Notes on Riemann Integration

2/23/2020 version, truncated

Contents

1 Riemann Integration

Integration (here meaning *definite* integration, not antidifferentiation[1]) is always about "adding stuff up". The "stuff" may be not be a finite, or even countable, set of numbers; the "sum" may be something like amount of space between the graph of a positive function $f : [a, b] \to \mathbf{R}$ ("area under a curve"), or the total mass or (electric) charge of a solid object for which we know the "mass density" or "charge density" at each point. Generally, the "stuff" is described in some way by a real-valued or vector-valued function on some subset of $\mathbf{R}^n$. Regardless of the dimension of the domain, or whether the function is real-valued vs. vector-valued, or which theory of integration is used (there are several progessively more general theories), the core idea that *integration is about adding stuff up* is always there, even when its presence isn't obvious. But to rigorize the vague "adding stuff up", one starts first with the simplest theory of integration, the subject of this chapter.

These notes develop the theory of the Riemann integral of a real-valued (and, later, a vector-valued) function f on a closed, bounded interval $[a, b] \subseteq \mathbf{R}$. The approach we use

---

[1] In higher mathematics, "definite integration" is the default meaning of "integration", unless context makes it clear that antidifferentiation is meant.

is very intuitive, rigorizing the "limit of Riemann sums" idea that's presented in Calculus 1 and that is used throughout physics. This approach has an additional advantage: it provides the most natural generalization to integration of vector-valued functions. For integration of real-valued functions, however, it is not the most efficient approach. A more efficient approach that (non-obviously) turns out to be essentially equivalent is discussed in Section 1.4, as optional reading for the interested student.

Throughout this chapter, when we use notation of the form "$[a, b]$", it is understood that $a, b \in \mathbf{R}$ and that $a < b$.

The symbol "▲" is used in these notes to mark the end of a definition, remark, or example, and some of the longer exercises. When all that is being defined is notation specific to these notes, sometimes we label that definition as "Notation 6.x" to avoid giving the impression that this notation is standard among mathematicians.

1.0 Overview of these notes

In this chapter, first we will define *(Riemann-)integrable real-valued function* f on an interval $[a, b]$, and the integral $\int_a^b f(x)\,dx$ (Section 1.1). From the definitions, based on *Riemann sums*, it is not obvious which functions on $[a, b]$ are integrable (for example, it is not obvious that continuous functions are integrable). Rather than trying to apply the definitions directly in many examples, which would be quite time-consuming, we first establish several general properties (starting in Section 1.2), such as the fact that the set $\mathcal{R}([a, b])$ of integrable functions on $[a, b]$ is a vector space and that the map $\mathcal{R}([a, b]) \to \mathbf{R}$ given by $f \mapsto \int_a^b f(x)\,dx$ is linear. In Section 1.3 we show that step-functions (Definition 1.42) are integrable, and develop an integrability criterion (for any $f : [a, b] \to \mathbf{R}$) based on step-functions. We use this in Section 1.5 to establish that continuous functions on $[a, b]$ are integrable. (As mentioned earlier, the intervening Section 1.4, optional reading, relates the approach taken in Sections 1.1–1.3 to a different approach preferred by many mathematicians. This approach involves something called *upper* and *lower integrals*)

The methods presented in Section 1.3 rely on the concept of *upper* and *lower sums* introduced there, as does the step-function-related integrability-criterion mentioned above (and therefore, also, our proof that "continuous implies integrable" in Section 1.5). However, *every important result that we prove using upper and lower sums can be proven without them.* We have used them in Sections 1.3 and 1.5 for two reasons: (1) they are very helpful for developing a visual understanding of integration of real-valued functions, and (2) upper and lower *sums* provide a bridge connecting the definitions of "integrable" and "integral" in Section 1.1 to those based on upper and lower *integrals* in Section 1.4. The approach to integration discussed in Section 1.4 relies *critically* on upper and lower sums; there they are not simply a dispensable convenience.

Section 1.6 establishes an additivity property reflecting the principle that "integration is about adding stuff up": if $a < c < b$ and f is integrable on $[a, b]$, then f is integrable on the subintervals $[a,c]$ and $[c,b]$, and $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ (the "amount of stuff" between a and b is the amount of stuff between a and c plus the amount of stuff between c and b). This is used in Section 1.7 to prove several theorems that go by the name "The Fundamental Theorem of Calculus", one of which is then used in Section 1.8 to prove the validity of the change-of-variable formula learned in Calculus 1.

In Section 1.9 we generalize from integrating real-valued functions on $[a,b]$ to integrating vector-valued functions on $[a,b]$. In this detour from real-valued functions, the generality and strengths of the Riemann-sum approach to the Riemann integral—i.e. the approach taken in Sections 1.1–1.3, as opposed to the approach taken in Section 1.4—become more evident: for any complete normed vector space $(V, \|\ \|)$, we can define "integrable function $f : [a,b] \to V$" and $\int_a^b f(x)\,dx$ *exactly* the way we did for real-valued functions (modulo obvious notational changes). In particular this applies to *any* finite-dimensional vector space with *any* norm. When $V = \mathbf{R}^n$, this definition of the vector-valued integral agrees with the one taught in Calculus 3, but now we see the Calc-3 definition as a corollary of something much more general. Our new definition of the vector-valued integral makes no reference to a basis of V (which the Calc-3 definition does implicitly or explicitly). We do not need V to be $\mathbf{R}^n$, or even to be finite-dimensional. Several of our results for real-valued functions generalize, without any change in the proofs, to V-valued functions, e.g. linearity of the map $f \mapsto \int_a^b f(x)\,dx$. In addition, once we establish an integrability criterion (Proposition 1.84) whose "$V = \mathbf{R}$" case eliminates any need for upper and lower sums (and, therefore, any need for the step-function-related integrability-criterion we used for real-valued functions) we are able to prove generalized versions, for V-valued functions, of other earlier results, e.g. the "triangle inequality for integrals" (equations (1.30) and (1.72)) and the integrability of continuous functions.

1.1 Definitions and first examples

Definition 1.1 (Partitions) A *partition* P of a closed, bounded interval $[a,b]$ is a finite set $\{x_0, x_1, \dots, x_N\}$, where $a = x_0 < x_1 < \cdots < x_N = b$. (Thus the number of points in P is $N+1 \geq 2$.) ▲

Remark 1.2 This use of the word "partition" is special to intervals. The student may be used to "partition of a set S" meaning a disjoint collection of subsets of S whose union is S. In the setting of Definition 1.1, the interval $[a,b]$ is the union of the subintervals $[x_{j-1}, x_j]$, $1 \leq j \leq N$, but these subintervals are not disjoint (unless $N = 1$): for $0 < j < N$, the point x_j lies in two of these subintervals, as the right endpoint of one and the left endpoint of another. We *could* express $[a,b]$ as the disjoint union of $N-1$ half-open intervals and one closed interval, e.g. $[x_0, x_1) \cup [x_1, x_2), \cup \cdots \cup [x_{N-2}, x_{N-1}) \cup [x_{N-1}, x_N]$ or $[x_0, x_1] \cup (x_1, x_2] \cup \cdots \cup (x_{N-2}, x_{N-1}] \cup (x_{N-1}, x_N]$, but the choice of which interval should include a given x_j (other than x_0 and x_N) would be asymmetric and artificial. For purposes of integration, it turns out that 1-point overlaps are irrelevant (we will see why later), so we allow ourselves to speak of $[a, b]$ as being "partitioned" into the closed subintervals $\{[x_{j-1}, x_j]\}$ even though this is not quite consistent with the set-theoretic notion. Finally, since the set of these closed subintervals is completely determined by the set of their endpoints, and vice-versa, the exceptional meaning of "partition" in the context of intervals doesn't cause a problem once one gets used to it.

Notation 1.3 For each partition $P = \{x_0, x_1, \ldots, x_N\}$ of an interval $[a, b]$, we define $\Delta_j(P) = x_j - x_{j-1}$, $1 \leq j \leq N$. When a single partition P is understood from context, we write simply Δ_j rather than $\Delta_j(P)$. ▲

Observe that for any partition P of $[a, b]$, we have

$$\sum_j \Delta_j(P) = b - a. \tag{1.1}$$

Definition 1.4 (Pointed partitions and Riemann sums) Let $a, b \in \mathbf{R}$, with $a < b$,[2] and let $P = \{x_0, x_1, \ldots, x_N\}$ be a partition of $[a, b]$.

1. We define the *mesh* of P, denoted $\text{mesh}(P)$ in these notes, to be $\max\{\Delta_j : 1 \leq j \leq N\}$.
2. A *pointing* T of P is a set $T = \{t_1, \ldots t_N\}$ such that $t_j \in [x_{j-1}, x_j]$ for each $j \in \{1, \ldots, N\}$. We call the pair (P, T) a *pointed partition* (of $[a, b]$). We define the mesh of (P, T) to be the mesh of P.
3. Given $f : [a, b] \to \mathbf{R}$ and a pointing $T = \{t_1, \ldots t_N\}$ of the partition P, the *Riemann sum* for f corresponding to the pointed partition (P, T) is

$$S(f; P, T) = \sum_{j=1}^{N} f(t_j)\Delta_j \,. \tag{1.2}$$

4. Given $f : [a, b] \to \mathbf{R}$, we will write

$$\begin{aligned} \mathcal{S}(f; P) &= \{S(f; P, T) : T \text{ is a pointing of } P\} \\ &= \text{the set of all Riemann sums of } f \text{ associated with the partition } P. \end{aligned}$$

---

[2] "Let $a, b \in \mathbf{R}$, with $a < b$" is read "Let a and b *be* in $\mathbf{R}$, with $a < b$," or , less commonly, "Let a and b *be* elements of $\mathbf{R}$, with $a < b$." (i) It is common in mathematical writing to allow a comma to stand for the word "and" when separating the last two items in a list, as this often aids readability. (ii) The symbol "$\in$" is one of several mathematical symbols that, depending on the structure of the sentence in which it appears, may be read as a preposition ("in", for "$\in$") or as some form of the verb "to be" followed by a preposition. In the sentence "Let $a, b \in \mathbf{R}$ satisfy the equation $a + b = 0$," the element-symbol is read simply as "in".

▲

Note that every partition $\{x_0, \ldots, x_N\}$ has a pointing (in fact, uncountably many); e.g. we can take $t_j = x_j$ for $1 \leq j \leq N$ (the "right-endpoint pointing").

For a general pointing of a partition, with notation as in Definition 1.4, we think of the point t_j as a "sample point" within the interval $[x_{j-1}, x_j]$, providing a "sample value" of f on this interval.

Remark 1.5 Observe that any interval $[a, b]$ as above has partitions of arbitrarily small mesh: given $\delta > 0$, let N be any positive integer such that $\Delta := \frac{b-a}{N} < \delta$, let $x_j = a + j\Delta$ for $0 \leq j \leq N$, and let P be the partition $\{x_0, x_1, \ldots, x_N\}$; we then have $\text{mesh}(P) < \delta$. Hence there also always exist pointed partitions of $[a, b]$ arbitrarily small mesh.▲

Definition 1.6 (Integrability) A function $f : [a, b] \to \mathbf{R}$ is *Riemann integrable* if there is a real number A such that for every $\epsilon > 0$ there exists $\delta > 0$ such that if (P, T) is any pointed partition of $[a, b]$ of mesh less than δ, then $|S(f; P, T) - A| < \epsilon$. More generally, if f is a real-valued function whose domain includes $[a, b]$, we say that f *is Riemann integrable on* $[a, b]$ (or *over* $[a, b]$) if $f|_{[a,b]}$ is Riemann integrable. ▲

Notation 1.7 For $a, b \in \mathbf{R}$ with $a < b$, we will let $\mathcal{R}([a, b])$ denote the set of all real-valued functions on $[a, b]$ that are Riemann-integrable. ▲

With notation as in Definition 1.6, suppose that A, A' are two real numbers satisfying the condition required of A in the definition. Let $\epsilon > 0$, and $\delta > 0$ be such that for every pointed partition of $[a, b]$ of mesh less than δ, we have $|S(f; P, T) - A| < \epsilon$ and $|S(f; P, T) - A'| < \epsilon$. Let (P_1, T_1) be a pointed partition of $[a, b]$ of mesh $< \delta$; such (P, T) exists by Remark 1.5. Then $|A' - A| \leq |A' - S(f; P_1, T_1)| + |S(f; P_1, T_1) - A| < 2\epsilon$. Since this is true for all $\epsilon > 0$, it follows that $A' - A = 0$, hence that $A' = A$. Therefore *if f is integrable on $[a, b]$, then there is a* unique *number A satisfying the condition in Definition 1.6.* Thus we can make the following definition:

Definition 1.8 (the Riemann integral) Let $f : [a, b] \to \mathbf{R}$ be a Riemann-integrable function. We define the *Riemann integral of f* to be the unique real number A satisfying the condition given in Definition 1.6. This number is denoted

$$\int_a^b f(x)\, dx, \int_a^b f(t)\, dt, \text{ etc.}; \tag{1.3}$$

any letter not reserved with another meaning can be used in place of the "dummy" variable x, t, etc., in the sample notation above. Since the name of the dummy variable does not affect the value of the integral, in these notes we will also use the notation

$$\int_a^b f \tag{1.4}$$

in place of (1.3). More generally, if f is a real-valued function on a domain that includes $[a,b]$, and f is integrable on $[a,b]$, we use the same notation (1.3), (1.4) for the Riemann integral of $f|_{[a,b]}$, and refer to the value of this integral as the *Riemann integral of f over* $[a,b]$.

Any conclusion of the form $\int_a^b f =$ [specific number] implicitly means "f is integrable on $[a,b]$ and $\int_a^b f =$ [that number]," if the integrability of f has not already been stated explicitly.

Finally, we define the phrase "$\int_a^b f$ exists" (or "$\int_a^b f(x)\,dx$ exists", etc. for any dummy variable), to mean that f is integrable on $[a,b]$. ▲

Definitions 1.6 and 1.8 give precise meaning to the notion that a definite integral is a "limit of Riemann sums". It is tempting to write, suggestively, that the value of $\int_a^b f(x)\,dx$ is "$\lim_{\mathrm{mesh}(P)\to 0} S(f;P,T)$", but this limit-notation cannot be interpreted literally. The quantity $S(f;P,T)$ is not the value of a function of $\mathrm{mesh}(P)$, or even the value of a function of P. For every $\delta > 0$, there are infinitely many partitions of $[a,b]$ of mesh δ, and for every partition there are infinitely many pointings. Thus for every value of $\mathrm{mesh}(P)$, there can be (and usually are) *infinitely many* values of Riemann sums of f associated with partitions of this mesh. In the notation "$\lim_{x\to x_0} g(x)$" for the limit at x_0 of a function $g : U \setminus \{x_0\} \to \mathbf{R}$, where $U \subseteq \mathbf{R}$, for each x there is *one and only one* number $g(x)$. However, there are a few ways to write the integral of an integrable function as a *true* limit. The following exercise gives one of these; Theorem 1.32, later in these notes, gives another.

Exercise 1.1 Let $f : [a,b] \to \mathbf{R}$ be a given function. Prove that f is Riemann integrable if and only if for every sequence $((P_n,T_n))_{n=1}^{\infty}$ of pointed partitions for which $\mathrm{mesh}(P_n) \to 0$, as $n \to \infty$, $\lim_{n\to\infty} S(f;P_n,T_n)$ exists, and that in the integrable case, the value of each such limit is $\int_a^b f$.

You may divide this exercise into the following two parts:

(a) Prove that if f is Riemann integrable then for any sequence $((P_n,T_n))_{n=1}^{\infty}$ of pointed partitions of $[a,b]$ for which $\mathrm{mesh}(P_n) \to 0$ as $n \to \infty$,

$$\lim_{n\to\infty} S(f;P_n,T_n) = \int_a^b f. \tag{1.5}$$

(Hence the integral can be evaluated by taking such a limit, *if you know ahead of time that f is integrable.*)

(b) Assume that for every sequence $((P_n,T_n))_{n=1}^{\infty}$ of pointed partitions of $[a,b]$ for which $\mathrm{mesh}(P_n) \to 0$ as $n \to \infty$, $\lim_{n\to\infty} S(f;P_n,T_n)$ exists. Prove that f is Riemann integrable on $[a,b]$, and that for every such sequence $((P_n,T_n))$, the equality (1.5) holds. ▲

Definitions 1.6 and 1.8 are very intuitive, and, as we shall see later, generalize naturally to the integration of vector-valued functions (functions $[a, b] \to V$, where V is a complete normed vector space). However, these definitions can be unwieldy at times; it can be a chore to show the integrability of functions that are any more complicated than the constant function in Example 1.13. In the interests of efficiency, we postpone presenting other examples until we have developed equivalent definitions that are (often) easier to work with. For now, however, we introduce some notation that will allow us to rewrite the definition of "$\int_a^b f = A$" more succinctly than in Definition 1.6:

Notation 1.9 (a) Let $\mathcal{P}([a, b])$ denote the set of partitions of $[a, b]$, and, for each $\delta > 0$, let $\mathcal{P}_\delta([a, b]) \subseteq \mathcal{P}([a, b])$ denote the set of partitions of $[a, b]$ of mesh less than δ.

(b) For each function $f : [a, b] \to \mathbf{R}$ and each $\delta > 0$, let

$$\begin{aligned} \mathcal{S}_\delta(f) &= \bigcup_{P \in \mathcal{P}_\delta([a,b])} \mathcal{S}(f; P) \\ &= \{\text{all Riemann sums of } f \text{ associated to partitions of mesh less than } \delta\}. \end{aligned}$$

▲

Definitions 1.6 and 1.8 can now be rewritten as follows:

Definition 1.10 (Redefinition of integrability and the integral) A function $f : [a, b] \to \mathbf{R}$ is *Riemann integrable* if there is a real number A such that for each $\epsilon > 0$ there exists $\delta > 0$ such that $\mathcal{S}_\delta(f) \subseteq B_\epsilon(A)$. If f is Riemann integrable, then the (automatically unique) such number A is called the *Riemann integral of f (over $[a, b]$)* and is denoted $\int_a^b f$ or as $\int_a^b f(x)\, dx$, $\int_a^b f(t)\, dt$, etc. ▲

In these notes we will not attempt to give an "if and only if" criterion for a function to be Riemann integrable. However, there is a simple *necessary* criterion:

Proposition 1.11 ("Integrable implies bounded") *If $f : [a, b] \to V$ is Riemann integrable, then f is bounded.*

Proof: Let $f \in \mathcal{R}([a, b])$, and let $A = \int_a^b f$. Let $\delta > 0$ be such that $\mathcal{S}_\delta(f) \subseteq B_1(A)$. Fix a partition $P = \{x_0, \ldots x_N\}$ of $[a, b]$ of mesh less than δ.

Assume that f is unbounded. Then f is unbounded on at least one of the intervals $I_j := [x_{j-1}, x_j]$, since there are only finitely many such intervals. Let $j_0 \in \{1, \ldots, N\}$ be such that f is unbounded on I_{j_0}. For each $n \in \mathbf{N}$, choose $z_n \in I_{j_0}$ such that $|f(z_n)| > n$; such z_n exist by the unboundedness assumption. For each $j \in \{1, \ldots, N\}$ with $j \neq j_0$, fix any number $t_j \in [x_{j-1}, x_j]$, let $T^{(n)}$ be the pointing $\{t_1^{(n)}, \ldots, t_N^{(n)}\}$ of P for which $t_j^{(n)} = \begin{cases} t_j \text{ if } j \neq j_0, \\ z_n \text{ if } j = j_0\,, \end{cases}$ and let $A' = \sum_{j \neq j_0} f(t_j)\Delta_j$. Then, using the triangle inequality,

$$\begin{aligned} |S(f; P, T^{(n)}) - A| = |f(z_n)\Delta_{j_0} + A' - A| &\geq |f(z_n)\Delta_{j_0}| - |A - A'| \\ &= |f(z_n)|\Delta_{j_0} - |A - A'| \\ &> n\Delta_{j_0} - |A - A'|. \end{aligned}$$

For n sufficiently large, $n\Delta_{j_0} - |A - A'| > 1$, implying that $S(f; P, T^{(n)}) \notin B_1(A)$, a contradiction.

Hence f is bounded. ■

An argument similar to the one preceding Definition 1.8 leads to a useful necessary criterion for integrability:

Proposition 1.12 If $f : [a, b] \to \mathbf{R}$ is Riemann integrable, then for every $\epsilon > 0$ there exists $\delta > 0$ such that if (P, T) and (Q, T') are pointed partitions of $[a, b]$ of mesh less than δ, we have $|S(f; P, T) - S(f; Q, T')| < \epsilon$.

