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

Steven Labalme

March 11, 2022

Contents

List of Figures

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

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

- 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.

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

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, $f_n - \epsilon_n \le f \le f_n + \epsilon_n$.
■ Thus, by Theorems 6.4 and 6.5,

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

■ It follows since

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

that f is integrable.

■ Hence,

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

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

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

• 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

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

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

- 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 $|f_n(x)| \le M_x$ 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 $n \in \mathbb{N}$, $|f_n(x)| \le M$.

- 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 8

Some Special Functions

8.1 Notes

• Plan:

1. Go over some of the hits in chapter 8.
2. Define sine.
3. Power series.
4. Exponential functions (log, sin, cos).

• Proposition (power series properties): If $\sum_{n=0}^{\infty} a_n x^n$ converges for all $|x| < R$, and $f : B_R(0) \to \mathbb{R}$ is defined by

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

then:

(a) f is continuous.

– From the root test, $\sum_{n=0}^{\infty} a_n x^n$ is in fact absolutely convergent on $(-R, R)$. Therefore, on any interval $[-R + \epsilon, R - \epsilon]$ $(0 < \epsilon < R)$, we have

$$|a_n x^n| \leq |a_n||R + \epsilon|^n$$

so by the M-test, $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-R + \epsilon, R - \epsilon]$. Then since $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-R + \epsilon, R - \epsilon]$, we have (a) since all $\sum_{n=0}^{N} a_n x^n$ are continuous.

(b) f is differentiable with $f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$.

– (b) follows similarly to (a) by uniform convergence.
– Note that $\limsup \sqrt[n]{|n a_n|} = \limsup \sqrt[n]{|a_n|}$ (since $\lim_{n \to \infty} \sqrt[n]{n} = 1$).
– Therefore, $\sum_{n=1}^{\infty} n a_n x^{n-1}$ converges on $(-R, R)$.

(c) More generally, f is infinitely differentiable with

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n x^{n-k}$$

– Now (c) follows as in the proof of (b).

(d) We have the identity

$$a_k = \frac{f^{(k)}(0)}{k!}$$

– (d) follows from (c) by plugging in zero.

- Note that historically, the analysis of power series motivated the development of all of the Chapter 7 theorems; we simply learned those first without motivation to present the proofs in an ordered manner.

- Aside: Consider the exponential function x^y for $x, y \in \mathbb{R}$ with $x \geq 0$.

  - We define it for natural numbers and integers fairly easily, then rationals, and then for reals as the supremum of exponentials of the entries in the Dedekind cut below $x \in \mathbb{R}$.

  - Under this definition, we can confirm our normal exponential rules and then that x^y is continuous.

- Recall that

$$\mathrm{e} = \sum_{n=0}^{\infty} \frac{1}{n!}$$

- So now we are going to construct $E(x)$, $L(x)$, $C(x)$, and $S(x)$ (which are just e^x, $\ln(x)$, $\cos(x)$, and $\sin(x)$).

- Define $E : \mathbb{C} \to \mathbb{C}$ by

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

  - By the proposition, it converges and is continuous for all $z \in \mathbb{C}$.

  - For the real numbers, E is differentiable. (E is also complex-differentiable, but we won't go into that).

- Proposition: $E(z)E(w) = E(z + w)$ for all $z, w \in \mathbb{C}$.

  - We have by the Cauchy product (Mertens' theorem) that

$$E(z)E(w) = \left(\sum_{n=0}^{\infty} \frac{z^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{w^n}{n!} \right)$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k w^{n-k}}{k!(n-k)!}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (z + w)^n$$

$$= E(z + w)$$

- Corollary: $E(z)E(-z) = E(0) = 1$ for all $z \in \mathbb{C}$.

- $E(x) > 0$ for $x \geq 0$.

  - It follows since $E(z + w) = E(z)E(w)$ that $E(x) > 0$ for all $x \in \mathbb{R}$.

- $\mathrm{d}E/\mathrm{d}x = E$; E is the unique, normalized ($E(0) = 1$) function such that this is true.

  - We can prove this from the power series definition.

- $E(x) \to \infty$ as $x \to \infty$ and $E(x) \to 0$ as $x \to -\infty$. (Also from the power series definition.)

- $0 \leq x_1 < x_2$ implies that $E(x_1) < E(x_2)$.

  - Either from $\mathrm{d}E/\mathrm{d}x = E > 0$ or from the power series definition.

  - It follows from $E(z + w) = E(z)E(w)$ that $x_1 < x_2$ implies $E(x_1) < E(x_2)$.

3/9:   • Plan:

1. Keep going with E, L, C, and S.

2. Prove the fundamental theorem of algebra.

• Define

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

– Recall that $E(z + w) = E(z)E(w)$.

• Theorem: $E(x) = e^x$ for all $x \in \mathbb{R}$.

– $E(1) = e^1$ (by definition).
– $E(n) = e^n$ (by $E(z + w) = E(z)E(w)$).
– $[E(p/q)]^q = E(p) = e^p$ (by $E(z + w) = E(z)E(w)$).
– $E(p/q) = e^{p/q}$ for all $p/q \in \mathbb{Q}$.
– $E(x) = e^x$ for all $x \in \mathbb{R}$ (since both LHS and RHS are continuous functions that agree on $\mathbb{Q}$).

• Briefly: $E : \mathbb{R} \to \mathbb{R}^+$ is a strictly increasing surjective function. Thus, we have an inverse function $L : \mathbb{R}^+ \to \mathbb{R}$.

• Theorem: L is differentiable (and therefore continuous).

– Since $E' = E > 0$ everywhere, we may apply the inverse function theorem at every point.

