| \int_{x_0}^x (f(t) - f(x_0)) \, dt \right|$$

    $$\leq \frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| \, dt$$

    $$< \epsilon$$

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

    $$\frac{1}{|x - x_0|} \int_{x_0}^x |f(t) - f(x_0)| \, dt < \frac{1}{|x - x_0|} \int_{x_0}^x \epsilon \, dt$$

    $$= \epsilon$$

- Applications:

    1. Theorem (MVT for integration): $f : [a, b] \to \mathbb{R}$ 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 $F, G : [a, b] \to \mathbb{R}$ 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

$$\int_a^b (FG)' = \int_a^b fG + \int_a^b Fg$$

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

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

3. Theorem (u-substitution).
   − Follows similarly from the chain rule and FTC.

- Integration of vector-valued functions.

- If $f : [a, b] \to \mathbb{R}^k$, 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 $\gamma : [a, b] \to \mathbb{R}^k$ 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 si 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

$$\Lambda(\gamma, P) = \sum_{i=1}^n |\gamma(x_i) - \gamma(x_{i-1})|$$

$$= \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} \gamma'(t) \, dt \right|$$

$$\leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |\gamma'(t)| \, dt$$

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

- 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.

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

  - Between 0 and 2π.

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

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

Chapter 7

Sequences and Series of Functions

7.1 Notes

- 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."
        - $f_n \to f$ pointwise implies basically nothing.
        - $f_n \to f$ 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 $f_n : X \to \mathbb{R}$ for all $n \in \mathbb{N}$ and $f : X \to \mathbb{R}$. *Denoted by $f_n \to f$.*

- Bad functions.
    - Consider $f_n : [0, 1] \to \mathbb{R}$ defined by $x \mapsto x^n$. Each f_n is continuous, but f is not (zero everywhere except $f(1) = 1$)[1].
    - Consider $f_n : \mathbb{R} \to \mathbb{R}$ 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 $f_m : \mathbb{R} \to \mathbb{R}$ defined by $x \mapsto \lim_{n \to \infty} \cos^{2n}(m!\pi x)$. 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 its 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 nth power just makes it spiky.

- Aside: Interchanging limits.
    - If all f_n are continuous, then $\lim_{x \to x_0} f_n(x) = f_n(x_0)$.

---

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

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

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

- 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 $f_n : X \to \mathbb{R}$ for all $n \in \mathbb{N}$ and $f : X \to \mathbb{R}$.

- Proposition (Cauchy criterion for uniform convergence): $f_n \to f$ 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 $f_n \to f$ 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: $f_n : X \to \mathbb{R}$, f_n continuous at $x_0 \in X$ for all n, and $f_n \to f$ 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:
    - $f_n \to f$ 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 $d(f, g) = \|f - g\|$.

- $f_n \to f \iff f$ is bounded.

  - $f_n \to f$ uniformly $\iff \lim_{n \to \infty} \sup |f_n(x) - f(x)| = 0 \iff f_n \to f$ 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: $f_n : [a, b] \to \mathbb{R}$ integrable and $f_n \to f$ uniformly implies f is integrable and

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

– 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 $f_n \to f$ uniformly, $\epsilon_n \to 0$ as $n \to \infty$.
■ By definition, $f_n - \epsilon_n \leq f \leq f_n + \epsilon_n$.
■ Thus, by Theorems 6.4 and 6.5,

$$\int_a^b (f_n - \epsilon_n) = \underline{\int} (f_n - \epsilon_n) \leq \underline{\int} f \leq \overline{f} \leq \int_a^b (f_n + \epsilon_n)$$

■ It follows since

$$0 \leq \overline{\int} f - \underline{\int} f \leq \int_a^b (f_n + \epsilon_n) - \int_a^b (f_n - \epsilon_n) = (b - a)...$$

that f is integrable.

■ Hence,

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

$$\left| \int_a^b f_n - \int_a^b f \right| \leq \epsilon_n$$

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

• Theorem: $f_n : [a, b] \to \mathbb{R}$, each f_n differentiable, $f_n \to f$ pointwise, and $(f_n)' \to g$ 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 $f_n \to f$ 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

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_N(x) - f_N(x_0)}{x - x_0} + \frac{f_N(x) - f_N(x_0)}{x - x_0} - f_N'(x_0) + f_N'(x_0) - g(x_0) \right|$$

$$\leq \left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_N(x) - f_N(x_0)}{x - x_0} \right| + \left| \frac{f_N(x) - f_N(x_0)}{x - x_0} - f_N'(x_0) \right| + |f_N'(x_0) - g(x_0)|$$

- ■ 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: $f_n : E \to \mathbb{R}$, $\{f_n\}$ 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

- Plan:

    1. Warm-up with matrices.
    2. The total derivatives of $f : \mathbb{R}^n \to \mathbb{R}^m$ ($n = m = 2$, i.e., $f : \mathbb{C} \to \mathbb{C}$).
    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 $\phi : V \to W$.