We omit the proof here, since this proposition is part of a more powerul result we will prove later (Theorem 1.32), and we want to get quickly to some simple examples of integrable and non-integrable functions. In the latter case, we will use Proposition 1.12 in its contrapositive form: for any given $f : [a, b] \to \mathbf{R}$, if there exists $\epsilon_0 > 0$ such that, for all $\delta > 0$, there exist pointed pointed partitions $(P, T), (Q, T')$ of $[a, b]$ of mesh less than δ for which $|S(f; P, T) - S(f; Q, T')| \geq \epsilon_0$, then f is not Riemann integrable.

Example 1.13 (an integrable function) For any $c \in \mathbf{R}$, the constant function $f : [a, b] \to \mathbf{R}$ given by $f(x) = c$ is integrable, and

$$\int_a^b c\, dx = c(b - a).$$

This follows from the fact that, as the student may check, every Riemann sum for f has the value $c(b - a)$. ▲

In particular, $\mathcal{R}([a, b])$ is nonempty!

Example 1.14 (a non-integrable function) Define $f : [a,b] \to \mathbf{R}$ by $f(x) = 1$ if $x \in \mathbf{Q}$ and $f(x) = 0$ if $x \notin \mathbf{Q}$. Let $P = \{x_0, \dots, x_N\}$ be a partition of $[a,b]$. For $1 \le j \le N$ choose $t_j, t_j' \in [x_{j-1}, x_j]$ such that $t_j \in \mathbf{Q}$ and $t_j' \notin \mathbf{Q}$. Let $T = \{t_1, \dots, t_N\}, T' = \{t_1', \dots, t_N'\}$. Then the Riemann sums of f corresponding to the pointed partitions T, T' respectively are Then, the corresponding Riemann sum is

$$S(f; P, T) = \sum_j f(t_j)\Delta_j = \sum_j \Delta_j = b - a$$

and

$$S(f; P, T') = \sum_j f(t_j')\Delta_j = \sum_j 0 = 0.$$

Hence $S(f; P, T) - S(f; P, T') = b - a$. Since this is true regardless of how small mesh(P) is, it follows that f is not Riemann integrable. (In the contrapositive form of Proposition 1.12 that we stated above, take $\epsilon_0 = b - a$, take δ arbitrary, and take $Q = P$.) ▲

For simplicity, **henceforth in these notes we will say simply use the word *integrable* to mean *Riemann integrable*, and refer to $\int_a^b f$ as the *integral* of f over** $[a,b]$**.** The student is cautioned that there are more general types of integrability—in particular, a type called *Lebesgue integrability*—and that usually when mathematicians say to each other (or to graduate students) that a function on $[a,b]$ is integrable, they mean Lebesgue-integrable. The analog of Proposition 1.11 is *false* for Lebesgue-integrable functions, and false even for functions for which we define an *improper integral* as in Calculus 2. Indeed, the fact that no unbounded function is Riemann integrable is viewed as a *weakness* of Riemann integration compared to Lebesgue integration. Nonetheless, studying Lebesgue integration without first studying Riemann integration can interfere with developing an intuitive understanding of *any* form of integration.

1.2 Linearity and order properties of the integral

Proposition 1.15 *Let $f, g : [a,b] \to \mathbf{R}$ and $c \in \mathbf{R}$. If both f and g are integrable, then so are $f + g$ and cf, and the following equalities hold:*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g. \tag{1.6}$$

$$\int_a^b cf = c\int_a^b f. \tag{1.7}$$

Proof: From the definition (1.2), we easily see that, for any pointed partition (P,T) of $[a,b]$, we have $S(f+g; P, T) = S(f; P, T) + S(g; P, T)$ and $S(cf; P, T) = cS(f; P, T)$.

Assume now that f and g are integrable, and let $A = \int_a^b f, C = \int_a^b g$. Let $\epsilon > 0$, and let $\delta_1, \delta_2 > 0$ be such that $\mathcal{S}_{\delta_1}(f) \subseteq B_\epsilon(A)$ and $\mathcal{S}_{\delta_2}(g) \subseteq B_\epsilon(C)$. Then for any pointed partition (P,T) of $[a,b]$ of mesh less than $\min\{\delta_1, \delta_2\}$, we have

$$\begin{aligned} |S(f+g;P,T) - (A+C)| &= |(S(f;P,T) - A) + (S(g;P,T) - C)| \\ &\leq |S(f;P,T) - A| + |S(g;P,T) - C| \\ &< 2\epsilon. \end{aligned}$$

It follows that $f+g$ is integrable and that (1.6) holds. Similarly, for any pointed partition (P,T) of $[a,b]$ of mesh less than δ_1,

$$|S(cf;P,T) - cA| = |cS(f;P,T) - cA| = |c|\,|S(f;P,T) - A| \leq |c|\epsilon,$$

from which the integrability of cf and the equality (1.7) follow. ■

Remark 1.16 The proof of Proposition 1.15 illustrated something that comes up in innumerable proofs. As you may have learned in MAA 4211, in proofs we are quite often in a situation of a form like the following: Statement 1 is true for all $x \in (0, \delta_1)$ (or for all $n > N_1$), statement 1 is true for all $x \in (0, \delta_2)$ (or for all $n > N_2$), ..., statement k is true for all $x \in (0, \delta_k)$ (or for all $n > N_k$). We then say "Let $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_k\}$" (or "Let $N = \max\{N_1, \ldots, N_k\}$"), and are then guaranteed that all k statements are true for all $x \in (0, \delta)$ (or for all $n \geq N$). As long as there are only finitely many statements involved (typically there are only two), this device always works. Once the student has had sufficient experience, he/she should not have trouble following proofs in which several of these intermediate steps are omitted. For example, in the proof of Proposition 1.15, we could have replaced the third and fourth sentences with, "Let $\epsilon > 0$, and let $\delta > 0$ be such that $\mathcal{S}_\delta(f) \subseteq B_\epsilon(A)$ and $\mathcal{S}_\delta(g) \subseteq B_\epsilon(C)$. Then for any pointed partition (P,T) of mesh less than δ, we have ... ." By the end of MAA 4211, students should definitely have had enough experience to be comfortable with such arguments (*but should always know how to justify them the "long way"*). So, **henceforth in these notes, we will use this device to shorten arguments whenever we can.** ▲

Observe that, since $\mathcal{R}([a,b])$ is nonempty, Proposition 1.15 can be phrased alternatively as follows:

Proposition 1.17 (Linearity of the integral) *The set $\mathcal{R}([a,b])$ is a vector space, and the map $f \mapsto \int_a^b f$ is a linear map $\mathcal{R}([a,b]) \to \mathbf{R}$.*

The integration-map $f \mapsto \int_a^b f$ also has the following "non-negativity" property:

Proposition 1.18 *Assume that* $f : [a,b] \to \mathbf{R}$ *is integrable and that* $f(x) \geq 0$ *for all* $x \in [a,b]$. *Then* $\int_a^b f \geq 0$.

Exercise 1.2 Prove Proposition 1.18.

Corollary 1.19 (Order property of the integral) *Assume that* $f, g : [a,b] \to \mathbf{R}$ *are integrable and that* $f(x) \geq g(x)$ *for all* $x \in [a,b]$. *Then* $\int_a^b f \geq \int_a^b g$.

Proof: Let $h = f - g$. Then $h \in \mathcal{R}([a,b])$ (by Proposition 1.17) and $h(x) \geq 0$ for all $x \in [a,b]$. Hence

$$0 \leq \int_a^b h = \int_a^b (f-g) = \int_a^b f - \int_a^b g,$$

and the result follows. ■

Remark 1.20 Note that Proposition 1.18 does *not* imply that if f is integrable over $[a,b]$ and $f(x) > 0$ for all $x \in [a,b]$, then $\int_a^b f > 0$. A similar observation applies to Corollary 1.19. Changing a non-strict inequality to a strict one in *hypotheses* does not mean that inequalities in *conclusions* become strict. (For example, given a convergent real-valued sequence $(a_n)_{n=1}^{\infty}$, it is true that if $a_n \geq 0$ for all n then $\lim_{n\to\infty} a_n \geq 0$, but it is *not* true that if $a_n > 0$ for all n then $\lim_{n\to\infty} a_n > 0$.) We will see later (Remark 1.54 following Exercise 1.6) that if f in Proposition 1.18 is assumed *continuous*, and $f(x) > 0$ for all $x \in [a,b]$, then indeed $\int_a^b f > 0$.

It is natural to ask whether pointwise-positivity of f *does* imply positivity of the integral if we assume only that f is integrable, not that f is continuous. In these notes, *we will leave this question open*. Students are invited to try either to prove that the integral is positive under these hypotheses, or to find a counterexample in which the hypotheses are met but $\int_a^b f = 0$. ▲

1.3 Upper and lower sums

Since unbounded functions are not (Riemann-)integrable (Proposition 1.11), we will simplify some parts of the presentation below by restricting attention to bounded functions.

Notation 1.21 We will write $\mathcal{B}([a,b])$ for the set of bounded real-valued functions on $[a,b]$. ▲

Thus, Proposition 1.11 can be written succinctly as: $\mathcal{R}([a,b]) \subseteq \mathcal{B}([a,b])$.

Definition 1.22 For each function $f : [a,b] \to \mathbf{R}$, $P \in \mathcal{P}([a,b])$, and $\delta > 0$, we define

$$\begin{aligned} U(f;P) &= \sup(\mathcal{S}(f;P)) \\ L(f;P) &= \inf(\mathcal{S}(f;P)), \\ U_\delta(f) &= \sup(\mathcal{S}_\delta(f)), \\ L_\delta(f) &= \inf(\mathcal{S}_\delta(f)). \end{aligned}$$

The quantities $U(f;P), L(f;P)$ are called, respectively, the *upper* and *lower sums of f with respect to P.* ▲

Observe that, trivially, in the setting of Definition 1.22 we have

$$\begin{aligned} L(f;P) &\leq U(f;P) \\ \text{and} \qquad L_\delta(f) &\leq U_\delta(f). \end{aligned} \tag{1.8}$$

Remark 1.23 Recall that, in general, the supremum of a nonempty set of real numbers can be ∞, and the infimum can be $-\infty$. A consequence of the upcoming Proposition 1.27 is that for $f \in \mathcal{B}([a,b])$, the upper and lower sums of f with respect to any partition are *finite* (i.e real numbers, never $\pm\infty$), and for any $\delta > 0$ so are the numbers $U_\delta(f)$ and $L_\delta(f)$. ▲

Lemma 1.24 *Let $\{X_\alpha : \alpha \in A\}$ be a collection of nonempty subsets X_α of $\mathbf{R}$ indexed by a nonempty set A. Then*

$$\begin{aligned} \sup\left(\bigcup\{X_\alpha : \alpha \in A\}\right) &= \sup\{\sup(X_\alpha) : \alpha \in A\} \\ \textit{and} \qquad \inf\left(\bigcup\{X_\alpha : \alpha \in A\}\right) &= \inf\{\inf(X_\alpha) : \alpha \in A\}. \end{aligned}$$

Exercise 1.3 Prove Lemma 1.24.

In the setting of Definition 1.22, applying Lemma 1.24 to the indexed collection $\{\mathcal{S}(f;P) : P \in \mathcal{P}_\delta([a,b])\}$, we have

$$\begin{aligned} U_\delta(f) &= \sup\{\sup(\mathcal{S}(f;P)) : P \in \mathcal{P}_\delta([a,b])\} \\ &= \sup\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}, \end{aligned} \tag{1.9}$$

and similarly

$$L_\delta(f) = \inf\{L(f;P) : P \in \mathcal{P}_\delta([a,b])\}. \tag{1.10}$$

Example 1.25 Let $f : [0,1] \to \mathbf{R}$ be the squaring function: $f(x) = x^2$. For each positive integer N, let $P_N = \{x_j := \frac{j}{N} : 0 \leq j \leq N\}$, a partition of $[0,1]$. Consecutive points of this partition are equally spaced: $\Delta_j(P_N) = \frac{1}{N}$ for each $j \in \{1, 2 \ldots, N\}$.[3] Let $T = \{t_1, \ldots, t_N\}$ be a pointing of P_N. Then

$$S(f; P_N, T) = \sum_{j=1}^{N} t_j^2 \ \Delta_j(P_N) = \sum_{j=1}^{N} t_j^2 \, \frac{1}{N} \ .$$

For the j^{th} term in the sum, we have $\frac{j-1}{N} = x_{j-1} \leq t_j \leq x_j = \frac{j}{N}$, implying $\frac{(j-1)^2}{N^2} \leq t_j^2 \leq \frac{j^2}{N^2}$. Hence

$$\sum_{j=1}^{N} \frac{(j-1)^2}{N^2} \frac{1}{N} \ \leq \ S(f; P_N, T) \ \leq \ \sum_{j=1}^{N} \frac{j^2}{N^2} \frac{1}{N} \ . \tag{1.11}$$

Morever, if we take T to be the "right-endpoint pointing" of P_N (i.e. $t_j = x_j$ for $1 \leq j \leq N$) then the value of $S(f; P_N, T)$ is exactly the rightmost sum in (1.11), while if we take T to be the "left-endpoint pointing" of P_N (i.e. $t_j = x_{j-1}$ for $1 \leq j \leq N$) then the value of $S(f; P_N, T)$ is exactly the leftmost sum in (1.11). Hence, using the fact that $\sum_{j=1}^{N} j^2 = \frac{N(N+1)(2N+1)}{6}$ (which is easily proven by induction, and which you may have learned in high school), it follows from (1.11) that

$$\begin{aligned} U(f; P_N) = \frac{1}{N^3} \sum_{j=1}^{N} j^2 \ &= \ \frac{1}{N^3} \frac{N(N+1)(2N+1)}{6} \\ &= \ \frac{1}{3} + \frac{1}{2N} + \frac{1}{6N^2} \end{aligned}$$

and that

$$\begin{aligned} L(f; P_N) = \frac{1}{N^3} \sum_{j=1}^{N} (j-1)^2 = \frac{1}{N^3} \sum_{j=1}^{N-1} j^2 \ &= \ \frac{1}{N^3} \frac{(N-1)N(2N-1)}{6} \\ &= \ \frac{1}{3} - \frac{1}{2N} + \frac{1}{6N^2} \ . \end{aligned}$$

▲

Remark 1.26 In Example 1.25, the fact that the supremum $U(f; P_N) = \sup(\mathcal{S}(f; P_N))$ and infimum $L(f; P_N) = \inf(\mathcal{S}(f; P_N))$ were achieved by, respectively, the right-endpoint and left-endpoint pointings of P_N, *was a consequence of having chosen the function f in this example to be monotone-increasing* on the interval of interest, $[0,1]$. *In this example,* $U(f; P_N)$ and $L(f; P_N)$ turned out to be the *maximal* and *minimal* Riemann sums of f for

---

[3] A partition with equally-spaced points is sometimes called a *regular partition.*

this partition. For a general function $f : [a,b] \to \mathbf{R}$ (bounded or otherwise), and partition P (with or without equally-spaced points) the values $U(f;P)$ and $L(f;P)$ may not be achieved by *any* pointings of P, let alone by the left-endpoint or right-endpoint pointings; there be no maximal or minimal Riemann sums. Never forget that "sup" and "inf" are *more general concepts* than "max" and "min", and that you cannot replace "sup" by "max" (or "inf" by "min") unless you have shown that the supremum (or infimum) of the set in question *lies in that set.* ▲

Proposition 1.27 *Let $f \in \mathcal{B}([a,b]$, and let $M = \sup\{f(x) : x \in [a,b]\}$ and $m = \inf\{f(x) : x \in [a,b]\}$. Then for each $P \in \mathcal{P}([a,b])$ and $S \in \mathcal{S}(f;P)$, we have*

$$m(b-a) \le S \le M(b-a). \tag{1.12}$$

Hence

$$\mathcal{S}(f;P) \subseteq [m(b-a), M(b-a)] \tag{1.13}$$

and

$$m(b-a) \le L(f;P) \le U(f;P) \le M(b-a). \tag{1.14}$$

Consequently, for each $\delta > 0$ and $P \in \mathcal{P}_\delta([a,b])$,

$$m(b-a) \ \le\ L_\delta(f) \ \le\ L(f;P) \ \le U(f;P) \ \le\ U_\delta(f) \ \le\ M(b-a). \tag{1.15}$$

Proof: Let $P \in \mathcal{P}([a,b])$ and let $S \in \mathcal{S}(f;P)$. Then $S = S(f;P,T)$ for some pointing T of P, so (1.12) follows from the Riemann-sum definition (1.2) and the fact that $\sum_j \Delta_j = b-a$. The first and third inequalities in (1.14) follow immediately from (1.12) and the definitions of $L(f;P)$ and $U(f;P)$, and the middle inequality is simply the trivially-true inequality (1.8). The inequalities $L_\delta(f) \le L(f;P)$ and $U(f;P) \le U_\delta(f)$ in (1.15) follow from (1.10) and (1.9), respectively, and the remaining inequalities follow from (1.14). ■

Proposition 1.28 *Let $f \in \mathcal{B}([a,b])$. Define functions $h_1, h_2 : (0,\infty) \to \mathbf{R}$ by*

$$\begin{aligned} h_1(\delta) &= L_\delta(f) \\ \textit{and} \qquad h_2(\delta) &= U_\delta(f). \end{aligned}$$

Then h_1 is monotone decreasing, h_2 is monotone increasing, and both functions are bounded.

Proof: For $\delta_1, \delta_2 \in \mathbf{R}$ with $\delta_1 < \delta_2$, every partition of mesh less than δ_1 also has mesh less than δ_2. Hence $\mathcal{P}_{\delta_1}([a,b]) \subseteq \mathcal{P}_{\delta_2}([a,b])$, implying that $\mathcal{S}_{\delta_1}(f) \subseteq \mathcal{S}_{\delta_2}(f)$. But for any nonempty subsets A, B of $\mathbf{R}$ with $A \subseteq B$, we have $\inf(A) \ge \inf(B)$ and $\sup(A) \le \sup(B)$. Hence $L_{\delta_1}(f) \ge L_{\delta_2}(f)$ and $U_{\delta_1}(f) \le U_{\delta_2}(f)$.

This proves the asserted monotonicity. Boundedness follows from (1.15). (In the notation of (1.15), the ranges of both h_1 and h_2 lie in $[m(b-a), M(b-a)]$.) ■

We will soon be considering certain limits as $\delta \to 0$. Notice that for a decreasing function h_1 on $(0,\infty)$, the value $h_1(\delta)$ *increases* as $\delta \to 0$; similarly, for an increasing function h_2 on $(0,\infty)$, the value $h_2(\delta)$ *decreases* as $\delta \to 0$.

Lemma 1.29 *Let $I \subseteq \mathbf{R}$ be an interval bounded from below, and let c be the left endpoint of $\bar{I}$ (the closure of I); equivalently, let $c = \inf(I)$.*

(i) If $h : I \setminus \{c\} \to \mathbf{R}$ is an increasing function that is bounded from below, then

$$\lim_{u\to c} h(u) = \inf(\mathrm{range}(h)). \tag{1.16}$$

(ii) If $h : I \setminus \{c\} \to \mathbf{R}$ is a decreasing function that is bounded from above, then

$$\lim_{u\to c} h(u) = \sup(\mathrm{range}(h)). \tag{1.17}$$

In particular, under the indicated hypotheses, the limits above exist.

Proof: Let $u_1 \in I \setminus \{c\}$ be such that $h(u_1) < \alpha + \epsilon$; such u_1 exists since (by definition of "inf") $\alpha + \epsilon$ is not a lower bound of range(h). Let $r = u_1 - c$; thus $r > 0$ and $u_1 = c + r$. Then for all u with $c < u < c + r$ we have $\alpha \leq h(u) \leq h(u_1) < \alpha + \epsilon$. Thus for all $u \in I \setminus \{c\}$ for which $|u - c| < r$, we have $|h(u) - \alpha| = h(u) - \alpha < \epsilon$. Since ϵ was arbitrary, this establishes that $\lim_{u\to c} h(u) = \alpha$.

This proves (i). Statement (ii) can be deduced by applying (i) to the function $-h$.
■

Corollary 1.30 *Let $f \in \mathcal{B}([a,b])$. Then $\lim_{\delta\to 0} L_\delta(f)$ and $\lim_{\delta\to 0} U_\delta(f)$ both exist, and*

$$\lim_{\delta\to 0} L_\delta(f) \;\leq\; \lim_{\delta\to 0} U_\delta(f). \tag{1.18}$$

Proof: Let $h_1, h_2 : (0,\infty) \to \mathbf{R}$ be the functions defined in Proposition 1.28. By the Proposition, each of these functions is monotone and bounded, so Lemma 1.29 implies that the limits in (1.18) exist. Since both these limits exist, and $L_\delta(f) \leq U_\delta(f)$ for each $\delta > 0$, the inequality (1.18) follows. ■

Another general lemma that will be used shortly is the following:

Lemma 1.31 *Let* $X \subseteq \mathbf{R}$ *be a nonempty, bounded set. Let* $r > 0$ *and assume that for all* $x_1, x_2 \in X$ *we have* $|x_1 - x_2| \leq r$. *Then* $\sup(X) - \inf(X) \leq r$.