• Now by the chain rule, $E(L(x)) = x$ for all $x \in \mathbb{R}^+$, so taking derivatives yields

$$E'(L(x))L'(x) = 1$$
$$E(L(x))L'(x) = 1$$
$$xL'(x) = 1$$
$$L'(x) = \frac{1}{x}$$

• Proposition:

1. $L(uw) = L(u) + L(w)$.
2. $L(x) = \int_1^x t^{-1} \, dt$.

• Trig functions:

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)] \qquad\qquad S(x) = \frac{1}{2i}[E(ix) - E(-ix)]$$

– You can use these definitions to prove trig identities, having derived them geometrically.

• Proposition: If $x \in \mathbb{R}$, then $C(x), S(x) \in \mathbb{R}$.

– Key observation: $E(\bar{z}) = \overline{E(z)}$.
– We have

$$\overline{C(x)} = \frac{1}{2}[\overline{E(ix)} + \overline{E(-ix)}]$$
$$= \frac{1}{2}[E(-ix) + E(ix)]$$
$$= C(x)$$

– Symmetric for $S(x)$.

- Note that we could equally well define C, S by

$$C(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \qquad\qquad S(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

- Proposition: $E(ix) = C(x) + iS(x)$.

- Proposition: C, S are differentiable with

$$C'(x) = -S(x) \qquad\qquad S'(x) = C(x)$$

- Proposition: For all $x \in \mathbb{R}$, $|E(ix)| = 1$.

   – We have that
   $$|E(ix)|^2 = E(ix)\overline{E(ix)} = E(ix)E(-ix) = E(0) = 1$$

   – Taking square roots of both sides yields the desired result.

- The above result proves that the imaginary axis maps onto the unit circle in the complex plane.

- We now define π and all that.

   – Goal: Show that for all $z \in \mathbb{C}$ with $|z| = 1$, there exists a unique $\theta \in [0, 2\pi)$ such that $e^{i\theta} = z$. Further, $E(ix)$ has period 2π.

- Proposition: $C(x)^2 + S(x)^2 = 1$.

   – Use $E(ix) = C(x) + iS(x)$ and $|E(ix)| = 1$.

- Proposition: There exists some positive number x such that $C(x) = 0$.

   – Suppose (contradiction): $C(x) > 0$ for all $x > 0$ (since $C(0) = 1$).
   – Thus, $S'(x) > 0$ for all $x > 0$.
   – Consequently, given $0 < x < y$,

   $$S(x)(y - x) < \int_x^y S(t)\,\mathrm{d}t = C(x) - C(y) \leq 2$$

   – But we can choose y large enough to make $S(x)(y - x) > 2$, a contradiction.

- $\boldsymbol{\pi}$: The real number such that $\pi/2$ is the unique smallest positive real number with $C(\pi/2) = 0$.

   – We know that a unique smallest number exists because since $C(0) = 1$ and C is continuous, there exists a neighborhood around 0 where C is nonzero.

- Proposition: $S(\pi/2) = 1$.

   – We have

   $$C(\pi/2)^2 + S(\pi/2)^2 = 1$$
   $$S(\pi/2) = \pm 1$$

   – Furthermore, since $S(0) = 0$ and $S'(x) = C(x)$ is positive on $[0, \pi/2)$, we know that S is increasing and thus $S(\pi/2) = +1$.

3/11:   - Analysis I:

   – $\mathbb{R}$.

- Metric spaces.
- Sequences and series (absolute vs. conditional).
- Continuity.

- Analysis II:

  - Differentiability (MVTs and all familiar derivative properties).
  - Integration (fundamental properties and FTC).
  - Sequences of functions (uniform vs. pointwise; uniform limit of cont. is cont., derivative of the limit is the limit of the derivatives, integral of the limit is the limit of the integrals, equicontinuity stuff).
  - More differentiation (but on $\mathbb{R}^n \to \mathbb{R}^m$; basic properties [relation between the total and partial derivative(s)], inverse function theorem, implicit function theorem).
  - Sharkovsky (won't be on the final).
  - exp, log, sin, and cos.

- The **fundamental domain** of $E(z)$ is $\mathbb{R} \times [0, 2\pi i] \subset \mathbb{C}$. Above or below that, $E(z)$ is just periodic.

  - Note that if $z = a + bi$, then $E(z) = E(a)E(bi)$, and we can calculate both of these terms.

- Proposition: Given $z \in \mathbb{C}$ with $|z| = 1$, there exists a unique $\theta \in [0, 2\pi)$ such that $e^{i\phi} = z$.

- Proves that the circumference of the unit circle is 2π as in Rudin (1976).

- Theorem (Fundamental theorem of algebra): $\mathbb{C}$ is algebraically closed. If $p(z) = \sum_{i=0}^n a_i z^i$ is a complex coefficient, complex polynomial, then there exists some $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.

  - Assume WLOG that $a_n = 1$.
  - Plan: look at $\mu = \inf_{z \in \mathbb{C}} |p(z)|$ and show that there exists some $z_0 \in \mathbb{C}$ such that $|p(z_0)| = \mu$, and then $\mu = 0$.
  - $\mu = \inf_{z \in \mathbb{C}} |p(z)|$.
    - We're going for "outside of a big enough ball, $p(z)$ is large [i.e., can't go back to zero]." Let the radius of this ball be R.
    - Just look at $z \in \mathbb{C}$ such that $|z| = R$.
    - We have

$$|p(z)| = |z^n + z_{n-1}z^{n-1} + \cdots + |a_0|$$
$$\geq R^n - |a_{n-1}|R^{n-1} - \cdots - |a_0|$$
$$= R^n[1 - |a_{n-1}|R^{-1} - \cdots - |a_0|R^{-n}]$$

    - As $R \to \infty$, $R^n \to \infty$ but everything else in the above expression goes to zero. This means that the overall expression goes to infinity (R^n dominates).
    - In particular, $|p(z)| > \mu + 100$ in the complement of some $\overline{B}_R(0)$.
    - Thus,

$$\mu = \inf_{z \in \overline{B}_R(0)} |p(z)|$$
$$= |p(z_0)|$$

    for some $z_0 \in \overline{B}_R(0)$ by the extreme value theorem.
  - $\mu = 0$.
    - Suppose (contradiction): $p(z_0) \neq 0$.

