4 Integration

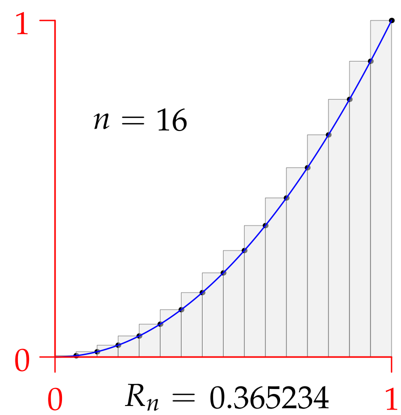
The theory of infinite series addresses the problem of summing infinitely many *finite* quantities. By contrast, integration is the business of summing infinitely many *infinitesimal* quantities. Mathematicians have attempted to do both for well over 2000 years, and the philosophical objections are just as old.[1] The development and increased application of calculus from the late 1600s spurred mathematicians to try to put the theory on a firmer footing, though from Newton and Leibniz it took another 150 years before Bernhard Riemann (1856) provided a thorough development of the integral.

32 The Riemann Integral

The basic idea behind Riemann integration is to approximate area using a sequence of rectangles whose width tends to zero. The following discussion is hopefully familiar.

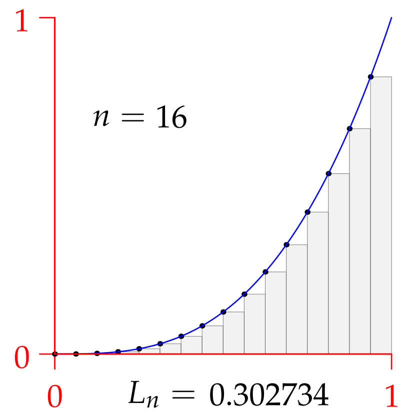
Example 4.1. Consider $f(x) = x^2$ defined on $[0, 1]$.
For each $n \in \mathbb{N}$, let $\Delta x = \frac{1}{n}$ and define $x_i = i\Delta x$.
Above each *subinterval* $[x_{i-1}, x_i]$, raise a rectangle of height $f(x_i) = x_i^2$.
The sum of the areas of these rectangles is the *Riemann sum with right-endpoints*[2]

$$R_n = \sum_{i=1}^{n} f(x_i)\Delta x = \sum_{i=1}^{n} \frac{i^2}{n^3} = \frac{n(n+1)(2n+1)}{6n^3}$$
$$= \frac{1}{3} + \frac{3n+1}{6n^2}$$



The *Riemann sum with left-endpoints* is defined similarly:

$$L_n = \sum_{i=1}^{n} f(x_{i-1})\Delta x = \sum_{i=1}^{n} \frac{(i-1)^2}{n^3} = \frac{1}{3} - \frac{3n-1}{6n^2}$$



Since f is increasing, the area A under the curve satisfies

$$L_n \le A \le R_n$$

and the squeeze theorem allows us to conclude that $A = \frac{1}{3}$.

The example contains the essential idea, but more flexibility is needed. To get further, we must properly define the concepts of partition and Riemann sum.

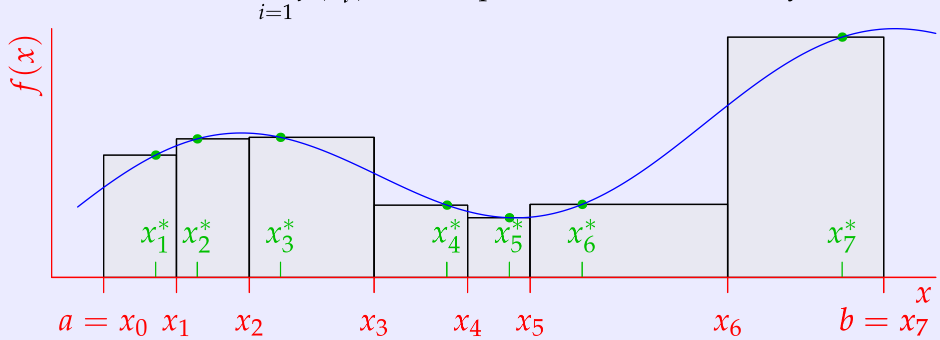
[1] Two of Zeno's ancient paradoxes are relevant here: Achilles and the Tortoise concerns a convergent infinite series, while the Arrow Paradox discusses a difficulty with integration by questioning whether time can be considered as a sum of instants. Perhaps the most famous contemporary criticism comes from Bishop George Berkeley, who gave his name to the Californian city and thus the first UC campus: in *The Analyst* (1734), Berkeley savaged the foundations of calculus, describing the infinitesimal increments required in Newton's theory of *fluxions* (derivatives) as merely the "ghosts of departed quantities."

[2] Now is a good time to review some identities: $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$, $\sum_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$, $\sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$

Definition 4.2. A *partition* $P = \{x_0, \ldots, x_n\}$ of an interval $[a, b]$ is a finite sequence such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

For each $1 \le i \le n$, define $\Delta x_i = x_i - x_{i-1}$. The *mesh* of the partition is $\mathrm{mesh}(P) := \max \Delta x_i$.
Choose a *sample point* x_i^* in each *subinterval* $[x_{i-1}, x_i]$.
If $f : [a, b] \to \mathbb{R}$, the *Riemann sum* $\sum_{i=1}^{n} f(x_i^*)\, \Delta x_i$ computes the area of a family of n rectangles.



In elementary calculus, one typically computes Riemann sums for *equally-spaced* partitions with *left*, *right* or *middle* sample points. The double freedom of partition & sample points makes applying the definition a challenge, so instead we consider two special families of rectangles.

Definition 4.3. Given a partition P of $[a, b]$ and a bounded function f on $[a, b]$, define

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x) \qquad U(f, P) = \sum_{i=1}^{n} M_i\, \Delta x_i$$

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x) \qquad L(f, P) = \sum_{i=1}^{n} m_i\, \Delta x_i$$

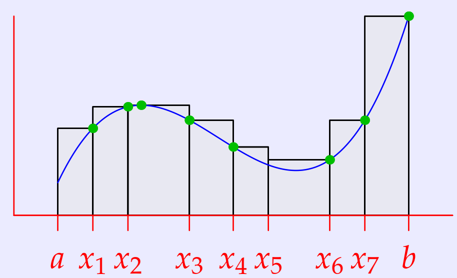
$U(f, P)$ and $L(f, P)$ are the *upper* and *lower Darboux sums* for f with respect to P. The *upper* and *lower Darboux integrals* are

$$U(f) = \inf U(f, P) \qquad L(f) = \sup L(f, P)$$

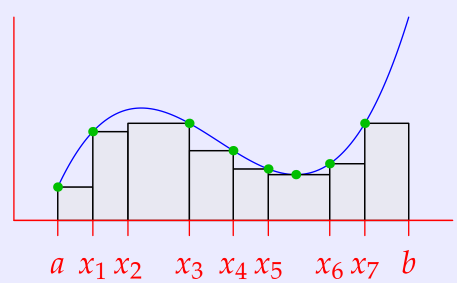
where the supremum and infimum are over all partitions.
We say that f is *(Riemann) integrable* on $[a, b]$ if $U(f) = L(f)$ and denote this value by

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\, \mathrm{d}x$$

If the interval is understood or irrelevant, it is common just to say that f is integrable and write $\int f$.



Upper Darboux sum $U(f, P)$



Lower Darboux sum $L(f, P)$

Intuitively, $L(f, P)$ is the sum of the areas of rectangles built on P which just fit under the graph of f. It is also the infimum of all Riemann sums on P. If f is discontinuous, then $L(f, P)$ need not be a Riemann sum; there might not be suitable sample points!

Examples 4.4. 1. We revisit Example 4.1 in this language.

Given a partition $Q = \{x_0, \ldots, x_n\}$ of $[0,1]$ and sample points $x_i^* \in [x_{i-1}, x_i]$, we compute the Riemann sum for $f(x) = x^2$

$$\sum_{i=1}^{n} f(x_i^*)\,\Delta x_i = \sum_{i=1}^{n} (x_i^*)^2 (x_i - x_{i-1})$$

Since f is increasing, we have $x_{i-1}^2 \le (x_i^*)^2 \le x_i^2$ on each interval, whence

$$L(f,Q) = \sum_{i=1}^{n} (x_{i-1})^2 (x_i - x_{i-1}) \le \sum_{i=1}^{n} (x_i^*)^2 (x_i - x_{i-1}) \le \sum_{i=1}^{n} (x_i)^2 (x_i - x_{i-1}) = U(f,Q)$$

The Darboux sums are therefore the Riemann sums for right and left endpoints.

If we take Q_n to be the partition with subintervals of equal width $\Delta x = \frac{1}{n}$, then

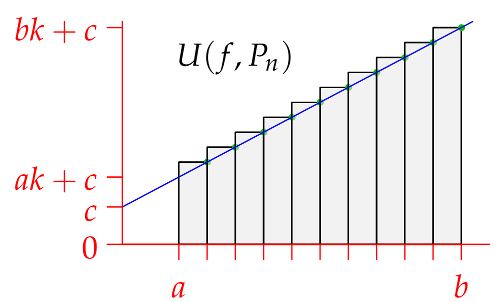
$$U(f) = \inf_P U(f,P) \le U(f,Q_n) = \sum_{i=1}^{n} \left(\frac{i}{n}\right)^2 \Delta x = R_n$$

is the right Riemann sum discussed originally. Similarly $L(f) \ge L_n$. Since L_n and R_n both converge to $\frac{1}{3}$ as $n \to \infty$, the squeeze theorem forces

$$L_n \le L(f) \le U(f) \le R_n \implies L(f) = U(f) = \frac{1}{3}$$

whence f is Riemann integrable on $[0,1]$ with $\int_0^1 x^2\,dx = \frac{1}{3}$.

