Notes and exercises from *Introduction to Real Analysis*

John Peloquin

Introduction

This document contains notes and exercises from [1].

Chapter 5

We provide an alternative proof of Cauchy-Schwarz for $\mathbb{R}^n$ (Proposition 2(v)).

Theorem (Cauchy-Schwarz)**.** *If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, then $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|$.*

Proof. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$. We must prove

$$\left| \sum_{k=1}^{n} x_k y_k \right| \leq \sqrt{\sum_{k=1}^{n} x_k^2} \sqrt{\sum_{k=1}^{n} y_k^2} \tag{1}$$

Let $B = 2\sum_{k=1}^{n} x_k y_k$, $A = \sum_{k=1}^{n} x_k^2$, and $C = \sum_{k=1}^{n} y_k^2$. Then (1) is equivalent to $(B/2)^2 - AC \leq 0$, which is equivalent to $D = B^2 - 4AC \leq 0$. We may assume $A \neq 0$, so D is just the discriminant of the quadratic polynomial $P(x) = Ax^2 + Bx + C$ and $D \leq 0$ if $P \geq 0$. Now

$$Ax^2 + Bx + C = \sum_{k=1}^{n} (xx_k)^2 + \sum_{k=1}^{n} 2(xx_k) y_k + \sum_{k=1}^{n} y_k^2 = \sum_{k=1}^{n} (xx_k + y_k)^2 \geq 0 \qquad \square$$

Chapter 13

We provide an alternative proof of Theorem 5.

Theorem. *If γ is a gauge on $I = [a, b]$, there exists a γ-fine tagged division of I.*

Proof. If not, bisect I. At least one of the two resulting subintervals must not have a γ-fine tagged division, lest we could combine γ-fine tagged divisions of both to obtain a γ-fine tagged division of I.

Continue bisecting recursively to obtain a descending chain $I = I_0 \supseteq \cdots \supseteq I_k \supseteq \cdots$ of subintervals with $l(I_k) \to 0$ such that I_k does not have a γ-fine tagged division. Let $x \in \bigcap_{k=0}^{\infty} I_k$. Then there exists $I_k \subseteq \gamma(x)$, so $\{(x, I_k)\}$ is a γ-fine tagged division of I_k—a contradiction. □

Remark. In the proof of the second part of the fundamental theorem of calculus (Theorem 28), note that $f - f(x)$ is absolutely integrable by Corollary 33.

We provide an alternative proof of the forward direction of Theorem 42.

Theorem. *If $f : [a, b] \to \mathbb{R}$ is integrable, then*

$$\int_a^b f = \lim_{c \to a^+} \int_c^b f$$

Proof. Note $\int_c^b f$ exists for all $c \in [a, b]$ by Theorem 21. Now

$$\begin{aligned}
\lim_{c \to a^+} \int_c^b f &= \lim_{c \to a^+} \left[\int_a^b f - \int_a^c f \right] && \text{by Theorem 18} \\
&= \int_a^b f - \lim_{c \to a^+} \int_a^c f \\
&= \int_a^b f - \int_a^a f && \text{by Theorem 27} \\
&= \int_a^b f
\end{aligned}$$

□

Remark. We used the continuity of the indefinite integral (Theorem 27), which was proved using Henstock's Lemma (Lemma 25). The proof of the forward direction of Theorem 42 in the book unnecessarily repeats some ideas from the proof of Henstock's Lemma. See also the proof of the forward direction of Theorem 14.6 in the book.

Our proof is natural because the theorem itself expresses the continuity of the integral when viewed as a function of the interval of integration. See also the remarks on the monotone convergence theorem below.

Chapter 14

Exercise (12)**.** Let $f : [a, \infty) \to \mathbb{R}$ be continuous and such that the indefinite integral $F(x) = \int_a^x f$ is bounded. Let $g : [a, \infty) \to \mathbb{R}$ be nonnegative, decreasing, and differentiable. Then $\int_a^\infty fg$ exists if either $\lim_{x\to\infty} g(x) = 0$ or $\int_a^\infty f$ exists.

Proof. For $b \in [a, \infty)$, fg is continuous on $[a, b]$, so $\int_a^b fg$ exists (Theorem 13.23). By the fundamental theorem of calculus (Theorem 13.28), $F' = f$ on $[a, b]$, so integration by parts (Proposition 13.17) yields

$$\int_a^b fg = \int_a^b F'g = F(b)g(b) - F(a)g(a) - \int_a^b Fg' = F(b)g(b) - \int_a^b Fg'$$

Now $\lim_{b\to\infty} g(b)$ exists since g is decreasing and bounded below by zero. If $\lim_{b\to\infty} g(b) = 0$, then boundedness of F implies $\lim_{b\to\infty} F(b)g(b) = 0$. If $\int_a^\infty f$ exists, then $\lim_{b\to\infty} F(b)$ exists by the continuity of the indefinite integral at ∞ (Theorem 6), so $\lim_{b\to\infty} F(b)g(b)$ exists.

We claim $\int_a^\infty Fg'$ exists. First observe

$$\int_a^\infty g' = \lim_{b\to\infty} \int_a^b g' = \lim_{b\to\infty} [g(b) - g(a)] = \lim_{b\to\infty} g(b) - g(a)$$

exists by the continuity of the indefinite integral, the fundamental theorem of calculus, and the limit of g. Since $g' \leq 0$,

$$\int_a^\infty |g'| = \int_a^\infty -g' = -\int_a^\infty g'$$

also exists by linearity of the integral. Write $|F| \leq M$. Then $|Fg'| \leq M|g'|$ and $\int_a^\infty M|g'| = M\int_a^\infty |g'|$ exists, so $\int_a^\infty Fg'$ exists by comparison (Corollary 7).

Finally, by continuity of the indefinite integral twice more,

$$\int_a^\infty fg = \lim_{b\to\infty} F(b)g(b) - \int_a^\infty Fg' \qquad \square$$

Remark. These are just Dirichlet's and Abel's tests for convergence of integrals, which are continuous versions of the corresponding discrete tests for infinite series of numbers (Exercises 4.22 and 4.25) and functions (Exercises 11.36 and 11.37). The proofs are essentially the same in the continuous and discrete cases, except integration by parts replaces summation by parts and absolute integrability replaces absolute convergence.

Remark. Instead of assuming that g is nonnegative and decreasing, we may assume only that g is bounded and monotone.

Chapter 15

Remark. The monotone convergence theorem (Theorem 1) shows continuity of the integral *when viewed as a function of the integrand.* Specifically, if $f_k \uparrow f$ (or $f_k \downarrow f$), then $\int_I f = \int_I \lim f_k = \lim \int_I f_k$, when this limit exists. Similarly, the results on "improper" integrals (Theorems 13.42 and 14.6) show continuity for the integral *when viewed as a function of the interval of integration.* Specifically, if $I_k \uparrow I$, then $\int_I f = \int_{\lim I_k} f = \lim \int_{I_k} f$, when this limit exists.

The proofs in the book for both types of continuity are similar, and involve combining infinitely many "approximating" gauges (in the former case for the functions f_k, and in the latter case for the intervals I_k) to construct a single "limit" gauge for f over I.

Remark. The dominated convergence theorem (Theorem 3) actually shows the *absolute integrability* of f, by comparison (Corollary 13.33 and Exercise 47).