- ■ Trick: Define

$$Q(z) = \frac{p(z + z_0)}{p(z_0)}$$

- ■ $Q(0) = 1$ and $Q(z) \geq 1$ (this second statement is what we're gonna contradict).
- ■ Write

$$Q(z) = 1 + b_k z^k + \cdots + b_n z^n$$

- ■ Goal: For r sufficiently small ($r > 0$), the right term below is less than $|b_k|$, so it is positive, meaning that Q is less than or equal to 1 minus a positive number, i.e., $Q < 1$, a contradiction.

$$|Q(re^{i\theta})| \leq 1 - r^k[|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|]$$

- – Proof of goal: There exists $\theta \in [0, 2\pi)$ such that $b_k e^{ik\theta} = -|b_k|$. Take r small enough so $r^k|b_k| < 1$, so by the triangle inequality together with the fact that $|1 + b_k r^k e^{ik\theta}| = 1 - r^k|b_k|$.

- • Corollary: Polynomials over $\mathbb{C}$ factor completely.

8.2   Chapter 8: Some Special Functions

• **Analytic function**: A function that can be represented by a power series.

- • Theorem 8.1: If

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$, then...

1. f converges uniformly on $[-R + \epsilon, R - \epsilon]$ for all $\epsilon > 0$;
2. f is continuous and differentiable on $(-R, R)$;
3. We have the identity

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$$

for all $|x| < R$.

- • Corollary: If f satisfies the hypotheses of Theorem 8.1, then f has derivatives of all orders in $(-R, R)$ given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} c_n x^{n-k}$$

In particular,

$$f^{(k)}(0) = k! c_k$$

for all $k \in \mathbb{N}_0$.

- • Note that there exist functions f that have derivatives of all orders at a point but cannot be expanded in a power series at that point (see Exercise 8.1).

- • Theorem 8.2: If $\sum c_n$ converges and

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

for $|x| < 1$, then

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} c_n$$

- Theorem 8.3: If $\{a_{ij}\}$ $(i, j \in \mathbb{N})$ is a double sequence, $\{b_i\}$ is defined by

$$b_i = \sum_{j=1}^{\infty} |a_{ij}|$$

for all $i \in \mathbb{N}$, and $\sum b_i$ converges, then

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$$

- Theorem 8.4: If

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$ and $a \in (-R, R)$, then f can be expanded in a power series about $x = a$ which converges in $|x - a| < R - |a|$ and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

for all $|x - a| < R - |a|$.

  - "This is an extension of Theorem 5.15 and is als known as Taylor's theorem" (Rudin, 1976, p. 176).

- Theorem 8.5: If $\sum a_n x^n, \sum b_n x^n$ converge on $S = (-R, R)$, E is the set of all $x \in S$ at which

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$$

and E has a limit point in S, then $a_n = b_n$ for $n \in \mathbb{N}_0$. Hence, the above equation holds for all $x \in S$.

- **E**: The function defined as follows for all $z \in \mathbb{C}$. *Given by*

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

- We have that $E(z + w) = E(z)E(w)$ and thus $E(z)E(-z) = E(z - z) = E(0) = 1$ for all $z, w \in \mathbb{C}$.

- Thus, $E(x) = 1/E(-x) > 0$ for all $x \in \mathbb{R}$.

- It follows since $E(x) \to \infty$ as $x \to \infty$ that $E(x) \to 0$ as $x \to -\infty$.

- $0 < x < y$ implies $E(x) < E(y)$.

- We have that
$$E'(z) = \lim_{h \to 0} \frac{E(z + h) - E(z)}{h} = E(z) \lim_{h \to 0} \frac{E(h) - 1}{h} = E(z)$$

- Rudin (1976) proves that $E(x) = e^x$ for all $x \in \mathbb{R}$ as in class.

- Theorem 8.6: Let e^x be defined on $\mathbb{R}$ as above. Then

  (a) e^x is continuous and differentiable for all x.
  (b) $(e^x)' = e^x$.
  (c) e^x is a strictly increasing function of x, and $e^x > 0$.
  (d) $e^{x+y} = e^x e^y$.

(e) $e^x \to \infty$ as $x \to \infty$ and $e^x \to 0$ as $x \to -\infty$.

(f) $\lim_{x \to \infty} x^n e^{-x} = 0$ for all n.

- Theorem 8.6f shows that e^x tends to infinity faster than any power of x.

- **L**: The inverse of E, implied to exist by the IVT since E is strictly increasing and differentiable on $\mathbb{R}$.

- Differentiating $L(E(x)) = x$ with the chain rule reveals that $L'(y) = 1/y$.

- Since $L(1) = L(E(0)) = 0$, the FTC implies that $L(y) = \int_1^y dx / x$.

- If $E(x) = u$ and $E(y) = v$, then

$$
\begin{aligned}
L(uv) &= L(E(x)E(y)) \\
&= L(E(x+y)) \\
&= x + y \\
&= L(u) + L(v)
\end{aligned}
$$

- We define $x^n = E(nL(x))$ for all $x > 0$ and $n \in \mathbb{N}$, which we can extend analogously to before to x^y for any $x > 0$ and $y \in \mathbb{R}$.

- In the same vein, we have that

$$
(x^\alpha)' = E(\alpha L(x)) \cdot \frac{\alpha}{x} = \alpha x^{\alpha - 1}
$$

- We also have $\lim_{x \to \infty} x^{-\alpha} \log x = 0$, i.e., that $\log x \to \infty$ slower than any positive power of x.