2. Suppose $f(x) = kx + c$ on the interval $[a,b]$ where $k > 0$. Take the evenly spaced partition P_n where $x_i = a + \frac{b-a}{n} i$ with $\Delta x_i = \frac{b-a}{n}$. Since f is increasing, the upper Darboux sum is again the Riemann sum with right-endpoints:



$$\begin{aligned}
U(f,P_n) = R_n &= \sum_{i=1}^{n} f(x_i)\Delta x \\
&= \frac{b-a}{n} \sum_{i=1}^{n} \frac{k(b-a)}{n} i + ak + c \\
&= \frac{b-a}{n} \left[\frac{k(b-a)}{n} \cdot \frac{1}{2} n(n+1) + (ak+c)n \right] \\
&\xrightarrow[n\to\infty]{} \frac{1}{2}k(b-a)^2 + (b-a)(ak+c) = \frac{k}{2}(b^2 - a^2) + c(b-a)
\end{aligned}$$

We similarly see that the lower Darboux sum is given by the Riemann sum with left endpoints, and that

$$L(f,P_n) = L_n = \frac{b-a}{n} \left[\frac{k(b-a)}{n} \cdot \frac{1}{2} n(n-1) + (ak+c)n \right] \xrightarrow[n\to\infty]{} \frac{k}{2}(b^2 - a^2) + c(b-a)$$

By the same argument as above, $L_n \le L(f) \le U(f) \le R_n$ and the squeeze theorem show that f is integrable with $\int_a^b f = \frac{k}{2}(b^2 - a^2) + c(b-a)$.

Following the examples, a few remarks are in order.

Riemann versus Darboux Definition 4.3 is really that of the *Darboux integral.* Riemann's definition is as follows: for $f[a,b] \to \mathbb{R}$ to be integrable with integral $\int_a^b f$ means

$$\forall \epsilon > 0,\ \exists \delta \text{ such that } \forall P, x_i^*,\ \text{mesh}(P) < \delta \implies \left| \sum_{i=1}^{n} f(x_i^*)\Delta x_i - \int_a^b f \right| < \epsilon$$

It can be shown that this is equivalent to the Darboux integral. We won't pursue Riemann's formulation further, except to observe that *if* a function is integrable and $\text{mesh}(P_n) \to 0$, then $\int_a^b f = \lim\limits_{n\to\infty} \sum\limits_{i=1}^{n} f(x_i^*)\Delta x_i$: this allows us to approximate integrals using any sample points we choose, hence why *right* endpoints ($x_i^* = x_i$) are so common in Freshman calculus.

Monotone Functions Darboux sums are particularly easy to compute for monotone functions. As in the examples, if f is increasing, then each $M_i = f(x_i)$, from which $U(f,P)$ is the Riemann sum with *right-endpoints.* Similarly, $L(f,P)$ is the Riemann sum with *left-endpoints.* The roles reverse if f is decreasing.

Area If f is positive and continuous,[3] the Riemann integral $\int_a^b f$ serves as a *definition* for the area under the curve $y = f(x)$. This should make intuitive sense:

1. In the second example where we have a straight line, we obtain the same value for the area by computing directly as the sum of a rectangle and a triangle!
2. If the area under the curve is to make sense, then, for any partition P, it plainly satisfies the inequalities

$$L(f,P) \leq \text{Area} \leq U(f,P)$$

But these are exactly the same as those satisfied by the integral itself:

$$L(f,P) \leq L(f) = \int_a^b f = U(f) \leq U(f,P)$$

In the examples we exhibited a sequence of partitions (P_n) where $U(f,P_n)$ and $L(f,P_n)$ both converged to the same limit. The next results develop some basic properties of partitions and make this process rigorous.

Lemma 4.5. *Suppose $f : [a,b] \to \mathbb{R}$ is bounded and suppose P, Q are partitions of $[a,b]$.*

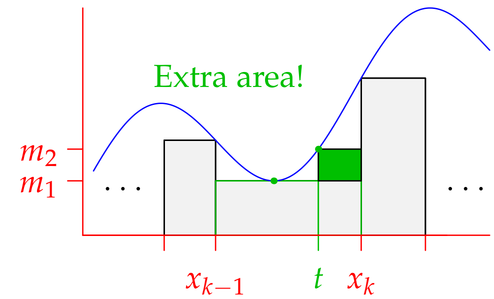
1. *If Q is a* refinement *of P, that is $P \subseteq Q$, then*

$$L(f,P) \leq L(f,Q) \leq U(f,Q) \leq U(f,P)$$

2. *For any partitions P, Q, we have $L(f,P) \leq U(f,Q)$*
3. *$L(f) \leq U(f)$*

[3] We'll see later (Theorem 4.16) that every continuous function is integrable.

Proof. 1. We prove inductively. First suppose that $Q = P \cup \{t\}$ contains exactly one additional point $t \in (x_{k-1}, x_k)$. Write



$$\begin{aligned} m_1 &= \inf\{f(x) : x \in [x_{k-1}, t]\} \\ m_2 &= \inf\{f(x) : x \in [t, x_{k-1}]\} \\ m &= \inf\{f(x) : x \in [x_{k-1}, x_k]\} = \min\{m_1, m_2\} \end{aligned}$$

The Darboux sums $L(f, P)$ and $L(f, Q)$ are identical except for the terms involving t; indeed

$$\begin{aligned} L(f, Q) - L(f, P) &= m_1(t - x_{k-1}) + m_2(x_k - t) - m(x_k - x_{k-1}) \\ &= (m_1 - m)(t - x_{k-1}) + (m_2 - m)(x_k - t) \geq 0 \end{aligned}$$

Since partitions are finite sets, by induction we see that $P \subseteq Q \implies L(f, P) \leq L(f, Q)$.

The argument for $U(f, Q) \leq U(f, P)$ is similar, and the middle inequality is trivial.

2. If P and Q are partitions, then $P \cup Q$ is a refinement of both P and Q. By part 1,

$$L(f, P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f, Q) \tag{$*$}$$

3. We leave this as an exercise.

■

Theorem 4.6 (Cauchy criterion for integrability). *Suppose $f : [a, b] \to \mathbb{R}$ is bounded.*

1. *f is integrable $\iff$ $\forall \epsilon > 0, \exists P$ such that $U(f, P) - L(f, P) < \epsilon$*
2. *f is integrable $\iff$ $\exists (P_n)_{n \in \mathbb{N}}$ such that $U(f, P_n) - L(f, P_n) \to 0$. Moreover, in such a case both sequences $U(f, P_n)$ and $L(f, P_n)$ converge to $\int_a^b f$.*

We call this a Cauchy criterion since integrability is demonstrated without mention of the integral!

Proof. 1. ($\Rightarrow$) Suppose f is integrable. Since $\inf U(f, Q) = \int f = \sup L(f, R)$, $\exists Q, R$ such that

$$U(f, Q) - \int f < \frac{\epsilon}{2} \quad \text{and} \quad \int f - L(f, R) < \frac{\epsilon}{2}$$

Let $P = Q \cup R$ and apply ($*$): $L(f, R) \leq L(f, P) \leq U(f, P) \leq U(f, Q)$. But then

$$U(f, P) - L(f, P) \leq U(f, Q) - L(f, R) = U(f, Q) - \int f + \int f - L(f, R) < \epsilon$$

($\Leftarrow$) For every partition, $L(f, P) \leq L(f) \leq U(f) \leq U(f, P)$. Thus

$$0 \leq U(f) - L(f) \leq U(f, P) - L(f, P) < \epsilon$$

Since this holds for all $\epsilon > 0$, we see that $U(f) = L(f)$.

2. This is an exercise.

■

Examples 4.7. 1. The freedom to choose a partition can be very useful. Consider $f(x) = \sqrt{x}$ on the interval $[0, b]$. We choose a partition that evaluates nicely when fed to this function:

$$P_n = \{x_0, \ldots, x_n\} \quad \text{where} \quad x_i = \left(\frac{i}{n}\right)^2 b$$
$$\implies \Delta x_i = x_i - x_{i-1} = \frac{b}{n^2}\left(i^2 - (i-1)^2\right) = \frac{(2i-1)b}{n^2}$$

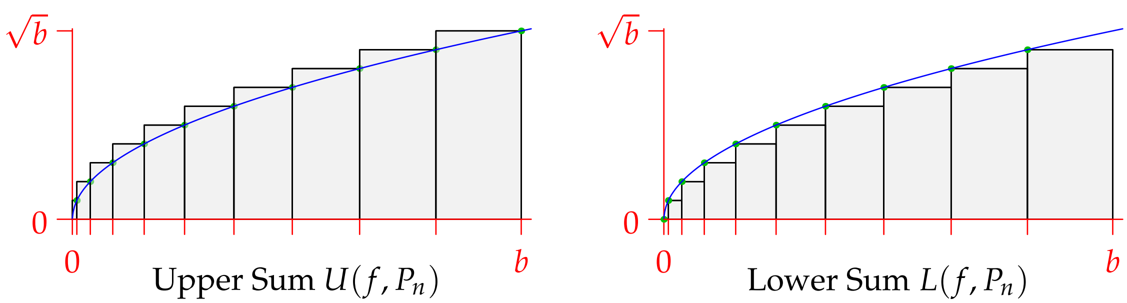
Since f is increasing on $[0, b]$, we see that

$$\begin{aligned} U(f, P_n) &= \sum_{i=1}^{n} f(x_i)\Delta x_i = \sum_{i=1}^{n} \frac{i\sqrt{b}}{n} \cdot \frac{(2i-1)b}{n^2} = \frac{b^{3/2}}{n^3} \sum_{i=1}^{n} 2i^2 - i \\ &= \frac{b^{3/2}}{n^3}\left[\frac{1}{3}n(n+1)(2n+1) - \frac{1}{2}n(n+1)\right] \xrightarrow[n\to\infty]{} \frac{2}{3}b^{3/2} \end{aligned}$$

Similarly

$$\begin{aligned} L(f, P_n) &= \sum_{i=1}^{n} f(x_{i-1})\Delta x_i = \sum_{i=1}^{n} \frac{(i-1)\sqrt{b}}{n} \cdot \frac{(2i-1)b}{n^2} = \frac{b^{3/2}}{n^3} \sum_{i=1}^{n} 2i^2 - 3i + 1 \\ &= \frac{b^{3/2}}{n^3}\left[\frac{1}{3}n(n+1)(2n+1) - \frac{3}{2}n(n+1) + n\right] \xrightarrow[n\to\infty]{} \frac{2}{3}b^{3/2} \end{aligned}$$

Since these limits are equal, we conclude that f is integrable and that $\int_0^b \sqrt{x}\,dx = \frac{2}{3}b^{3/2}$.