Remark. The theme of "domination" recurs frequently, for example in:

- The comparison test for absolute integrability (Corollary 13.33)
- The comparison tests for "improper" integrability (Corollary 13.43(ii) and Corollary 14.7(ii))
- The dominated convergence theorem and its many applications

Remark. In the proof of the general Leibniz differentiation rule for integrals (Corollary 9), to see that $D_1 G$ is continuous, let $\boldsymbol{v_0} = (x_0, y_0, z_0)$ and $\boldsymbol{v} = (x, y, z)$ and observe that

$$\begin{aligned}
\lim_{\boldsymbol{v}\to\boldsymbol{v_0}} D_1 G(\boldsymbol{v}) &= \lim_{\boldsymbol{v}\to\boldsymbol{v_0}} \int_y^z D_1 f(x,t)\,dt \\
&= \lim_{\boldsymbol{v}\to\boldsymbol{v_0}} \left[\int_y^{y_0} D_1 f(x,t)\,dt + \int_{y_0}^{z_0} D_1 f(x,t)\,dt + \int_{z_0}^{z} D_1 f(x,t)\,dt \right] \\
&= \lim_{\boldsymbol{v}\to\boldsymbol{v_0}} \int_y^{y_0} D_1 f(x,t)\,dt + \lim_{\boldsymbol{v}\to\boldsymbol{v_0}} \int_{y_0}^{z_0} D_1 f(x,t)\,dt + \lim_{\boldsymbol{v}\to\boldsymbol{v_0}} \int_{z_0}^{z} D_1 f(x,t)\,dt \\
&= \lim_{x\to x_0} \int_{y_0}^{z_0} D_1 f(x,t)\,dt \qquad (1) \\
&= \int_{y_0}^{z_0} D_1 f(x_0,t)\,dt \qquad (2) \\
&= D_1 G(\boldsymbol{v_0})
\end{aligned}$$

where (1) follows from the continuity of indefinite integrals (Theorem 13.27, see above) and (2) follows from Theorem 5. $\square$

We provide an alternative proof of Proposition 47.

Proposition. *Let $f : \mathbb{R} \to \mathbb{R}$ and let $\{E_i\}$ be a family of pairwise disjoint subsets of $\mathbb{R}$ over each of which f is absolutely integrable. Then f is absolutely integrable over $E = \bigcup_{i=1}^{\infty} E_i$ if and only if $\sum_{i=1}^{\infty} \int_{E_i} |f| < \infty$, in which case*

$$\int_E |f| = \sum_{i=1}^{\infty} \int_{E_i} |f| \quad \textit{and} \quad \int_E f = \sum_{i=1}^{\infty} \int_{E_i} f$$

Proof. We decompose f into positive and negative parts (Proposition 13.40(i)).

Since f is absolutely integrable over each E_i, f^+ and f^- are integrable over each E_i. Let $S_n = \bigcup_{i=1}^{n} E_i$, so $\{S_n\}$ is increasing and $E = \bigcup_{i=1}^{\infty} S_i$. By disjoint additivity (Proposition 12(ii)), f^+ and f^- are integrable over each S_n with

$$\int_{S_n} f^+ = \sum_{i=1}^{n} \int_{E_i} f^+ \quad \text{and} \quad \int_{S_n} f^- = \sum_{i=1}^{n} \int_{E_i} f^-$$

By monotone convergence (Proposition 16), f^+ and f^- are integrable over E, and equivalently f is absolutely integrable over E, if and only if

$$\lim_{n\to\infty} \int_{S_n} f^+ = \sum_{i=1}^{\infty} \int_{E_i} f^+ < \infty \quad \text{and} \quad \sum_{i=1}^{\infty} \int_{E_i} f^- < \infty \tag{1}$$

in which case

$$\int_E f^+ = \sum_{i=1}^{\infty} \int_{E_i} f^+ \quad \text{and} \quad \int_E f^- = \sum_{i=1}^{\infty} \int_{E_i} f^-$$

Now if (1) holds, then

$$\sum_{i=1}^{\infty} \int_{E_i} |f| = \sum_{i=1}^{\infty} \int_{E_i} [f^+ + f^-] = \int_E f^+ + \int_E f^- = \int_E |f| < \infty$$

and similarly

$$\sum_{i=1}^{\infty} \int_{E_i} f = \sum_{i=1}^{\infty} \int_{E_i} [f^+ - f^-] = \int_E f^+ - \int_E f^- = \int_E f$$

Conversely, if $\sum_{i=1}^{\infty} \int_{E_i} |f| < \infty$, then

$$\sum_{i=1}^{n} \int_{E_i} f^+ \le \sum_{i=1}^{n} \int_{E_i} |f| \le \sum_{i=1}^{\infty} \int_{E_i} |f| < \infty$$

and similarly for $\sum_{i=1}^{n} \int_{E_i} f^-$, so (1) holds. $\square$

Remark. This result gives us *countable additivity* of the integral over arbitrary (disjoint) sets, provided that we have absolute integrability over each of the sets and the sum of these integrals is finite.

Exercise (1)**.** The monotone convergence theorem (Theorem 1) does not hold for the Riemann integral.

Proof. Let $I = [0,1]$, $J = \mathbb{Q} \cap I = \{r_1, r_2, \ldots\}$, and $J_k = \{r_1, \ldots, r_k\}$. Each C_{J_k} has only finitely many discontinuities and so is R-integrable over I with $R\int_0^1 C_{J_k} = 0$ and hence $\lim_{k\to\infty} R\int_0^1 C_{J_k} = 0$. Moreover, $C_{J_k} \uparrow C_J$. However, C_J is not R-integrable. (See also Example 12.4.) □

Exercise (5)**.** The dominated convergence theorem (Theorem 3) does not hold for the Riemann integral.

Proof. By the proof of Exercise 1, since $|C_{J_k}| \le 1$ and $R\int_0^1 1 = 1$. □

Exercise (6 (Bounded convergence theorem))**.** Let I be a compact interval and suppose $f_k, f : I \to \mathbb{R}$ with $|f_k| \le M$ and f_k integrable for all $k \ge 1$. If $f_k \to f$, then f is integrable and $\int_I f = \lim_{k\to\infty} \int_I f_k$.

Proof. By the dominated convergence theorem, since $\int_I M < \infty$. □

Remark. This result does not hold for unbounded intervals I. For example, consider $I = [0, \infty)$ and $f_k(x) = e^{-x/k}$. Then $|f_k| \le 1$ on I and $\int_0^\infty f_k = k < \infty$ for all $k \ge 1$. Also $f_k \to 1$ on I. However, $\int_0^\infty 1 = \infty$. See also Exercise 32.

Exercise (7)**.** The bounded convergence theorem (Exercise 6) does not hold for the Riemann integral.

Proof. By the proof of Exercise 5. □

Exercise (9 (Laplace transform))**.** Let $f : [0, \infty) \to \mathbb{R}$ be continuous and define

$$\mathscr{L}\{f\}(x) = \int_0^\infty e^{-xt} f(t)\, dt$$

If f is of exponential order—that is, if there exist $M > 0$ and $a \in \mathbb{R}$ such that $|f(t)| \le Me^{at}$ for all $t > 0$—then $\mathscr{L}\{f\}$ defines a continuous function for $x > a$.

Proof. Define $g(x,t) = e^{-xt} f(t)$. For fixed $x > a$, $\int_0^c g(x,t)\, dt$ exists for all $c \ge 0$ by continuity of g in t. Also $|g(x,t)| \le Me^{(a-x)t}$ for $t > 0$ and $\int_0^\infty e^{(a-x)t}\, dt < \infty$, so $\mathscr{L}\{f\}(x) = \int_0^\infty g(x,t)\, dt$ exists by comparison.

For fixed $t > 0$, g is continuous in x, so by the continuity of the integral (Theorem 5), $\mathscr{L}\{f\}$ is continuous for $x > a$. □

Exercise (12)**.** $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Proof. We have

$$\begin{aligned}
\Gamma\left(\tfrac{1}{2}\right) &= \int_0^\infty t^{-1/2} e^{-t}\, dt \\
&= 2\int_0^\infty \frac{e^{-t}}{2\sqrt{t}}\, dt \\
&= 2\int_0^\infty e^{-u^2}\, du && \text{by substitution of } u = \sqrt{t} \\
&= 2\left(\frac{\sqrt{\pi}}{2}\right) = \sqrt{\pi} && \text{by Example 8} \qquad \square
\end{aligned}$$

Exercise (21)**.** The monotone additivity property (Proposition 16) does not hold for the Riemann integral.

Proof. Let $f = 1 \geq 0$. Let $E = \mathbb{Q} \cap [0,1] = \{r_1, r_2, \ldots\}$ and $E_k = \{r_1, \ldots, r_k\}$, so $E_k \uparrow E$. Then $R\int_{E_k} f = 0$, so $\lim_{k\to\infty} R\int_{E_k} f = 0$, but f is not R-integrable over E. □

Exercise (22)**.** The countable additivity property (Proposition 17) does not hold the Riemann integral.