- We define

$$
C(x) = \frac{1}{2}[E(ix) + E(-ix)] \qquad\qquad S(x) = \frac{1}{2i}[E(ix) - E(-ix)]
$$

- We know that $E(\bar{z}) = \overline{E(z)}$, so $C(x), S(x)$ are real for real x.

- Also, $E(ix) = C(x) + iS(x)$.

- We have $|E(ix)| = 1$ for all $x \in \mathbb{R}$.

- We have $C(0) = 1$ and $S(0) = 0$.

- We have

$$
C'(x) = -S(x) \qquad\qquad S'(x) = C(x)
$$

- Rudin (1976) proves, as in class, that there exist positive numbers x for which $C(x) = 0$.

- A smallest positive number such that $C(x) = 0$ exists since $f^{-1}(\{0\})$ is closed as the preimage of a closed set under a continuous function.

- We can prove as in class that $C(\pi/2) = 0$ and $S(\pi/2) = 1$. It follows that

$$
E(i\frac{\pi}{2}) = i
$$

so that, by the addition formula, $E(2\pi i) = 1$, and hence $E(z + 2\pi i) = E(z)$ by the addition formula for all $z \in \mathbb{C}$.

- Theorem 8.7:

(a) E is periodic with period $2\pi i$.

  (b) C, S are periodic with period 2π.

  (c) $0 < t < 2\pi$ implies that $E(it) \neq 1$.

  (d) $z \in \mathbb{C}$ with $|z| = 1$ implies there is a unique $t \in [0, 2\pi)$ with $E(it) = z$.

- Calculating the circumference of a circle.

  - Consider the curve $\gamma : [0, 2\pi] \to \mathbb{C}$ defined by

  $$\gamma(t) = E(it)$$

  - This is a simple closed curve in the plane whose range is exactly the unit circle in the plane.

  - Thus, since $\gamma'(t) = iE(it)$, the length of γ (i.e., the circumference of the unit circle) is

  $$\int_0^{2\pi} |\gamma'(t)| \, \mathrm{d}t = 2\pi$$

  - This shows that π has the same geometric significance in analysis with which it was originally defined in geometry.

- Similarly, we can consider the triangle with vertices at $z_1 = 0$, $z_2 = \gamma(t_0)$, and $z_3 = C(t_0)$ to recover the original geometric definition of $C(t)$.

  - We can do the same with S.

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 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ 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. $|A\mathbf{x}| \leq \|A\| |x|$.
    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** (function $\mathbf{f}$ at $\mathbf{x}_0$): A function $\mathbf{f} : U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$) such that to $\mathbf{x}_0 \in U$ there corresponds some linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

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

- Proposition: Let $\mathbf{f} : U \to \mathbb{R}^m$ 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 $\mathbf{f}, \mathbf{g} : U \to \mathbb{R}^m$ 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 $\mathbf{f} : U \to \mathbb{R}^m$ and $\mathbf{g} : V \to \mathbb{R}^k$, 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 $f_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq m$, and $1 \leq j \leq n$. *Denoted by $(\partial f_i / \partial x_j)(\mathbf{x}_0)$, $(D_j f_i)(\mathbf{x}_0)$. Given by*

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x_0}) = \lim_{t \to 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{e}_j) - f_i(\mathbf{x}_0)}{t}$$

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

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

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

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

- Theorem: Let $\mathbf{f} = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$, 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 $\mathbf{x}_0$, $\mathbf{f}(t\mathbf{e}_j)/t \to 0$ as $t \to 0$.
  - Additionally, $\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j)/t = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$.
  - Therefore,

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

  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, $\mathbf{f} : E \to \mathbb{R}^n$ 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 $\mathbf{f}|_U : U \to V$ 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 $\varphi_{\mathbf{y}} : U \to \mathbb{R}^n$ 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

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

    - ■ 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 $\mathbf{y}_0 \in f(U)$ be such that $\mathbf{y}_0 = f(\mathbf{p}_0)$.
  - – 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 $\varphi_{\mathbf{y}} : \overline{B} \to \overline{B}$ 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.
- φ_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!

2/25:
- Plan:

  1. A point on the IFT.
  2. Commuting partials.
  3. Implicit function theorem.

- Subtle point: Last time, in the proof of the IFT, we first found the $U \subset E$ and prove that $\mathbf{f}|_U$ is injective, and then we proved that $\mathbf{f}(U)$ is open.

- The properties of $\varphi_{\mathbf{y}}$.

  - – $\varphi_{\mathbf{y}}(U) \subset U$.
  - – $\varphi_{\mathbf{y}}$ is a contraction since $|\varphi_{\mathbf{y}}(\mathbf{x}_1) - \varphi_{\mathbf{y}}(\mathbf{x}_2)| \leq \frac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2|$.
  - – $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ iff $\varphi_{\mathbf{y}}(\mathbf{x}) = \mathbf{x}$ (fixed points for this contraction mapping).

- Commuting partials.

- When does the following hold?

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

- Simple answer: Not often, but with enough regularity, yes.

- Theorem: Given $f : E \to \mathbb{R}$ where $E \subset \mathbb{R}^n$, we say that f is C^2 (or of class C^2)

- **Class C^2** (function f): A function $f : E \to \mathbb{R}$ (where $E \subset \mathbb{R}^n$) such that all partials $\partial^2 f / \partial x_j \partial x_i$ exist and are continuous for all points in E. *Denoted by $\boldsymbol{f \in C^2}$.*

- Lemma (MVT): If $E \subset \mathbb{R}^2$ open, $f : E \to \mathbb{R}$, $\partial f / \partial x$, $\partial^2 f / \partial y \partial x$ exist for all $(x, y) \in E$, $Q = [a, a+h] \times [b, b+k] \subset E$, and

$$\Delta(f, Q) = f(a + h, b + k) - f(a + h, b) + f(a, b + k) - f(a, b)$$

then there exists $(x_0, y_0) \in Q$ such that

$$\Delta(f, Q) = hk \frac{\partial^2}{\partial y \partial x}(x_0, y_0)$$

- Proof idea: We reduce to the goal of the 1D MVT.
- Define $u(t) = f(t, b + k) - f(t, b)$. Then u is differentiable by the sum and scalar multiple rules.
- It follows that

$$\begin{aligned}
\Delta(f, Q) &= u(a + h) - u(a) \\
&= hu'(x_0) \\
&= h \left[\frac{\partial f}{\partial x} - \frac{\partial f}{\partial x} \right] \\
&= hk \frac{\partial^2}{\partial y \partial x}(x_0, y_0)
\end{aligned}$$

- Theorem: If $f \in C^2$, then

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

for all $1 \leq i, j \leq n$.