2. We finish this section with the classic example of a non-integrable function. Let $f : [a, b] \to \mathbb{R}$ to be the indicator function of the irrational numbers,

$$f(x) = \begin{cases} 1 & \text{if } x \notin \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{Q} \end{cases}$$

Since any interval of positive length contains both rational and irrational numbers, we see that

$$\sup\{f(x) : x \in [x_{i-1}, x_i]\} = 1 \quad \text{and} \quad \inf\{f(x) : x \in [x_{i-1}, x_i]\} = 0$$

for *any* partition $P = \{x_0, \ldots, x_n\}$. We conclude that

$$U(f, P) = \sum_{i=1}^{n}(x_i - x_{i-1}) = b - a \implies U(f) = b - a \quad \text{and}$$
$$L(f, P) = 0 \implies L(f) = 0$$

Since the upper and lower integrals are unequal, f is not Riemann integrable.

As any freshman calculus student can attest, if you can find an anti-derivative, then the fundamental theorem of calculus (Section 34) makes evaluating integrals far easier. For instance, you are probably desperate to write

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{2}{3}x^{3/2} = x^{1/2} \implies \int_0^b \sqrt{x}\,\mathrm{d}x = \frac{2}{3}x^{3/2}\Big|_0^b = \frac{2}{3}b^{3/2}$$

rather than computing Riemann/Darboux sums as in the previous example! In most practical cases, however, no easy-to-compute anti-derivative exists, so the best we can do is approximate integrals by evaluating Riemann sums for progressively finer partitions. Thankfully computers excel at such tedious work!

Exercises 32 1. For each function on the given interval, use partitions to find the upper and lower Darboux integrals. Hence prove that the function is integrable and compute its integral.

(a) $f(x) = x^3$ on $[0, b]$ for any $b > 0$.

(b) $g(x) = \sqrt[3]{x}$ on $[0, b]$.

(*Hint: mimic Example 4.7.1*)

2. Repeat question 1 for the following two functions. You *cannot* simply compute Riemann sums for left and right endpoints and take limits: why not?

(a) $h(x) = x(2-x)$ on $[0, 2]$

(*Hint: choose a partition with* $2n$ *points such that* $x_n = 1$ *and observe that* $h(2-x) = h(x)$)

(b) $k(x) = \begin{cases} 2x & \text{if } x \le 1 \\ 5-x & \text{if } x > 1 \end{cases}$ on $[0, 3]$.

(*Hint: this time try a partition with* $3n$ *points…*)

3. Let $f(x) = x$ for rational x and $f(x) = 0$ for irrational x.

(a) Calculate the upper and lower Darboux integrals for f on the interval $[0, b]$.

(b) Is f integrable on $[0, b]$?

4. Prove part 3 of Lemma 4.5: $L(f) \le U(f)$.

5. Prove part 2 of Theorem 4.6.

$$f \text{ is integrable} \iff \exists (P_n)_{n\in\mathbb{N}} \text{ such that } U(f, P_n) - L(f, P_n) \to 0.$$

Moreover, both $U(f, P_n)$ and $L(f, P_n)$ converge to $\int f$.

6. (a) Reread Definition 4.3. What happens if we allow $f : [a, b] \to \mathbb{R}$ to be *unbounded*?

(b) Read "*Riemann versus Darboux*" on page 4. Explain why being *Riemann* integrable also forces f to be bounded.

7. (If you like coding) Write a short program to estimate $\int_a^b f(x)\,\mathrm{d}x$ using Riemann sums. This can be very simple (equal partitions with right endpoints), or more complex (random partition and sample points given a mesh). Apply your program to estimate $\int_0^5 \sin(x^2 e^{-\sqrt{x}})\,\mathrm{d}x$.

33 Properties of the Riemann Integral

The rough take-away of this long section is that everything you think is integrable probably is! There will not be many examples since we have not established many explicit values for integrals.

Theorem 4.8 (Linearity). *If f, g are integrable and k, l are constant, then $kf + lg$ is integrable and*

$$\int kf + lg = k\int f + l\int g$$

Example 4.9. Thanks to examples in the previous section, we can now calculate, for instance

$$\int_0^2 5x^3 - 3\sqrt{x}\,\mathrm{d}x = 5 \cdot \frac{1}{4} \cdot 2^4 - 3 \cdot \frac{2}{3} \cdot 2^{3/2} = 20 - 4\sqrt{2}$$

Proof. Suppose $\epsilon > 0$ is given. By Theorem 4.6 part 3, there exist partitions R, S such that

$$U(f,R) - L(f,R) < \frac{\epsilon}{2} \quad \text{and} \quad U(g,S) - L(g,S) < \frac{\epsilon}{2}$$

By Theorem 4.6 part 1, if $P := R \cup S$, then both inequalities are satisfied by P. On each subinterval,

$$\inf f(x) + \inf g(x) \le \inf(f(x) + g(x)) \quad \text{and} \quad \sup(f(x) + g(x)) \le \sup f(x) + \sup g(x)$$

since the individual suprema/infima could be 'evaluated' at different places. Thus

$$L(f,P) + L(g,P) \le L(f+g,P) \le U(f+g,P) \le U(f,P) + U(g,P)$$

whence $U(f+g,P) - L(f+g,P) < \epsilon$ and $f+g$ is integrable. Moreover,

$$\int (f+g) - \int f - \int g \le \Big(U(f,P) - \int f\Big) + \Big(U(g,P) - \int g\Big) < \epsilon$$

Using lower Darboux integrals similarly, we see that

$$-\epsilon < \int (f+g) - \int f - \int g < \epsilon$$

Since this holds for all $\epsilon > 0$, we conclude that $\int (f+g) = \int f + \int g$.
That kf is integrable with $\int kf = k\int f$ is an exercise. Put these together for the result. ∎

Corollary 4.10 (Changing endvalues). *Suppose g is integrable on $[a,b]$ and that $f : [a,b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ on (a,b). Then f is integrable on $[a,b]$ and $\int_a^b f = \int_a^b g$.*

Definition 4.11 (Integration on an open interval). A *bounded* function $f : (a,b) \to \mathbb{R}$ is *integrable* if it has an integrable extension $g : [a,b] \to \mathbb{R}$ where $f(x) = g(x)$ on (a,b). In such a case, we define $\int_a^b f := \int_a^b g$.

The Corollary (its proof is an exercise) shows that the choice of extension is irrelevant.

Theorem 4.12 (Basic Comparisons). *Suppose f and g are integrable on $[a,b]$.*

1. *If $f(x) \le g(x)$, then $\int f \le \int g$.*
2. *If $m \le f(x) \le M$ then $m(b-a) \le \int_a^b f \le M(b-a)$.*
3. *fg is integrable.*
4. *$|f|$ is integrable and $\left|\int f\right| \le \int |f|$.*
5. *$\max(f,g)$ and $\min(f,g)$ are integrable.*

Part 3 is *not* integration by parts and does *not* tell us how $\int fg$ relates to $\int f$ and $\int g$!

Proof. 1. Since $g(x) - f(x) \ge 0$ is integrable, $L(g-f,P) \ge 0$ for all partitions P, and so

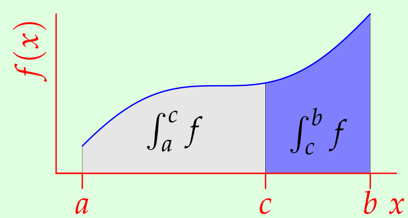
$$0 \le L(g-f) = \int g - f = \int g - \int f$$

2. Apply part 1 twice.
3. This is an exercise.
4. The integrability is an exercise. For the comparison, apply part 1 to $-|f| \le f \le |f|$.
5. $\max(f,g) = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|$, etc.