- If we pick bases $N_1, \ldots, N_n$ of V and $w_1, \ldots, w_m$ 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. $|A\mathbf{x}| \leq \|A\| |x|$.
        2. $\|A\| < \infty$ and all $A : \mathbb{R}^n \to \mathbb{R}^m$ 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** (multivariate function f at $\mathbf{x}_0$): A function $f : U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$) such that to $\mathbf{x}_0 \in U$ there corresponds some linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{|f(\mathbf{x}_0 - \mathbf{h}) - f(\mathbf{x}_0) - A\mathbf{h}|}{|\mathbf{h}|} = 0$$

- **Total derivative** (of f multivariate at $\mathbf{x}_0$): The linear transformation A in the above definition. *Denoted by* $\boldsymbol{f'(\mathbf{x}_0)}$.

- "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 f is differentiable at $\mathbf{x}_0 \in U$ and A, B are both derivatives of f at $\mathbf{x}_0$. Then $A = B$.

     – Proposition: Differentiable implies continuous.

     – Proposition: Sum rule, product rule, quotient rule.

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 $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ is independent in X, then X has a basis containing $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$.

- 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 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 $A : \mathbb{R}^n \to \mathbb{R}^m$ uniformly continuous.

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

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

    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 $\|B - A\| \cdot \left\|A^{-1}\right\| < 1$ implies $B \in \Omega$.

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

$$\|B - A\| \cdot \left\|A^{-1}\right\| < 1$$

$$\beta \cdot \frac{1}{\alpha} < 1$$

$$\beta < \alpha$$

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

$$\alpha|\mathbf{x}| = \alpha|A^{-1}A\mathbf{x}| \leq \alpha\left\|A^{-1}\right\| \cdot |Ax| = |A\mathbf{x}| \leq |(A - B)\mathbf{x}| + |B\mathbf{x}| \leq \beta|\mathbf{x}| + |B\mathbf{x}|$$

$$(\alpha - \beta)|\mathbf{x}| \leq |B\mathbf{x}|$$

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. $\qquad\square$

(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 $\|B - A\| \cdot \left\|A^{-1}\right\| < 1$, so $B \in \Omega$ by part (a), as desired.

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

$$(\alpha - \beta)|B^{-1}\mathbf{y}| \leq |BB^{-1}\mathbf{y}| = |\mathbf{y}|$$

$$\left|B^{-1}\left(\frac{\mathbf{y}}{|\mathbf{y}|}\right)\right| \leq (\alpha - \beta)^{-1}$$

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

$$B^{-1} - A^{-1} = B^{-1}I - IA^{-1}$$

$$= B^{-1}AA^{-1} - B^{-1}BA^{-1}$$

$$= B^{-1}(A - B)A^{-1}$$

implies by Theorem 9.7c that

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

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

- 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 $a_{11}, \ldots, a_{mn}$ 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 $p \to A_p$ 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 $\mathbf{f} : E \to \mathbb{R}^m$ where $E \subset \mathbb{R}^n$, then $\mathbf{f}' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$.

- $\mathbf{f}$ 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 $A' : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$, as expected.

- Theorem 9.15 (Chain Rule): E open in $\mathbb{R}^n$, $\mathbf{f} : E \to \mathbb{R}^m$ differentiable at $\mathbf{x}_0 \in E$, $I \supset \mathbf{f}(E)$ open in $\mathbb{R}^m$, and $\mathbf{g} : I \to \mathbb{R}^k$ differentiable at $\mathbf{f}(\mathbf{x}_0)$ implies $\mathbf{F} : E \to \mathbb{R}^k$ 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. $\qquad\square$

- **Components** (of $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$): The real functions $f_1, \ldots, f_m$ 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 $f_i(\mathbf{x}) = f(\mathbf{x}) \cdot \mathbf{u}_i$ $(1 \leq i \leq m)$, where $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is the standard basis of $\mathbb{R}^m$.

- Defines partial derivatives.

- Theorem 9.17: $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \to \mathbb{R}^m$ 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, $\mathbf{f} : E \to \mathbb{R}^m$ 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 $\mathbf{a}, \mathbf{b} \in E$.

---

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

- Corollary: If, in addition, $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in E$, then $\mathbf{f}$ is constant.

- **Continuously differentiable** (mapping $\mathbf{f} : E \to \mathbb{R}^m$): A function $\mathbf{f} : E \to \mathbb{R}^m$ such that $\mathbf{f}' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. *Also known as $\mathscr{C}'$-mapping. Denoted by $\mathbf{f} \in \mathscr{C}'(E)$.*

- Theorem 9.21: Let $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \to \mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}'(E)$ iff the partial derivatives $D_j f_i$ $(1 \leq i \leq m; 1 \leq j \leq n)$ exist and are continuous on E.