Proof: Let $\epsilon > 0$, and let $x_1, x_2 \in X$ be such that $x_1 > \sup(X) - \epsilon$ and $x_2 < \inf(X) + \epsilon$; such x_1, x_2 exist since $\sup(X)$ and $\inf(X)$ are, respectively, the *least* upper bound and *greatest* lower bound of X. Then $\sup(X) < x_1 + \epsilon$ and $\inf(X) > x_2 - \epsilon$, so $\sup(X) - \inf(X) < x_1 + \epsilon - (x_2 - \epsilon) = x_1 - x_2 + 2\epsilon \leq r + 2\epsilon$. Since ϵ was arbitrary, the result follows. ■

We can now recast integrability in terms of the limits in Corollary 1.30:

Theorem 1.32 *For each* $f \in \mathcal{B}([a,b])$, *the following are equivalent:*

(i) f *is integrable over* $[a,b]$.

(ii) $\lim_{\delta\to 0} L_\delta(f) \;=\; \lim_{\delta\to 0} U_\delta(f)$.

(iii) $\lim_{\delta\to 0}(U_\delta(f) - L_\delta(f)) = 0$.

(iv) For every $\epsilon > 0$ *there exists* $\delta > 0$ *such that for all* $S_1, S_2 \in \mathcal{S}_\delta(f)$, *we have* $|S_2 - S_1| < \epsilon$.

In the integrable case,

$$\int_a^b f \;=\; \lim_{\delta\to 0} L_\delta(f) \;=\; \lim_{\delta\to 0} U_\delta(f). \tag{1.19}$$

Proof: Let $f \in \mathcal{B}([a,b])$.

<u>(i) $\Longrightarrow$ (ii), plus last sentence of Proposition.</u>

Assume that f is integrable over $[a,b]$, and let $A = \int_a^b f$. Let $\epsilon > 0$, and let $\delta > 0$ be such that $\mathcal{S}_\delta(f) \subseteq (A-\epsilon, A+\epsilon)$. Then the infimum $L_\delta(f)$ of $\mathcal{S}_\delta(f)$ and supremum $U_\delta(f)$ of $\mathcal{S}_\delta(f)$ lie in $[A-\epsilon, A+\epsilon]$, implying that $|L_\delta(f) - A| \leq \epsilon$ and $|U_\delta(f) - A| \leq \epsilon$. Since ϵ was arbitrary, it follows that $\lim_{\delta\to 0} L_\delta(f) = A = \lim_{\delta\to 0} U_\delta(f)$. This implies both statement (ii) and the last sentence of the Proposition.

<u>(ii) $\Longleftrightarrow$ (iii)</u>

By Corollary 1.30, both $\lim_{\delta\to 0} U_\delta(f)$ and $\lim_{\delta\to 0} L_\delta(f)$ exist. Hence $\lim_{\delta\to 0}(U_\delta(f) - L_\delta(f)) = \lim_{\delta\to 0} U_\delta(f) - \lim_{\delta\to 0} L_\delta(f)$. The equivalence of (ii) and (iii) is immediate from this equality.

<u>(ii) $\Longrightarrow$ both (i) and (iv)</u>

Assume the equality in (ii) holds, and let A be the (common) value of the indicated limits. Let $\epsilon > 0$. Let $\delta > 0$ be such that $|L_\delta(f) - A| < \epsilon$ and $|U_\delta(f) - A| < \epsilon$; such δ exists since $\lim_{\delta\to 0} L_\delta(f) = A = \lim_{\delta\to 0} U_\delta(f)$. Then for all $S \in \mathcal{S}_\delta(f)$,

$$A - \epsilon < L_\delta(f) \leq S \leq U_\delta(f) < A + \epsilon,$$

implying that $\mathcal{S}_\delta(f) \subseteq (A - \epsilon, A + \epsilon)$. Hence $\int_a^b f = A$, so (i) is true. Furthermore, for all $S_1, S_2 \in \mathcal{S}_\delta(f)$ we have $|S_2 - S_1| < 2\epsilon$. Since ϵ was arbitrary, this establishes (iv).

<u>(iv) $\implies$ (iii)</u>

Assume that (iv) holds. Let $\delta_0 > 0$ be such that for all $S_1, S_2 \in \mathcal{S}_{\delta_0}(f)$, $|S_2 - S_1| < \epsilon$. Then by Lemma 1.31, $0 \leq U_{\delta_0}(f) - L_{\delta_0}(f) = \sup(\mathcal{S}_{\delta_0}(f)) - \inf(\mathcal{S}_{\delta_0}(f) \leq \epsilon$. The monotonicities of the functions $\delta \mapsto L_\delta(f), \delta \mapsto U_\delta(f)$ established in Proposition 1.28 imply that for all $\delta \in (0, \delta_0]$, we have

$$L_{\delta_0}(f) \leq L_\delta(f) \leq U_\delta(f) \leq U_{\delta_0}(f).$$

Hence for all such δ, $0 \leq U_\delta(f) - L_\delta(f) \leq U_{\delta_0}(f) - L_{\delta_0}(f) \leq \epsilon$. Since, by Corollary 1.30, $\lim_{\delta\to 0} U_\delta(f)$ and $\lim_{\delta\to 0} L_\delta(f)$ both exist, so does $\lim_{\delta\to 0}(U_\delta(f) - L_\delta(f))$, and by the basic order-property established in MAA 4211 for limits of real-valued functions,

$$0 \leq \lim_{\delta\to 0}(U_\delta(f) - L_\delta(f)) \leq \epsilon. \tag{1.20}$$

Since ϵ was arbitrary, (1.20), this implies that $\lim_{\delta\to 0}(U_\delta(f) - L_\delta(f)) = 0$. ■

Remark 1.33 Since every Riemann-integrable function is bounded, the (previously unproven) Proposition 1.12 amounts to the "(i) $\implies$ (iv)" implication in Theorem 1.32. ▲

Remark 1.34 Statement (iv) in Theorem 1.32 can be thought of, loosely, as a "Cauchy criterion for the convergence of Riemann sums" (with "convergence of Riemann sums" interpreted heuristically, since the set of Riemann sums of a function f on $[a, b]$ is not a sequence). ▲

Remark 1.35 The equivalence of (i) and (iv) in Theorem 1.32 can be proven without any use of upper and lower sums. We will give such a proof later, when we discuss integration of vector-valued functions. ▲

The following characterization of upper and lower sums, worthwhile for its own sake, simplifies our work when we apply Theorem 1.32 to compute integrals or prove integrability.

Proposition 1.36 *Let* $f \in \mathcal{B}([a,b])$*, and let* $P = \{x_0, \ldots, x_N\}$ *be a partition of* $[a,b]$*. For* $1 \leq j \leq N$*, let*

$$m_j = \inf\{f(x) : x_{j-1} \leq x \leq x_j\} \quad \textit{and} \quad M_j = \sup\{f(x) : x_{j-1} \leq x \leq x_j\}.$$

Then

$$L(f;P) = \sum_{j=1}^{N} m_j \Delta_j \quad \textit{and} \quad U(f;P) = \sum_{j=1}^{N} M_j \Delta_j\,. \tag{1.21}$$

Proof: Since $f(x) \leq M_j$ for all $x \in [x_{j-1}, x_j], 1 \leq j \leq N$, it is clear that for any pointing T of P we have $S(f;P,T) \leq \sum_j M_j \Delta_j$, so $\sum_{j=1}^{N} M_j \Delta_j$ is an upper bound for $\mathcal{S}(f;P)$.

Now let $\epsilon > 0$. For each $j \in \{1, 2, \ldots, N\}$, let $t_j \in [x_{j-1}, x_j]$ be such that $f(t_j) > M_j - \frac{\epsilon}{b-a}$; such t_j exists by the definition of M_j. Let $T = \{t_1, \ldots, t_N\}$. Then T is a pointing of P, and

$$\begin{aligned} S(f;P,T) = \sum_{j=1}^{N} f(t_j)\Delta_j > \sum_j \left(M_j - \frac{\epsilon}{b-a}\right)\Delta_j \quad &= \left(\sum_j M_j \Delta_j\right) - \frac{\epsilon}{b-a} \sum_j \Delta_j \\ &= \left(\sum_j M_j \Delta_j\right) - \epsilon. \end{aligned}$$

Hence no number smaller than $\sum_j M_j \Delta_j$ is an upper bound for $\mathcal{S}(f;P)$. Thus $\sum_j M_j \Delta_j$ is the least upper bound (= supremum) of $\mathcal{S}(f;P)$, yielding the second equality in (1.21). A similar argument (left to the student) establishes the first equality. ■

Definition 1.37 Let A be a set and let $B \subseteq A$. The *characteristic function*[4] *of* B *(viewed as a subset* A*)* is the function $\chi_B : A \to \mathbf{R}$ defined by

$$\chi_B(p) = \begin{cases} 1 \text{ if } p \in B, \\ 0 \text{ if } p \notin B. \end{cases}$$

▲

For example, the function in Example 1.14 is simply the restriction of $\chi_{\mathbf{Q}}$ to $[a,b]$ (regarding $\mathbf{Q}$ as a subset of $\mathbf{R}$). The characteristic function of any interval (including a one-point interval) is an example of a *step function*; see Definition 1.42.

The next example and our proof of the next proposition illustrate how Theorem 1.32 can be used.

---

[4] In some areas of mathematics, such as probability, characteristic functions are called *indicator functions*, and the notation $\mathbf{1}_B$ is used instead of χ_B.

Example 1.38 Fix $c \in [a,b]$ and let $f = \chi_{\{c\}} : [a,b] \to \mathbf{R}$. Let $\delta > 0$ and let $P = \{x_0, \dots x_N\})$ be a partition of $[a,b]$ of mesh less than δ. With notation as in Proposition 1.36, $m_j = 0$ for every j, and $M_j = 0$ unless $c \in [x_{j-1}, x_j]$; the latter can happen for at most two values of j. When nonzero, the value of M_j is 1. Hence, using (1.21), we have

$$0 = L(f;P) \le U(f;P) < 2\delta.$$

Since this holds for all $P \in \mathcal{P}_\delta([a,b])$, it follows that

$$0 \le L_\delta(f) \le U_\delta(f) \le 2\delta.$$

Using the Squeeze Theorem and Theorem 1.32, we conclude that

$$\int_a^b f = 0.$$

▲

Proposition 1.39 Suppose that $a \le c < d \le b$. Then the characteristic function $\chi_{(c,d)} : [a,b] \to \mathbf{R}$ is integrable, and

$$\int_a^b \chi_{(c,d)} = d - c.$$

Proof: To streamline notation in this proof, let $f = \chi_{(c,d)}$. We will show that $\lim_{\delta \to 0} L_\delta(f) = \lim_{\delta \to 0} U_\delta(f) = d - c$. For this, it suffices to restrict attention to δ less than any fixed, positive number; in particular, to $\delta < d - c$.

Let $\delta \in (0, d-c)$, let $P = \{x_0, \dots, x_N\}$ be a partition of $[a,b]$ of mesh less than δ, and let j_1, j_2 be the unique indices in $\{1, \dots, N\}$ such that $c \in [x_{j_1-1}, x_{j_1})$ and $d \in (x_{j_2-1}, x_{j_2}]$. Then $x_{j_1-1} \le c < d \le x_{j_2}$, so $j_1 - 1 < j_2$; equivalently, $j_1 \le j_2$. If $j_2 = j_1$ then $\delta > x_{j_1} - x_{j_1-1} = x_{j_2} - x_{j_1-1} > \delta$, a contradiction, so in fact we have $j_1 < j_2$; equivalently, $j_1 \le j_2 - 1$. Hence $c < x_{j_1} \le x_{j_2-1} < d$, so x_j lies in (c,d) if $j_1 \le j < j_2$.

For $j \in \{1, \dots, N\}$, let m_j, M_j be as in Proposition 1.36; observe that each of these numbers is either 0 or 1. Then for each $j \in \{1, \dots, N\}$ we have the following:

$$\begin{array}{ll} \text{If } j < j_1 \text{ or } j > j_2 & \text{then } [x_{j-1}, x_j] \cap (c,d) = \emptyset, \text{ so } m_j = M_j = 0. \\ \text{If } j_1 < j < j_2 & \text{then } [x_{j-1}, x_j] \subseteq (c,d), \text{ so } m_j = M_j = 1. \\ \text{If } j = j_1 \text{ or } j = j_2 & \text{then } 0 \le m_j \le M_j \le 1. \end{array}$$

(As will be seen, more precise information about the indices j_1, j_2 is unnecessary, so we do not waste time on that.) Therefore, using (1.21),

$$\begin{aligned} U(f;P) \le \sum_{j=j_1}^{j_2} \Delta_j(P) = x_{j_2} - x_{j_1-1} &= (x_{j_2} - d) + (d - c) + (c - x_{j_1-1}) \\ &< (x_{j_2} - x_{j_2-1}) + (d - c) + (x_{j_1} - x_{j_1}) \\ &< \delta + (d - c) + \delta \\ &= d - c + 2\delta, \end{aligned}$$

and, since $x_{j_2-1} > x_{j_2} - \delta \geq d - \delta$ and $x_{j_1} < x_{j_1-1} + \delta \leq c + \delta$,

$$L(f;P) \geq \sum_{j=j_1+1}^{j_2-1} \Delta_j = x_{j_2-1} - x_{j_1} \geq (d-\delta) - (c+\delta) = (d-c) - 2\delta.$$

(For the case in which $j_1 + 1 > j_2 - 1$—which can happen only if $j_1 + 1 = j_2$, since $j_1 < j_2$—recall that the notation "$\sum_{j=m}^{n}$" means "the sum over all j satisfying $m \leq j \leq n$ if this index-set is nonempty, and 0 if this index-set is empty.")

Since the above inequalities hold for all $P \in \mathcal{P}_\delta([a,b])$, it follows that

$$(d-c) - 2\delta \leq L_\delta(f) \leq U_\delta(f) \leq (d-c) + 2\delta.$$

The result now follows from the Squeeze Theorem and Theorem 1.32. ■

Proposition 1.40 *Let $f, g : [a,b] \to \mathbf{R}$. Assume that f is integrable and that g differs from f at only finitely many points (i.e. that there are only finitely many $x \in [a,b]$ for which $g(x) \neq f(x)$). Then g is integrable, and $\int_a^b g = \int_a^b f$.*

Proof: Let $x_1, \ldots, x_n$ be the values of x for which $g(x) \neq f(x)$ (we may assume there is at least one such value, since otherwise $g = f$ and the conclusion is trivial). Let $h = g - f$. Then $h(x) = 0$ for all $x \notin \{x_1, \ldots, x_n\}$, so h is a linear combination of the functions $\chi_{\{x_1\}}, \ldots, \chi_{\{x_n\}}$; specifically, $h = \sum_i c_i \chi_{\{x_i\}}$ where $c_i = h(x_i)$. By Proposition 1.17 and Example 1.38, h is integrable and

$$\int_a^b h = \sum_i c_i \int_a^b \chi_{\{x_i\}} = \sum_i c_i \cdot 0 = 0.$$

But $g = f + h$, so g is the sum of two integrable functions. Using Proposition 1.17, the conclusion follows. ■

Corollary 1.41 *Let $I \subseteq [a,b]$ be an interval, and let $c \leq d$ be the left and right endpoints, respectively, of $\bar{I}$. Then $\chi_I : [a,b] \to \mathbf{R}$ is integrable and $\int_a^b \chi_I = d - c$.*

Corollary 1.41 follows easily from Proposition 1.39, Example 1.38, and "linearity of the integral" (Proposition 1.17); it would follow from Propositions 1.40 and 1.39 except for the special case $c = d$. This corollary is also a special case of the upcoming Proposition 1.44 (and is proven essentially the same way), so we omit writing a separate proof here.

Definition 1.42 A function $f : [a, b] \to \mathbf{R}$ is a *step function* if there exists a partition $\{x_0, \ldots x_N\}$ of $[a, b]$ such that for each $j \in \{1, \ldots, N\}$, f is constant on the open interval (x_{j-1}, x_j). ▲

Lemma 1.43 *If $f : [a, b] \to \mathbf{R}$ is a step-function, then f is a linear combination of characteristic functions of intervals.*

Proof: Let $f : [a, b] \to \mathbf{R}$ be a step-function, let $P = \{x_0, \ldots x_N\} \in \mathcal{P}([a, b])$ be such that for each $j \in \{1, \ldots, N\}$, $f|_{(x_{j-1}, x_j)}$ is constant, and for each such j let c_j denote the (constant) value of $f|_{(x_{j-1}, x_j)}$. Then, as is easily verified,

$$f = \sum_{j=1}^{N} c_j \chi_{(x_{j-1}, x_j)} + \sum_{j=0}^{N} f(x_j) \chi_{\{x_j\}}, \tag{1.22}$$

a linear combination of characteristic functions of intervals. ■

Exercise 1.4 Prove the converse of Lemma 1.43. (Note that in the phrase "linear combination of characteristic functions of intervals", it is not given that the intervals do not overlap.)

Proposition 1.44 *Let $f : [a, b] \to \mathbf{R}$ be a step-function, and let P, N and $c_1, \ldots, c_N$ be as in the proof of Lemma 1.43. Then f is integrable and*

$$\int_a^b f = \sum_{j=1}^{N} c_j \Delta_j(P).$$

Proof: This follows from equation 1.22, Proposition 1.17, and Example 1.38. ■

Proposition 1.45 ("Step-function lemma") *A function $f \in \mathcal{B}([a, b])$ is integrable if and only if for each $\epsilon > 0$, there exists a partition P of $[a, b]$ such that*

$$U(f; P) - L(f; P) < \epsilon. \tag{1.23}$$

Remark 1.46 The strength of Proposition 1.45 is that *there is no reference to the mesh of the partition P*. This makes the Proposition much simpler to apply than many of our results up till now. ▲

Proof of Proposition 1.45:

($\Longrightarrow$) Assume that f is integrable. Then by Theorem 1.32, $\lim_{\delta\to 0}(U_\delta(f)-L_\delta(f))=0$. Let $\epsilon>0$, and let $\delta>0$ be such that $U_\delta(f)-L_\delta(f)<\epsilon$; such δ exists since the above limit is 0. Let P be any partition of mesh less than δ. Then, by (1.15), we have $U(f;P)-L(f;P)\le U_\delta(f)-L_\delta(f)<\epsilon$.

($\Longleftarrow$) Assume that for each $\epsilon>0$, there exists a partition P of $[a,b]$ such that $U(f;P)-L(f;P)<\epsilon$.

Let $\epsilon>0$, and let $P=\{x_0,\dots,x_N\}\in\mathcal{P}([a,b])$ be such that $U(f;P)-L(f;P)<\epsilon$. For $1\le j\le N$ let M_j and m_j be as in Proposition 1.36. Let $m=\inf\{f(x):x\in[a,b]\}$ and $M=\sup\{f(x):x\in[a,b]\}$ Define functions $f_1,f_2:[a,b]\to\mathbf{R}$ by

$$\begin{aligned} f_1 &= \sum_{j=1}^N m_j\chi_{(x_{j-1},x_j)}+m\chi_P=\sum_{j=1}^N m_j\chi_{(x_{j-1},x_j)}+\sum_{j=0}^N m\chi_{\{x_j\}} \\ f_2 &= \sum_{j=1}^N M_j\chi_{(x_{j-1},x_j)}+M\chi_P=\sum_{j=1}^N M_j\chi_{(x_{j-1},x_j)}+\sum_{j=0}^N M\chi_{\{x_j\}}. \end{aligned}$$

Observe also that

$$f_1(x)\le f(x)\le f_2(x) \tag{1.24}$$

for every $x\in[a,b]$. Furthermore, f_1 and f_2 are step functions, hence are integrable, and from Proposition 1.44 we have

$$\int_a^b f_1=\sum_{j=1}^N m_j\Delta_j(P)=L(f;P) \tag{1.25}$$

and

$$\int_a^b f_2=\sum_{j=1}^N M_j\Delta_j(P)=U(f;P). \tag{1.26}$$

Let $\delta>0$ be such that

$$\mathcal{S}_\delta(f_1)\subseteq(L(f;P)-\epsilon,L(f;P)+\epsilon) \tag{1.27}$$

and

$$\mathcal{S}_\delta(f_2)\subseteq(U(f;P)-\epsilon,U(f;P)+\epsilon); \tag{1.28}$$

such δ exists by (1.25)–(1.26). Let (Q,T) be any pointed partition of $[a,b]$ of mesh less than δ. From (1.24) and the definition of “Riemann sum”, it is immediate that $S(f_1;Q,T)\le S(f;Q,T)\le S(f_2;Q,T)$. But $S(f_1;Q,T)\in\mathcal{S}_\delta(f_1)$ and $S(f_2;Q,T)\in\mathcal{S}_\delta(f_2)$. Thus, using (1.27)–(1.28), we have

$$L(f;P)-\epsilon<S(f_1;Q,T)\le S(f;Q,T)\le S(f_2;Q,T)<U(f;P)+\epsilon. \tag{1.29}$$

Now let $S_1, S_2 \in \mathcal{S}_\delta(f)$. By (1.29), both S_1 and S_2 lie in the interval $(L(f;P)-\epsilon, U(f;P)+\epsilon)$. Hence

$$|S_2 - S_1| < (U(f;P)+\epsilon) - (L(f;P)-\epsilon) = (U(f;P) - L(f;P)) + 2\epsilon < 3\epsilon$$

(using our initial hypothesis). Theorem 1.32 (specifically, the implication "(iii) $\Longrightarrow$ (i)") therefore implies that f is integrable. ■

Remark 1.47 Our proof shows that Proposition 1.45 is equivalent to the more easily visualized statement: A function $f \in \mathcal{B}([a,b])$ is integrable if and only if for each $\epsilon > 0$, there exist step-functions $f_1, f_2 : [a,b] \to \mathbf{R}$ such that $f_1(x) \le f(x) \le f_2(x)$ for all $x \in [a,b]$ (f is "squeezed" between f_1 and f_2) such that $\int_a^b (f_2 - f_1) = \int_a^b f_2 - \int_a^b f_1 < \epsilon$. ▲

Exercise 1.5 For any real-valued function f, the *positive part of* f, denoted f_+, and *negative part of* f, denoted f_-, are defined by $f_+(x) = \max\{f(x), 0\}$ and $f_-(x) = -\min\{f(x), 0\}$. (Thus both f_+ and f_- are non-negative, and $f = f_+ - f_-$ [why?].)