Proof. As in the proof of Exercise 21, except let $E_k = \{r_k\}$. □

Exercise (23)**.** The collection $\mathcal{N} = \{E \subseteq \mathbb{R} \mid \int_E 1 = 0\}$ of null sets is closed under countable unions, countable intersections, differences, and subsets.

Proof. First observe that if $E, F \in \mathcal{N}$, so $\int_E 1 = 0 = \int_F 1$, then $\int_{E\cup F} 1$ and $\int_{E\cap F} 1$ exist by absolute integrability (Proposition 14) and

$$\int_{E\cup F} 1 + \int_{E\cap F} 1 = \int_E 1 + \int_F 1 = 0$$

by additivity (Proposition 12). It follows that $\int_{E\cup F} 1 = 0 = \int_{E\cap F} 1$ by positivity of the integral, so $E \cup F \in \mathcal{N}$ and $E \cap F \in \mathcal{N}$. By induction, $\mathcal{N}$ is closed under finite unions and intersections.

Suppose $\{E_k\}_{k=1}^\infty \subseteq \mathcal{N}$. Define $S_n = \bigcup_{k=1}^n E_k$ and $S = \bigcup_{k=1}^\infty E_k$. By the above, $S_n \in \mathcal{N}$ for all $n \geq 1$, so $\int_{S_n} 1 = 0$ for all $n \geq 1$. Now $S_n \uparrow S$, so by monotone convergence (Proposition 16), $\int_S 1 = 0$ and $S \in \mathcal{N}$. Similarly if $T_n = \bigcap_{k=1}^n E_k$ and $T = \bigcap_{k=1}^\infty E_k$, then $T_n \in \mathcal{N}$ for all $n \geq 1$ and $T_n \downarrow T$, so $T \in \mathcal{N}$. Therefore $\mathcal{N}$ is closed under countable unions and intersections.

Now if $E, F \in \mathcal{N}$, then by the above and additivity,

$$\int_{E-F} 1 = \int_E 1 - \int_{E\cap F} 1 = 0$$

so $E - F \in \mathcal{N}$ and $\mathcal{N}$ is closed under differences.

Finally, if $E \in \mathcal{N}$ and $F \subseteq E$, then

$$\int_F 1 = \int_E 1 - \int_{E-F} 1 = 0$$

so $F \in \mathcal{N}$ and $\mathcal{N}$ is closed under subsets. □

Exercise (26)**.** Let $f, g : I \to \mathbb{R}$. Suppose f is absolutely integrable and there exists a sequence $\{s_n\}$ of step functions[1] with $|s_n| \leq M$ for all $n \geq 1$ and $s_n \to g$. Then fg is absolutely integrable.

Proof. Observe that fs_n is integrable and

$$|fs_n| \leq |f||s_n| \leq M|f|$$

which is integrable. Also $fs_n \to fg$. Therefore fg is absolutely integrable by the dominated convergence theorem. □

Exercise (31 (Improved dominated convergence theorem))**.** Let $f, f_k, g : I \to \mathbb{R}$. Suppose f_k, g are integrable with $|f_k| \leq g$ for all $k \geq 1$ and $f_k \to f$. Then

$$\lim_{k\to\infty} \int_I |f_k - f| = 0$$

Proof. By the dominated convergence theorem, f is integrable, hence $f_k - f$ is integrable, and $\lim_{k\to\infty} \int_I (f_k - f) = 0$. Also

$$|f_k - f| \leq |f_k| + |f| \leq g + g = 2g$$

so $f_k - f$ is absolutely integrable by comparison. Now $|f_k - f| \to 0$, so by the dominated convergence theorem again,

$$\lim_{k\to\infty} \int_I |f_k - f| = \int_I \lim_{k\to\infty} |f_k - f| = \int_I 0 = 0$$ □

[1] A step function is of the form $s = \sum_{k=1}^n a_k C_{I_k}$ where $a_k \in \mathbb{R}$ and I_k is an interval for $1 \leq k \leq n$.

Exercise (32)**.** Let $f_n(t) = n/(t^2+n^2)$ for $t \in \mathbb{R}$. Then $0 \le f_n \le 1$ and $\int_{-\infty}^{\infty} f_n = \pi$ for all $n \ge 1$, and $f_n \to 0$.

Proof. For all $t \in \mathbb{R}$ and $n \ge 1$,

$$0 < \frac{n}{t^2+n^2} \le \frac{n}{n^2} = \frac{1}{n} \le 1$$

Also

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{n}{t^2+n^2}\,dt &= \int_{-\infty}^{\infty} \frac{n^2}{n^2u^2+n^2}\,du \qquad \text{by substitution of } u = t/n \\
&= \int_{-\infty}^{\infty} \frac{1}{u^2+1}\,du \\
&= \tan^{-1} u\Big|_{-\infty}^{\infty} \\
&= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi
\end{aligned}$$

Finally,

$$\lim_{n\to\infty} \frac{n}{t^2+n^2} = \lim_{n\to\infty} \frac{1/n}{t^2(1/n^2)+1} = 0$$ □

Remark. In this example,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n = \pi \ne 0 = \int_{-\infty}^{\infty} \lim_{n\to\infty} f_n$$

This does not contradict the dominated convergence theorem because $\{f_n\}$ is not dominated by an integrable function over all of $\mathbb{R}$. Similarly this does not contradict the bounded convergence theorem because $\mathbb{R}$ is unbounded.

Exercise (33)**.** Let $f, f_k, g : \mathbb{R} \to \mathbb{R}$. Suppose f_k, g are integrable with $|f_k| \le g$ for all $k \ge 1$ and $f_k \to f$. Let F_k be the indefinite integral of f_k, and F that of f, so $F_k(x) = \int_0^x f_k$ and $F(x) = \int_0^x f$. Then $F_k \to F$ uniformly on $\mathbb{R}$.

Proof. By the dominated convergence theorem (Exercise 31), $\int_{\mathbb{R}} |f_k - f| \to 0$. Given $\epsilon > 0$, choose N such that $\int_{\mathbb{R}} |f_k - f| < \epsilon$ for all $k \ge N$. Then

$$|F_k(x) - F(x)| = \left| \int_0^x (f_k - f) \right| \le \int_0^x |f_k - f| \le \int_{\mathbb{R}} |f_k - f| < \epsilon$$

for all $x \in \mathbb{R}$ and $k \ge N$, hence $F_k \to F$ uniformly on $\mathbb{R}$. □

Exercise (34)**.** Let $f_k = C_{[k,k+1/k)}$. Then $f_k \to 0$ on $\mathbb{R}$ and $\int_{\mathbb{R}} f_k \to 0$, but $\{f_k\}$ is not dominated by an integrable function on $\mathbb{R}$.

Proof. It is clear that $f_k \to 0$ and $\int_{\mathbb{R}} f_k = 1/k \to 0$. Suppose $g : \mathbb{R} \to \mathbb{R}$ is integrable and $f_k \leq g$ for all $k \geq 1$. Then

$$\frac{1}{k} = \int_k^{k+1/k} f_k \leq \int_k^{k+1/k} g$$

for all $k \geq 1$, so

$$\sum_{k=1}^{n} \frac{1}{k} \leq \sum_{k=1}^{n} \int_k^{k+1/k} g \leq \int_{\mathbb{R}} g$$

for all $n \geq 1$. But then $\infty = \sum_{k=1}^{\infty} 1/k \leq \int_{\mathbb{R}} g$, so g is not integrable after all. □

Remark. In this example,

$$\lim_{k\to\infty} \int_{\mathbb{R}} f_k = 0 = \int_{\mathbb{R}} \lim_{k\to\infty} f_k$$

This shows that domination is not a necessary condition for interchanging limit and integral.

Exercise (40 (Fatou))**.** Let $f, f_k : I \to \mathbb{R}$. Suppose $f_k \geq 0$ and f_k is integrable for all $k \geq 1$ with $f_k \to f$ and $\liminf_{k\to\infty} \int_I f_k < \infty$. Then f is integrable and

$$\int_I f \leq \liminf_{k\to\infty} \int_I f_k$$

Proof. Let $g_k = \inf_{m\geq k} f_m$. Note g_k is well defined since $f_m \geq 0$ for all $m \geq 1$. Also $g_k \uparrow f$ (in other words, $f = \liminf_{k\to\infty} f_k$).