- Idea.

- To make life easy, take $n = 2$. Then we just need the right kind of mean value theorem (the one in the lemma).

- Proof.

- Follows from the lemma as $h, k \to 0$.
- See Theorem 15.3 in Labalme (2021).

2/28:   - Plan:

1. Implicit function theorem (end of Chapter 9).
2. Sharkovsky's theorem.
3. Go back and talk about Chapter 8 material.

- Theorem (Implicit Function Theorem; informal): Given a nice system of equations

$$f_1(c_1, \ldots, x_n, y_1, \ldots, y_m) = 0$$

$$\vdots$$

$$f_n(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$$

and a particular solution $(\mathbf{a}, \mathbf{b})$, we can solve for $\mathbf{y} = (y_1, \ldots, y_m)$ locally at $(\mathbf{a}, \mathbf{b})$.

- Theorem (Implicit Function Theorem): If $E \subset \mathbb{R}^{n+m}$, $\mathbf{f} : E \to \mathbb{R}^n$ continuously differentiable, $(\mathbf{a}, \mathbf{b}) \in E$ with $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$, $A = \mathbf{f}'(\mathbf{a}, \mathbf{b})$, $A_\mathbf{x} : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\mathbf{x} \mapsto \mathbf{f}'(\mathbf{a}, \mathbf{b})(\mathbf{x}, \mathbf{0})$ invertible, and $A_\mathbf{y} : \mathbb{R}^m \to \mathbb{R}^n$ defined by $\mathbf{y} \mapsto \mathbf{f}'(\mathbf{a}, \mathbf{b})(\mathbf{0}, \mathbf{y})$, then there exists $U \subset \mathbb{R}^{n+m}$, $W \subset \mathbb{R}^m$ with $(\mathbf{a}, \mathbf{b}) \in U$, $\mathbf{b} \in W$ such that:

  1. For every $\mathbf{y} \in W$, there exists a unique $\mathbf{x} \in \mathbb{R}^n$ such that $(\mathbf{x}, \mathbf{y}) \in U$ and $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.

  2. There is a continuously differentiable function $\mathbf{g} : W \to \mathbb{R}^n$ such that

$$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0}$$

  for all $\mathbf{y} \in W$ and

$$\mathbf{g}'(\mathbf{b}) = -(A_\mathbf{x})^{-1} A_\mathbf{y}$$

- Example: Consider $f : \mathbb{R}^{1+1} \to \mathbb{R}$ defined by $(x, y) \mapsto x^2 + y^2 - 1$.

  - Then $f^{-1}(\{0\})$ is the unit circle.
  - $Df = \begin{bmatrix} 2x & 2y \end{bmatrix}$, $A_x = \begin{bmatrix} 2x \end{bmatrix}$, and $A_y = \begin{bmatrix} 2y \end{bmatrix}$.

- Idea:

  - Use the inverse function theorem, and apply it to $A_\mathbf{x}$.
  - Goal: Find U and W; from this, $\mathbf{g}$ follows uniquely (though you also technically need to show continuous differentiability).
  - Define $F : E \to \mathbb{R}^{n+m}$ by $F(\mathbf{x}, \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y})$. Claim: $F'(\mathbf{a}, \mathbf{b})$ is invertible. Apply the inverse function theorem to F.

- Proof left to us.

- The goal of Sharkovsky's theorem is to understand the iterates of f, i.e., $x, f(x), f^2(x), f^3(x), \ldots$. This fits in thematically with the contraction mapping theorem.

- **Periodic** (point $p \in I$): A point $p \in I$ for which $f^m(p) = p$ for some $m \in \mathbb{N}$.

- **Period** (of $p \in I$ periodic): The least number m such that $f^m(p) = p$.

- **Fixed point**: A periodic point of period 1.

- Example: $f : [0, 1] \to [0, 1]$ defined by $f(x) = 1 - x$ has periodic points of period 2 everywhere on its domain save $1/2$, which has period 1.

- Theorem: If f has a point p of period 3, then f has points of all other periods.

- **Sharkovsky ordering**: All of the odd numbers, then 2^1 times the odd numbers, then 2^2 times the odd numbers, then continuing for $n \to \infty$, and then 2^n as large as possible all the way down to 1.

  - If n comes before m in the Sharkovsky ordering, we write $n \triangleright m$.

- Sharkovsky's theorem: If $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$ continuous satisfies $f(A) \subset A$ and has a period m point with $m \triangleright l$, then f has a period l point.

- Example: Logistic maps $g_b : [0, 1] \to [0, 1]$ defined by $x \mapsto bx(1 - x)$ for $b \in [1, 4]$.

3/2:
- Theorem (Li & Yorke): If f has a period 3 point, then there exists an uncountable set $S \subset A$ such that for all $p, q \in S$,

$$\liminf_{n \to \infty} |f^n(p) - f^n(q)| = 0 \qquad \limsup_{n \to \infty} |f^n(p) - f^n(q)| > 0$$

- Plan:

  1. Prove the warm-up to Sharkovsky.

- Theorem: If f continuous has a period 3 point, then f has points of all other periods.

- Notation.