Parts (a) and (b) below can be done in either order: whichever part you do first, you may use to help you do the other part quickly. (But, obviously, you may not resort to circular reasoning.) To see how the result of (b) can be used to help with (a), compare the function f_+ with $|f| + f$.

(a) Prove that if f is integrable on $[a,b]$, then so are f_+ and f_-.

(b) Prove that if f is integrable on $[a,b]$ then so is $|f|$ (the function $x \mapsto |f(x)|$), and

$$\left| \int_a^b f \right| \le \int_a^b |f|. \tag{1.30}$$

The inequality (1.30) may be thought of as a "triangle inequality for integrals". ▲

1.4 Upper and lower integrals

This section is optional reading.

Definition 1.48 For any $f : [a,b] \to \mathbf{R}$, we define the *lower* and *upper Riemann integrals* of f over $[a,b]$ to be

$$\underline{\int_a^b} f = \sup\{L(f;P) : P \in \mathcal{P}([a,b])\},$$
$$\overline{\int_a^b} f = \inf\{U(f;P) : P \in \mathcal{P}([a,b])\},$$

respectively. We will frequently omit “Riemann” from this terminology, and may write the lower and upper integrals using dummy-variable notation, e.g. “$\underline{\int}_a^b f(x)\,dx$” for “$\underline{\int}_a^b f$.” ▲

In words: the lower integral is the *supremum* of *lower sums*, while the the upper integral is the *infimum* of *upper sums*

Now consider any fixed, arbitrary, $f \in \mathcal{B}([a,b])$. Using Proposition 1.28 and Lemma 1.29, we can express the limits of $U_\delta(f)$ and $L_\delta(f)$ (as $\delta \to 0$) as follows:

$$\begin{aligned}
\lim_{\delta\to 0} U_\delta(f) &= \inf_{\delta>0}(\sup\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}); \\
\lim_{\delta\to 0} L_\delta(f) &= \sup_{\delta>0}(\{\inf\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}).
\end{aligned}$$

For each $\delta > 0$ and $P \in \mathcal{P}_\delta([a,b])$, we have $U(f;P) \geq L(f;P)$. Hence, using Lemma 1.24,

$$\begin{aligned}
\inf_{\delta>0}(\sup\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}) &\geq \inf_{\delta>0}(\inf\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}) \\
&= \inf\left(\bigcup_{\delta>0}\{\{U(f;P) : P \in \mathcal{P}_\delta([a,b])\}\right) \\
&= \inf\{U(f;P) : P \in \mathcal{P}([a,b]\} \\
&= \overline{\int}_a^b f.
\end{aligned}$$

Thus

$$\begin{aligned}
\lim_{\delta\to 0} U_\delta(f) &\geq \overline{\int}_a^b f, && (1.31) \\
\text{and similarly} \qquad \lim_{\delta\to 0} L_\delta(f) &\leq \underline{\int}_a^b f. && (1.32)
\end{aligned}$$

We will show that the inequalities (1.31)–(1.32) can be sharpened to equalities when f is integrable, but some preliminary work is needed first.

Definition 1.49 Let P and Q denote partitions of $[a,b]$. We say Q is a *refinement* of P, or that Q *refines* P, if $P \subseteq Q$. The *common refinement* of P and Q is $P \cup Q$. ▲

Lemma 1.50 *Let* $f : [a,b] \to \mathbf{R}$.

(i) Let P and Q be partitions of $[a,b]$, and assume that Q refines P. Then

$$L(f;P) \leq L(f;Q) \leq U(f;Q) \leq U(f;P). \tag{1.33}$$

(ii) For any *partitions P, Q of $[a,b]$,*

$$L(f;P) \leq U(f;Q). \tag{1.34}$$

(iii) The upper and lower integrals satisfy

$$\underline{\int_a^b} f(x)\,dx \;\leq\; \overline{\int_a^b} f(x)\,dx. \tag{1.35}$$

Sketch of proof. (i) The middle inequality in (1.33) is simply (1.8). The first and third inequalities can be reduced to the case in which $P = \{a,b\}$ and $Q = \{a,c,b\}$, where the result is quickly established by comparing the Riemann sums associated with Q with those associated with P.

(ii) Let P and Q be partitions of $[a,b]$, and let R be their common refinement. Applying (i) twice, we obtain

$$L(f;P) \leq L(f;R) \leq U(f;R) \leq U(f;Q).$$

Hence (1.34) holds.

(iii) For each partition Q, taking the supremum over all partitions P in (1.34) yields

$$\underline{\int_a^b} f(x)\,dx \leq U(f;Q).$$

Taking the infimum over Q then yields (1.35). ■

Theorem 1.51 *A bounded function $f : [a,b] \to \mathbf{R}$ is integrable if and only if*

$$\underline{\int_a^b} f \;=\; \overline{\int_a^b} f. \tag{1.36}$$

In the integrable case,

$$\underline{\int_a^b} f \;=\; \overline{\int_a^b} f = \int_a^b f. \tag{1.37}$$

Proof: Let $f \in \mathcal{B}([a,b])$.

First suppose that f is integrable. Then by Proposition 1.11, f is bounded, so our analysis leading to (1.31)–(1.32) applies. These inequalities, together with (1.35), yield

$$\lim_{\delta\to 0} L_\delta(f) \le \underline{\int_a^b} f \ \le \ \overline{\int_a^b} f \ \le \ \lim_{\delta\to 0} U_\delta(f). \tag{1.38}$$

But from Theorem 1.32, since f is integrable, the leftmost and rightmost expressions in (1.38) are equal to each other and to $\int_a^b f$. Hence

$$\int_a^b f = \lim_{\delta\to 0} L_\delta(f) = \underline{\int_a^b} f \ = \ \overline{\int_a^b} f \ = \ \lim_{\delta\to 0} U_\delta(f).$$

This proves that the upper and lower integrals are equal, and establishes (1.37).

Conversely, suppose that $\underline{\int_a^b} f \ = \overline{\int_a^b} f$, and let A denote the value of these quantities. Let $\epsilon > 0$. Let P, Q be partitions of $[a,b]$ such that $L(f;P) > A-\epsilon$ and $U(f;Q) < A+\epsilon$; such partitions exist by the definition of lower and upper integrals. Let R be the common refinement of P and Q. Then, as in the proof of Lemma 1.50(ii), we have $L(f;P) \le L(f;R) \le U(f;R) \le U(f;Q)$. Hence

$$A-\epsilon < L(f;R) \le U(f;R) < A+\epsilon,$$

implying $U(f;R) - L(f;R) < 2\epsilon$. Since ϵ was arbitrary, Proposition 1.45 then implies that f is integrable. ■

Among the implications of Theorem 1.51 is that *if* $f : [a,b] \to \mathbf{R}$ is integrable, then the inequalities in (1.31) and (1.32) can be replaced by equalities. The student may well wonder whether equality holds in (1.31) and (1.32) even without the assumption of integrability. The answer is yes (this is one of several unrelated results each of which is sometimes given the name "Darboux's Theorem"), but the proof is not obvious, and we do not give it in these notes. We refer the interested student to [10, Section 18.2,Theorem VIII].

Remark 1.52 (Two approaches to the Riemann integral) Because Theorem 1.51 is true, equation (1.36) can be taken as the *definition* of "a bounded function f is (Riemann) integrable on $[a,b]$", in place of Definition 1.6, without changing either the set of functions being called "integrable" or the values of their integrals. If we use (1.36) to *define* what "integrable" means, then Theorem 1.51 yields the second sentence of Definition 1.6 as a *theorem* rather than a definition. Many (probably most) textbooks use this alternate definition of integrability, often phrased without any mention of Riemann sums (taking (1.21) to be the *definition* of $L(f;P)$ and $U(f;P)$). This approach has several advantages—for example, the definition of integrability is *much* simpler (there is no ϵ or δ; the mesh of a partition is never even mentioned), and many proofs can be done more efficiently.

However, there are also disadvantages[5] of using (1.36) instead of Definition 1.6 to define integrability. The chief *mathematical* disadvantage of the approach based on (1.36) is that the generalization to integrals of vector-valued functions is less natural (especially for functions taking values in an infinite-dimensional vector space). The other potential disadvantages are primarily pedagogical. One is that unless a proof of Theorem 1.51 is provided, the notion of "$\int_a^b f(x)\,dx$" in this approach does not clearly reduce to the notion that students learn in Calculus 1 (and again in Calculus 3, generalized to definite integrals of functions of two or three variables)—a notion that is completely correct, but that is usually not given a precise statement in Calculus 1-2-3 because students are not yet equipped to understand or appreciate the precise statement. Definition 1.6 is *exactly* the Calculus 1-2-3 notion of integrability, just defined precisely. It is this notion, rather than equation (1.36), on which all quantities defined through integrals in physics and other sciences are based.[6] Without Theorem 1.51, it is not clear that the "upper integral = lower integral" definition of integrability leads to the same notions of integration, or values of integrals, conceptualized in Calculus 1 (whether or not (1.21) is used to define upper and lower sums). Thus, some mathematicians find presentations of the Riemann integral that take (1.36) as definition, but do not include a proof of Theorem 1.51 (e.g. the presentation in [7]), to be unsatisfying. But when presentations that take (1.36) as definition *do* include a proof of Theorem 1.51 (as in [6, Theorem 6.14] and [10, Section 18.2]), some of the efficiency initially gained from the upper-integral/lower-integral definition is lost.

When (1.21) is used to define upper and lower sums, in addition to using (1.36) to define integrability, there is another efficiency-gain (the need to prove Proposition 1.36 is avoided), but offsetting are additional pedagogical disadvantages. One is that all connection to Riemann sums has been removed (unless prominent mention is made elsewhere in the presentation), putting even more distance between the integral defined this way and the integral as conceptualized in Calculus 1-2-3 and in the sciences. Another is that, using (1.21), we cannot even *define* upper and lower sums (and therefore upper and lower integrals), even within the extended reals, without restricting attention to

---

[5] Most of the "disadvantages" referred to here are pedagogical in nature, so should properly be called "disadvantages *in the opinion of the author of these notes*", but repeating such a mouthful in place of "disadvantages" everywhere would have made this discussion hard to read. The judgment of what is pedagogically better or worse is highly subjective. Every instructor is forced to make pedagogical choices, some aspect of which would be deemed disadvantageous by instructors making different choices. No criticism is intended of anyone whose pedagogical choices are different from those of the author of these notes.

[6] It is exactly these quantities in physics, including vector-valued integrals, for which the notion of "integral" was originally developed; calculus was invented in order to provide a mathematical description of physical laws. Examples of quantities in physics that, to this day, are understood by starting with Riemann sums, include work, centers of mass, moments of inertia, hydrostatic force, all line integrals in electricity and magnetism (E&M), and all flux integrals in E&M and in fluid dynamics.

functions that are at least *semi-bounded*: bounded above or bounded below. (In contrast, upper and lower sums as defined in Definition 1.22 *always* exist in the extended reals; no boundedness assumptions are needed.) Usually, to simplify presentations based on (1.21) and (1.36), a restriction is made to functions that are *bounded*, not just semi-bounded. Thus one loses Proposition 1.11. Instead of the non-integrability of unbounded functions being a consequence of the *concept* of the Riemann integral, unbounded functions are simply removed from consideration from the start (and students may reasonably wonder, "Why?"). Thus one of the chief deficiencies of the Riemann integral (as compared with the Lebesgue integral) cannot be *demonstrated*; it's been defined away. ▲

1.5 The integrability of continuous functions

Theorem 1.53 (Continuous functions are integrable) *If $f : [a,b] \to \mathbf{R}$ is continuous, then f is integrable on $[a,b]$.*

Proof: Let $f : [a,b] \to \mathbf{R}$ be a continuous function. Since $[a,b]$ is compact, f is bounded. Therefore, by Proposition 1.45, to prove that f is integrable it suffices to show that for each $\epsilon > 0$, there exists a partition P of $[a,b]$ such that $U(f;P) - L(f;P) < \epsilon$.

Let $\epsilon > 0$. Recall from MAA 4211 that every continuous function on a compact space is uniformly continuous. Since $[a,b]$ is compact and f is continuous, it follows that f is uniformly continuous. Let $\delta > 0$ be such that if $x, y \in [a,b]$ and $|x-y| < \delta$, then $|f(x) - f(y)| < \epsilon/(b-a)$; such δ exists since f is uniformly continuous. Let $P = \{x_0, \dots, x_N\} \in \mathcal{P}_\delta([a,b])$; such P exists by Remark 1.5.

Recall also from MAA 4211 the Extreme Value Theorem: every continuous real-valued function on a compact space attains a maximum value and a minimum value. In particular, this applies to $f|_{[x_{j-1},x_j]}$ for each $j \in \{1,\dots,N\}$. For each such j let m_j, M_j denote, respectively, the minimum and maximum values of $f|_{[x_{j-1},x_j]}$, and let $x_j', x_j'' \in [x_{j-1}, x_j]$ be such that $f(x_j') = m_j$ and $f(x_j'') = M_j$. Then, for each $j \in \{1,\dots,N\}$, we have $|x_j' - x_j''| \leq \mathrm{mesh}(P) < \delta$, so

$$M_j - m_j = f(x_j'') - f(x_j') < \frac{\epsilon}{b-a}\ .$$

But by Proposition 1.36, $L(f;P) = \sum_j m_j \Delta_j$ and $U(f;P) = \sum_j M_j \Delta_j$ (where $\Delta_j = \Delta_j(P)$). Hence

$$U(f;P) - L(f;P) = \sum_j (M_j - m_j)\Delta_j < \sum_j \frac{\epsilon}{b-a}\Delta_j = \frac{\epsilon}{b-a}\sum_j \Delta_j = \epsilon.$$

Since ϵ was arbitrary, we conclude from Proposition 1.45 that f is integrable. ■

Exercise 1.6 (a) Assume that $f : [a, b] \to \mathbf{R}$ is continuous, that $f(x) \geq 0$ for all $x \in [a, b]$, and that $f(x) > 0$ for some $x \in [a, b]$. Prove that $\int_a^b f > 0$.

(b) Assume that $f, g : [a, b] \to \mathbf{R}$ are continuous, that $f(x) \geq g(x)$ for all $x \in [a, b]$, and that $f(x) > g(x)$ for some $x \in [a, b]$. As a corollary of part (a), prove that $\int_a^b f > \int_a^b g$.

Remark 1.54 In particular, if $f : [a, b] \to \mathbf{R}$ is continuous and $f(x) > 0$ for all $x \in [a, b]$, then $\int_a^b f > 0$. (See Remark 1.20.) ▲

1.6 Additivity of the integral

The Riemann integral has an additivity property (unrelated to linearity) expressed by the following proposition.

Proposition 1.55 (Additivity of the integral) *Suppose $a < c < b$. A function $f : [a, b] \to \mathbf{R}$ is integrable on $[a, b]$ if and only if it is integrable on both $[a, c]$ and $[c, b]$. In the integrable case,*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{1.39}$$

Remark 1.56 As mentioned in Section 1.0, equation (1.39) reflects the principle that "integration is about adding stuff up": the "amount of stuff" between a and b is the "amount of stuff" between a and c plus the "amount of stuff" between c and b.▲

Proof of Proposition 1.55: Let $f_1 = f|_{[a,c]}$ and $f_2 = f|_{[c,b]}$.

First suppose that f_1 and f_2 are integrable, and let A and C, respectively, denote their integrals. Since f_1 and f_2 are integrable, they are bounded; hence so is f. Let $M > 0$ be such that $|f(x)| \leq M$ for all $x \in [a, b]$.

Let $\epsilon > 0$. Let $\delta_0 > 0$ be such $\mathcal{S}_{\delta_0}(f_1) \subseteq B_\epsilon(A)$ and $\mathcal{S}_{\delta_0}(f_2) \subseteq B_\epsilon(C)$; such δ_0 exists by the assumed integrability of f_1 and f_2. Let $\delta = \min\{\delta_0, \frac{\epsilon}{4M}\}$. Then $\mathcal{S}_\delta(f_1) \subseteq \mathcal{S}_\epsilon(f_1) \subseteq B_\epsilon(A)$ and $\mathcal{S}_\delta(f_2) \subseteq \mathcal{S}_\epsilon(f_2) \subseteq B_\epsilon(C)$ (since $\delta \leq \delta_0$) and $4M\delta \leq \epsilon$, facts we will use later.

Let $(P, T) = (\{x_0, \ldots, x_N\}, \{t_1, \ldots, t_N\})$ be a pointed partition of $[a, b]$ of mesh less than δ. Since $x_0 = a < x_1 \leq x_N = b$, the sets $\{j \in \{0, \ldots, N-1\} : x_j < c\}$ and $\{j \in \{1, \ldots, N\} : x_j > c\}$ are nonempty, so we may define

$$\begin{aligned} j' &= \max\{j \in \{0, \ldots, N-1\} : x_j < c\}, && (1.40) \\ j'' &= \min\{j \in \{1, \ldots, N\} : x_j > c\} && (1.41) \end{aligned}$$

(thus the value of $j'' - j'$ is either 2 or 1, accordingly as c is or is not an element of P). Define partitions P', P'' of $[a, c], [c, b]$, respectively, by

$$\begin{aligned} P' = (P \cap [a, c]) \cup \{c\} &= \{x_0, \ldots, x_{j'}, c\}, && (1.42) \\ P'' = \{c\} \cup (P \cap [c, b]) &= \{c, x_{j''}, \ldots, x_N\}; && (1.43) \end{aligned}$$

observe that $\operatorname{mesh}(P')$ and $\operatorname{mesh}(P'')$ are at most $\operatorname{mesh}(P)$, hence are less than δ. Define pointings T', T'' of P', P'', respectively, by

$$\begin{aligned} T' &= \{t_1, \ldots, t_{j'}, c\}, \\ T'' &= \{c, t_{j''+1}, \ldots, t_N\}. \end{aligned}$$

Observe that every term except possibly the last (respectively, first) in the sum defining $S(f_1; P', T')$ (resp., $S(f_2; P'', T'')$) is a term in the sum defining $S(f; P, T)$. Similarly, if $1 \leq j \leq j'$ or $j'' < j \leq N$, the term $f(t_j)\Delta_j(P)$ in the sum defining $S(f; P, T)$ is a term in either the sum defining $S(f_1; P', T')$ or the sum defining $S(f_2; P'', T'')$ (but not both). Note that since j'' is either $j' + 1$ or $j' + 2$, the only $j \in \{1, \ldots, N\}$ not satisfying $j \leq j'$ or $j > j''$ are $j' + 1$ and j'' (which are equal if $c \notin P$, and differ by 1 otherwise). Hence

$$\begin{aligned} &S(f; P, T) - (S(f_1; P', T') + S(f_2; P'', T'')) \\ &\quad = \begin{cases} f(t_{j''})(x_{j''} - x_{j'}) - f(c)(c - x_{j'}) - f(c)(x_{j''} - c) & \text{if } c \notin P \\ \\ f(t_{j'+1})(c - x_{j'}) + f(t_{j''})(x_{j''} - c) - f(c)(c - x_{j'}) - f(c)(x_{j''} - c) & \text{if } c \in P. \end{cases} \end{aligned} \tag{1.44}$$

Observe that the first line of (1.44) simplifies to $(f(t_{j''}) - f(c))(x_{j''} - x_{j'})$, while the second line simplifies to $(f(t_{j'+1}) - f(c))(c - x_{j'}) + (f(t_{j''}) - f(c))(x_{j''} - c)$. The numbers $c - x_{j'}$ and $x_{j''} - c$ lie in the interval $(0, \delta)$, as does $x_{j''} - x_{j'}$ in the "$c \notin P$" case. Since $|f(x) - f(y)| \leq |f(x)| + |f(y)| \leq M$ for all $x, y \in [a, b]$, and $2M < 4M$, the triangle inequality then shows that whichever line of (1.44) applies, we have

$$|S(f; P, T) - (S(f_1; P', T') + S(f_2; P'', T''))| \leq 4M\delta.$$

By the triangle inequality and the definition of δ, this implies that

$$\begin{aligned} |S(f; P, T) - (A + C)| &\leq |S(f; P, T) - (S(f_1; P', T') + S(f_2; P'', T''))| \\ &\quad + |S(f_1; P, T) - A| + |S(f_2; P, T) - C| \\ &< 4M\delta + 2\epsilon \\ &\leq 3\epsilon. \end{aligned} \tag{1.45}$$

Since (P, T) was an arbitrary pointed partition of mesh less than δ, the inequality (1.45) shows that $\mathcal{S}_\delta(f) \subseteq B_{3\epsilon}(A + C)$. Since ϵ was arbitrary, this establishes that f is integrable and that $\int_a^b f = A + C$, as desired.