We claim g_k is integrable for all $k \geq 1$. Indeed, let $g_{k,m} = f_k \wedge \cdots \wedge f_{k+m}$. Then $g_{k,m}$ is integrable (Proposition 14) and $\int_I g_{k,m} \geq 0$ for all $m \geq 1$. Also $g_{k,m} \downarrow g_k$, hence $\int_I g_{k,m} \downarrow$, so by the monotone convergence theorem (Theorem 1), g_k is integrable as claimed.

We have $\int_I g_k \uparrow$. We claim that $\{\int_I g_k\}$ is bounded above. Indeed, for any $1 \leq k \leq n$ we have $g_k = \inf_{m\geq k} f_m \leq f_n$, so $\int_I g_k \leq \int_I f_n$, and hence

$$\int_I g_k \leq \inf_{n\geq k} \int_I f_n \leq \sup_{k\geq 1} \inf_{n\geq k} \int_I f_n = \liminf_{k\to\infty} \int_I f_k < \infty$$

It follows from the monotone convergence theorem that f is integrable and

$$\int_I f = \lim_{k\to\infty} \int_I g_k \leq \liminf_{k\to\infty} \int_I f_k$$ □

Remark. We see from the proof that the conclusion can be stated as

$$\int_I \liminf_{k\to\infty} f_k \le \liminf_{k\to\infty} \int_I f_k$$

so this result allows interchange of limit inferior and integral (with inequality) under certain conditions. The proof is essentially one half of the proof of the dominated convergence theorem (Theorem 3).

Exercise (46)**.** If $f : \mathbb{R} \to \mathbb{R}$ is absolutely integrable and uniformly continuous, then $\lim_{|x|\to\infty} f(x) = 0$.

Proof. If $\lim_{|x|\to\infty} f(x) \ne 0$, then $\lim_{x\to\infty} f(x) \ne 0$ or $\lim_{x\to-\infty} f(x) \ne 0$. Suppose $\lim_{x\to\infty} f(x) \ne 0$ (the other case is similar). Then there is $\epsilon > 0$ and a sequence $x_k \to \infty$ with $x_{k+1} - x_k \ge 1$ such that $|f(x_k)| \ge \epsilon$ for all $k \ge 1$. By the uniform continuity of f, there is $0 < \delta < 1/2$ such that $|f(x)| \ge \epsilon/2$ whenever $|x - x_k| < \delta$. Now

$$\int_{x_k-\delta}^{x_k+\delta} |f| \ge \int_{x_k-\delta}^{x_k+\delta} \frac{\epsilon}{2} = \epsilon\delta$$

But since the intervals $[x_k - \delta, x_k + \delta]$ are pairwise disjoint, this means

$$\int_{\mathbb{R}} |f| \ge \sum_{k=1}^{\infty} \int_{x_k-\delta}^{x_k+\delta} |f| \ge \sum_{k=1}^{\infty} \epsilon\delta = \infty$$

by additivity, contradicting that f is absolutely integrable. □

Remark. This result does not hold for continuous functions. Consider $f : \mathbb{R} \to \mathbb{R}$ a nonnegative, continuous bump function with bumps of height 1 and area $1/k^2$ at all integers $\pm k \ne 0$, and zero in between. Then f is absolutely integrable since $\sum 2/k^2 < \infty$, but $\lim_{|x|\to\infty} f(x) \ne 0$ since $f(\pm k) = 1$ for all integers $k \ne 0$. Note f is not uniformly continuous since the bumps get arbitrarily steep as $|x| \to \infty$.

Chapter 16

Exercise (10)**.** Let $f(x) = \sin(x)/x$ on $I = [0, \infty)$. Then f is integrable, but f^+ is not integrable (Example 14.10). Let $E_n = [2n\pi, 2n\pi + \pi]$ and $E = \bigcup_{n=0}^{\infty} E_n$. Then E is measurable (Proposition 4), but f is not integrable over E since $f C_E = f^+$.

Remark. This example shows that the class of integrable functions is not closed under restriction to measurable subsets.

Exercise (13)**.** Let E be a nonmeasurable subset of $\mathbb{R}$. By definition, there exists a compact interval $I \subseteq \mathbb{R}$ such that $E \cap I$ is not integrable. Let $f = C_{E\cap I} - C_{E^c \cap I}$. Then $|f| = C_I$ is integrable, but f is not integrable, because $C_{E\cap I} = (f + |f|)/2$ is not integrable.

Remark. This example shows that absolute integrability of a function f is a stronger property than integrability of $|f|$.

Chapter 17

Remark. Fubini's theorem for double integrals (Theorem 12) is a continuous version of the corresponding discrete result for double series (Exercise 4.16), with absolute integrability replacing absolute convergence.

Remark. By "truncating", we can extend results about functions with bounded domains or ranges to functions with unbounded domains or ranges through limit processes. For example, for f nonnegative and integrable, we might let $f_n = f \wedge nC_{[-n,n]}$ be a "truncated" version of f with $f_n \uparrow f$, and then extend an integration result for f_n on $[-n, n]$ to a similar result for f through one of the convergence theorems for integrals. This technique is used in the proofs of the Fubini theorems and many other results.

Example (14)**.** Let $f(x, y) = e^{-xy} - 2e^{-2xy}$. Then for $y > 0$,

$$g(y) = \int_1^\infty f(x, y)\,dx = \left[\frac{-e^{-xy}}{y} + \frac{e^{-2xy}}{y}\right]_{x=1}^{\infty} = \frac{e^{-y} - e^{-2y}}{y}$$

We claim $\int_0^\infty g$ exists. First, $\int_c^1 g$ exists for all $0 < c \le 1$ by continuity of g. Also $g(y) \to 1$ as $y \to 0$ (by L'Hospital), so g is bounded on $(0, 1]$ and $\int_0^1 g$ exists by comparison. Note $g(y) > 0$ for $y > 0$ since e^z is strictly increasing, so $\int_0^1 g > 0$. Integration by parts yields

$$\int_1^\infty g = \int_1^\infty \frac{1}{y^2}\left(\frac{e^{-2y}}{2} - e^{-y}\right) dy + C$$

which exists by comparison with $\int_1^\infty 1/y^2\,dy$.

Now for $x \ge 1$,

$$h(x) = \int_0^1 f(x, y)\,dy = \frac{e^{-2x} - e^{-x}}{x}$$

Observe that $\int_1^c h$ exists for all $c \geq 1$ by continuity of h, and $|h| = g$ on $[1,\infty)$, so $\int_1^\infty h$ exists by comparison. Since $h(x) < 0$ for $x \geq 1$, $\int_1^\infty h < 0$. By the above,

$$-\infty < \int_1^\infty \int_0^1 f(x,y)\,dy\,dx < 0 < \int_0^1 \int_1^\infty f(x,y)\,dx\,dy < \infty$$

so these two iterated integrals exist but are not equal.

Exercise (6)**.** $\int_{\mathbb{R}^2} C_{[0,1]} = 0$.

Proof. We prove something stronger. Let $f(x,y) = (1/x)C_{(0,1]}(x)$. We claim that $\int_{\mathbb{R}^2} f = 0$, from which the result follows by comparison (setting $f(0,0) = 1$). We appeal directly to the definition of the integral.

Let $\epsilon > 0$. For each $n \geq 0$, let S_n be an open rectangle (strip) of area $\epsilon/4^{n+1}$ surrounding the interval $I_n = (1/2^{n+1}, 1/2^n]$. Define a gauge γ on $\mathbb{R}^2$ as follows: if $x \in I_n$, then $\gamma(x) \subseteq S_n$, otherwise $\gamma(x)$ is arbitrary. Let $\mathscr{D}$ be a γ-fine tagged division of $\mathbb{R}^2$. Let $\mathscr{D}_n = \{(x,I) \in \mathscr{D} \mid x \in I_n\}$. Then

$$S(f,\mathscr{D}_n) = \sum_{(x,I)\in\mathscr{D}_n} \frac{1}{x} v(I) < 2^{n+1} v(S_n) = \frac{\epsilon}{2^{n+1}}$$

Therefore

$$S(f,\mathscr{D}) = \sum_{n=0}^{\infty} S(f,\mathscr{D}_n) \leq \sum_{n=0}^{\infty} \frac{\epsilon}{2^{n+1}} = \epsilon \qquad \square$$

Remark. This proof shows that, although $\int_{[0,1]} 1/x = \infty$, $\int_{[0,1]^2} 1/x = 0$—that is, although there is infinite *area* under the curve $1/x$ over the unit interval in $\mathbb{R}$, there is no *volume* under the curve over the unit square in $\mathbb{R}^2$.