∎

Theorem 4.13 (Domain splitting). *Suppose that $f : [a,b] \to \mathbb{R}$ and let $c \in (a,b)$. If f is integrable on both $[a,c]$ and $[c,b]$, then it is integrable on $[a,b]$ and*

$$\int_a^b f = \int_a^c f + \int_c^b f$$

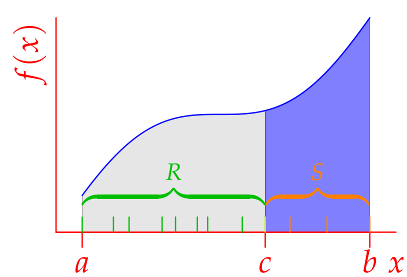


In light of this result, it is conventional to allow integral limits to be reversed:

$$\int_b^a f := -\int_a^b f \quad \text{is consistent with} \quad \int_a^a f = 0$$

Proof. Let $\epsilon > 0$ be given, then $\exists R, S$ partitions of $[a,c]$, $[c,b]$ such that

$$U(f,R) - L(f,R) < \frac{\epsilon}{2}, \qquad U(f,S) - L(f,S) < \frac{\epsilon}{2}$$



Choose $P = R \cup S$ to partition $[a,b]$, then

$$U(f,P) - L(f,P) = U(f,R) + U(f,S) - L(f,R) - L(f,S) < \epsilon$$

Moreover

$$\int_a^b f - \int_a^c f - \int_c^b f \le U(f,P) - L(f,R) - L(f,S) = U(f,P) - L(f,P) < \epsilon$$

The other side is similar.

∎

Example 4.14. If $f(x) = \sqrt{x}$ on $[0, 1]$ and $f(x) = 1$ on $[1, 2]$, then

$$\int_0^2 f = \int_0^1 \sqrt{x}\,\mathrm{d}x + \int_1^2 1\,\mathrm{d}x = \frac{2}{3} + 1 = \frac{5}{3}$$

Monotonic & Continuous Functions

We are now in a position to establish the integrability of two large classes of functions.

Definition 4.15. A function $f : [a, b] \to \mathbb{R}$ is:

Monotonic if it is either *increasing* ($x < y \implies f(x) \le f(y)$) or *decreasing*.

Piecewise monotonic if there is a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ such that f is monotonic on each open subinterval (x_{k-1}, x_k).

Piecewise continuous if there is a partition such that f is *uniformly* continuous on each (x_{k-1}, x_k).

Theorem 4.16. *If f is* monotonic *or* continuous *on $[a, b]$, then it is integrable.*

Examples 4.17. 1. Since sine is continuous, we can approximate via a sequence of Riemann sums

$$\int_0^\pi \sin x\,\mathrm{d}x = \frac{\pi}{n} \lim_{n\to\infty} \sum_{i=1}^n \sin\frac{\pi i}{n}$$

Evaluating this limit is another matter entirely, one best handled in the next section...

2. Similarly, $e^{\sqrt{x}}$ can be integrated and therefore approximated via Riemann sums:

$$\int_0^1 e^{\sqrt{x}}\,\mathrm{d}x = \frac{1}{n} \lim_{n\to\infty} \sum_{i=1}^n \exp\sqrt{\frac{i}{n}} = \lim_{n\to\infty} \sum_{i=1}^n \frac{2j-1}{n} \exp\frac{j}{n}$$

Both sums use right endpoints; the first has equal subintervals and the second is analogous to Example 4.7.1. These limits would typically be estimated using a computer.

Proof. Suppose $f : [a, b] \to \mathbb{R}$ is continuous. Since $[a, b]$ is closed and bounded, f is *uniformly* continuous. Let $\epsilon > 0$ be given:

$$\exists \delta > 0 \text{ such that } \forall x, y \in [a, b],\ |x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a}$$

Let P be a partition with mesh $P < \delta$. Since f attains its bounds on each $[x_{i-1}, x_i]$ (extreme value theorem),

$$\exists x_i^*, y_i^* \in [x_{i-1}, x_i] \quad \text{such that} \quad M_i - m_i = f(x_i^*) - f(y_i^*) < \frac{\epsilon}{b-a}$$

from which

$$U(f, P) - L(f, P) < \sum_{i=1}^n \frac{\epsilon}{b-a}(x_i - x_{i-1}) = \epsilon$$

The monotonicity argument is an exercise. ■

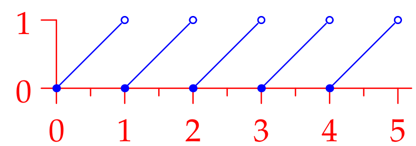
Corollary 4.18. *Piecewise continuous and* bounded *piecewise monotonic functions are integrable.*

Proof. If f is piecewise continuous, then the restriction of f to (x_{k-1}, x_k) has a continuous extension $g_k : [x_{k-1}, x_k] \to \mathbb{R}$; integrable by Theorem 4.16. By Corollary 4.10, f is integrable on $[x_{k-1}, x_k]$ with $\int_{x_{k-1}}^{x_k} f = \int_{x_{k-1}}^{x_k} g_k$. Several applications of Theorem 4.13 finish things off:

$$\int_a^b f = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f$$

The argument for piecewise monotonicity is similar. ■

Example 4.19. The 'fractional part' function $f(x) = x - \lfloor x \rfloor$ is both piecewise continuous and piecewise monotone on any bounded interval. It is therefore integrable on any $[a, b]$.



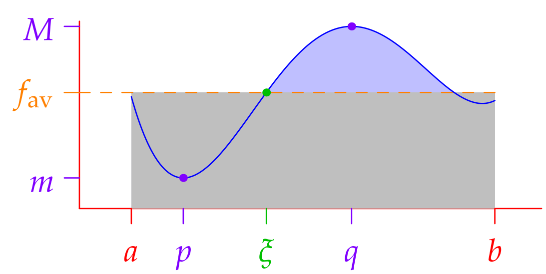
We finish with the final incarnation of the intermediate value theorem.

Corollary 4.20 (IVT for integrals). *If f is continuous on $[a, b]$, then $\exists \xi \in (a, b)$ for which*

$$f(\xi) = \frac{1}{b-a} \int_a^b f$$

Proof. Since f is continuous, it is integrable on $[a, b]$. By the extreme value theorem it is also bounded and attains its bounds: $\exists p, q \in [a, b]$ such that

$$f(p) := \inf_{x \in [a,b]} f(x), \qquad f(q) = \sup_{x \in [a,b]} f(x)$$



Applying Theorem 4.12, part 2, with $m = f(p)$ and $M = f(q)$, we see that

$$(b-a)f(p) \le \int_a^b f \le (b-a)f(q)$$

Now divide by $b - a$ and apply the usual intermediate value theorem for f to see that the required ξ exists between p and q. ■

In the picture, when f is positive and continuous, the grey area equals that under the curve; imagine levelling off the blue hill with a bulldozer... The notation $f_{\text{av}} = \frac{1}{b-a} \int_a^b f$ is short for the average value of f on $[a, b]$: to see why this interpretation is sensible, approach $\int f$ via a sequence of Riemann sums on equally-spaced partitions P_n, then

$$\frac{1}{b-a} \int_a^b f = \lim_{n \to \infty} \sum_{i=1}^n f(x_i^*) \Delta x = \lim_{n \to \infty} \frac{f(x_1^*) + \cdots + f(x_n^*)}{n}$$

is the limit of a sequence of *averages* of equally-spaced samples $f(x_i^*)$.

What can/can't be integrated? (non-examinable)

We now know a great many examples of integrable functions: essentially

- Piecewise continuous & monotonic functions are integrable.
- Linear combinations, products, absolute values, maximums and minimums of (already) integrable functions.

After so many positive integrability conditions, it is reasonable to ask precisely which functions are Riemann integrable. There is a precise answer, though it is quite tricky to understand.

Theorem 4.21 (Lebesgue). *Suppose $f : [a,b] \to \mathbb{R}$ is bounded. Then*

$$f \text{ is Riemann integrable} \iff \text{it is continuous except on a set of measure zero}$$

Naïvely, the *measure* of a set is the sum of the lengths of its maximal subintervals; though unfortunately this doesn't make for a very useful definition.[4] Any countable subset has measure zero; Lebesgue's result is almost as if we can extend Corollary 4.18 to allow for infinite sums. Indeed you might have encountered a function which is continuous only on the irrationals; such a function is Riemann integrable. There are also some uncountable sets with measure zero such as Cantor's middle-third set: if f is the indicator function of Cantor's set

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathcal{C} \\ 0 & \text{otherwise} \end{cases}$$

then f is continuous except on $\mathcal{C}$, and is Riemann integrable with $\int_0^1 f(x)\,\mathrm{d}x = 0$.

Exercises 33

1. Explain why $\int_0^{2\pi} x^2 \sin^8(e^x)\,\mathrm{d}x \le \frac{8}{3}\pi^3$
2. If f is integrable on $[a,b]$ prove that it is integrable on any interval $[c,d] \subseteq [a,b]$.
3. We complete the proof of Theorem 4.8 (linearity of integration).
   (a) Suppose $k > 0$, let $A \subseteq \mathbb{R}$ and define $kA := \{kx : x \in A\}$. Prove that $\sup kA = k \sup A$ and $\inf kA = k \inf A$.
   (b) If $k > 0$ prove that kf is integrable on any interval and that $\int kf = k \int f$.
   (c) How should you modify your argument if $k < 0$?
4. Give an example of an integrable but *discontinuous* function on a closed bounded interval $[a,b]$ for which the conclusion of the Intermediate Value Theorem for Integrals is *false*.

[4] Formally, the *length* of an open interval (a,b) is $b-a$ and a set $A \subseteq \mathbb{R}$ has *measure zero* if

$$\forall \epsilon > 0,\ \exists \text{ open intervals } I_n \text{ such that } A \subseteq \bigcup_{n=1}^{\infty} I_n \text{ and } \sum_{i=1}^{\infty} \operatorname{length}(I_n) < \epsilon$$

More generally, the *measure* of a set (subject to a technical condition) is the infimum of the sum of the lengths of any countable collection of open covering intervals. A rigorous discussion of *measure theory* is properly a matter for graduate analysis. Somewhat surprisingly, there exist sets with positive measure that contain no subintervals, and even sets which are non-measurable!

5. Explicitly compute the value of the integral $\int_{1/2}^{15/2} x - \lfloor x \rfloor \, \mathrm{d}x$ (recall Example 4.19).