Conversely, suppose that f is integrable on $[a, b]$. Then f is bounded. Let $\epsilon > 0$. Let P be a partition of $[a, b]$ such that $U(f; P) - L(f; P) < \epsilon$; such P exists by Proposition 1.45. Define indices j', j'' and partitions P', P'' (of $[a, c]$ and $[c, b]$, respectively) just as in (1.40)–(1.41) and (1.42)–(1.43). For each $j \in \{1, \ldots, N\}$, define m_j and M_j as in Proposition 1.36. Let $M' = \sup(f([x_{j'}, c]))$ and $m' = \inf(f([x_{j'}, c]))$; since $[x_{j'}, c] \subseteq [x_{j'}, x_{j'+1}]$ we have $M' \leq M_{j'+1}$ and $m' \geq m_{j'+1}$. Then, applying Proposition 1.36,

$$
\begin{aligned}
U(f_1;P') - L(f_1;P') &= \sum_{j=1}^{j'} (M_j - m_j)\Delta_j(P) + (M' - m')(c - x_{j'}) \\
&\leq \sum_{j=1}^{j'} (M_j - m_j)\Delta_j(P) + (M_{j'+1} - m_{j'+1})(x_{j'+1} - x_{j'}) \\
&= \sum_{j=1}^{j'+1} (M_j - m_j)\Delta_j(P) \\
&\leq \sum_{j=1}^{N} (M_j - m_j)\Delta_j(P) \\
&= U(f;P) - L(f;P) \\
&< \epsilon.
\end{aligned}
$$

Similarly, $U(f_2;P'') - L(f_2,P'') < \epsilon$. Since ϵ was arbitrary, Proposition 1.45 implies that f_1 and f_2 are integrable. ■

Corollary 1.57 *Let* $f : [a,b] \to \mathbf{R}$.

(a) The function f *is integrable if and only if its restriction to each closed subinterval of* $[a,b]$ *is integrable.*

(b) Suppose f *is integrable,* n *is a positive integer, and* $a < c_1 < c_2 \cdots < c_n < b$. *Then*

$$\int_a^b f = \int_a^{c_1} f + \int_{c_1}^{c_2} f + \cdots + \int_{c_n}^{b} f.$$

Exercise 1.7 Prove Corollary 1.57.

Definition 1.58 Let $a, b \in \mathbf{R}$, with $a \leq b$, and let f be a real-valued function on $[a,b]$.

(i) We define $\int_a^a f = 0$, and say that this integral exists.

(ii) If $b > a$, we say that $\int_b^a f$ exists if and only if $\int_a^b f$ exists, in which case we define $\int_b^a f = -\int_a^b f$.

▲

Corollary 1.59 *Let $a, b \in \mathbf{R}$ (with the possibilities $a < b$, $a = b$, $a > b$ all allowed).*

(i) Let $c \in \mathbf{R}$. Then $\int_a^b c = c(b-a)$.

(ii) Suppose that f is integrable on the closed interval with endpoints a and b, and that $|f(x)| \leq M$ for every x in this interval. Then

$$\left| \int_a^b f \right| \leq M|b-a|. \tag{1.46}$$

Proof: In view of Definition 1.58, it suffices to establish (i) and (ii) in the case $a < b$, so let us assume $a < b$. Then (i) follows from Example 1.13. For (ii), note that every Riemann sum of f over $[a, b]$ lies in the interval $[-M(b-a), M(b-a)]$. The definition of $\int_a^b f$ then implies that $\int_a^b f$ also lies in this interval. Therefore (1.46) holds. ■

Note that in Proposition 1.55, if both integrals on the right-hand side of equation (1.39) exist, then so does the integral on the left-hand side, while if the integral on the left-hand side exists, so do both of the integrals on the right-hand side. Hence if any two of the three integrals written in equation (1.39) exist, so does the third, and the equation holds true. Observe also that equation (1.39) can be rewritten as

$$\int_c^b f = \int_a^b f - \int_a^c f,$$

which, using Definition 1.58, can be further rewritten as

$$\int_c^b f = \int_c^a f + \int_a^b f, \tag{1.47}$$

which differs from (1.39) only by a permutation of the letters a, b, c. Furthermore, $\int_c^a f$ exists if and only if $\int_a^c f$ exists, so two of the three integrals in (1.39) exist if and only if two of the three integrals in (1.47) exist. However, if $a < c < b$, then in (1.47) we do not have $c < a < b$; equation (1.47) holds even though the limits of integration do not have the same order-relation as in Proposition 1.55. Pushing these ideas a little further leads to the following:

Corollary 1.60 *Let $a, b, c \in \mathbf{R}$ and let f be a real-valued function defined on an interval that includes a, b, and c. (No ordering or distinctness of a, b, c is assumed.) Then if any two of the three integrals $\int_a^c f$, $\int_a^b f$, $\int_b^c f$ exists, so does the third, and*

$$\int_a^c f = \int_a^b f + \int_b^c f; \tag{1.48}$$

equivalently,

$$\int_a^c f - \int_b^c f = \int_a^b f. \tag{1.49}$$

Exercise 1.8 Prove Corollary 1.60. (Do not forget to handle the cases in which two or three of the numbers a, b, c are equal.)

1.7 The Fundamental Theorem of Calculus

There are essentially two different, but closely related, theorems that go by the name "The Fundamental Theorem of Calculus"[7] (or, more historically, "The Fundamental Theorem of *Integral* Calculus"; this longer name is more descriptive but is rarely used anymore). One of these involves the integral of a derivative, and the other the derivative of an integral. More precisely, this is the primary distinction between the *conclusions* of these theorems. For each of these types of conclusions, there are actually more than one theorem, differing in their hypotheses. Of the various theorems that go by the name "The Fundamental Theorem of Calculus", we will prove the two that are of the greatest use in calculus, and refer to each of these two as "part of the Fundamental Theorem of Calculus".[8] Later, in optional reading for the student, we discuss some of the other, related, theorems that are sometimes called "The Fundamental Theorem of Calculus", and discuss the nomenclature for all these theorems.

The following simple lemma is needed for the statement of the first theorem we will prove.

Lemma 1.61 *Let $U \subseteq \mathbf{R}$ be an open interval, $f : U \to \mathbf{R}$ a continuous function, and $a, b \in U$. Then $\int_a^b f$ exists.*

Proof: If $a = b$ then, by definition, the integral exists and is 0. If $a \neq b$ then the restriction of f to $[\min\{a, b\}, \max\{a, b\}]$ is continuous, so by Theorem 1.53, the integral of f over this interval exists. If $a < b$ we are done; if $a > b$ the result follows from Definition 1.58. ■

This lemma assures us that the function F in the theorem below is indeed well-defined.

---

[7] Since there is more than one theorem called "The Fundamental Theorem of Calculus", it is tempting to refer to these theorems collectively as "The Fundamental Theorem*s* of Calculus". We choose not do so in these notes, however, since that terminology can give the impression that this group of theorems contains *all* the theorems that are fundamental to calculus, when in fact "The Fundamental Theorem of Calculus" is simply a historical name for one theorem and its relatives.

[8] Needless to say, we could state a single theorem that has each of these two theorems as a part, but the only motivation would be that the two theorems share a name. Combining them into a single theorem, each part of which has its own hypotheses, would be rather artificial, and would be inconvenient for the proofs.

Theorem 1.62 ("part of" The Fundamental Theorem of Calculus) *Let f be a continuous real-valued function on an open interval $U \subseteq \mathbf{R}$, and let $a \in U$. Define $F : U \to \mathbf{R}$ by*

$$F(x) = \int_a^x f(t)\,dt.$$

Then F is differentiable and $F' = f$.

Proof: Fix $x_0 \in U$, and let $x \in U$. Corollary 1.59(i) implies that $\int_{x_0}^x f(x_0)\,dt = f(x_0)(x - x_0)$ (here we are integrating the *constant* function $t \mapsto f(x_0)$). Using Corollary 1.60 in the form (1.49), we also have $F(x) - F(x_0) = \int_{x_0}^x f(t)dt$. Hence

$$F(x) - F(x_0) - f(x_0)(x - x_0) = \int_{x_0}^x f(t)\,dt - \int_{x_0}^x f(x_0)\,dt = \int_{x_0}^x (f(t) - f(x_0))\,dt.$$

Therefore for all $x \in U$ with $x \neq x_0$, we have

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| = \left|\frac{F(x) - F(x_0) - f(x_0)(x - x_0)}{x - x_0}\right| = \frac{\left|\int_{x_0}^x (f(t) - f(x_0))\,dt\right|}{|x - x_0|}. \tag{1.50}$$

Now let $\epsilon > 0$, and let $\delta > 0$ be such that for all $x \in U$ with $|x - x_0| < \delta$ we have $|f(x) - f(x_0)| < \epsilon$; such δ exists since f is continuous at x_0. Then for all $x \in B_\delta(x_0) \setminus \{x_0\}$, (1.50) and Corollary 1.59(ii) imply that

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| \leq \frac{\epsilon|x - x_0|}{|x - x_0|} = \epsilon.$$

Hence $\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$. Thus F is differentiable at x_0, and $F'(x_0) = f(x_0)$. Since x_0 was arbitrary, we are done. ■

An *antiderivative* of a function f on an open set U is a differentiable function F such that $F' = f$. An immediate corollary of Theorem 1.62 is:

Corollary 1.63 *Every continuous real-valued function f on an open interval has an antiderivative on that interval.*

Proof: Fix any $a \in U$. Then the function $x \mapsto \int_a^x f(t)\,dt$ is an antiderivative of f on U. ■

Theorem 1.64 ("part of" The Fundamental Theorem of Calculus) *Let $U \subseteq \mathbf{R}$ be an open interval, let $f : U \to \mathbf{R}$ be a continuous function, and let F be an antiderivative of f on U. Then for all $a, b \in U$,*

$$\int_a^b f(t)dt = F(b) - F(a). \tag{1.51}$$

Observe that Theorem 1.64 can be stated equivalently as follows:

Theorem 1.65 ("part of" The Fundamental Theorem of Calculus) *Let $U \subseteq \mathbf{R}$ be an open interval, and let $F : U \to \mathbf{R}$ be a differentiable function whose derivative F' is continuous. Then for all $a, b \in U$,*

$$\int_a^b F'(t)\,dt = F(b) - F(a). \tag{1.52}$$

Proof of Theorem 1.64: Fix $a \in U$, and define $G : U \to \mathbf{R}$ by $G(x) = \int_a^x f(t)\,dt$. By Theorem 1.62, $G' = f$. But by hypothesis, $F' = f$. Recall the following consequence of the Mean Value Theorem: If two differentiable functions H_1, H_2 on an open interval have identical derivatives, then $H_2 - H_1$ is constant (on that interval). Hence $G - F$ is constant. Therefore for all $x \in U$,

$$G(x) - F(x) = G(a) - F(a) = 0 - F(a) = -F(a),$$

so $G(x) = F(x) - F(a)$. Thus for any $b \in U$, $\int_a^b f(t)\,dt = G(b) = F(b) - F(a)$. ■

Remark 1.66 Theorem 1.64 (and even a stronger version) can be proven without the use of Theorem 1.62; the earlier theorem simply affords us a proof of Theorem 1.64 than is shorter than other proofs. See Theorem 1.69 and Exercise 1.10 later. ▲

Exercise 1.9 Evaluate $\lim_{n\to\infty} \frac{1}{n}\left[\left(\frac{1}{n}\right)^6 + \left(\frac{2}{n}\right)^6 + \left(\frac{3}{n}\right)^6 + \cdots + \left(\frac{n}{n}\right)^6\right]$.

Problems like the exercise above were common in high-school math-team competitions when the writer of these notes was in high school. Usually, students were given 2 minutes or so to solve such a problem. The trick is to recognize the sequence whose limit is being taken as a sequence of Riemann sums for an appropriate function over an appropriate interval, then use Exercise 1.1 and the Fundamental Theorem of Calculus.

Remark 1.67 "True" integration refers to what we call the "definite integral" in Calculus 1; it's about adding stuff up. This is true whether we are talking about the Riemann integral, a generalization called the Riemann-Stieltjes integral, improper integrals, or the Lebesgue integral. Nothing in the *concept* of integration involves differentiation. Archimedes already had this concept of integration as "adding up stuff" nearly two millennia before derivatives and integrals were defined, when he realized that the area inside a circle could be computed as the limit as $n \to \infty$ of the area of an inscribed regular n-gon. The Fundamental Theorem of Calculus (FTC) relates two completely distinct concepts: *integration* and *antidifferentiation*. Because we are able to compute antiderivatives of so many familiar functions, the FTC is a key tool in the computation of (definite) integrals.

It is *because* of the Fundamental Theorem of Calculus that antiderivatives are also called by a name, "indefinite integrals", that involves the word "integral". If you learned indefinite integration before definite integration, you may have received the false impression that "integration" always *means* "antidifferentiation". In this case, when learning about the Riemann-sum definition of the integral (either in Calculus 1 or in Advanced Calculus), you may have wondered, "What does this have to do with integration?" But you should now realize that this is the wrong question. Once you understand what integration actually means, but before you learn the FTC, the right question is "What does *antidifferentiation* have to do with integration?" This question is answered by the FTC.

If you learned indefinite integration before definite integration, another question you may have asked yourself is, "Where does this symbol '$\int$' " come from?" It comes from definite integration. The history of the symbol is that "$\int$" is an elongated S, the "S" standing for "sum". The reason that the same symbol is used for antiderivatives is, again, the FTC. ▲

Remark 1.68 Observe that the last statement of Theorem 1.62 can be written as

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x), \tag{1.53}$$

a statement about "the derivative of an integral" (more precisely, the derivative of a function defined by an integral in the specific way above), whereas equation (1.52)—"$\int_a^b F'(t)\,dt = F(b) - F(a)$"—is a statement about the integral of a derivative. Although equations (1.52) and (1.53) look different, they are actually equivalent once we know Corollary 1.63 (every continuous real-valued function on an open interval has an antiderivative). By simply changing notation, we can rewrite (1.52) as $\int_a^x F'(t)\,dt = F(x) - F(a)$. Since F in this equation is assumed differentiable, and $F(a)$ is just a constant, the right-hand side of this equation is differentiable in x; hence so is the left-hand side. Thus, given (1.52), we deduce that $\frac{d}{dx}\int_a^x F'(t)\,dt = \frac{d}{dx}(F(x) - F(a)) = F'(x)$. Since F' was assumed continuous in the hypotheses leading to (1.52), and since, every continuous function f : (open interval) $\to \mathbf{R}$ has *some* antiderivative F, there is no loss of generality if we replace F' (which was assumed continuous in (1.52)) in this last equation by an arbitrary continuous function f. But this yields (1.53).

Conversely, our proof of Theorem 1.64 shows that (1.53) implies (1.51), hence also implies (1.52). This equivalence is the reason that both Theorem 1.64 and Theorem 1.62 are often referred to by the same name, "The Fundamental Theorem of Calculus." However, as written, Theorem 1.62 is a stronger theorem than Theorem 1.64, since it implies that every continuous real-valued function on an open interval has an antiderivative, which cannot be deduced from Theorem 1.64. ▲

The remainder of this section is optional reading. (However, if you're wondering why Theorems 1.62 and Theorem 1.64 were given their names in these notes, the answer is contained in Remark 1.71.)

In Theorem 1.65, we assumed that the integrand F' was continuous. This hypothesis can be weakened to the assumption that F' is merely *integrable* over the appropriate interval, thereby obtaining the following stronger theorem (which can also reasonably be called "the Fundamental Theorem of Calculus"):

Theorem 1.69 *Let $U \subseteq \mathbf{R}$ be an open interval, let $a, b \in U$, and let $F : U \to \mathbf{R}$ be a differentiable function whose derivative F' is integrable over the interval with endpoints a and b. Then equation (1.52) holds.*

Exercise 1.10 (Optional.) Prove Theorem 1.69. *Hint*: It suffices to prove the result in the case $a < b$ (why?). Assume $a < b$. For any partition $P = \{x_0, \dots, x_N\}$ of $[a, b]$, observe that $F(b) - F(a) = \sum_{j=1}^{N}(F(x_j) - F(x_{j-1}))$. Apply the Mean Value Theorem to F on each interval $[x_{j-1}, x_j]$, deduce that (for any partition P), $L(f; P) \le F(b) - F(a) \le U(f; P)$. Now apply an appropriate result that we proved earlier.

Theorem 1.62 can also be strengthened:

Theorem 1.70 *If $f \in \mathcal{R}([b, c])$, and $a \in [b, c]$, then the function $F : [b, c] \to \mathbf{R}$ defined by*

$$F(x) = \int_a^x f(t)\,dt$$

is continuous. If, in addition, f is continuous at $x_0 \in (b, c)$, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.

Observe that the proof we gave of Theorem 1.62 actually proves the second assertion in Theorem 1.70; we simply replace the *arbitrary* point x_0 in that proof by the *specific* point x_0 in the statement of Theorem 1.70. What is new in Theorem 1.70 is really the *first* assertion: that if we assume only that f is *integrable* (rather than *continuous*) on a closed, bounded interval containing a, we can still deduce something about the function F, namely that it is continuous.

Exercise 1.11 Prove the first assertion in Theorem 1.70.

Most textbooks on introductory or advanced calculus state only Theorems 1.64 and 1.62; usually Theorems 1.69 and 1.70 are stated only in more advanced textbooks on analysis.

Remark 1.71 (Naming the theorems in this section) For purposes of this Remark, minor differences in the statements of the theorems under discussion are ignored.

If you ask different mathematicians (or even the same mathematician at different times), "What is the Fundamental Theorem of Calculus?" you will get different answers. The answer the mathematician chooses to give may also depend on the level of who's asking the question. In some textbooks, Theorem 1.64 is called the Fundamental Theorem of Calculus (FTC) or the Fundamental Theorem of Integral Calculus (FTIC), and Theorem 1.62 is stated but not given a name (e.g. [8, 11]). In other textbooks, exactly the opposite is true: Theorem 1.62 is called the FTC, and Theorem 1.64 is stated but not given a name (e.g. [5]).

This is not where the name-discrepancies end. The first edition of Apostol's *analysis* textbook [2] calls Theorem 1.64 the FTIC, and states and proves a generalized version[9] of Theorem 1.70, but does not give a name to the latter theorem. The second edition of Apostol's analysis textbook, [3], calls Theorem 1.64 the *Second* FTIC. In this edition, Apostol still does not give his version of Theorem 1.70 a name, but says afterwards that part (iii) of this theorem—the only part that gives a relation between integration and differentiation—is "sometimes called the *first* FTIC" in the special case of the pure Riemann integral. Apostol's *calculus* textbook [1] calls Theorem 1.62 the First FTC, and Theorem 1.64 the Second FTC. Some textbooks implicitly (but never explicitly) combine Theorems 1.62 and 1.64 into one theorem[10], by calling Theorem 1.62 the "FTC, part 1", and call Theorem 1.64 the "FTC, part 2" (e.g. [4, 9]). Some textbooks effectively reverse this numbering, calling Theorem 1.62 the "FTC–*Second* Form" and Theorem 1.64 the "FTC–*First* Form". (Thus, among authors calling Theorems 1.62 and 1.64 parts or forms of the same theorem, there is inconsistency about which part/form is the first, and which is the second.) In [10] (Taylor & Mann), no theorem is given a name that includes "FTC"; Theorems 1.64 and 1.62 are stated but not given names. Rudin [6, 7] states only the stronger versions of Theorems 1.64 and 1.62 (Theorems 1.69 and 1.70), calls Theorem 1.69 the FTC, and does not give a name to Theorem 1.70.