Exercise (9 (Fubini))**.** Let $\alpha, \beta : [a,b] \to \mathbb{R}$ be continuous with $\alpha \leq \beta$ and

$$E = [a,b] \times [\alpha,\beta] = \{(x,y) \mid a \leq x \leq b,\ \alpha(x) \leq y \leq \beta(x)\}$$

If $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous, then f is integrable over E and

$$\int_E f = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x,y)\,dy\,dx$$

Proof. Let $m = \min\alpha$ and $M = \max\beta$. Then $m \leq \alpha \leq \beta \leq M$, so $E \subseteq I = [a,b] \times [m,M]$. Note E is compact in $\mathbb{R}^2$ and f is continuous on E, so fC_E is bounded on I and is also the limit of a sequence of step functions on I (Lemma 9). By

Fubini's theorem for bounded functions on bounded intervals (Theorem 11), fC_E is integrable over I, so f is integrable over E, and

$$\int_E f = \int_a^b \int_m^M f C_E \, dy \, dx$$

But for fixed $x \in [a,b]$, $\int_m^M f C_E \, dy = \int_{\alpha(x)}^{\beta(x)} f \, dy$ by definition of E, so

$$\int_E f = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x,y) \, dy \, dx \qquad \square$$

Exercise (13).

$$\int_{[0,1]^2} \frac{1}{x+y} = 2\log 2 \qquad \int_{[0,1]^2} \frac{1}{x^2+y^2} = \infty$$

Proof. Let $f(x,y) = 1/(x+y)$. Let $E_n = [1/n, 1]^2$, so $E_n \subseteq E_{n+1}$ and $\bigcup_{n=1}^\infty E_n = (0,1]^2$. Then f is continuous and integrable over E_n, and by Fubini,

$$\begin{aligned}
\int_{E_n} f &= \int_{1/n}^1 \int_{1/n}^1 f(x,y) \, dx \, dy \\
&= \int_{1/n}^1 \left[\log(y+1) - \log(y+1/n)\right] dy \\
&= 2\log 2 + \frac{2\log 2}{n} - \frac{2\log n}{n} - 2\left(\frac{n+1}{n}\right)\log\left(\frac{n+1}{n}\right) \\
&\to 2\log 2 \quad \text{as } n \to \infty
\end{aligned}$$

Since $f > 0$ on $(0,1]^2$, monotone convergence implies f is integrable over $(0,1]^2$, hence also over $[0,1]^2$ (see Exercise 6), and $\int_{[0,1]^2} f = 2\log 2$.

Let $g(x,y) = 1/(x^2+y^2)$. For $x, y > 0$, $0 < x^2 + y^2 < x^2 + 2xy + y^2 = (x+y)^2$, so $h(x,y) = 1/(x+y)^2 < g(x,y)$. Now

$$\begin{aligned}
\int_{E_n} h &= \int_{1/n}^1 \int_{1/n}^1 h(x,y) \, dx \, dy \\
&= \int_{1/n}^1 \left(\frac{1}{y+1/n} - \frac{1}{y+1}\right) dy \\
&= 2\log\left(\frac{n+1}{n}\right) - \log\left(\frac{2}{n}\right) - \log 2 \\
&\to \infty \quad \text{as } n \to \infty
\end{aligned}$$

By comparison, $\int_{E_n} g \to \infty$ as $n \to \infty$, so g is not integrable over $[0,1]^2$. $\square$

Exercise (16 (Generalized dominated convergence theorem))**.** Let $f_k, f, g : \mathbb{R}^p \to \mathbb{R}$. Suppose f_k, g are integrable with $|f_k| \le g$ almost everywhere for each $k \ge 1$, and $f_k \to f$ pointwise almost everywhere. Then f is integrable and

$$\int_{\mathbb{R}^p} f = \lim_{k\to\infty} \int_{\mathbb{R}^p} f_k$$

Proof. Let $E_0 = \{x \in \mathbb{R}^p \mid \lim_{k\to\infty} f_k(x) \neq f(x)\}$ and $E_k = \{x \in \mathbb{R}^p \mid |f_k(x)| > g(x)\}$, so E_k is null for all $k \ge 0$ and $E = \bigcup_{k=0}^{\infty} E_k$ is also null. Define $\overline{f_k} = f_k C_{E^c}$, $\overline{f} = f C_{E^c}$, and $\overline{g} = g C_{E^c}$. Since $\overline{f_k} = f_k$ and $\overline{g} = g$ almost everywhere, $\overline{f_k}$ and $\overline{g}$ are integrable and $\int_{\mathbb{R}^p} \overline{f_k} = \int_{\mathbb{R}^p} f_k$ (Lemma 9). Also $|\overline{f_k}| \le \overline{g}$ and $\overline{f_k} \to \overline{f}$. By the dominated convergence theorem, $\overline{f}$ is integrable and

$$\int_{\mathbb{R}^p} \overline{f} = \lim_{k\to\infty} \int_{\mathbb{R}^p} \overline{f_k} = \lim_{k\to\infty} \int_{\mathbb{R}^p} f_k$$

But $f = \overline{f}$ almost everywhere, so f is integrable and $\int_{\mathbb{R}^p} f = \int_{\mathbb{R}^p} \overline{f}$. □

Exercise (18)**.** Let $f : G \to \mathbb{R}$ and $g : H \to \mathbb{R}$ be absolutely integrable limits of step functions and $h(x, y) = f(x)g(y)$. Then h is absolutely integrable over $I = G \times H$ and $\int_I h = \int_G f \cdot \int_H g$.

Proof. Note h is the limit of step functions on $\mathbb{R}^2$ and

$$\begin{aligned}
\int_H \int_G |h(x, y)|\, dx\, dy &= \int_H \int_G |f(x)||g(y)|\, dx\, dy \\
&= \int_H |g(y)| \int_G |f(x)|\, dx\, dy \\
&= \int_G |f(x)|\, dx \int_H |g(y)|\, dy
\end{aligned}$$

which exists by assumption. By Fubini (Theorem 12), h is absolutely integrable over I, and by a similar computation, $\int_I h = \int_G f \cdot \int_H g$. □

Exercise (24)**.** $I = \int_0^{\infty} e^{-x^2}\, dx = \sqrt{\pi}/2$.

Proof. Note I exists by comparison with e^{-x} and $I \geq 0$ since $e^{-x^2} \geq 0$. Now

$$\begin{aligned}
I^2 &= \int_0^\infty e^{-x^2}\,dx \int_0^\infty e^{-y^2}\,dy \\
&= \int_0^\infty \int_0^\infty e^{-(x^2+y^2)}\,dx\,dy && \text{by Exercise 18} \\
&= \int_0^{\pi/2} \int_0^\infty e^{-r^2}\,r\,dr\,d\theta && \text{by polar substitution} \\
&= \int_0^{\pi/2} \frac{1}{2}\,d\theta = \frac{\pi}{4}
\end{aligned}$$

It follows that $I = \sqrt{\pi}/2$. □

Chapter 18

Remark. Convolution is commutative (Exercise 1) and associative (Exercise 2), and if φ_n is a delta sequence, then it is an "approximate identity" element since $f * \varphi_n \approx f$ under appropriate conditions (Theorem 7).

Remark. In Example 2, note

$$\varphi_n(t) = \frac{1}{\pi}\frac{1/n}{t^2 + 1/n^2} = \frac{n}{\pi[(nt)^2 + 1]}$$

so if $\varphi(t) = 1/[\pi(t^2 + 1)]$, then $\varphi_n(t) = n\varphi(nt)$.