6. We prove and extend Corollary 4.10. Suppose f is integrable on $[a,b]$.

   (a) If $g : [a,b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ for all $x \in (a,b)$, prove that g is integrable and $\int_a^b g = \int_a^b f$.

   (*Hint: consider $h = f - g$ and show that $\int h = 0$*)

   (b) Now suppose $g : [a,b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ for all $x \in [a,b]$ except at finitely many points. Prove that g is integrable and $\int_a^b g = \int_a^b f$.

7. Show that an increasing function on $[a,b]$ is integrable and thus complete Theorem 4.16.

   (*Hint: Choose a partition P with* mesh $P < \frac{\epsilon}{f(b)-f(a)}$)

8. Suppose f and g are integrable on $[a,b]$.

   (a) Define $h(x) = (f(x))^2$. We know:

   - f is bounded: $\exists K$ such that $|f(x)| \le K$ on $[a,b]$.
   - Given $\epsilon > 0$, $\exists P$ such that $U(f,P) - L(f,P) < \frac{\epsilon}{2K}$. For each subinterval $[x_{i-1}, x_i]$, let

   $$M_i = \sup f(x), \qquad m_i = \inf f(x), \qquad \overline{M}_i = \sup h(x), \qquad \overline{m}_i = \inf h(x)$$

   Prove that $\overline{M}_i - \overline{m}_i \le 2(M_i - m_i)K$ and use this to conclude that h is integrable.

   (b) Prove that fg is integrable.

   (*Hint: $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$*)

   (c) Prove that $U(|f|, P) - L(|f|, P) \le U(f,P) - L(f,P)$ for any partition P. Hence conclude that $|f|$ is integrable.

   (*One can extend these arguments—it's a bit harder!—to show that if j is continuous, then $j \circ f$ is integrable. Parts (a) and (c) correspond, respectively, to $j(x) = x^2$ and $j(x) = |x|$.*)

9. (Hard) Let $f(x) = \begin{cases} x & \text{if } x \neq 0 \text{ and } \sin\frac{1}{x} > 0 \\ -x & \text{if } x \neq 0 \text{ and } \sin\frac{1}{x} < 0 \\ 0 & \text{if } x = 0 \end{cases}$

   (a) Show that f is not piecewise continuous on $[0,1]$.

   (b) Show that f is not piecewise monotonic on $[0,1]$.

   (c) Show that f is integrable on $[0,1]$.

   (*Hint: given ϵ, hunt for a suitable partition to make $U(f,P) - L(f,P) < \epsilon$ by considering $[0, x_1]$ differently to the other subintervals*)

   (d) Make a similar argument to show that $g = \sin\frac{1}{x}$ is integrable on $(0,1]$, where

   $$g(x) = \begin{cases} \sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

   Note that neither argument *evaluates* the integrals!

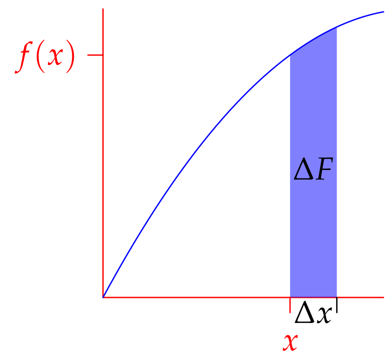
34 The Fundamental Theorem of Calculus

The key result linking integration and differentiation is usually presented in two parts.[5] While there are significant subtleties, the rough statements are as follows:

Part I Differentiation reverses integration: $\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x)$

Part II Integration reverses differentiation: $\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a)$

These facts seemed intuitively obvious to early practitioners of calculus: indeed, given a continuous positive function f:



- Let $F(x)$ denote the area under the curve between 0 and x.
- A small increase Δx results in the area increasing by ΔF.
- $\Delta F \approx f(x)\Delta x$ is approximately the area of a rectangle, whence $\frac{\Delta F}{\Delta x} \approx f(x)$. This is part I.
- $F(b) - F(a) \approx \sum \Delta F_i \approx \sum f(x_i)\Delta x_i$. Since $F' = f$, this is part II.

In fact when Leibniz introduced the symbols $\int$ and d in the late 1600's, it was partly to reflect the fundamental theorem.[6] If you're happy with non-rigorous notions of limit, rate of change, area, and (infinite) sums, the above is all you need!

Of course, we are very much concerned with the details: What must we assume regarding f and F, and how are these properties used in the proof?

Theorem 4.22 (FTC, part I). *Suppose f is integrable on $[a, b]$. For any $x \in [a, b]$, define*

$$F(x) := \int_a^x f(t)\,\mathrm{d}t$$

Then:

1. *F is uniformly continuous on $[a, b]$;*
2. *If f is continuous at $c \in [a, b]$, then F is differentiable at c with $F'(c) = f(c)$.*

As ever, the condition at $c = a$ should be *right*-continuous and the conclusion *right*-differentiable, etc.

Compare this with the naïve version above where we assumed f was continuous. We now require only the *integrability* of f, and its continuity at *one point* for the full result.

[5] We follow the traditional numbering; some authors reverse these.

[6] $\int$ is a stylized S for *sum*, while d stands for *difference*. Given a sequence $F = (F_0, F_1, F_2, \ldots, F_n)$, construct a new sequence of *differences*

$$\mathrm{d}F = (F_1 - F_0, F_2 - F_1, \ldots, F_n - F_{n-1})$$

which can then be summed:

$$\int \mathrm{d}F = (F_1 - F_0) + (F_2 - F_1) + \cdots (F_n - F_{n-1}) = F_n - F_0 \qquad (*)$$

Viewing a function as an 'infinite sequence' of values spaced along an interval, dF becomes a sequence of *infinitesimals* and $(*)$ is essentially the fundamental theorem: $\int \mathrm{d}F = F(b) - F(a)$. It is the conception of a function that is suspect here, not the essential relationship between sums and differences.

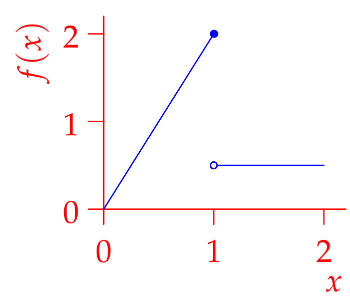
Examples 4.23. You should have seen many examples in an elementary calculus course.

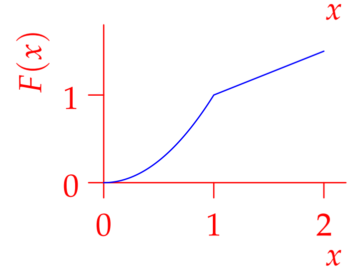
1. Since $f(x) = \sin^2(x^3 - 7)$ is continuous on any bounded interval, we conclude that
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_4^x \sin^2(t^3-7)\,\mathrm{d}t = \sin^2(x^3-7)$$
If one follows Theorem 4.13 and its resulting conventions, then this is valid for all $x \in \mathbb{R}$.

2. The chain rule permits more complicated examples. For instance, since $f(t) = \sin\sqrt{t}$ is continuous on its domain $[0,\infty)$ and $y(x) = x^2+3$ has range $[3,\infty) \subseteq \mathrm{dom}(f)$, we have
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{x^2+3} \sin\sqrt{t}\,\mathrm{d}t = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \sin\sqrt{t}\,\mathrm{d}t = 2x\sin\sqrt{x^2+3}$$

3. For a final positive example, observe that
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{\sin x}^{e^x} \tan(t^2)\,\mathrm{d}t = e^x\tan(e^{2x}) - \cos x\tan(\sin^2 x)$$
To evaluate this, one first chooses any constant a and writes
$$\int_{\sin x}^{e^x} = \int_a^{e^x} + \int_{\sin x}^a = \int_a^{e^x} - \int_a^{\sin x}$$
before differentiating. This is valid provided $\sin x$, e^x and a all lie in the same subinterval of
$$\mathrm{dom}\tan(t^2) = \mathbb{R}\setminus\{\pm\sqrt{\tfrac{\pi}{2}}, \pm\sqrt{\tfrac{3\pi}{2}}, \pm\sqrt{\tfrac{5\pi}{2}}, \ldots\}$$
Since $|\sin x| \le 1 < \sqrt{\frac{\pi}{2}}$, this requires
$$\left|e^{2x}\right| < \frac{\pi}{2} \iff x < \frac{1}{2}\ln\frac{\pi}{2}$$
Choosing $a = 1$ would certainly suffice.

4. Now consider why the theorem requires continuity. The piecewise continuous function
$$f : [0,2] \to \mathbb{R} : x \mapsto \begin{cases} 2x & \text{if } x \le 1 \\ \frac{1}{2} & \text{if } x > 1 \end{cases}$$
has a jump discontinuity at $x = 1$. We can still compute
$$F(x) = \begin{cases} \int_0^x 2t\,\mathrm{d}t = x^2 & \text{if } x \le 1 \\ \int_0^1 2t\,\mathrm{d}t + \int_1^x \frac{1}{2}\mathrm{d}t = \frac{1}{2}(x+1) & \text{if } x > 1 \end{cases}$$
This is continuous, indeed uniformly so. However the discontinuity of f results in F having a *corner* and thus being *non-differentiable* at $x = 1$. Indeed, $F'(x) = f(x)$ for all $x \ne 1$; that is, at all values of x where f is continuous.





Proving FTC I Neither half of the theorem is particularly difficult once you write down what you know and what you need to prove. Here are the key ingredients:

1. F uniformly continuous means controlling the size of

$$|F(y) - F(x)| = \left|\int_a^y f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right| = \left|\int_x^y f(t)\,\mathrm{d}t\right| \leq \int_x^y |f(t)|\ \mathrm{d}t$$

But the boundedness of f allows us to bound this last integral...