What all the theorems whose textbook names include "FTC" (when the theorems are named at all) have in common is that they are theorems about, and only about,

---

[9] In [2] and [3], Apostol works with a generalized version of the Riemann integral called the Riemann-Stieltjes integral. In some cases, but not all, he specifically states what various theorems reduce to for the Riemann integral. A similar comment applies to Rudin [6, 7].

[10] Presumably, authors' reasons for never doing this explicitly are similar to those mentioned in footnote 8.

an "inverse" relationship of differentiation and integration (their conclusions are purely about the derivative of an integral or the integral of a derivative). Sometimes you may see "FTC" included in the name of Theorem 1.70, but this is less conventional than leaving the theorem un-named, because the first assertion of Theorem 1.70 has nothing to do with a relation between integration and differentiation. (Apostol's treatment in [3], in which he says only that *part* of his un-named version of Theorem 1.70 is called the First FTIC, is more conventional.)

The upshot is that students should take none of these name-variants as gospel. It is okay to call any of Theorems 1.62, 1.64, 1.69, and 1.70 the FTC, or part of the FTC, or a form or version of the FTC.

The writer of these notes generally regards Theorem 1.65 (equivalently, Theorem 1.64) as the "true" FTC, but when there is a need to refer to the (extremely important) Theorem 1.62, he uses language that implicitly combines Theorem 1.62 and Theorem 1.64 into one grand FTC; hence the name given here to Theorem 1.62. (This is similar to the approach that calls Theorems 1.64 and 1.62 "parts 1 and 2"—in whatever order—of the FTC.) Theorem 1.65 has a place in mathematics that is rather more special than that of Theorem 1.62, in several respects:

- Theorem 1.65 is the first of several important theorems, covered in a traditional Calculus 3 course, that have a certain formal similarity that is actually very deep. Other theorems in this collection are the "Fundamental Theorem of Line Integrals", Green's Theorem, Stokes's Theorem, and the Divergence Theorem. Each of these theorems pertains to integrating a suitably defined derivative "$d\omega$" of a suitably defined object ω over a "nice" n-dimensional set S with $(n-1)$-dimensional boundary ∂S ($n = 1, 2$, or 3), and makes a statement of the form

  $$n\text{-dimensional integral of } d\omega \text{ over } S = (n-1)\text{-dimensional integral of } \omega \text{ over } \partial S.$$

  (For these purposes, the 0-dimensional integral of a function $F : [a, b] \to \mathbf{R}$ is simply $F(b) - F(a)$.) This collection of theorems generalizes to a single theorem, also called Stokes's Theorem, that holds for all $n \geq 1$ (not just $n = 1, 2, 3$). The FTC (in the form 1.65) is simultaneously a special case of this generalized version of Stokes's Theorem, and a key step in its proof. This more general Stokes's Theorem is extremely important on its own, but is also the inspiration for a large subject in algebraic topology called "homology and cohomology theory".

- As discussed in Remark 1.67, Theorem 1.64 (equivalent to Theorem 1.65) is the reason that we use the symbol "$\int$" for antiderivatives.

▲

1.8 Change of variable

Definition 1.72 Let $U \subseteq \mathbf{R}$ be open. A function $g : U \to \mathbf{R}$ is *continuously differentiable* if g is differentiable and g' is continuous. ▲

Proposition 1.73 (Change-of-variable in one-dimensional integrals) *Let $U, I \subseteq \mathbf{R}$ be open intervals, $f : U \to \mathbf{R}$ a continuous function, $\varphi : I \to U$ a continuously differentiable function. Then for any $a, b \in I$,*

$$\int_{\varphi(a)}^{\varphi(b)} f = \int_a^b (f \circ \varphi)\, \varphi' \; ; \tag{1.54}$$

equivalently, in "dummy-variable notation",

$$\int_{\varphi(a)}^{\varphi(b)} f(u)\, du = \int_a^b f(\varphi(x))\, \varphi'(x)\, dx. \tag{1.55}$$

Proof: Fix $a \in I$. Define $F : U \to \mathbf{R}$ by $F(y) = \int_{\varphi(a)}^{y} f$. Then, by Theorem 1.62, F is differentiable and $F' = f$. Define $G : I \to \mathbf{R}$ by $G = F \circ \varphi$. Then G is the composition of differentiable functions, so G is differentiable and $G' = (F' \circ \varphi)\varphi' = (f \circ \varphi)\varphi'$. The function $(f \circ \varphi)\varphi'$ is continuous (why?), so by Theorem 1.62, the function $H : I \to \mathbf{R}$ defined by $H(x) = \int_a^x (f \circ \varphi)\varphi'$ is differentiable, and $H' = (f \circ \varphi)\varphi' = G'$. Since I is an interval, "$H' = G'$" implies that $G - H$ is constant (see the proof of Theorem 1.64). Hence for all $x \in I$,

$$G(x) - H(x) = G(a) - H(a) = F(\varphi(a)) - H(a) = \int_{\varphi(a)}^{\varphi(a)} f - \int_a^a (f \circ \varphi)\varphi' = 0 - 0 = 0.$$

Hence $G(x) = H(x)$ for all $x \in I$. In particular, $G(b) = H(b)$, which is exactly equation (1.54). ■

Remark 1.74 Writing the result of Proposition 1.73 in the form (1.55) explains the terminology "change of variable"; we think of the dummy variables in (1.55) as being related by the equation $u = \varphi(x)$. However, observe that in this proposition, φ need not be one-to-one. This is a remarkable feature of the "one-dimensional" change-of-variables formula that is not shared by the change-of-variables formula for multiple integrals (the last topic we will study in this course, if we complete the syllabus). ▲

Remark 1.75 (Helpfulness of Leibniz notation when changing variables) In the Leibniz notation for the derivative of a function f, names are chosen for the independent and dependent variables—say x and y, respectively, related by the equation $y = f(x)$;

sometimes we simply write "$y = y(x)$." With this choice of variables, the Leibniz notation for $f'(x)$ is $\frac{dy}{dx}$ (in which we must remember that the right-hand side is not actually a fraction with real numbers in numerator and denominator). In some situations, this notation can lead to problems; in others, it is extremely helpful. The change-of-variables formula(s) in Proposition 1.73 is an instance in which the Leibniz notation is a truly marvelous mnemonic device. In place of introducing a *name* φ for the functional relation between u and x that we're thinking of when we write the formula (1.55), we simply write "$u(x)$" in place of $\varphi(x)$ on the right-hand side, and write $\frac{du}{dx}$ in place of $\varphi'(x)$. In the limits of integration on the left-hand side, instead of writing "$\varphi(a)$" and "$\varphi(b)$", we could write $u(a)$ and $u(b)$, but—since we are thinking of this as a change of variables—we often write "$u = u(a)$" and "$u = u(b)$" instead. For the sake of symmetry, we often use similar notation for the limits of integration on the right-hand side. Equation (1.55) then becomes

$$\int_{u=u(a)}^{u=u(b)} f(u)\, du = \int_{x=a}^{x=b} f(u(x)) \frac{du}{dx}\, dx,$$

or, even more familiarly,

$$\int_{x=a}^{x=b} f(u(x)) \frac{du}{dx}\, dx = \int_{u=u(a)}^{u=u(b)} f(u)\, du. \tag{1.56}$$

In other words, if we simply pretend that $\frac{du}{dx}$ in (1.56) is a true fraction, whose denominator can be cancelled by the "dx" appearing to its right, then it appears "obvious" that the left-hand side of (1.56) equals the right-hand side. While this logic for equating the left-hand side with the right-hand side is completely bogus, it does allow us to *remember* (1.54) and (1.55)—*which we have rigorously proven*—more easily. This is a tremendous benefit, and both student and seasoned mathematician alike have no reason to be embarrassed by relying on the above "abuse of notation" (pretending that $\frac{du}{dx}$ is a fraction, etc.) to help *remember* (1.54) and (1.55). Just keep in mind that a valid *proof* is needed to deduce that (1.56) is correct; "proof by abuse of notation" (or "proof by misunderstanding notation") is not a valid method of proof. ▲

Remark 1.76 As the student will recall from Calculus 1, Proposition 1.73 is a useful tool for the evaluation of integrals. Reasonable names for this tool are "integration by substitution" and "changing variables in the integral".[11] Calculus 1-2-3 tomes currently on the market usually call this technique, and its analog for indefinite integrals, by the abysmal name "u-substitution". You will not find this terminology in older, "classic" textbooks such as [1–3, 8, 11], or in Rosenlicht [5]. In older books, the technique is named according to the *concept* of substitution, rather than a *letter* that is commonly used

---

[11] However, once we learn about changing variables in multiple integrals—a topic at the end of this course, if we complete the syllabus—we will see that "changing variables" is not a great description of (1.55) unless φ restricts to a bijection from the interval with endpoints a, b to the interval with endpoints $\varphi(a), \varphi(b)$.

in substitutions. Calling this technique "u-substitution" is like calling every function $f : (\text{subset of } \mathbf{R}) \to \mathbf{R}$ an "x-function". ▲

1.9 Integration of vector-valued functions

This section is an expanded version of Rosenlicht's homework problem VI.6, a problem that illustrates the generality and several strengths of the Riemann-sum approach to the Riemann integral.

Throughout this section, $(V, \| \;\|)$ denotes a complete normed vector space[12], with the associated metric d. Usually we will write simply V rather than $(V, \| \;\|)$, with understanding that V has been given a fixed norm $\| \;\|$ for which the metric space (V, d) is complete. We will write 0_V for the zero element of V. For $c \in \mathbf{R}$ and $v \in V$, we define "vc" to mean cv. Open balls in V will generally be denoted by notation of the form "$B_\epsilon(v)$", but in situations in which balls in V and balls in $\mathbf{R}$ both enter the discussion, we put an appropriate superscript V or $\mathbf{R}$ on the "B".

We will extend the theory of the Riemann integral from the realm of real-valued functions to the realm of vector-valued functions, by which we mean functions from an interval $[a, b]$ to a (complete, normed) vector space V. We do not assume that V is finite-dimensional, except where noted. However, the case $V = \mathbf{R}^n$ (with, say, the Euclidean norm) is an important special case, and it is very helpful to keep this case in mind when trying to grasp what various definitions, propositions, etc., are saying.

Definition 1.77 (Riemann sums) Let $f : [a, b] \to V$ be a function and let $(P, T) = (P, \{t_1, \ldots t_N\})$ be a pointed partition of $[a, b]$. The *Riemann sum* for f corresponding to (P, T) is

$$S(f; P, T) = \sum_{j=1}^{N} f(t_j)\Delta_j(P)\,. \tag{1.57}$$

As we did for real-valued functions, we will write

$$\mathcal{S}(f; P) \quad = \quad \{S(f; P, T) : T \text{ is a pointing of } P\},$$

and for each $\delta > 0$, write

$$\mathcal{S}_\delta(f) = \bigcup \{\mathcal{S}(f; Q) : Q \in \mathcal{P}_\delta([a, b])\}.$$

▲

---

[12] A complete normed vector space is called a *Banach space*, but to help the student keep in mind the important features we are assuming of our $(V, \| \;\|)$, we will stick to the self-descriptive term "complete normed vector space".

Note that there is *no difference* between the definitions (1.2) and (1.57) of Riemann sums, except that in (1.57) the function f is taking its values in V rather than $\mathbf{R}$. The same definition would work with V replace by *any* vector space (whether or not normed or complete); all that is needed for the definition (1.57) is the vector-space structure on V. For the next definition, we need only a little more: the metric structure on V given by a norm. This definition could be written exactly as Definition 1.6, simply replacing the absolute-value symbols by norm-symbols, but we will use our notation "$\mathcal{S}_\delta(f)$" to state the definition more efficiently (as we did for real-valued functions in Definition 1.10).

Definition 1.78 (Integrability) A function $f:[a,b]\to V$ is *(Riemann) integrable* if there is a vector $A\in V$ such that for each $\epsilon>0$ there exists $\delta>0$ such that $\mathcal{S}_\delta(f)\subseteq B_\epsilon(A)$. More generally, if f is a V-valued function whose domain includes $[a,b]$, we say that f *is integrable on* $[a,b]$ (or *over* $[a,b]$) if $f|_{[a,b]}$ is integrable. ▲

We continue our convention (for these notes) that "integrable" means "Riemann integrable" and that all integrals we discuss are Riemann integrals.

If there exist distinct $A,A'\in V$ both satisfying the condition satisfied by A in Definition 1.78, then for $\epsilon=\|A-A'\|/2$ and $S\in V$ we cannot have both $\|S-A\|<\epsilon$ and $\|S-A'\|<\epsilon$ (the triangle inequality would lead to a contradiction). Therefore if, just as for real-valued functions, if f is integrable on $[a,b]$ then there is a *unique* $A\in V$ satisfying the condition in Definition 1.78. Thus we can define the integral of f exactly as in Definition 1.8, just with $\mathbf{R}$ replaced by V:

Definition 1.79 Let $f:[a,b]\to V$ be integrable. We define the *integral of f* to be the unique $A\in V$ satisfying the condition given in Definition 1.6, and denote this A as $\int_a^b f$ or as $\int_a^b f(x)dx$, etc. for any dummy variable. More generally, if f is a V-valued function on a domain that includes $[a,b]$, and f is integrable on $[a,b]$, we use the notation $\int_a^b f$ (or $\int_a^b f(x)dx$, etc.) for the integral of $f|_{[a,b]}$, and refer to the value of this integral as *the integral of f over* $[a,b]$. We define the phrase "$\int_a^b f$ exists" (or "$\int_a^b f(x)\,dx$ exists", etc.) to mean that f is integrable on $[a,b]$. ▲

Notation 1.80 We let $\mathrm{Func}([a,b],V)$ denote the set of *all* functions $[a,b]\to V$, and let $\mathcal{R}([a,b],V)\subseteq\mathrm{Func}([a,b],V)$ denote the set of integrable functions from $[a,b]$ to V.

The set $\mathrm{Func}([a,b],V)$ is itself a vector space, with zero element the constant function $x\mapsto 0_V$, and with the vector-space operations defined through pointwise operations: for $f,g\in\mathrm{Func}([a,b],V)$ and any $c\in\mathbf{R}$, we define elements $f+g$ and cf of $\mathrm{Func}([a,b],V)$ by $(f+g)(x):=f(x)+g(x)$ and $(cf)(x):=cf(x)$ for all $x\in[a,b]$.

Remark 1.81 If $\dim(V) = 0$, then $V = \{0_V\}$ and $\mathrm{Func}([a, b], V)$ contains only the constant function $x \mapsto 0_V$. All Riemann sums of this function have the value 0_V. Hence this function is integrable, and the value of the integral is 0_V.

Thus, in a discussion of integrating vector-valued functions, the 0-dimensional vector space is not interesting. We have not excluded it from our discussion, though, since a restriction of the form "Assume $\dim(V) \geq 1$" might give the impression that something goes wrong if $\dim(V) = 0$, rather than that this case is simply uninteresting. ▲

Exercise 1.12 Recall that two norms $\|\ \|_1, \|\ \|_2$ on V are called *equivalent* if there exist real numbers $c_1, c_2 > 0$ such that for all $v \in V$ we have $\|v\|_2 \leq c_1\|v\|_1$ and $\|v\|_1 \leq c_2\|v\|_2$. Show that if the given norm $\|\ \|$ on V is replaced by any equivalent norm, neither the set $\mathcal{R}([a, b], V)$ nor the value of any integral changes. ▲

For Exercises 1.13, 1.14, and 1.15 below, you simply need to go through the proofs of the corresponding statements for real-valued functions, and observe that if you replace absolute-value symbols (if they occur at all) by norm-symbols, the same arguments work verbatim.

Proposition 1.82 ("Integrable implies bounded") *If $f : [a, b] \to V$ is integrable, then f is bounded.*

Exercise 1.13 Prove Proposition 1.82.

Exercise 1.14 Establish the analog of Example 1.13 for V-valued functions: For any $v \in V$, the constant function $f : [a, b] \to V$ given by $f(x) = v$ is integrable, and

$$\int_a^b v\, dx = (b - a)v.$$

Proposition 1.83 (linearity of the integral) *The set $\mathcal{R}([a, b], V)$ is a vector space (a vector subspace of* $\mathrm{Func}([a, b], V)$*), and the map $\mathcal{R}([a, b], V) \to V$ defined by $f \mapsto \int_a^b f$ is linear.*

Exercise 1.15 Prove Proposition 1.83.

Since a general vector space is not an ordered set (statements such as "$v < w$" for $v, w \in V$ are *meaningless* unless $V = \mathbf{R}$), there are no analogs of Proposition 1.18 or Corollary 1.19 for V-valued functions (for general V). For the same reason, there are no analogs of upper and lower sums. However, we used upper and lower sums only as a tool to simplify proofs and to aid in visualization of certain facts. Most facts about integrable real-valued functions that do not *explicitly* (i.e. in their statements, not just their proofs) rely on the fact that **R** is ordered, *do* generalize to V-valued functions. For some of these facts, we will have to use a different proof-strategy, since we often used the fact that **R** is ordered as a crutch to simplify proofs (and often to gain useful insight!). The proofs of results such as "continuous implies integrable", given in this section for V-valued functions, would have worked just as well earlier for **R**-valued functions.

To get rid of our reliance on upper and lower sums in various proofs, we need to establish Theorem 1.32's "(i) $\Longleftrightarrow$ (iv)" implication, a "Cauchy-like" criterion for integrability, in a way that does not use the order structure of **R** (in particular, a way that does involve statement (ii) or (iii) of that theorem). We do this in the following proposition.

Proposition 1.84 *Let* $f \in \mathrm{Func}([a,b],V)$*. Then* f *is integrable if and only if, for each* $\epsilon > 0$*, there exists* $\delta > 0$ *such that for all* $S_1, S_2 \in \mathcal{S}_\delta(f)$*, we have* $\|S_1 - S_2\| < \epsilon$*.*

Proof: First assume that f is integrable. Let $A = \int_a^b f$ and let $\epsilon > 0$. Let $\delta > 0$ be such that $\mathcal{S}_\delta(f) \subseteq B_{\epsilon/2}(A)$. Then for all $S_1, S_2 \in \mathcal{S}_\delta(f)$ we have

$$\|S_1 - S_2\| = d(S_1, S_2) \leq d(S_1, A) + d(A, S_2) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This proves the "only if" assertion of the proposition.

Conversely, assume that for each $\epsilon > 0$, there exists $\delta > 0$ such that for all $S_1, S_2 \in \mathcal{S}_\delta(f)$, we have $\|S_1 - S_2\| < \epsilon$. For each $n \in \mathbf{N}$, select a Riemann sum $S^{(n)} \in \mathcal{S}_{1/n}(f)$. Let $\epsilon > 0$, and let δ be such that for all $S_1, S_2 \in \mathcal{S}_\delta(f)$, we have $\|S_1 - S_2\| < \epsilon$. Let $N \in \mathbf{N}$ be any integer greater than $1/\delta$. Then for all $n, m \geq N$ the Riemann sums $S^{(n)}, S^{(m)}$ both lie in $\mathcal{S}_\delta(f)$, so $\|S^{(n)} - S^{(m)}\| < \epsilon$. Therefore the sequence $\vec{S} := (S^{(n)})_{n=1}^{\infty}$ in (V, d) is Cauchy. Since (V, d) is complete, this sequence converges; let $A_{\vec{S}}$ denote its limit. (The notation indicates a possibility that we have not yet eliminated but will shortly, namely that this limit could depend on the choice of the sequence $\vec{S}$.)

Again let $\epsilon > 0$ be arbitrary, and now let $\delta > 0$ be such that for all $S_1, S_2 \in \mathcal{S}_\delta(f)$, we have $\|S_1 - S_2\| < \frac{\epsilon}{2}$. Let $N \in \mathbf{N}$ be such that $N > \frac{1}{\delta}$ and $\|S^{(N)} - A_{\vec{S}}\| < \frac{\epsilon}{2}$; such N exists since $(S^{(n)})$ converges to $A_{\vec{S}}$. For every $S \in \mathcal{S}_\delta(f)$ we then have

$$d(S, A_{\vec{S}}) \leq d(S, S^{(N)}) + d(S^{(N)}, A_{\vec{S}}) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence f is integrable. (Furthermore, $A_{\vec{S}} = \int_a^b f$, so the value of this limit is independent of the choice of sequence $\vec{S}$). ■

Exercise 1.16 Show that the statements in Exercise 1.1 for functions $f : [a,b] \to \mathbf{R}$ also hold for functions $f : [a,b] \to V$.