  – We say "I covers J" and write $I \to J$ when $I, J \subset A$ are closed intervals with $f(I) \supset J$.

- Lemma 1: If $[a, b] \to [a, b]$, then f continuous has a fixed point in $[a, b]$.

  – Consider $f(x) - x$ on $[a, b]$.
  – Either $f(a) = a$, $f(b) = b$, or we can invoke the IVT.
  – Alternatively, since $[a, b] \subset f([a, b])$, there exist a_0, b_0 with $f(a_0) = a$ and $f(b_0) = b$. Thus, $f(a_0) - a_0 \leq 0$ and $f(b_0) - b_0 \geq 0$, so by the IVT, there is some zero of $f(x) - x$ on $[a, b]$, as desired.

- Lemma 2: p has period m iff p is a fixed point of f^m and not a fixed point of f^i for $i < m$.

- Lemma 3: Suppose we have a loop of intervals $J_0 \to J_1 \to J_2 \to \cdots \to J_{n-1} \to J_0 \to \cdots$. Then there is a fixed point $p \in J_0$ of f^n such that $f^i(p) \in J_i$ for all $0 \leq i < n$.
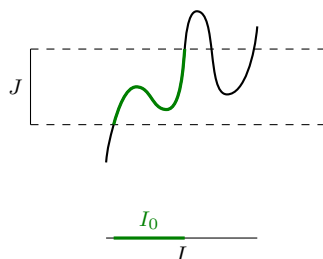

Figure 9.1: Loop mapping.

  – Isn't this obvious? No – there's an issue, namely that $f(J_i) \not\subset J_{i+1}$. We can solve this though by noting that if $I \to J$, then there exists some subinterval $I_0 \subset I$ such that $f(I_0) \subset J$ and $I_0 \to J$ (i.e., $f(I_0) = J$).
  – You can use this idea to pull a J_i' out of each J_i for which set equality holds.
  – Then by the previous lemma, there exists a fixed point p of f^n in J_0'. Then $f(p) \in J_1' \subset J_1, \ldots, f^{n-1}(p) \in J_{n-1}' \subset J_{n-1}$.

- Notation.

  – In the case of Lemma 3, we say that p is **following** the cycle $J_0 \to \cdots \to J_{n-1} \to J_0 \to \cdots$.
  – We call $J_0 \to \cdots \to J_{n-1} \to J_0 \to \cdots$ **elementary** if it is only followed by points of period n.

- Lemma: Let q have period m and let $\mathcal{O} = \{q, f(q), \ldots, f^{m-1}(q)\}$. Let $J_0 \to \cdots \to J_{n-1} \to J_0 \to \cdots$ and suppose

  (i) all endpoints of the J_i are in $\mathcal{O}$;

(ii) the loop is not followed by any point in $\mathcal{O}$;

(iii) The interior of J_0, int (J_0) is disjoint from the other J_i.

Then the loop is elementary, and so f has a period n point.

    – Suppose p follows the loop.

    – Then by (i) and (ii), p is not an endpoint of J_0, so if $f^i(p) = p \in J_0$ for $i < n$, then int $(J_0) \cap J_i$, contradicting (iii).

3/4:
- Plan:

  1. Finish 3 implies all $m \in \mathbb{N}$.
  2. Hint at the rest of Sharkovsky.
  3. Power series/fun with the exponential function (Chapter 8).

- Theorem (warm-up Sharkovsky): If $f : I \to \mathbb{R}$ continuous, $f(I) \subset I$, and f has a period 3 point, then f has a period n point for all $n \in \mathbb{N}$.

  – Consider the three point loop.

  – Let I_1 be the interval from p_0 to p_1, and let I_2 be the interval from p_1 to p_2. We can choose p_0, p_1, p_2 WLOG such that $I_1 \cap I_2 = \{p_1\}$.

  – We have $I_1 \to I_2$, $I_2 \to I_2$, and $I_2 \to I_1$.

  – We may construct

$$I_1 \underbrace{\to I_2}_{\substack{m-1 \\ \text{times}}} \to I_1$$

  to yield a point of period m by the last lemma.

  – Period 1 point: $I_2 \to I_2$ apply Lemma 1.

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

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

    $$\|B - A\| \cdot \|A^{-1}\| < 1$$
    $$\beta \cdot \frac{1}{\alpha} < 1$$
    $$\beta < \alpha$$

    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

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

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

    To prove that $A \mapsto A^{-1}$ is continuous, it will suffice to show that $\|B^{-1} - A^{-1}\| \to 0$ 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 $\mathbf{u} \in \mathbb{R}^n$, $\|B^{-1}\|$ 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

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

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 $\mathbb{R}^m$.

---

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

- 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 $\mathbf{a}, \mathbf{b} \in E$.

- 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 $\mathscr{C}^1$-mapping. Denoted by $\mathbf{f} \in \mathscr{C}^1(E)$.*

- Theorem 9.21: Let $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \to \mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}^1(E)$ 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 $\varphi : X \to X$ 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 $n = 0, 1, 2, \ldots$. 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 $\lim_{n\to\infty} x_n = x$. Therefore, since φ is Lipschitz continuous, we have that

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

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, $\mathbf{f} : E \to \mathbb{R}^n$ a $\mathscr{C}^1$-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 $\mathbf{y}$.) Now since $\mathbf{f} \in \mathscr{C}^1$ 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

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

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

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

---

[2]How do we motivate this definition?

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, $\mathbf{y} \in f(\overline{B}) \subset \mathbf{f}(U) = V$, as desired. $\qquad\square$

(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 $\mathbf{g} \in \mathscr{C}^1(V)$.