2. $F'(c) = f(c)$ means showing that $\lim_{x\to c} \frac{F(x)-F(c)}{x-c} = f(c)$, which means controlling the size of

$$\left|\frac{F(x) - F(c)}{x - c} - f(c)\right| = \left|\frac{1}{x-c}\int_c^x f(t)\,\mathrm{d}t - f(c)\right|$$

The trick here will be to bring the *constant* $f(c)$ inside the integral as $\frac{1}{x-c}\int_c^x f(c)\,\mathrm{d}t$ so that what we really have to control is the size of $\frac{1}{|x-c|}\int_c^x |f(t) - f(c)|\ \mathrm{d}t$. This is where the continuity of f comes in...

Proof. 1. Since f is integrable, it is bounded: $\exists M > 0$ such that $|f(x)| \leq M$ for all x.

Let $\epsilon > 0$ be given and define $\delta = \frac{\epsilon}{M}$. Then, for any $x, y \in [a, b]$,

$$\begin{aligned} 0 < y - x < \delta \implies |F(y) - F(x)| &= \left|\int_x^y f(t)\,\mathrm{d}t\right| \leq \int_x^y |f(t)|\ \mathrm{d}t && \text{(Theorem 4.12, part 4)} \\ &\leq M(y - x) && \text{(Theorem 4.12, part 2)} \\ &< M\delta = \epsilon \end{aligned}$$

We conclude that F is uniformly continuous on $[a, b]$.

2. Let $\epsilon > 0$ be given. Since f is continuous at c, $\exists\delta > 0$ such that, for all $t \in [a, b]$,

$$|t - c| < \delta \implies |f(t) - f(c)| < \frac{\epsilon}{2}$$

Now for all $x \in [a, b]$ (except c),

$$\begin{aligned} 0 < |x - c| < \delta \implies \left|\frac{F(x) - F(c)}{x - c} - f(c)\right| &= \left|\frac{1}{x-c}\int_c^x f(t) - f(c)\,\mathrm{d}t\right| && \text{(Theorem 4.8)} \\ &\leq \frac{1}{|x-c|}\int_c^x |f(t) - f(c)|\,\mathrm{d}t && \text{(Theorem 4.12)} \\ &\leq \frac{1}{|x-c|}\frac{\epsilon}{2}|x - c| = \frac{\epsilon}{2} < \epsilon \end{aligned}$$

Clearly $\lim_{x\to c} \frac{F(x)-F(c)}{x-c} = f(c)$, and so F is differentiable at c with $F'(c) = f(c)$. ∎

The Fundamental Theorem, part II As with part I, the *formulaic* part of the result should be familiar, though we are more interested in the assumptions and where they are needed.

Theorem 4.24 (FTC, part II). *Suppose g is continuous on $[a,b]$, differentiable on (a,b), and that g' is integrable*[7] *on (a,b). Then,*

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

Part II is often expressed in terms of *anti-derivatives*: F being an anti-derivative of f if $F' = f$. Combined with FTC I, we recover the familiar '$+c$' result and a simpler version of the fundamental theorem often seen in elementary calculus.

Corollary 4.25. *Let f be continuous on $[a,b]$.*

- *If F is an anti-derivative of f, then $\int_a^b f = F(b) - F(a)$.*
- *Every anti-derivative has the form $F(x) = \int_a^x f(t)\,\mathrm{d}t + c$ for some constant c.*

Examples 4.26. Again, basic examples should be familiar.

1. Plainly $g(x) = x^2 + 2x^{3/2}$ is continuous on $[1,4]$ and differentiable on $(1,4)$ with derivative $g'(x) = 2x + 3\sqrt{x}$; this last is continuous (and thus integrable) on $(1,4)$. We conclude that

$$\int_1^4 2x + 3\sqrt{x}\,\mathrm{d}x = x^2 + 2x^{3/2}\Big|_1^4 = (16+16) - (1+2) = 29$$

2. If $g(x) = \sin(3x^2)$, then $g'(x) = 6x\cos(3x^2)$. Certainly g satisfies the hypotheses of the theorem on any bounded interval $[a,b]$. We conclude

$$\int_a^b 6x\cos(3x^2)\,\mathrm{d}x = \sin(3b^2) - \sin(3a^2)$$

   Moreover, every anti-derivative of $f(x) = 6x\cos(3x^2)$ has the form $F(x) = \sin(3x^2) + c$.

3. Recall Example 4.23.4 where we saw that the discontinuity of f led to the *non-differentiability* of $F(x) = \int_0^x f(t)\,\mathrm{d}t$ at $x = 1$. The function F therefore fails the hypotheses of FTC II on the interval $[0,2]$.

   However, except at $x = 1$, F is an anti-derivative of f and moreover $\int_0^2 f(x)\,\mathrm{d}x = F(2) - F(0)$, so we *appear* to have the formulaic conclusion of FTC II, though this is tautological given the definition of F!

   The way out of this conundrum is to note that other anti-derivatives $\hat{F}$ of f exist (except at $x = 1$), and which fail to satisfy the conclusion. For instance

$$\hat{F}(x) = \begin{cases} x^2 & \text{if } x < 1 \\ \frac{1}{2}x & \text{if } x > 1 \end{cases} \implies \hat{F}(2) - \hat{F}(0) = 1 \neq \frac{3}{2} = \int_0^2 f(x)\,\mathrm{d}x$$

[7] See Definition 4.11 if you're unsure what it means for g' to be integrable on a bounded *open* interval.

Proving FTC II See Exercise 10 for a relatively easy proof when $g' = f$ is continuous. For the real McCoy, we can only rely on the *integrability* of g': the trick is to use the mean value theorem to write $g(b) - g(a)$ as a Riemann sum over a suitable partition.

Proof. Let $\epsilon > 0$ be given and choose a partition P such that $U(g', P) - L(g', P) < \epsilon$. Since g satisfies the mean value theorem on each subinterval of the partition P, we see that

$$\exists \xi_i \in (x_{i-1}, x_i) \quad \text{such that} \quad g'(\xi_i) = \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}$$

from which

$$g(b) - g(a) = \sum_{i=1}^{n} g(x_i) - g(x_{i-1}) = \sum_{i=1}^{n} g'(\xi_i)(x_i - x_{i-1})$$

This is a Riemann sum for g' associated to the partition P, hence,

$$L(g', P) \le g(b) - g(a) \le U(g', P)$$

However we also have $L(g', P) \le \int_a^b g' \le U(g', P)$. Since these hold for all ϵ, the proof is complete.

■

While we certainly used the integrability of g' in the proof, it might seem strange that we assumed it at all: shouldn't every derivative be integrable? Perhaps surprisingly, the answer is no! If you want a challenge, look up the *Volterra function*, which is differentiable everywhere, but whose derivative is *non-integrable* (on, for instance, $[0, 1]$)!

The Rules of Integration

If one wants to *evaluate* an integral, rather than merely show it exists, there are really only two options:

1. Evaluate Riemann sums and take limits: often difficult if not impossible to do explicitly.
2. Use FTC II. The problem now becomes the finding of *anti-derivatives*, for which the core method is essentially *guess and differentiate.* To obtain general rules, we attempt to reverse the rules of differentiation.

Integration by Parts First consider the product rule: the product $g = uv$ of two differentiable functions is differentiable with $g' = u'v + uv'$. Now apply Theorems 4.8, 4.12 and FTC II.

Corollary 4.27 (Integration by Parts). *Suppose u, v are continuous on $[a, b]$, differentiable on (a, b), and that u', v' are integrable on (a, b). Then*

$$\int_a^b u'(x)v(x)\,\mathrm{d}x = u(b)v(b) - u(a)v(a) - \int_a^b u(x)v'(x)\,\mathrm{d}x$$

This is significantly less useful than the product rule since it is only capable of transforming the integral of a product into another such integral.

Examples 4.28. You should have seen myriad examples in a previous course. With practice, there is no need to explicitly state u and v.

1. Let $u(x) = x$ and $v'(x) = \cos x$. Then $u'(x) = 1$ and $v(x) = \sin x$, whence

$$\begin{aligned}\int_0^{\pi/2} x\cos x\,\mathrm{d}x &= [x\sin x]_0^{\pi/2} - \int_0^{\pi/2}\sin x\,\mathrm{d}x = \frac{\pi}{2}\sin\frac{\pi}{2} - 0 - [-\cos x]_0^{\pi/2} \\ &= \frac{\pi}{2} + \cos\frac{\pi}{2} - \cos 0 = \frac{\pi}{2} - 1\end{aligned}$$

2. Let $u(x) = \ln x$ and $v'(x) = 1$. Then $u'(x) = \frac{1}{x}$ and $v(x) = x$, whence

$$\begin{aligned}\int_e^{e^2}\ln x\,\mathrm{d}x &= [x\ln x]_e^{e^2} - \int_e^{e^2}\frac{x}{x}\,\mathrm{d}x = e^2\ln e^2 - e\ln e - [x]_e^{e^2} \\ &= 2e^2 - e - e^2 + e = e^2\end{aligned}$$

Change of Variables/Substitution We now turn our attention to the chain rule. If $g(x) = F\big(u(x)\big)$, where F and u are differentiable, then g is differentiable with

$$g'(x) = \frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\mathrm{d}F}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} = F'\big(u(x)\big)u'(x)$$

Now integrate both sides; the only issue is what assumptions are needed to invoke FTC II.