In Example 3, note

$$\varphi_n(t) = \sqrt{n/\pi}\, e^{-nt^2} = \frac{\sqrt{n}}{\sqrt{\pi}} e^{-(\sqrt{n}\,t)^2}$$

so if $\varphi(t) = e^{-t^2}/\sqrt{\pi}$, then $\varphi_n(t) = \sqrt{n}\,\varphi(\sqrt{n}\,t)$.

Therefore both of these examples are instances of vertical stretching and horizontal shrinking of a base function, like Examples 4 and 5.

Exercise (3)**.** Let $f, g : \mathbb{R} \to \mathbb{R}$. Suppose f is bounded and continuous and g is continuous and absolutely integrable. Then $f * g$ exists and is bounded and continuous, with $\sup|f * g| \leq \sup|f| \int_{-\infty}^{\infty} |g|$.

Proof. Let $h(x, y) = f(x - y)g(y)$. For $x \in \mathbb{R}$, $\int_a^b h(x, y)\,dy$ exists for all $a, b \in \mathbb{R}$ by continuity of f and g. Let $M = \sup|f|$. Then $|h(x, y)| \leq M|g(y)|$ for all $y \in \mathbb{R}$ and

$M|g|$ is integrable, so $f * g(x) = \int_{-\infty}^{\infty} h(x,y)\,dy$ exists by comparison. Similarly $\int_{-\infty}^{\infty} |h(x,y)|\,dy$ exists, so $|f * g(x)| \le M \int_{-\infty}^{\infty} |g|$. Since x was arbitrary, it follows that $\sup|f * g| \le M \int_{-\infty}^{\infty} |g|$. Finally, since h is also continuous in x for each $y \in \mathbb{R}$, it follows that $f * g$ is continuous. □

Remark. Contrary to what the book asks us to prove, $f * g$ is not necessarily integrable. For example, if $f(x) = 1$ and $g(y) = e^{-y^2}$, then $f * g(x) = \sqrt{\pi}$, which is not integrable.

Chapter 22

Exercise (11)**.** A normed linear space $(X, \|\cdot\|)$ is complete if and only if, when $\{x_k\} \subseteq X$ and $\sum \|x_k\|$ converges, then $\sum x_k$ converges.[2]

Proof. Suppose X is complete and $\{x_k\} \subseteq X$ with $\sum \|x_k\| < \infty$. By the Cauchy criterion for series in $\mathbb{R}$ (Proposition 4.5), for any $\epsilon > 0$ there exists N such that for all $n \ge m \ge N$,

$$\left\| \sum_{k=m}^{n} x_k \right\| \le \sum_{k=m}^{n} \|x_k\| < \epsilon$$

Therefore $\sum x_k$ is Cauchy in X, and hence convergent by completeness of X.

Conversely, suppose every absolutely convergent series in X converges and $\{x_k\} \subseteq X$ is Cauchy. We claim $\{x_k\}$ has a convergent subsequence $\{x_{k_j}\}$. Indeed, choose x_{k_j} such that $\|x_{k_j} - x_{k_{j-1}}\| < 2^{-j+1}$ for $j > 1$. Set $s_1 = x_{k_1}$ and $s_j = x_{k_j} - x_{k_{j-1}}$ for $j > 1$. Then

$$\sum_{j=1}^{\infty} \|s_j\| \le \|x_{k_1}\| + \sum_{j=2}^{\infty} 2^{-j+1} = \|x_{k_j}\| + 1$$

so $\sum s_j$ is absolutely convergent and hence convergent. But $\sum_{i=1}^{j} s_i = x_{k_j}$, so $\{x_{k_j}\}$ converges. It follows that $\{x_k\}$ also converges, to the same value, since $\{x_k\}$ is Cauchy. Therefore X is complete. □

Remark. This result helps us better understand the proof of the Riesz-Fischer theorem on the completeness of $\mathscr{L}^1$ (Theorem 7). In light of this result, to prove that $\mathscr{L}^1$ is complete it suffices to show that every absolutely convergent series of functions in $\mathscr{L}^1$ converges in $\mathscr{L}^1$.

[2] See also Proposition 6.5 in [3].

In the proof in the book, the series is $f = \sum_{k=1}^{\infty}(f_{n_{k+1}} - f_{n_k})$. We see that $f \in \mathscr{L}^1$, and that the partial sums converge to f under $\|\cdot\|_1$, by comparison with $g = \sum_{k=1}^{\infty}|f_{n_{k+1}} - f_{n_k}|$ and an appeal to the dominated convergence theorem. We see that $g \in \mathscr{L}^1$ by an appeal to the monotone convergence theorem.

Chapter 26

Remark. The Arzela-Ascoli theorem (Theorem 11) shows that any bounded, equicontinuous family of (real-valued) functions on a compact metric space is totally bounded. Indeed, if $\mathscr{F}$ is such a family, then $\overline{\mathscr{F}}$ is closed, bounded, and equicontinuous (Exercise 2), hence compact by Arzela-Ascoli, and hence totally bounded (Theorem 25.6).

Exercise (6)**.** The integral operator $T : \mathscr{C}[a,b] \to \mathscr{C}[a,b]$ given by $T(f)(x) = \int_a^x f$ is compact.

Proof. Note that $T(f) \in \mathscr{C}[a,b]$ for $f \in \mathscr{C}[a,b]$ by integrability of continuous functions and continuity of indefinite integrals. Also T is linear by linearity of the integral, and T is continuous since $\|T(f)\|_\infty \le \|f\|_\infty(b-a)$.

Let $E \subseteq \mathscr{C}[a,b]$ be bounded. We know $T(E)$ is bounded by continuity of T. We claim $T(E)$ is equicontinuous. It then follows that $\overline{T(E)}$ is closed, bounded, and equicontinuous (Exercise 2), so $\overline{T(E)}$ is compact by Arzela-Ascoli, and T is compact.

Let $\epsilon > 0$. Let $\|f\|_\infty \le M$ for all $f \in E$ and set $\delta = \epsilon/M$. Then for all $f \in E$ and $a \le x \le y \le b$ with $y - x < \delta$,

$$|T(f)(y) - T(f)(x)| = \left|\int_x^y f\right| \le \int_x^y |f| \le M(y-x) < M\delta = \epsilon$$

Therefore $T(E)$ is equicontinuous as claimed. □

Exercise (12)**.** Let X be a metric space with $D \subseteq X$ dense and suppose $f_k : X \to \mathbb{R}$ is continuous. If $\{f_k\}$ converges uniformly on D, $\{f_k\}$ converges uniformly on X.

Proof. We first claim $\{f_i\}$ converges pointwise on X. Let $x \in X$. By density of D in X, $x = \lim_j x_j$ for some $\{x_j\} \subseteq D$. Let $a_{ij} = f_i(x_j)$. Then $\lim_j a_{ij} = f_i(x)$ exists for all i by continuity of the f_i on X. Also $\lim_i a_{ij}$ exists uniformly for all j by uniform convergence of the f_i on D. Therefore $\lim_i f_i(x)$ exists as claimed, and $\lim_i f_i(x) = \lim_j \lim_i f_i(x_j)$ (Proposition 11.5).

Let $f(x) = \lim_i f_i(x)$ for all $x \in X$. We claim $f_i \to f$ uniformly on X. Let $\epsilon > 0$. Choose N such that $|f(x) - f_i(x)| < \epsilon$ for all $i \geq N$ and $x \in D$. For $x \in X$, write $x = \lim_j x_j$ with $\{x_j\} \subseteq D$. Then

$$|f(x_j) - f_i(x_j)| < \epsilon$$

for all $i \geq N$ and $j \geq 1$. By the above, $f(x) = \lim_j f(x_j)$, so letting $j \to \infty$ yields

$$|f(x) - f_i(x)| \leq \epsilon$$

for all $i \geq N$. Therefore $f_i \to f$ uniformly as claimed. □

Remark. The limit f is continuous (Proposition 2).

Remark. It is an immediate corollary that if $\sum_{k=1}^{\infty} f_k$ converges uniformly on D, then $\sum_{k=1}^{\infty} f_k$ converges uniformly on X.