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, $\mathbf{k} \to \mathbf{0}$ implies that $\mathbf{h} \to \mathbf{0}$, 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 $\mathbf{g} : V \to U$ is continuous, $\mathbf{f}' : U \to L(\mathbb{R}^n)$ is continuous, and $M \mapsto M^{-1} : L(\mathbb{R}^n) \to L(\mathbb{R}^n)$ is continuous. But we have the first condition since differentiability implies continuity and $\mathbf{g}$ is differentiable, we have the second condition since $\mathbf{f} \in \mathscr{C}^1$ by hypothesis, and we have the third condition by Theorem 9.8b, as desired. $\qquad\square$

- Theorem 9.25: $E \subset \mathbb{R}^n$ open, $\mathbf{f} :: E \to \mathbb{R}^n$ a $\mathscr{C}^1$-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).

3/7:
- Motivating the Implicit Function Theorem.

  - Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is continuously differentiable.
  - Consider the equation $f(x, y) = 0$. If (a, b) satisfies $f(a, b) = 0$ and $\partial f/\partial y \neq 0$, then we can solve for y in terms of x near (a, b).

- Why we must require that $\partial f/\partial y \neq 0$:

  ■ Suppose that $f(x,y)$ is constant in some neighborhood of (a,b). Then there are infinitely many values of y for each x. Consider the line in the plane $x = b$. Since f is constant in a neighborhood of b, the intersection of this line with $\{(x,y) : f(x,y) = 0\}$ will be a segment.

  ■ As another example, consider $f(x,y) = x^2 + y^2 - 1$. $(1,0)$ satisfies $f(1,0) = 0$, but $\mathrm{d}f/\mathrm{d}y = 2y$, so $(\mathrm{d}f/\mathrm{d}y)|_{(x,y)=(1,0)} = 0$. And indeed, for any x slightly less than 1, there are two values of y (those on the upper and lower hemicircles) for which $f(x,y) = 0$, i.e., there is no one-to-one mapping.

- $(\mathbf{x}, \mathbf{y})$: The vector $(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}$, where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$.

- Every $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ can be split into two linear transformations A_x, A_y defined by

$$A_x \mathbf{h} = A(\mathbf{h}, \mathbf{0}) \qquad\qquad A_y \mathbf{k} = A(\mathbf{0}, \mathbf{k})$$

  for all $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{k} \in \mathbb{R}^m$.

  - It follows that

$$A(\mathbf{h}, \mathbf{k}) = A_x \mathbf{h} + A_y \mathbf{k}$$

- Theorem 9.27: If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and A_x invertible, then to every $\mathbf{k} \in \mathbb{R}^m$ there corresponds a unique $\mathbf{h} \in \mathbb{R}^n$ such that $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$. This $\mathbf{h}$ can be computed from $\mathbf{k}$ via the formula

$$\mathbf{h} = -(A_x)^{-1} A_y \mathbf{k}$$

  Proof. Let $\mathbf{k} \in \mathbb{R}^m$ be arbitrary. Since $A - x$ is invertible, the vector $\mathbf{h} = -(A_x)^{-1} A_y \mathbf{k}$ is well-defined and unique. It follows that

$$A(\mathbf{h}, \mathbf{k}) = A_x \mathbf{h} + A_y \mathbf{k} = \mathbf{0}$$

  as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

- Theorem 9.28 (Implicit Function Theorem): If $E \subset \mathbb{R}^{n+m}$ is open, $\mathbf{f} : E \to \mathbb{R}^n \in \mathscr{C}^1$, $(\mathbf{a}, \mathbf{b}) \in E$ satisfies $\mathbf{f}(\mathbf{a}, \mathbf{b}) = 0$, $A = \mathbf{f}'(\mathbf{a}, \mathbf{b})$, and A_x is invertible, then there exist

$$U \subset \mathbb{R}^{n+m} \qquad\qquad\qquad W \subset \mathbb{R}^m$$

  open with $(\mathbf{a}, \mathbf{b}) \in U$ and $\mathbf{b} \in W$ such that to every $\mathbf{y} \in W$ there corresponds a unique $\mathbf{x}$ such that

$$(\mathbf{x}, \mathbf{y}) \in U \qquad\qquad\qquad \mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$$

  and if $\mathbf{g}(\mathbf{y}) = \mathbf{x}$, then $\mathbf{g} : W \to \mathbb{R}^n \in \mathscr{C}^1$, $\mathbf{g}(\mathbf{b}) = \mathbf{a}$, $\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0}$ for all $\mathbf{y} \in W$, and $\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1} A_y$.

3/8: *Proof.* Define $\mathbf{F} : E \to \mathbb{R}^{n+m}$ by

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y})$$

for all $(\mathbf{x}, \mathbf{y}) \in E$. Since $\mathbf{f} \in \mathscr{C}^1$ and $\mathbf{y} \mapsto \mathbf{y}$ is a $\mathscr{C}^1$-mapping as the identity function, $\mathbf{F} \in \mathscr{C}^1$.

To prove that $\mathbf{F}'(\mathbf{a}, \mathbf{b}) = B$ is invertible, Theorem 9.5 tells us that it will suffice to show that it is 1-1. To do so, we will verify that $B(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ iff $(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. The forward case follows since B is a linear transformation. For the reverse case, however, we must first note that $B(\mathbf{x}, \mathbf{y}) = (A(\mathbf{x}, \mathbf{y}), \mathbf{y})$, which follows from the fact that $\mathbf{F}$ acts like $\mathbf{f}$ on its first n dimensions and then like the identity transformation thereafter (recall that the total derivative of a linear transformation is itself). With this established, the equation $B(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ tells us that $A(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ and $\mathbf{y} = \mathbf{0}$. Combining these two results, we have that $A(\mathbf{x}, \mathbf{0}) = \mathbf{0}$. But by Theorem 9.27, it follows that $\mathbf{x} = -(A_x)^{-1} A_y \mathbf{0} = \mathbf{0}$. Therefore, $(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, as desired.