Theorem 4.29 (Substitution Rule). *Suppose we have two continuous functions: $u : [a,b] \to \mathbb{R}$ and $f : \mathrm{range}(u) \to \mathbb{R}$. Suppose also that u is differentiable on (a,b) with integrable derivative u'. Then*

$$\int_a^b f\big(u(x)\big)\,u'(x)\,\mathrm{d}x = \int_{u(a)}^{u(b)} f(u)\,\mathrm{d}u$$

This is the famous 'u-sub' formula from elementary calculus.

Proof. We leave as an exercise the verification that both integrals exist. We may also assume that $\mathrm{range}(u)$ is an interval of positive length[8] for otherwise both integrals are trivially zero.
Choose any $c \in \mathrm{range}(u)$ and define

$$F : \mathrm{range}(u) \to \mathbb{R} \ \text{ by } \ F(v) := \int_c^v f(t)\,\mathrm{d}t$$

Since f is continuous, by FTC I we see that F is differentiable with $F'(u) = f(u)$. But now

$$\begin{aligned}\int_a^b f\big(u(x)\big)u'(x)\,\mathrm{d}x &= \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}x}F\big(u(x)\big)\right]\mathrm{d}x && \text{(chain rule)} \\ &= F\big(u(b)\big) - F\big(u(a)\big) && \text{(FTC II)} \\ &= \int_{u(a)}^{u(b)} f(u)\,\mathrm{d}u\end{aligned}$$

■

[8] By the intermediate and extreme value theorems, $\mathrm{range}(u)$ is already a closed bounded interval.

Examples 4.30. Reading the theorem is bad enough; its *application* often requires significant creativity in order to recognize a suitable substitution.[9]

1. To evaluate the integral $\int_0^{\sqrt{\pi}} 2x\sin x^2\,\mathrm{d}x$, consider the substitution $u(x) = x^2$ defined on $[0,\sqrt{\pi}]$. Certainly u is continuous, and its derivative $u'(x) = 2x$ is integrable on $(0,\sqrt{\pi})$. Finally $f(u) = \sin u$ is continuous on $\operatorname{range}(u) = [0,\pi]$. The hypotheses are satisfied, whence
$$\int_0^{\sqrt{\pi}} 2x\sin x^2\,\mathrm{d}x = \int_0^{\sqrt{\pi}} f\big(u(x)\big)u'(x)\,\mathrm{d}x = \int_0^{\pi} \sin u\,\mathrm{d}u = -\cos u\Big|_0^{\pi} = 2$$

2. For the following integral with $f(u) = \frac{1}{u^2+1}$, we make the substitution $u(x) = x^2 - 2$. Note that $u : [\sqrt{2},\sqrt{3}] \to [0,1]$ and that $u'(x) = 2x$ is integrable; moreover, $f(u)$ is continuous on $\operatorname{range}(u) = [0,1]$. We conclude that
$$\int_{\sqrt{2}}^{\sqrt{3}} \frac{2x}{x^4-4x^2+5}\,\mathrm{d}x = \int_{\sqrt{2}}^{\sqrt{3}} \frac{2x}{(x^2-2)^2+1}\,\mathrm{d}x = \int_0^1 \frac{1}{u^2+1}\,\mathrm{d}u = \arctan u\Big|_0^1 = \frac{\pi}{4}$$

3. The hypotheses on u really are all that is necessary. In particular, u doesn't need to be left-/right-differentiable at the endpoints of $[a,b]$! For instance, with $f(u) = u^2$ and $u(x) = \sqrt{x}$ on $[0,4]$, we easily verify
$$\frac{8}{3} = \int_0^4 \frac{1}{2}\sqrt{x}\,\mathrm{d}x = \int_0^4 \frac{x}{2\sqrt{x}}\,\mathrm{d}x = \int_0^4 f\big(u(x)\big)u'(x)\,\mathrm{d}x = \int_0^2 f(u)\,\mathrm{d}u = \int_0^2 u^2\,\mathrm{d}u = \frac{8}{3}$$

4. Sloppy use of the substitution rule might lead to utter nonsense. For instance, consider the 'substitution' $u = x^2$ in the following:
$$\int_{-1}^2 \frac{1}{x}\,\mathrm{d}x = \int_{-1}^2 \frac{1}{2x^2}2x\,\mathrm{d}x = \int_1^4 \frac{1}{2u}\,\mathrm{d}u = \frac{1}{2}(\ln 4 - \ln 1) = \ln 2$$
   Of course the left hand integral does not exist since $\frac{1}{x}$ is undefined at $0 \in (-1,2)$, so the conclusion is false. In the language of the substitution rule, $f(u) = \frac{1}{2u}$ is not continuous on $\operatorname{range}(u) = [0,4]$: it is not even defined at $u = 0$! You are very unlikely to make precisely this mistake since the first integral is so clearly undefined, but for more complicated functions...

[9] Hence the old adage, "Differentiation is a science; integration an art." To illustrate via an example, consider the function $f(x) = \tan(e^x\cos(3x^2) + 4x^3)$. The product and chain rules allow one to explicitly compute the derivative
$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{1}{1+(e^x\cos(3x^2)+4x^3)^2}\left(e^x\cos(3x^2) - 6xe^x\sin(3x^2) + 12x^2\right)$$
By contrast, the integration analogues (integration by parts/substitution) are essentially useless in attempting to find an *explicit* anti-derivative facilitating the integration of the same function via FTC II; for instance, the integral
$$\int_0^1 \tan(e^x\cos(3x^2)+4x^3)\,\mathrm{d}x$$
is likely impossible to evaluate explicitly and can only be approximated (e.g. via Riemann sums).

Exercises 34 1. Calculate the following limits:

(a) $\lim_{x \to 0} \dfrac{1}{x} \displaystyle\int_0^x e^{t^2} \, \mathrm{d}t$ (b) $\lim_{h \to 0} \dfrac{1}{h} \displaystyle\int_3^{3+h} e^{t^2} \, \mathrm{d}t$

2. Let $f(t) = \begin{cases} 0 & \text{if } t < 0 \\ t & \text{if } 0 \le t \le 1 \\ 4 & \text{if } t > 1 \end{cases}$

   (a) Determine the function $F(x) = \int_0^x f(t) \, \mathrm{d}t$.
   (b) Sketch F. Where is F continuous?
   (c) Where is F differentiable? Calculate F' at the points of differentiability.

3. Let f be a continuous function on $\mathbb{R}$.

   (a) Define $F(x) = \int_{x-1}^{x+1} f(t) \, \mathrm{d}t$. Carefully show that F is differentiable on $\mathbb{R}$ and compute F'.
   (b) Repeat for the function $G(x) = \int_0^{\sin x} f(t) \, \mathrm{d}t$.

4. Recall Examples 4.23.4 and 4.26.3. Find all anti-derivatives F of f on $[0, 1) \cup (1, 2]$. How many satisfy $\int_0^2 f(x) \, \mathrm{d}x = F(2) - F(0)$?

5. Consider integration by parts. Plainly $\int_a^x u'(t)v(t) \, \mathrm{d}t$ is an anti-derivative of $u'(x)v(x)$ by FTC I: what does integration by parts say is another?

6. Use change of variables to integrate $\int_0^1 x\sqrt{1 - x^2} \, \mathrm{d}x$

7. Use integration by parts and the substitution rule to evaluate $\int_0^b \arcsin x \, \mathrm{d}x$ for any $b < 1$.

8. Use integration by parts to evaluate $\int_0^b x \arctan x \, \mathrm{d}x$for any $b > 0$

9. Check that the assumptions of int by subs guarantee that both integrals are well-defined (i.e. that $(f \circ u)u'$ and f are integrable on the required intervals.

10. We prove a simpler version of the fundamental theorem of calculus.

   (a) Suppose f is continuous on $[a, b]$ and define $F(x) = \int_a^x f(t) \, \mathrm{d}t$. For any $c, x \in [a, b]$ where $c \neq x$, prove that

   $$m \le \frac{F(x) - F(c)}{x - c} \le M$$

   where m, M are the maximum and minimum values of $f(t)$ on the closed interval bounded by c, x. Make sure to explain why m, M exist, and use this to deduce that $F'(c) = f(c)$.
   (b) Suppose f is continuous on $[a, b]$ and that F is any anti-derivative of f on a, b (that is, $F' = f$). Use part (a) and the mean value theorem to prove that $\int_a^b f(t) \, \mathrm{d}t = F(b) - F(a)$.

36 Improper Integrals

The Riemann integral has several limitations. Even allowing for functions to be integrable on open intervals (Exercise 32.6), the definition of $\int_a^b f(x)\,\mathrm{d}x$ requires the following:

- That (a, b) be a *bounded* interval.
- That f be *bounded* on (a, b).

There is a natural way to extend the Riemann integral to unbounded intervals and functions: *limits.*

Definition 4.31. Suppose $f : [a, b) \to \mathbb{R}$ satisfies the following properties:

- f is integrable on every closed bounded subinterval $[a, t] \subseteq [a, b)$.
- Either $b = \infty$, or b is finite and f is unbounded at b,

The *improper integral* of f on $[a, b)$ is defined to be

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{t\to b^-} \int_a^t f(x)\,\mathrm{d}x$$

This is *convergent* or *divergent* in the same manner as the limit.
If an integral is improper at its lower limit then $\int_a^b f(x)\,\mathrm{d}x := \lim_{s\to a^+} \int_s^b f(x)\,\mathrm{d}x$.
If an integral is improper at both ends, choose any $c \in (a, b)$ and define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{s\to a^+} \int_s^c f(x)\,\mathrm{d}x + \lim_{t\to b^-} \int_c^t f(x)\,\mathrm{d}x$$

provided *both* one-sided improper integrals exist and the limit sum makes sense.