Chapter 28

Remark. In the proof of Lemma 3, in order to construct f intuitively we want to "normalize" g at s and t by dividing by $g(s)$ and $g(t)$, and then scale the result by a and b to yield those values at s and t, respectively. Since $g(s) \neq g(t)$, we can divide by $g(s) - g(t) \neq 0$, so

$$f(x) = \frac{h(x)}{g(s) - g(t)}$$

To ensure that $f(s) = a$, we must have $h(s) = a[g(s) - g(t)]$, which suggests that $h(x)$ should contain the term $a[g(x) - g(t)]$. Similarly to ensure that $f(t) = b$, we must have $h(t) = b[g(s) - g(t)]$, which suggests that $h(x)$ should contain the term $b[g(s) - g(x)]$. Adding these terms together, we get

$$f(x) = \frac{a[g(x) - g(t)] - b[g(x) - g(s)]}{g(s) - g(t)}$$

which satisfies $f(s) = a$ and $f(t) = b$ as desired.

Chapter 29

Remark. We see that it is best to view the derivative of a function $f : X \to Y$ between normed linear spaces X and Y as a continuous linear transformation from X to Y (Definition 1). The following are special cases:

- If $X = Y = \mathbb{R}$, such transformations are uniquely represented by numbers in $\mathbb{R}$ under the scalar product, so we identify derivatives with numbers (Definition 9.1).
- If $X = \mathbb{R}^n$ and $Y = \mathbb{R}$, such transformations are uniquely represented by (gradient) vectors in $\mathbb{R}^n$ under the dot product, so we identify derivatives with vectors (Definitions 10.4 and 10.5 and Propositions 10.3 and 10.6).
- If $X = \mathbb{R}$ and $Y = \mathbb{R}^m$, such transformations are uniquely represented by (tangent) vectors in $\mathbb{R}^m$ under the scalar product, so we identify derivatives with vectors (Definition 10.10). This is also true if Y is an arbitrary normed linear space (Example 5).
- If $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, such transformations are uniquely represented by $m \times n$ (Jacobian) matrices over $\mathbb{R}$ under the matrix product, so we can identify derivatives with matrices.

Dieudonné considers such definitions "slavish subservience to the shibboleth of numerical interpretation" ([2], p. 141).

Remark. We provide additional examples of derivatives from matrix calculus. Let $\mathbb{R}^{n\times n}$ be the space of $n \times n$ matrices over $\mathbb{R}$ identified with $L(\mathbb{R}^n, \mathbb{R}^n)$ under the standard basis and the operator norm induced by $\|\cdot\|_1$ on $\mathbb{R}^n$. For $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, let A_{ij} denote the (i, j)-th entry of $\boldsymbol{A}$. Observe that

$$|A_{ij}| \le \sum_{k=1}^{n} |A_{kj}| = \|\boldsymbol{A}\boldsymbol{e_j}\| \le \|\boldsymbol{A}\|$$

Conversely, if $\boldsymbol{x} \in \mathbb{R}^n$ with $\|\boldsymbol{x}\| \le 1$, then

$$\|\boldsymbol{A}\boldsymbol{x}\| = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} A_{ij} x_i \right| \le \sum_{i,j} |A_{ij}|$$

and hence $\|\boldsymbol{A}\| \le \sum_{i,j} |A_{ij}|$. Let $\boldsymbol{E_{ij}} \in \mathbb{R}^{n\times n}$ denote the (i, j)-th standard basis matrix with (i, j)-th entry equal to 1 and all other entries equal to 0.

- The derivative of a scalar-valued function of matrices $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ at $\boldsymbol{A}$ is naturally identified with the $n \times n$ (gradient) matrix

$$\nabla f(\boldsymbol{A}) = [df(\boldsymbol{A})(\boldsymbol{E_{ij}})]$$

Indeed, for $\boldsymbol{H} \in \mathbb{R}^{n\times n}$,

$$\begin{aligned} df(\boldsymbol{A})(\boldsymbol{H}) &= df(\boldsymbol{A})(\sum_{i,j} \boldsymbol{E_{ij}} H_{ij}) \\ &= \sum_{i,j} df(\boldsymbol{A})(\boldsymbol{E_{ij}}) H_{ij} \\ &= \sum_{i,j} \nabla f(\boldsymbol{A})_{ij} H_{ij} \\ &= \nabla f(\boldsymbol{A}) : \boldsymbol{H} \end{aligned}$$

where $\nabla f(\boldsymbol{A}) : \boldsymbol{H}$ denotes the Frobenius product of $\nabla f(\boldsymbol{A})$ and $\boldsymbol{H}$. Note that $df(\boldsymbol{A})(\boldsymbol{E_{ij}}) = d_{ij} f(\boldsymbol{A})$, the (i, j)-th partial derivative of f at $\boldsymbol{A}$, by a generalization of Proposition 16. Therefore we may write

$$\frac{d}{d\boldsymbol{A}} f(\boldsymbol{A}) = df(\boldsymbol{A}) = \nabla f(\boldsymbol{A}) = \left[\frac{\partial}{\partial A_{ij}} f(\boldsymbol{A}) \right]$$

For example, the trace function $\mathrm{Tr} : \mathbb{R}^{n\times n} \to \mathbb{R}$ is linear, hence $d\,\mathrm{Tr}(\boldsymbol{A}) = \mathrm{Tr}$. Since $\mathrm{Tr}(\boldsymbol{E_{ij}}) = \delta_{ij}$, it follows that $\nabla\,\mathrm{Tr}(\boldsymbol{A}) = \boldsymbol{I}$.

- The derivative of a matrix-valued function of scalars $\boldsymbol{F} : \mathbb{R} \to \mathbb{R}^{n\times n}$ at $x \in \mathbb{R}$ is naturally identified with the (tangent) matrix

$$\boldsymbol{F}'(x) = [F'_{ij}(x)]$$

where $F_{ij} : \mathbb{R} \to \mathbb{R}$ is the (i, j)-th component function of $\boldsymbol{F}$. Indeed, it follows from the inequalities above that $\boldsymbol{T} : \mathbb{R} \to \mathbb{R}^{n\times n}$ is the derivative of $\boldsymbol{F}$ at x if and only if each T_{ij} is the derivative of F_{ij} at x. So for $h \in \mathbb{R}$,

$$d\boldsymbol{F}(x)(h) = d\boldsymbol{F}(x)(1)h = [dF_{ij}(x)(1)]h = [F'_{ij}(x)]h = \boldsymbol{F}'(x)h$$

Therefore we may write

$$\frac{d}{dx}\boldsymbol{F} = \left[\frac{d}{dx} F_{ij} \right]$$

For example, taking $x = F_{ij}$ we have $d\boldsymbol{F}/dF_{ij} = \boldsymbol{E_{ij}}$.

Remark. We see that many results about differentiable functions on $\mathbb{R}$ can be extended to differentiable functions $f : X \to Y$ on more general normed linear spaces X by constructing differentiable curves $\gamma : [a,b] \to X$ and applying the original results (via the chain rule) to the composites $f \circ \gamma$. For example, see the proof of the mean value theorem on $\mathbb{R}^n$ (Theorem 10.11).

Remark. A function $f : X \to Y$ between the normed linear spaces X and Y is differentiable at x_0 if and only if there is $T \in L(X,Y)$ such that the function

$$\delta(x) = \begin{cases} \|f(x) - f(x_0) - T(x - x_0)\| / \|x - x_0\| & \text{if } x \neq x_0 \\ 0 & \text{if } x = x_0 \end{cases}$$

is continuous at x_0, in which case $T = df(x_0)$. We call δ the *differential quotient (d.q.)* for f at x_0.