Exercise 1.17 Prove that the analog of Proposition 1.55, "Additivity of the integral", holds for V-valued functions. The first half of the proof of Proposition 1.55 can be mimicked fairly easily. For the second half, which made use of upper and lower sums, you will need to figure out how to use Proposition 1.84 in place of Proposition 1.45, the "Step-function lemma".

Definition 1.85 Let V^* denote the set of *continuous* linear transformations from V to$\mathbf{R}$. (V^* is called the *dual space* or *continuous dual* of V.)

The handout "Some notes on normed vector spaces" (http://dgarchive.com/classes/4212_s19/misc_handouts/normed_vector_spaces.pdf) proves, among other things, several facts we will need concerning linear transformations from V to $\mathbf{R}$. We collect these here into a proposition so that we may refer to them easily:

Proposition 1.86

(a) A linear transformation $\xi : V \to \mathbf{R}$ is continuous if and only if there exists a real number K such that

$$|\xi(v)| \leq K\|v\| \quad \text{for all } v \in V. \tag{1.58}$$

(b) If V is finite-dimensional*, then every linear transformation from V to $\mathbf{R}$ is continuous.*

(c) If V is finite-dimensional*, then any two norms on V are equivalent.*

We mention a few things in passing:

- The "normed vector spaces" handout actually proves (a) and (b) for linear transformations from V to any normed vector space, not just $\mathbf{R}$.
- Facts (b) and (c) are false if is V infinite-dimensional.
- For fact (a), all that we will need is the "only if" part.

One additional fact that we will use, not mentioned in the "normed vector spaces" handout, is that a finite-dimensional vector space, endowed with any norm, is complete. This follows from facts proven (one hopes) in MAA 4211: (i) If d_1, d_2 are equivalent metrics on a set E ("equivalence" being defined the same way as for norms on vector spaces), then (E, d_1) is complete if and only if (E, d_2) is complete. (ii) If two norms on a vector space are equivalent, so are their associated metrics. (iii) If V has finite dimension $n \geq 1$, and $\| \ \|$ is the ℓ^∞ (or the ℓ^2) norm determined by some choice of basis, then V is complete with respect to the associated metric.

Thus, every finite-dimensional normed vector space is a complete normed vector space.

Returning to general V (not necessarily finite-dimensional): given any $f \in \mathrm{Func}([a,b],V)$ and any $\xi \in V^*$, the composition $\xi \circ f$ is a real-valued function on $[a,b]$. The next proposition relates the integrability, and the integrals, of the V-valued function f and the real-valued function $\xi \circ f$. Before we state the proposition, the student should do the following easy exercise relating *Riemann sums* of the V-valued function f and the real-valued function $\xi \circ f$.

Exercise 1.18 Show that for any $f \in \mathrm{Func}([a,b],V)$ and $\xi \in V^*$, and any pointed partition (P,T) of $[a,b]$,

$$\xi\left(S(f;P,T)\right) = S(\xi \circ f;P,T). \tag{1.59}$$

Proposition 1.87 *If $f \in \mathcal{R}([a,b],V)$, then for every $\xi \in V^*$ we have $\xi \circ f \in \mathcal{R}([a,b])$, and*

$$\xi\left(\int_a^b f\right) = \int_a^b \xi \circ f. \tag{1.60}$$

Proof: Let $f \in \mathcal{R}([a,b],V)$, let $\xi \in V^*$, and let $A = \int_a^b f$. Let $K > 0$ be such that (1.58) is satisfied. Let $\epsilon > 0$, let $\epsilon_1 = \frac{\epsilon}{K}$, and let $\delta > 0$ be such that $\mathcal{S}_\delta(f) \subseteq B^V_{\epsilon_1}(A)$.

Now let (P,T) be a pointed partition of $[a,b]$ of mesh less than δ. Then $S(f;P,T) \in B^V_{\epsilon_1}(A)$, and, using equation (1.59),

$$\begin{aligned} |S(\xi \circ f;P,T) - \xi(A)| = |\xi\left(S(f;P,T)\right) - \xi(A)| &= |\xi\left(S(f;P,T) - A\right)| \\ &\leq K\|S(f;P,T) - A\| \\ &< K\epsilon_1 = \epsilon, \end{aligned}$$

so $S(\xi \circ f;P,T) \in B^{\mathbf{R}}_\epsilon(\xi(A))$.

Hence $\mathcal{S}_\delta(\xi \circ f) \subseteq B^{\mathbf{R}}_\epsilon(A)$. Since ϵ was arbitrary, it follows that $\xi \circ f \in \mathcal{R}([a,b])$ and that 1.60 holds. ■

Results much stronger than Proposition 1.87 are true if V is finite-dimensional. We show one of these next, and deduce as a corollary that if V is finite-dimensional, the "if …then" in Proposition 1.87 can be strengthened to "if and only if". For the next few pages, to make visually clear which objects are elements of V and which are real numbers, we will use boldface for elements of V and for V-valued functions. (However, **R** still denotes the reals!)

A function $\mathbf{f} : [a,b] \to \mathbf{R}^n$ is often written in the form $(f_1,\ldots,f_n)$, where the f_i are real-valued functions on $\mathbf{R}^n$. We can also write $(f_1,\ldots,f_n)$ as $\sum_{i=1}^n f_i\mathbf{e}_i$, where $\{\mathbf{e}_i\}_{i=1}^n$ is the standard basis of $\mathbf{R}^n$ ($\mathbf{e}_i$ is the vector whose i^{th} coordinate is 1, and all of whose other coordinates are 0). The *coordinate functions determined by the basis* $\{\mathbf{e}_i\}_{i=1}^n$ are exactly the usual coordinate functions $\{x_i : \mathbf{R}^n \to \mathbf{R}\}_{i=1}^n$ (the functions defined by $x_i(a_1, a_2, \ldots, a_n) = a_i$). Observe that $f_i = x_i \circ \mathbf{f}$. The student should keep this concrete example in mind when reading the next proposition, while remembering that $\mathbf{R}^n$ is just *one example* of an n-dimensional vector space, and that the standard basis of $\mathbf{R}^n$ is just *one example* of a basis of $\mathbf{R}^n$.

Proposition 1.88 *Assume that V has finite dimension $n \geq 1$ and let $\{\mathbf{v}_i\}_{i=1}^n$ be a basis of V. Let $\mathbf{f} \in \mathrm{Func}([a,b], V)$, and let $f_1, \ldots, f_n$ be the unique real-valued functions on $[a,b]$ defined by writing $\mathbf{f}$ pointwise in terms of a basis:*

$$\mathbf{f}(x) = \sum_{i=1}^n f_i(x)\, \mathbf{v}_i \quad \text{for all } x \in [a,b]. \tag{1.61}$$

Then the V-valued function $\mathbf{f}$ is integrable if and only if each of the real-valued functions f_i is integrable. In the integrable case,

$$\int_a^b (f_1\mathbf{v}_1 + \cdots + f_n\mathbf{v}_n) = \left(\int_a^b f_1\right)\mathbf{v}_1 + \cdots + \left(\int_a^b f_n\right)\mathbf{v}_n \ . \tag{1.62}$$

Proof: Let $\{\xi_i : V \to \mathbf{R}\}_{i=1}^n$ be the coordinate functions on V determined by the basis $\{\mathbf{v}_i\}_{i=1}^n$. (Thus $\xi_i(\sum_j a_j\mathbf{v}_j) = a_i$, $\mathbf{w} = \sum_{i=1}^n \xi_i(\mathbf{w})\mathbf{v}_i$ for all $\mathbf{w} \in V$, and $f_i = \xi_i \circ \mathbf{f}$ for $1 \leq i \leq n$.) Then for each $i \in \{1, \ldots, n\}$, the function ξ_i is a linear transformation $V \to \mathbf{R}$, so by Proposition 1.86(b), ξ_i is continuous (hence an element of V^*).

First assume that $\mathbf{f}$ is integrable on $[a,b]$, and let $\mathbf{A} = \int_a^b \mathbf{f}(x)\, dx$. For $1 \leq i \leq n$ let $A_i = \xi_i(\mathbf{A})$; thus $\mathbf{A} = \sum_{i=1}^n A_i\mathbf{v}_i$. Let $\epsilon > 0$, let $\epsilon_1 = \epsilon/K$, and let $\delta > 0$ be such that $\mathcal{S}_\delta(\mathbf{f}) \subseteq B^V_{\epsilon_1}(\mathbf{A})$.

By Proposition 1.87, for $1 \leq i \leq n$ the real-valued function $\xi_i \circ \mathbf{f}$ is integrable, and

$$A_i = \xi_i\left(\int_a^b \mathbf{f}\right) = \int_a^b \xi_i \circ \mathbf{f} = \int_a^b f_i \ .$$

Since $\mathbf{A} = \sum_{i=1}^m A_i\mathbf{v}_i$, this establishes (1.62).

We have now shown that if $\mathbf{f}$ is integrable on $[a,b]$, then (i) each component function f_i is integrable on $[a,b]$, and (ii) the equality (1.62) holds. For the converse of the integrability implication, assume now that f_i is integrable on $[a,b]$ for $1 \leq i \leq n$.

Let $A_i = \int_a^b f_i$, $1 \leq i \leq n$, and let $\mathbf{A} = \sum_{i=1}^n A_i\mathbf{v}_i$. Let $\epsilon > 0$, let $C = \sum_{i=1}^n \|\mathbf{v}_i\|$, and let $\epsilon_1 = \epsilon/C$. For $1 \leq i \leq n$ let $\delta_i > 0$ be such that $\mathcal{S}_{\delta_i}(f_i) \subseteq B^{\mathbf{R}}_{\epsilon_1}(A_i)$, and let $\delta = \min\{\delta_i : 1 \leq i \leq n\}$.

Let (P,T) be a pointed partition of $[a,b]$ of mesh less than δ, and let $\mathbf{S} = S(\mathbf{f};P,T)$. Then, again using Exercise 1.18,

$$\begin{aligned}
\mathbf{S} - \mathbf{A} &= \sum_{i=1}^{n} \xi_i(\mathbf{S})\mathbf{v}_i - \sum_{i=1}^{n} A_i\mathbf{v}_i \\
&= \sum_{i=1}^{n} \left[\xi_i(S(\mathbf{f};P,T)) - A_i\right]\mathbf{v}_i \\
&= \sum_{i=1}^{n} \left[S(\xi_i \circ \mathbf{f};P,T) - \ A_i\right]\mathbf{v}_i \\
&= \sum_{i=1}^{n} \left(S(f_i;P,T) - A_i\right)\mathbf{v}_i \ .
\end{aligned}$$

Hence

$$\begin{aligned}
\|\mathbf{S} - \mathbf{A}\| &\leq \sum_{i=1}^{n} \|\left(S(f_i;P,T) - A_i\right)\mathbf{v}_i\| \\
&= \sum_{i=1}^{n} |S(f_i;P,T) - A_i| \ \|\mathbf{v}_i\| \\
&< \sum_{i=1}^{n} \epsilon_1\|\mathbf{v}_i\| \\
&= \epsilon_1 C = \epsilon.
\end{aligned}$$

Therefore we have produced $\delta > 0$ such that $\|\mathbf{S} - \mathbf{A}\| < \epsilon$ for all $\mathbf{S} \in \mathcal{S}_\delta(\mathbf{f})$. Since ϵ was arbitrary, it follows that $\mathbf{f}$ is integrable on $[a,b]$. ■

Corollary 1.89 *Assume that V is finite-dimensional and let $\mathbf{f} \in \mathrm{Func}([a,b],V)$. Then $\mathbf{f} \in \mathcal{R}([a,b],V)$ if and only if for every $\xi \in V^*$ we have $\xi \circ \mathbf{f} \in \mathcal{R}([a,b])$.*

Proof: The "only if"part of the implication follows from Proposition 1.87. For the "if" part, assume that for every $\xi \in V^*$ we have $\xi \circ \mathbf{f} \in \mathcal{R}([a,b])$. If $\dim(V) = 0$ then trivially $\mathbf{f} \in \mathcal{R}([a,b],V)$, so assume that $n := \dim(V) \geq 1$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of V, and, as in the proof of Proposition 1.88, let $\{\xi_i : V \to \mathbf{R}\}_{i=1}^n$ be the corresponding coordinate functions on V. Then for each i, we have $\xi_i \in V^*$, so (by our hypothesis) $\xi_i \circ \mathbf{f}$ is integrable. But $\xi_i \circ \mathbf{f}$ is exactly the function f_i in the statement of Proposition 1.88. Hence that Proposition implies that $\mathbf{f}$ is integrable. ■

Remark 1.90 Equation (1.62) *formally* looks very similar to

$$\int_a^b \left(\sum_{i=1}^m c_i f_i \right) = \sum_{i=1}^m c_i \int_a^b f_i \qquad (\text{where } c_1, \dots c_m \in \mathbf{R}), \tag{1.63}$$

just with the real constants c_i in (1.63) replaced by "vector constants" $\mathbf{v}_i$ that happen to form a basis of V. But (1.62) and (1.63) are really very different statements. For real-valued functions $f_1, \dots, f_m$, the equality (1.63) is one version of the statement that (i) $\mathcal{R}([a,b])$, the set of integrable *real-valued* functions on $[a,b]$, is a vector space and that (ii) "$\int_a^b$" is a linear map $\mathcal{R}([a,b]) \to \mathbf{R}$. The only meaning of "$\int_a^b$" in equation (1.63) is integration of a *real-valued* function on $[a,b]$. The number of functions m is arbitrary; it's not related to the dimension of anything (unlike the n in (1.62)). The corresponding statement for V-valued functions is *not* (1.62); it's that (i) $\mathcal{R}([a,b],V)$ is a vector space and that (ii) the map $\int_a^b : \mathcal{R}([a,b],V) \to V$ is linear:

$$\int_a^b \left(\sum_{j=1}^m c_j \mathbf{f}_j(x) \right) dx = \sum_{j=1}^m c_j \int_a^b \mathbf{f}_j(x)\ dx \tag{1.64}$$

for all integers $m > 0$, all $\mathbf{f}_1, \dots, \mathbf{f}_m \in \mathcal{R}([a,b],V)$, and all $c_1, \dots, c_m \in \mathbf{R}$.

In (1.64), like in (1.63), the notation "$\int_a^b$" has only one meaning, but in (1.64) the meaning is integration of a <u>V-valued</u> function on $[a,b]$.

Equation (1.62) may be *interpreted, informally*, as saying that the basis vectors $\mathbf{v}_j$ behave as "vector constants" that can be pulled through the integral sign "just like" scalar constants (real numbers). But the "just like" is inaccurate. As noted above, in equation (1.63) the notation "$\int_a^b$" has the same meaning on both sides of the equation; it is a *single* operator (fancy name for function) on *one* vector space, $\mathcal{R}([a,b],V)$. In (1.62), the same notation "$\int_a^b$" is used for *two different operators*, the one on the left-hand side having domain $\mathcal{R}([a,b],V)$, and the one on the right-hand side having domain $\mathcal{R}([a,b])$. The operators are conceptually similar, but they have very different domains. It is important to keep in mind that while "Vector constants can be pulled through the integral sign just like scalar constants" is something that could be conjectured, or even expected, before proving anything, there is no such thing as "proof by analogy".

We will say more about equation (1.62) later in Remark 1.97, after establishing some more results. ▲

Together, the next two propositions generalize Exercise 1.5 from real-valued functions to V-valued functions (with V assumed finite-dimensional).

Proposition 1.91 *Assume that V is finite-dimensional. If $\mathbf{f} : [a,b] \to V$ is integrable, then the real-valued function $x \mapsto \|\mathbf{f}(x)\|$ is integrable.*

Proof: Let $g : [a, b] \to \mathbf{R}$ denote the function $x \mapsto \|\mathbf{f}(x)\|$.

If $\dim(V) = 0$ then $\mathbf{f}$ is the constant function 0_V and g is the constant function 0, which is integrable.

Assume now that $n := \dim(V) \geq 1$ and let $\{\mathbf{v}_i\}_{i=1}^n$ be a basis of V. Since every element of V is a unique linear combination $\sum_i a_i\mathbf{v}_i$, we can define a function $\| \ \|_1 : V \to \mathbf{R}$ by $\|\sum_i a_i\mathbf{v}_i\|_1 = \sum_i |a_i|$. As the student may easily show, $\| \ \|_1$ is a norm on V. (The proof is virtually identical to the proof that the ℓ^1-norm on $\mathbf{R}^n$ is a norm.)

As in Proposition 1.88, let $f_1, \dots f_n$ be the component-functions of $\mathbf{f}$ determined by this basis, i.e. the unique real-valued functions such that $\mathbf{f} = \sum_i f_i\mathbf{v}_i$. By Proposition 1.88, each component-function f_i is integrable, hence also bounded (thus Theorem 1.32 applies to f_i).

By Proposition 1.86(3), the norm $\| \ \|_1$ is equivalent to the given norm $\| \ \|$ on V. Let $c > 0$ be such that for all $\mathbf{v} \in V$, $\|\mathbf{v}\| \leq c\|\mathbf{v}\|_1$.

Let $\epsilon > 0$. For each $i \in \{1, \dots, n\}$ let $\delta_i > 0$ be such that $U_{\delta_i}(f_i) - L_{\delta_i}(f_i) < \frac{\epsilon}{cn}$; such δ_i exist by the "(i) $\Longrightarrow$ (iii)" implication of Theorem 1.32. Let $\delta = \min\{\delta_1, \dots, \delta_n\}$. Then for each $P \in \mathcal{P}_\delta([a, b])$ and each $i \in \{1, \dots, n\}$ we have

$$U(f_i; P) - L(f_i; P) \leq U_\delta(f_i) - L_\delta(f_i) \leq U_{\delta_i}(f_i) - L_{\delta_i}(f_i) < \frac{\epsilon}{cn} .$$

Let $P = \{x_0, \dots, x_N\} \in \mathcal{P}_\delta([a, b])$. For $1 \leq i \leq n$ and $1 \leq j \leq N$ let $M_{i,j} = \sup\{f_i(x) : x \in [x_{j-1}, x_j]\}$ and $m_{i,j} = \inf\{f_i(x) : x \in [x_{j-1}, x_j]\}$. Observe that for any $s, t \in [x_{j-1}, x_j]$, and any $i \in \{1, \dots, n\}$, we have

$$|f_i(s) - f_i(t)| \leq M_{i,j} - m_{i,j}. \tag{1.65}$$

Let $T = \{t_1, \dots, t_N\}$ and $T' = \{t'_1, \dots, t'_N\}$ be arbitrary pointings of P. Then, using the triangle inequality in the form $\|\mathbf{a}\| - \|\mathbf{b}\| \leq \|\mathbf{a} - \mathbf{b}\|$, we have

$$\begin{aligned}
S(g;P,T)-S(g;P,T') &= \sum_{j=1}^{N}(\|\mathbf{f}(t_j)\|-\|\mathbf{f}(t'_j)\|)\Delta_j \\
&\leq \sum_{j=1}^{N}\|\mathbf{f}(t_j)-\mathbf{f}(t'_j)\|\,\Delta_j \\
&\leq \sum_{j=1}^{N}c\|\mathbf{f}(t_j)-\mathbf{f}(t'_j)\|_1\,\Delta_j \\
&= c\sum_{j=1}^{N}\left(\sum_{i=1}^{n}|f_i(t_j)-f_i(t'_j)|\right)\Delta_j \\
&= c\sum_{i=1}^{n}\left(\sum_{j=1}^{N}|f_i(t_j)-f_i(t'_j)|\Delta_j\right) \\
&\leq c\sum_{i=1}^{n}\left(\sum_{j=1}^{N}(M_{i,j}-m_{i,j})\Delta_j\right) \qquad \text{(using (1.65))} \\
&= c\sum_{i=1}^{n}(U(f_i;P)-L(f_i;P)) \\
&< c\sum_{i=1}^{n}\frac{\epsilon}{cn} \\
&= \epsilon.
\end{aligned}$$

Thus $S(g;P,T)-S(g;P,T')<\epsilon$ for all pointings T,T' of P. Taking the supremum over T and then the infimum over T', we deduce that $U(g;P)-L(g;P)\leq\epsilon$. Since ϵ was arbitrary, it follows from Proposition 1.45 that g is integrable. ■

We are done restricting attention to finite-dimensional V for now, so we resume using non-boldface letters for elements of V and for V-valued functions.