Having established that $E \subset \mathbb{R}^{n+m}$ is open, $\mathbf{F} : E \to \mathbb{R}^{n+m} \in \mathscr{C}^1$, $\mathbf{F}'(\mathbf{a}, \mathbf{b})$ is invertible, and $(\mathbf{0}, \mathbf{b}) = \mathbf{F}(\mathbf{a}, \mathbf{b})$, we have by Inverse Function Theorem that there exist open sets $U, V \in \mathbb{R}^{n+m}$ with $(\mathbf{a}, \mathbf{b}) \in U$ and $(\mathbf{0}, \mathbf{b}) \in V$ such that $\mathbf{F}$ is 1-1 on U and $\mathbf{F}(U) = V$. Let

$$W = \{\mathbf{y} \in \mathbb{R}^m : (\mathbf{0}, \mathbf{y}) \in V\}$$

The openness of W follows from the fact that V is open (and, technically, that the projection operator is continuous). It also clearly follows from the definition of W and the fact that $(\mathbf{0}, \mathbf{b}) \in V$ that $\mathbf{b} \in W$. Now let $\mathbf{y} \in W$ be arbitrary. Since $\mathbf{F}$ is 1-1 on U and $\mathbf{F}(U) = V$, there exists a unique $(\mathbf{x}, \mathbf{y}) \in U$ such that $\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{y})$. Since $\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y})$, it follows that $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.

We formalize the notion that to every $\mathbf{y} \in W$ we can identify a unique $\mathbf{x} \in \mathbb{R}^n$ by the above rule by defining the function $\mathbf{g} : W \to \mathbb{R}^n$ such that $\mathbf{g}(\mathbf{y}) = \mathbf{x}$ for all $\mathbf{y} \in W$. It clearly follows by the definition of the function that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$ and that $\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0}$ for all $\mathbf{y} \in W$.

To verify that $\mathbf{g} \in \mathscr{C}^1$, note that the IVT also implies that if $\mathbf{G} : V \to U$ is such that $\mathbf{G}(\mathbf{F}(\mathbf{x}, \mathbf{y})) = (\mathbf{x}, \mathbf{y})$, then $\mathbf{G} \in \mathscr{C}^1$. Since $\mathbf{G}(\mathbf{0}, \mathbf{y}) = (\mathbf{g}(\mathbf{y}), \mathbf{y})$, it follows that $\mathbf{g} \in \mathscr{C}^1$.

Since $\mathbf{g} \in \mathscr{C}^1$, we know that $\mathbf{g}'(\mathbf{b})$ is well-defined.

By the definition of $\mathbf{f}$ and $\mathbf{G}$, we know that $\mathbf{f}(\mathbf{G}(\mathbf{0}, \mathbf{y})) = \mathbf{0}$ for all $\mathbf{y} \in W$. To take the derivative of the above expression with respect to changes in $\mathbf{y}$ only, define $\boldsymbol{\Phi} : W \to \mathbb{R}^{n+m}$ by $\boldsymbol{\Phi}(\mathbf{y}) = \mathbf{G}(\mathbf{0}, \mathbf{y})$. Thus, since $\mathbf{f} \circ \boldsymbol{\Phi}$ is constant on its domain, the chain rule implies that

$$\mathbf{f}'(\boldsymbol{\Phi}(\mathbf{y})) \cdot \boldsymbol{\Phi}'(\mathbf{y}) = 0$$

for all $\mathbf{y} \in W$. In particular,

$$\begin{aligned}
0 &= \mathbf{f}'(\boldsymbol{\Phi}(\mathbf{b})) \cdot \boldsymbol{\Phi}'(\mathbf{b}) \\
&= \mathbf{f}'(\mathbf{a}, \mathbf{b}) \cdot (\mathbf{g}(\mathbf{b}), \mathbf{b})' \\
&= A \cdot (\mathbf{g}'(\mathbf{b}), I) \\
&= A_x \mathbf{g}'(\mathbf{b}) + A_y I \\
\mathbf{g}'(\mathbf{b}) &= -(A_x)^{-1} A_y
\end{aligned}$$

as desired. $\qquad\square$

- Rudin (1976) discusses a number of topics from linear algebra.

- **Second-order partial derivative** (of f): A partial derivative of one of the partial derivatives of f, if it exists. *Denoted by $\boldsymbol{D_{ij}f}$. Given by*

$$D_{ij}f = D_i D_j f$$

- **Class $\mathscr{C}^2$** (function f): A function f for which $D_{ij}f$ is continuous on E for all $1 \le i, j \le n$.

- Theorem 9.40: If $E \subset \mathbb{R}^2$ open, $f : E \to \mathbb{R}$, $D_1 f, D_{21} f$ exist at every $(x, y) \in E$, $Q \subset E$ a closed rectangle with sides parallel to the coordinate axes having (a, b) and $(a+h, b+k)$ as opposite vertices (for $h, k \neq 0$), and

$$\Delta(f, Q) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)$$

then there exists $(x, y) \in \operatorname{int} Q$ such that

$$\Delta(f, Q) = hk(D_{21}f)(x, y)$$

- Theorem 9.41: If $E \subset \mathbb{R}^2$ is open, $f : E \to \mathbb{R}$, $D_1 f, D_{21} f, D_2 f$ exist on E, and $D_{21} f$ is continuous at $(a, b) \in E$, then $D_{12} f$ exists at (a, b) and

$$(D_{12}f)(a, b) = (D_{21}f)(a, b)$$

- Corollary: $D_{21}f = D_{12}f$ if $f \in \mathscr{C}^2(E)$.

References

Labalme, S. (2021). *Calculus and analytic geometry (Thomas) notes.* Retrieved February 25, 2022, from https://github.com/shadypuck/CAAGThomasNotes/blob/master/Notes/notes.pdf

Rudin, W. (1976). *Principles of mathematical analysis* (A. A. Arthur & S. L. Langman, Eds.; Third). McGraw-Hill.