Theorem 4.13 says that the choice of c for a doubly-improper integral is irrelevant.

Many properties of the Riemann integral transfer to improper integrals, though not all. For example, part 1 of Theorem 4.12 extends:

Theorem 4.32. *If $0 \le f(x) \le g(x)$ on $[a, b)$, then $\int_a^b f \le \int_a^b g$, whenever the integrals exist (standard or improper). In particular:*

- $\int_a^b f = \infty \implies \int_a^b g = \infty$
- $\int_a^b g$ *converges* $\implies \int_a^b f$ *converges to a value* $\le \int_a^b g$.

We leave some of the detail to Exercise 36.7.

Examples 4.33. 1. $\int_0^t x^2\,dx = \frac{1}{3}t^3$ for any $t > 0$. Clearly

$$\int_0^\infty x^2\,dx = \lim_{t\to\infty} \frac{1}{3}t^3 = \infty$$

More formally, the improper integral $\int_0^\infty x^2\,dx$ diverges to infinity.

2. With $f(x) = x^{-4/3}$ defined on $[1,\infty)$,

$$\int_1^\infty x^{-4/3}\,dx = \lim_{t\to\infty} \int_1^t x^{-4/3}\,dx = \lim_{t\to\infty} \left[-3x^{-1/3}\right]_1^t = \lim_{t\to\infty} 3 - 3t^{-1/3} = 3$$

3. Consider $f(x) = |x|\, e^{-x^2/2}$ on $(-\infty,\infty)$. On a bounded interval $[0,t)$, we have

$$\int_0^t f(x)\,dx = \int_0^t xe^{-x^2/2}\,dx = \left[-e^{-x^2/2}\right]_0^t = 1 - e^{-t^2/2} \xrightarrow[t\to\infty]{} 1$$

By symmetry, we conclude that

$$\int_{-\infty}^\infty |x|\, e^{-x^2/2}\,dx = 1 + 1 = 2$$

This example is important in probability: multiplying by $\frac{1}{\sqrt{2\pi}}$, we have computed the the expectation of $|X|$ when X is a normally-distributed random variable

$$\mathbb{E}(|X|) = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\, |x|\, e^{-x^2/2}\,dx = \sqrt{\frac{2}{\pi}}$$

4. If $t \in [0,1)$, we can use our knowledge of derivatives $\frac{d}{dx}\sin^{-1} x = \frac{1}{\sqrt{1-x^2}}$ to evaluate

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{t\to 1^-} \int_0^t \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{t\to 1^-} \sin^{-1} t = \frac{\pi}{2}$$

and that, moreover $\int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,dx = \pi$. By comparison, we see that

$$\frac{1}{\sqrt{1-x^4}} \le \frac{1}{\sqrt{1-x^2}} \implies \int_{-1}^1 \frac{1}{\sqrt{1-x^4}}\,dx \le \int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,dx = \pi$$

5. Improper integrals need not exist. For instance,

$$\lim_{t\to\infty} \int_0^t \sin x\,dx = \lim_{t\to\infty} 1 - \cos t$$

diverges by oscillation.

Exercises 36 1. Use your answers from the previous section to decide whether the improper integrals $\int_0^1 \arcsin x \, \mathrm{d}x$ and $\int_0^\infty x \arctan x \, \mathrm{d}x$ exist. If so, what are their values?

2. Let p be a positive constant. Prove the following:

$$\int_0^1 \frac{1}{x^p} \, \mathrm{d}x = \begin{cases} \frac{1}{1-p} & \text{if } p < 1 \\ \infty & \text{if } p \geq 1 \end{cases} \qquad \int_1^\infty \frac{1}{x^p} \, \mathrm{d}x = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p \leq 1 \end{cases}$$

3. Explain why $\int_a^b f(x) \, \mathrm{d}x = \lim\limits_{t \to b^-} \int_a^t f(x) \, \mathrm{d}x$ holds, even when f is integrable on $[a, b]$.

4. State a version of integration by parts modified for when $\int_a^b u'(x)v(x) \, \mathrm{d}x$ is an improper integral. Now evaluate $\int_0^\infty xe^{-4x} \, \mathrm{d}x$.

5. What is wrong with the following calculation?

$$\int_{-\infty}^{\infty} x \, \mathrm{d}x = \lim_{t \to \infty} \frac{1}{2}x^2 \Big|_{-t}^{t} = \lim_{t \to \infty} \frac{1}{2}(t^2 - t^2) = \lim_{t \to \infty} 0 = 0$$

6. Prove or disprove: if $\int f$ and $\int g$ are convergent improper integrals, so is $\int fg$.

7. Prove part of Theorem 4.32. Suppose $0 \leq f(x) \leq g(x)$ for all $x \in [a, b)$, and that $\int_a^b g$ is a convergent improper integral. Prove that $\int_a^b f$ converges and that $\int_a^b f \leq \int_a^b g$.

Generalizing the Riemann Integral (non-examinable)

In the 1890's, Thomas Stieltjes[10] offered a generalization of the Riemann integral.

Definition 4.34. Let α be a monotonically increasing function on an interval $[a, b]$. Given a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ and a function f, define the differences

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

The *upper/lower Darboux–Stieltjes sums/integrals* are defined analogously to the pure Riemann case:

$$U(f, P, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i \qquad L(f, P, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$$
$$U(f, \alpha) = \inf U(f, P, \alpha) \qquad L(f, \alpha) = \sup L(f, P, \alpha)$$

f is *Riemann–Stieltjes integrable* of class $\mathcal{R}(\alpha)$ if $U(f, \alpha) = L(f, \alpha)$: we denote this value $\int_a^b f(x)\,\mathrm{d}\alpha$.

The standard Riemann integral corresponds to $\alpha(x) = x$. It is the ability to choose other functions α that makes the Riemann–Stieltjes integral both powerful and applicable.

Standard Properties Most results in sections 32 and 33 hold with suitable modifications, as does the discussion of improper integrals. For instance,

$$f \in \mathcal{R}(\alpha) \iff \exists P \text{ such that } U(f, P, \alpha) - L(f, P, \alpha) < \epsilon$$

The result regarding piecewise continuity of f is a notable exception: if f and α are simultaneously piecewise continuous then f might not lie in $\mathcal{R}(\alpha)$.

Weighted integrals If α is differentiable, then we obtain a standard Riemann integral

$$\int_a^b f(x)\,\mathrm{d}\alpha = \int_a^b f(x)\alpha'(x)\,\mathrm{d}x$$

weighted so that $f(x)$ contributes more when α is increasing rapidly.

Probability If $\alpha(a) = 0$ and $\alpha(b) = 1$, then α may be viewed as a *probability distribution function.*Its derivative α' is the corresponding *probability density function.* For example:

1. The *uniform distribution* on $[a, b]$ has $\alpha = \frac{1}{b-a}(x - a)$ so that

$$\int_a^b f(x)\,\mathrm{d}\alpha = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x$$

Since α' is constant, the integrals weigh all values of x *uniformly.*

2. The standard *normal distribution* has $\alpha(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,\mathrm{d}t$. The fact that $\alpha' = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is maximal when $x = 0$ reflects the fact that a normally distributed variable is clustered near its mean.

In all cases, $\int f(x)\,\mathrm{d}\alpha = \mathbb{E}(f(X))$ computes an expectation (see, for instance, Example 4.33.3).

[10]Stieltjes was Dutch; for the pronunciation try 'steelchez.'

Non-differentiable α A major flexibility comes when we allow α to be non-differentiable, or even discontinuous! For example, given a partition $Q = \{s_0, \ldots, s_n\}$ of $[a, b]$, and a positive sequence $(c_k)_{k=1}^n$, define

$$\alpha(x) = \begin{cases} 0 & \text{if } x = a \\ \sum\limits_{i=1}^{k} c_i & \text{if } x \in (s_{k-1}, s_k] \end{cases}$$

This is an increasing step function on $[a, b]$. The Riemann–Stieltjes integral becomes a weighted sum

$$\int_a^b f(x)\,\mathrm{d}\alpha = \sum_{i=1}^{n} c_i f(s_i)$$

Taking instead an infinite sequence $(s_n) \subseteq [a, b]$ results in an infinite series, which helps explain why so many results for series and integrals look similar!

This also touches on probability. For example, let $p \in [0, 1]$, $n \in \mathbb{N}$, and $s_k = k$ on the interval $[0, n]$. If $c_k = \binom{n}{k} p^k (1-p)^{n-k}$, then

$$\int f(x)\,\mathrm{d}\alpha = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f(x) = \mathbb{E}(f(X))$$

is the expectation of $f(X)$ when $X \sim B(n, p)$ is a binomially distributed random variable.

Integrals and Convergence

The *Lebesgue integral* is another common generalization. Its main purpose is to permit the transfer of integrability to the limit of a sequence of integrable functions.[11] To see the problem, consider the sequence

$$f_n : [0, 1] \to \mathbb{R} : x \mapsto \begin{cases} 1 & \text{if } x = \frac{p}{q} \in \mathbb{Q} \text{ with } q \le n \\ 0 & \text{otherwise} \end{cases}$$

Each f_n is piecewise continuous and thus Riemann integrable with $\int_0^1 f_n(x)\,\mathrm{d}x = 0$. However, the pointwise limit of f_n is the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

which is *not* Riemann integrable. In the Lebesgue theory, the limit f turns out to be integrable with integral 0, so that

$$\lim_{n \to \infty} \int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \lim_{n \to \infty} f_n(x)\,\mathrm{d}x$$

Recall that the interchange of limits and integrals would be automatic *if* the convergence $f_n \to f$ were *uniform*: of course the convergence isn't uniform here.

[11] Recall how *uniform convergence* does this for continuity.