We provide an alternative proof of the chain rule using the previous remark:

Theorem (Chain rule)**.** *Let X, Y, and Z be normed linear spaces. Let $f : X \to Y$, $g : Y \to Z$, and $h = g \circ f : X \to Z$. If f is differentiable at x_0 and g is differentiable at $y_0 = f(x_0)$, then h is differentiable at x_0 and $dh(x_0) = dg(y_0)df(x_0)$.*

Proof. Let $T = df(x_0)$, $U = dg(y_0)$, and $V = UT$. Let δ_f be the d.q. for f at x_0 and δ_g the d.q. for g at y_0. We claim

$$\delta_h(x) = \begin{cases} \|h(x) - h(x_0) - V(x - x_0)\| / \|x - x_0\| & \text{if } x \neq x_0 \\ 0 & \text{if } x = x_0 \end{cases}$$

is continuous at x_0, from which the theorem follows. Observe

$$\begin{aligned} h(x) - h(x_0) - V(x - x_0) = g(f(x)) - g(f(x_0)) - U(f(x) - f(x_0)) \\ + U(f(x) - f(x_0) - T(x - x_0)) \end{aligned}$$

For x close but not equal to x_0, $\|f(x) - f(x_0)\| / \|x - x_0\| \leq \|T\| + 1$, so

$$\delta_h(x) \leq (\|T\| + 1)\delta_g(f(x)) + \|U\|\delta_f(x)$$

Now $\delta_g \circ f$ and δ_f continuously vanish at x_0, so δ_h does also. □

Remark. If a function f is differentiable at a point x_0, then change in f *at* x_0 is well approximated by $df(x_0)$. If f is *continuously* differentiable at x_0, then change in f *between any two points sufficiently close to* x_0 is well approximated by $df(x_0)$ (Corollary 15).

Remark. The proof of the continuity of df in Proposition 17 is flawed because (assuming $df = d_1 f = d_2 f = 0$ outside of D) we have

$$\begin{aligned} df &: X \times Y \to L(X \times Y, Z) \\ d_1 f \circ I_1 &: X \to L(X, Z) \\ d_2 f \circ I_2 &: Y \to L(Y, Z) \end{aligned}$$

so the functional equation $df = d_1 f \circ I_1 + d_2 f \circ I_2$ makes no sense.

By Proposition 16,

$$df(x, y)(h, k) = d_1 f(x, y)(h) + d_2 f(x, y)(k)$$

for all $(x, y), (h, k) \in X \times Y$, so

$$df(x, y) = d_1 f(x, y) \circ P_1 + d_2 f(x, y) \circ P_2$$

for all $(x, y) \in X \times Y$, where $P_1 \in L(X \times Y, X)$ and $P_2 \in L(X \times Y, Y)$ are just the projections $P_1(h, k) = h$ and $P_2(h, k) = k$. Define $\mathscr{P}_1 : L(X, Z) \to L(X \times Y, Z)$ by $\mathscr{P}_1(\varphi) = \varphi \circ P_1$ and $\mathscr{P}_2 : L(Y, Z) \to L(X \times Y, Z)$ by $\mathscr{P}_2(\psi) = \psi \circ P_2$. Then

$$df = \mathscr{P}_1 \circ d_1 f + \mathscr{P}_2 \circ d_2 f \tag{1}$$

Now $\mathscr{P}_1$ is linear and

$$\|\mathscr{P}_1(\varphi)\| = \|\varphi \circ P_1\| \le \|\varphi\| \|P_1\| = \|\varphi\|$$

so $\mathscr{P}_1$ is also continuous, and similarly $\mathscr{P}_2$ is linear and continuous. It follows from (1) that df is continuous. For the converse, see Exercise 15 below.

Exercise (15)**.** If $f : D \subseteq X \times Y \to Z$ is continuously differentiable on D, then $d_1 f$ and $d_2 f$ exist and are continuous on D.

Proof. By Proposition 16, $d_1 f$ and $d_2 f$ exist and

$$df(x, y)(h, k) = d_1 f(x, y)(h) + d_2 f(x, y)(k)$$

In particular,

$$df(x, y)(h, 0) = d_1 f(x, y)(h) \qquad \text{and} \qquad df(x, y)(0, k) = d_2 f(x, y)(k)$$

Letting $I_1 \in L(X, X \times Y)$ and $I_2 \in L(Y, X \times Y)$ be the injections $I_1(h) = (h, 0)$ and $I_2(k) = (0, k)$, we have

$$df(x, y) \circ I_1 = d_1 f(x, y) \qquad \text{and} \qquad df(x, y) \circ I_2 = d_2 f(x, y)$$

Define $\mathscr{I}_1 : L(X \times Y, Z) \to L(X, Z)$ by $\mathscr{I}_1(\varphi) = \varphi \circ I_1$ and $\mathscr{I}_2 : L(X \times Y, Z) \to L(Y, Z)$ by $\mathscr{I}_2(\psi) = \psi \circ I_2$. Then $\mathscr{I}_1$ and $\mathscr{I}_2$ are linear and continuous and (assuming $df = d_1 f = d_2 f = 0$ outside of D)

$$\mathscr{I}_1 \circ df = d_1 f \qquad \text{and} \qquad \mathscr{I}_2 \circ df = d_2 f$$

so $d_1 f$ and $d_2 f$ are continuous. □

Chapter 30

Remark. In the proof of the inverse function theorem (Theorem 1), after the initial reductions have been made so that $x_0 = y_0 = 0$, $f : X \to X$, and $df(0) = I$, we view f as a local "perturbation of the identity" near the origin, $f = I + p$, so that a local inverse g of f satisfies $g = I - p \circ g$ and is therefore a fixed point of the transformation $\mathscr{T} : \varphi \mapsto I - p \circ \varphi$ on an appropriate function space.

In seeking a function space in which we can apply the Banach fixed point theorem to $\mathscr{T}$, we are naturally led to consider closed subspaces of bounded continuous functions on X, since such spaces are complete. We know about the local behavior of p near the origin. More specifically, since f is continuously differentiable at the origin, for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$\|p(x_1) - p(x_2)\| \le \epsilon \|x_1 - x_2\| \qquad \text{for } \|x_1\|, \|x_2\| \le \delta$$

and, taking $x_2 = 0$, $\|p(x_1)\| \le \epsilon \|x_1\|$ for $\|x_1\| \le \delta$ (Corollary 29.15). Therefore we restrict our attention to subspaces of functions defined near the origin. If φ is such a function, then

$$\|\mathscr{T}(\varphi)(x)\| \le \|x\| + \|p(\varphi(x))\|$$

so if $\|x\| \le \delta/2$ and $\|\varphi(x)\| \le \delta$, then $\|\mathscr{T}(\varphi)(x)\| \le (1/2 + \epsilon)\delta$ by the above, and $\|\mathscr{T}(\varphi)(x)\| < \delta$ if we choose $\epsilon < 1/2$, say $\epsilon = 1/3$. This suggests the subspace

$$S = \{\varphi \in \mathscr{BC}(B(0, \delta/2), X) \mid \|\varphi\| \le \delta\}$$

By the above, S is closed under $\mathscr{T}$, and S is contracted by $\mathscr{T}$ since

$$\|\mathscr{T}(\varphi)(x) - \mathscr{T}(\psi)(x)\| = \|p(\varphi(x)) - p(\psi(x))\| \le \epsilon \|\varphi(x) - \psi(x)\|$$

Therefore $\mathscr{T}$ has a unique fixed point g, which is the desired local inverse of f. We see that g has nice properties: it is differentiable, and continuously so if f is, with derivative equal to the inverse of the derivative of f.

Remark. Recall that an indexed family $\mathscr{F} = \{S_x \mid x \in X\}$ of sets is equivalently represented by the indexing function $\pi : \coprod_{x\in X} S_x \to X$ where

$$\coprod_{x\in X} S_x = \bigcup_{x\in X} (\{x\} \times S_x) = \{(x,s) \mid x \in X \wedge s \in S_x\}$$

represents the disjoint union over $\mathscr{F}$ and $\pi(x,s) = x$. Indeed, for $x \in X$, S_x is obtained by projecting $\pi^{-1}(x)$. Moreover, if $S_x \neq \emptyset$ for all $x \in X$, then a choice function for $\mathscr{F}$ is obtained by projecting a right inverse of π.

In the proof of the implicit function theorem (Theorem 4), this idea is used. The function $F(x, y) = (x, f(x, y))$ is an indexing function, where $F^{-1}(x,0)$ is the set of all y such that $f(x, y) = 0$. The local inverse function for F is therefore a local choice function for the indexed family, and yields the implicit function φ such that $f(x,\varphi(x)) = 0$. Since the inverse function has the nice property of being continuously differentiable, so does φ.

References

[1] DePree, J. D. and C. W. Swartz. *Introduction to Real Analysis.* Wiley, 1988.

[2] Dieudonné, J. *Foundations of Modern Analysis.* Academic Press, 1960.

[3] Royden, H. L. *Real Analysis*, 3rd ed. Prentice Hall, 1988.