Proposition 1.92 ("Triangle inequality for integrals") *Let $f\in\mathcal{R}([a,b],V)$, and let $\|f(\cdot)\|:[a,b]\to\mathbf{R}$ denote the function $x\mapsto\|f(x)\|$. Then*

$$\left\|\int_a^b f\right\|\leq\lim_{\delta\to0}U_\delta(\|f(\cdot)\|). \tag{1.66}$$

Hence if $\|f(\cdot)\|$ is integrable,

$$\left\|\int_a^b f(x)\,dx\right\|\leq\int_a^b\|f(x)\|\,dx. \tag{1.67}$$

Our nickname "triangle inequality for integrals" really refers only to inequality (1.67). The reason for this nickname is discussed later in item 6 of Remark 1.97.

Proof of Proposition 1.92: Let us write $A = \int_a^b f$ and $g = \|f(\cdot)\|$.

Let $\epsilon > 0$, and let $\delta_1 > 0$ be such that $\mathcal{S}_{\delta_1}(f) \subseteq B^V_\epsilon(A)$. Let $\delta \in (0, \delta_1]$, let $P \in \mathcal{P}_\delta([a,b])$, and let $T = \{t_1, \ldots, t_N\}$ be a pointing of P. Then, using the triangle inequality,

$$\begin{aligned} \|S(f;P,T)\| = \left\| \sum_{j=1}^N f(t_j)\Delta_j \right\| &\leq \sum_{j=1}^N \|f(t_j\| \, \Delta_j \\ &= S(g;P,T) \\ &\leq U(g;P) \quad \text{(by the definition of } U(g;P)\text{)} \\ &\leq U_\delta(g) \quad \text{(by the definition of } U_\delta(g)\text{)}. \end{aligned}$$

But, since $\delta \leq \delta_1$, $S(f;P,T)$ lies in $\mathcal{S}_\delta(f)$, so $\|S(f;P,T) - A\| < \epsilon$. Hence

$$\|A\| \leq \|A - S(f;P,T)\| + \|S(f;P,T)\| < \epsilon + \|S(f;P,T)\| \leq \epsilon + U_\delta(g);$$

i.e. $\|A\| < U_\delta(g) + \epsilon$. Hence, by an order-property of limits of real-valued functions,

$$\|A\| \leq \lim_{\delta \to 0} U_\delta(g) + \epsilon.$$

Since ϵ was arbitrary, we conclude that $\|A\| \leq \lim_{\delta\to 0} U_\delta(g)$, which is (1.66).

If g is integrable, then $\lim_{\delta \to 0} U_\delta(g) = \int_a^b g$ (by Theorem 1.32), so (1.66) reduces to (1.67) in this case. ■

Observe that, by Proposition 1.91, if V is finite-dimensional, then under the hypotheses of Proposition 1.92 the function $\|f(\cdot)\|$ is automatically integrable, so the stronger conclusion (1.67) holds. We record this fact later in Corollary 1.96.

Remark 1.93 (Optional reading, intended for students who have read Section 1.4.) In view of the Darboux theorem mentioned in Section 1.4, we can alternatively write (1.66) as

$$\left\| \int_a^b f(x)\,dx \right\| \leq \overline{\int_a^b} \|f(x)\|\,dx.$$

▲

Proposition 1.94 ("Continuous implies integrable") *If $f : [a,b] \to V$ is continuous, then f is integrable.*

Proof: Let f be a continuous function from $[a,b]$ to V. Since $[a,b]$ is compact, f is uniformly continuous. Let $\epsilon > 0$, and let $\delta > 0$ be such that if $x, y \in [a,b]$ and $|x-y| < \delta$ then $\|f(x) - f(y)\| < \epsilon_1 := \frac{\epsilon}{2(b-a)}$.

Let $P_1 = \{x_0, \ldots, x_{N_1}\}$, $P_2 = \{y_0, \ldots, y_{N_2}\} \in \mathcal{P}_\delta([a,b])$. Let $P = P_1 \cup P_2 = \{z_0, \ldots, z_N\}$; then $P \in \mathcal{P}_\delta([a,b])$ as well. For $1 \leq j \leq N_1$ let $i_j \in \{1, 2, \ldots N\}$ be the index for which $x_j = z_{i_j}$. (It is helpful to draw a diagram of the interval $[a,b]$ to follow the proof from this point on.) Let $T_1 = \{t_1, \ldots, t_{N_1}\}$, $T = \{s_1, \ldots, s_N\}$ be pointings of P_1, P respectively. Then

$$S(f;P_1,T_1) - S(f;P,T) = \sum_{j=1}^{N_1} f(t_j)(x_j - x_{j-1}) - \sum_{i=1}^{N} f(s_i)(z_i - z_{i-1})$$

$$= \sum_{j=1}^{N_1} f(t_j)\left(\sum_{i=i_{j-1}+1}^{i_j} (z_i - z_{i-1})\right) - \sum_{j=1}^{N_1}\left(\sum_{i=i_{j-1}+1}^{i_j} f(s_i)(z_i - z_{i-1})\right)$$

$$= \sum_{j=1}^{N_1}\left(\sum_{i=i_{j-1}+1}^{i_j} (f(t_j) - f(s_i))(z_i - z_{i-1})\right).$$

Note that in the expression "$f(t_j) - f(s_i)$" on the last line, we have $i_{j-1} \leq i-1 < i \leq i_j$, implying $x_{j-1} \leq z_{i-1} \leq s_i \leq z_i \leq x_j$. Thus s_i lies in $[x_{j-1}, x_j]$, as does t_j. Since $x_j - x_{j-1} < \delta$, we have $|t_j - s_i| < \delta$, implying $\|f(t_j) - f(s_i)\| < \epsilon_1$. Therefore, applying the iterated triangle inequality, we have

$$\begin{aligned}
\|S(f;P_1,T_1) - S(f;P,T)\| &\leq \sum_{j=1}^{N_1}\left\|\left(\sum_{i=i_{j-1}+1}^{i_j} (f(t_j) - f(s_i))(z_i - z_{i-1})\right)\right\| \\
&\leq \sum_{j=1}^{N_1}\left(\sum_{i=i_{j-1}+1}^{i_j} \|(f(t_j) - f(s_i))(z_i - z_{i-1})\|\right) \\
&= \sum_{j=1}^{N_1}\left(\sum_{i=i_{j-1}+1}^{i_j} (z_i - z_{i-1})\|(f(t_j) - f(s_i))\|\right) \\
&< \sum_{j=1}^{N_1}\left(\sum_{i=i_{j-1}+1}^{i_j} (z_i - z_{i-1})\epsilon_1\right) \qquad (1.68) \\
&= \epsilon_1 \sum_{i=1}^{N}(z_i - z_{i-1}) \qquad (1.69) \\
&= \epsilon_1(b-a) \qquad (1.70) \\
&= \frac{\epsilon}{2}\ . \qquad (1.71)
\end{aligned}$$

Thus, $\|S(f;P_1,T_1) - S(f;P,T)\| < \frac{\epsilon}{2}$. Similarly, $\|S(f;P_2,T_2) - S(f;P,T)\| < \frac{\epsilon}{2}$. Hence

$$\begin{aligned} \|S(f;P_1,T_1) - S(f;P_2,T_2)\| &\leq \|S(f;P_1,T_1) - S(f;P,T)\| + \|S(f;P,T) - S(f;P_2,T_2)\| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Therefore for any $S_1, S_2 \in \mathcal{S}_\delta(f)$ we have $\|S_1 - S_2\| < \epsilon$. Since ϵ was arbitrary, it follows from Proposition 1.84 that f is integrable on $[a,b]$. ■

Remark 1.95 The argument above gives a second proof that continuous *real*-valued functions on $[a,b]$ are integrable, without relying on the Proposition 1.45 (the "Step-function lemma") or the Extreme Value Theorem. ▲

Corollary 1.96 *Let $f \in \mathcal{R}([a,b],V)$. If either (a) V is finite-dimensional or (b) f is continuous, then the real-valued function $x \mapsto \|f(x)\|$ is integrable, and*

$$\left\| \int_a^b f(x)\, dx \right\| \leq \int_a^b \|f(x)\|\, dx. \tag{1.72}$$

Proof: Let g denote the function $x \mapsto \|f(x)\|$. Under hypothesis (a), Proposition 1.91 implies that g is integrable. Under hypothesis (b), g is the composition of the continuous function $f : [a,b] \to V$ with the continuous function $\| \ \| : V \to \mathbf{R}$. (Students: why is the latter function is continuous?) Hence g is continuous, so by Theorem 1.53, g is integrable.

Thus, under either hypothesis (a) or (b), the function g is integrable. Therefore Proposition 1.92 shows that the inequality (1.72) holds. ■

Remark 1.97 It is worthwhile to revisit the Calculus 3 version of (1.62):

$$\int_a^b (f_1(t)\mathbf{i} + f_2(t)\mathbf{j} + f_3(t)\mathbf{k})\, dt = \left(\int_a^b f_1(t)\, dt\right)\mathbf{i} + \left(\int_a^b f_2(t)\, dt\right)\mathbf{j} + \left(\int_a^b f_3(t)\, dt\right)\mathbf{k}. \tag{1.73}$$

In Calc 3 we are taught that equation (1.73) is the *definition* of the integral on the left-hand side. *Now that we have proven Proposition 1.88,* we see that for the case $V = \mathbf{R}^3$, with any norm (since they are all equivalent), the definition given in Calc 3 is equivalent to the loftier Definition 1.79. But the loftier definition, while requiring more sophistication and more work, has several advantages:

1. It shows *from the start* that we can define integrals of V-valued functions for any finite-dimensional vector space, not just $\mathbf{R}^n$.

2. It shows *from the start* that, given a finite-dimensional vector space V, we do not need to introduce a basis of V in order to define integrals of V-valued functions.

3. By defining the integral without reference to a basis, the loftier definition *guarantees* that the value of the integral is independent of the choice of basis (a fact that needs to be *proven*, even for the case $V = \mathbf{R}^3$, if we use the Calc 3 definition).

4. It directly incorporates the principle that *integration is about adding stuff up* (the "stuff density" being the function we're integrating), which equation (1.73), taken as a definition, does not.

5. It is an *elegant* generalization of the definition of integrals of real-valued functions: essentially nothing changed in passing from Definitions 1.6 and 1.8 to Definitions 1.78 and 1.79; all we had to do was to replace absolute-value symbols by norm symbols.

6. It enables us to prove Proposition 1.92 very easily, and to show *why*, for finite-dimensional V, we obtain the inequality in Corollary 1.96 *for any norm whatsoever on* V. We saw that the inequality (1.66) follows simply from applying the (iterated) triangle inequality to Riemann sums. The stronger inequality (1.72)—simply (1.67) stated a second time—then followed (for finite-dimensional V) as soon as we showed that the pointwise norm of an integrable V-valued function is an integrable real-valued function, which we saw is true for *any* norm on V. Inequality (1.67) or (1.72) can be viewed as *the iterated triangle inequality generalized from finite sums to integrals* (hence the nickname we have given to Proposition 1.92).

   With the generalized Calc 3 definition, the result of Corollary 1.96 in the finite-dimensional case can also be proven, with a little cleverness but not much difficulty, for a *Euclidean* norm on V—one that comes from an inner product. A standard argument (starting from the generalized Calc 3 definition), presented in [6,7], makes use of the Cauchy-Schwartz inequality for inner products to obtain (1.72). Unfortunately, this argument obscures the fundamental reason *why* (1.72) *ought* to be true. In addition, not every norm on $\mathbf{R}^n$, or on a general finite-dimensional vector space, is Euclidean (an ℓ^2 norm); it need not even be an ℓ^p-norm for *any* p. Thus, with a general norm on V, the argument based on the Cauchy-Schwartz inequality does not yield even the *integrability* of the pointwise norm of an integrable V-valued function, let alone the inequality (1.72).

7. The loftier definition tells us (after proving Proposition 1.88) *why* equation (1.73) *should* be true; that it's not just a definition introduced for convenient bookkeeping.

8. The loftier definition does not even require V to be finite-dimensional; it requires only that V be a complete normed vector space. (We have seen several examples of

infinite-dimensional complete normed vector spaces in MAA 4211–4212: the space $\ell^\infty(\mathbf{R})$, and the space $C(X)$ for any compact metric space X of infinite cardinality.) The Calc 3 definition does not generalize to infinite-dimensional V.

The advantages listed above of the loftier definition are only one side of the coin, however. Even at levels more advanced that of MAA 4211–4212, there are good textbooks (such as [6,7]) and good teachers who prefer the "generalized Calc 3 definition" of an $\mathbf{R}^n$-valued function (with Calc 3's $\mathbf{R}^3$ replaced by $\mathbf{R}^n$), and who assume boundedness of the integrand (the function being integrated) from the start. This approach defines-away the need to prove Propositions 1.82, 1.87, and 1.88, and thereby enables other results to be written down sooner[13], albeit at the expense of some insight and generality. (Additional time is saved, in this approach, by asserting and proving Proposition 1.91 only for the Euclidean norm, rather than for an *arbitrary* norm.) In this approach, the fact that the integral of an integrable $\mathbf{R}^n$-valued function has a basis-independent characterization is something *proven*, rather than something that drops out of the definition of "integral". This is an instance of something quite common in mathematics: often there are two (or more) approaches to the same topic, with some *theorems* in approach A being *definitions* in approach B and vice-versa. ▲

The Fundamental Theorem of Calculus (FTC) also generalizes to V-valued functions, but to state or prove this we first need to define *derivative* for such functions:

Definition 1.98 Let $U \subseteq \mathbf{R}$ be open and let $f : U \to V$ be a function. For $x_0 \in U$, if $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ exists, we say that f is differentiable at x_0, denote this limit $f'(x_0)$, and call this limit the *derivative of f at x_0*. If f is differentiable at every point of U, we simply say that f is *differentiable*. If f is differentiable and the function $f' : U \to V$ defined by $x \mapsto f'(x)$ is continuous, we say that f is *continuously differentiable*. ▲

Remark 1.99 Later in the course we will have a different, higher-level definition of "derivative of a function at a point", related but inequivalent to the one above (even for real-valued functions). The same functions will end up being differentiable at any given point, but what type of object the derivative is will change. ▲

Exercise 1.19 Let f, U, and x_0 be as in Definition 1.98. Prove that if f is differentiable at x_0, then f is continuous at x_0.

Given $f : [a,b] \to \mathbf{R}$, where $a \leq b$, we define the existence and (when the integral exists) the value of $\int_b^a f$ just as in Definition 1.58, simply replacing "real-valued" and "0" by "V-valued" and "0_V" respectively.

---

[13] As a practical matter, this can be very important, since semesters have finite length! But several times in his mathematical life, the author of these notes has been grateful that he learned the "invariant" definition, i.e. Definition 1.79, early on, so his preference is to expose students to this approach.

The next two exercises involve no more than looking back at some of our work on real-valued functions, replacing "real-valued" with "V-valued", replacing some absolute-value symbols with norm symbols, and (for proofs) seeing that the same arguments as before still work.

Exercise 1.20 State and prove the generalizations of Corollary 1.59, Corollary 1.60 and Lemma 1.61 to V-valued functions. (Practically no work is needed for this.)

Exercise 1.21 State and prove the generalizations of Theorem 1.62 ("part of" the FTC) and Corollary 1.63, to V-valued functions. (In Corollary 1.63, the definition of "antiderivative" is generalized the obvious way.)

For V-valued functions, the generalization of the "other part" of the FTC—Theorem 1.64 or the equivalent Theorem 1.65—is still true, but we cannot argue just as before, because our proof made use of the Mean Value Theorem. *The vector-valued-function analog of the Mean Value Theorem is* **false**, even for $V = \mathbf{R}^2$. For example, consider the function $\mathbf{f} : \mathbf{R} \to \mathbf{R}^2$ given by $\mathbf{f}(t) = (\cos t, \sin t)$. The restriction of f to $[0, 2\pi]$ is continuous, and is differentiable on $(0, 2\pi)$. However, $\frac{f(2\pi)-f(0)}{2\pi-0} = \mathbf{0}$ (the zero vector in $\mathbf{R}^2$). But $\mathbf{f}'(t) = (-\sin t, \cos t)$, which is never $\mathbf{0}$.

To prove the generalization of Theorem 1.64 for V-valued functions, the following lemma is needed:

Lemma 1.100 *Let $U \subseteq \mathbf{R}$ be an open interval, $f : U \to V$ a differentiable function. If f' is identically zero, then f is constant.*

Earlier this semester, we proved Lemma 1.100 in the special case $V = \mathbf{R}$, but relied on the Mean Value Theorem for that proof. Since the MVT is false for V-valued functions, a different approach is needed. When V is finite-dimensional the proof is not much harder than in the real-valued case; we re-express a vector-valued function $\mathbf{f}$ in terms of finitely many real-valued functions to which we can apply the MVT:

Exercise 1.22 Assume that V has finite dimension $n \geq 1$ and let $\{\mathbf{v}_i\}_{i=1}^n$ be a basis of V. Let $U \subseteq \mathbf{R}$ be an open interval, Let $\mathbf{f} \in \mathrm{Func}(U, V)$, and let $f_1, \dots, f_n$ be the unique real-valued functions on U for which $\mathbf{f} = \sum_{i=1}^n f_i \mathbf{v}_i$, i.e. for which equation (1.61) holds.

(a) Let $x_0 \in U$. Prove that if $\mathbf{f}$ is differentiable at x_0, then each of the component functions f_i is differentiable at x_0.

(b) Prove that if $\mathbf{f}$ is continuously differentiable, then each of the component functions f_i is continuously differentiable.

(c) Prove that if $\mathbf{f}$ is differentiable and $\mathbf{f}'$ is identically zero, then $\mathbf{f}$ is constant.

(The converses of (a), (b), and (c) are all true; they just are not part of this exercise.)▲

Without the assumption of finite-dimensionality, a good approach to Lemma 1.100 is to recast the statement in differential-equations terms: Given an open interval $U \subseteq \mathbf{R}$, a point $t_0 \in U$, and an element $v_0 \in V$, the constant function $y : U \to V$ defined by $y(t) = v_0$ (for all $t \in U$) is the unique solution of the initial-value problem

$$\frac{dy}{dt} = 0_V, \quad y(t_0) = v_0 \ . \tag{1.74}$$

on U. Stated this way, the lemma follows from the "Fundamental Theorem of ODEs" for V-valued functions, part of which asserts the uniqueness of solutions to initial-value problems of the form $\frac{dy}{dt} = g(t, y), \quad y(t_0) = v_0$, when g is a "nice enough" V-valued function on some neighborhood of $(t_0, v_0) \in \mathbf{R} \times V$. Time permitting, we will see a precise statement and proof this theorem later in the semester. *We will not prove Lemma 1.100 in these notes* (other than the finite-dimensional case, Exercise 1.22(c)). For now, the student is asked to take on faith that the function $(t, v) \mapsto 0_V$ meets the criterion of being "nice enough" to apply this theorem, and therefore that the constant function $t \mapsto v_0$ is indeed the unique solution of the initial-value problem (1.74).

Exercise 1.23 (a) Show that the uniqueness statement above (for the IVP (1.74), with arbitrary initial-condition point $(t_0, v_0) \in U \times V$) is, indeed, equivalent to Lemma 1.100.

(b) Assuming Lemma 1.100, prove the analog of Theorem 1.64 for V-valued functions.

1.10 Index for notation and terminology

References

[1] T. M. Apostol, *Calculus*, 2nd ed., vol. 1, John Wiley & Sons, 1967. 38, 41

[2] T. M. Apostol, *Mathematical Analysis*, Addison-Wesley, 1957. 38, 41

[3] T. M. Apostol, *Mathematical Analysis*, 2nd ed., Addison-Wesley, 1974. 38, 39, 41

[4] J. Hass, M. D. Weir, and G. B Thomas, *University Calculus*, alternate edition, Pearson Addison-Wesley, 2008. 38

[5] M. Rosenlicht, *Introduction to Analysis*, Scott, Foresman and Co., 1968. Reprinted by Dover, 1986. 38, 41

[6] W. Rudin, *Principles of Mathematical Analysis*, 2nd ed., McGraw-Hill, 1964. 27, 38, 56, 57

[7] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, 1976. 27, 38, 56, 57

[8] A. Schwartz, *Calculus and Analytic Geometry*, 2nd ed., Holt, Rinehart, and Winston, 1967. 38, 41

[9] J. Stewart, *Calculus: Early Transcendentals*, 6th ed., Thomson Brooks/Cole, 2008. 38

[10] A. E. Taylor and W. R. Mann, *Advanced Calculus*, 2nd ed., Xerox College Publishing, 1972. 26, 27, 38

[11] G. B. Thomas, *Calculus and Analytic Geometry*, alternate edition, Addison-Wesley, 1972. 38, 41

[12] D. G. Zill and W. S. Wright, *Calculus: Early Transcendentals*, 4th ed., Jones and Bartlett, 2011.