Chapter 7: The Riemann Integral

When the derivative is introduced, it is not hard to see that the limit of the difference quotient should be equal to the slope of the tangent line, or when the horizontal axis is time and the vertical is distance, equal to the instantaneous velocity. But the integral is not so easily interpreted: Why is the area under the curve in any way related to the antiderivative? It really was a stroke of genius by Newton and Leibnitz to see that connection.

In fact, integration, in the sense of area or volume, was invented first, many centuries earlier: Archimedes was involved in trying to compute the volume of a wine cask. The derivative was a much harder concept. In fact, Zeno's paradoxes show how hard the Greeks found the concept of limit in general. Differentiation could really only be recognized and investigated after Descartes invented coordinate geometry. As a result, a few calculus texts, seeking a "historical" approach to the subject, cover integration before differentiation.

Definition. Let f be a <u>bounded</u> function on a <u>compact</u> interval $[a, b]$. (We'll see later why the underlined words are important. For now, just note that we are not assuming that f is even continuous.)

(a) A *partition* P of $[a, b]$ is a finite subset of $[a, b]$ containing a and b. If Q is another partition and $P \subseteq Q$, then Q is a *refinement* of P.

(b) Write $P = \{x_0, x_1, x_2, \ldots, x_n\}$ where $a = x_0 < x_1 < x_2 < \cdots < x_n = b$. Then

$$U(f, P) = \sum_{i=1}^{n} \sup\{f(x) : x_{i-1} \le x \le x_i\}(x_i - x_{i-1})$$
$$L(f, P) = \sum_{i=1}^{n} \inf\{f(x) : x_{i-1} \le x \le x_i\}(x_i - x_{i-1})$$

are the *upper* and *lower sums* with respect to f and P.

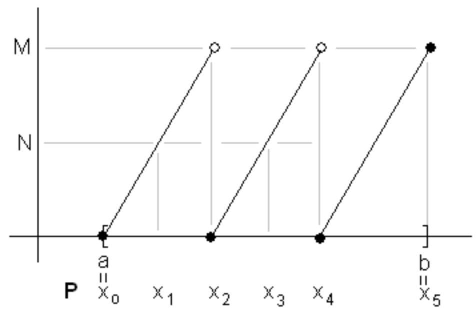
(c) The *upper* and *lower integrals* of f on $[a, b]$ are $U(f) = \inf\{U(f, P)\}$ and $L(f) = \sup\{L(f, P)\}$, where the inf and sup are taken over all partitions P of $[a, b]$.

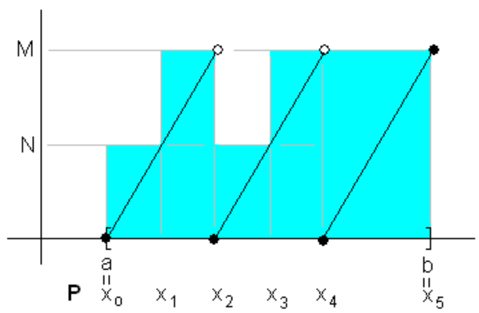
(d) If $U(f) = L(f)$, then f is *integrable* on $[a, b]$, and their common value is denoted $\int_a^b f$ or $\int_a^b f(x)\,dx$.

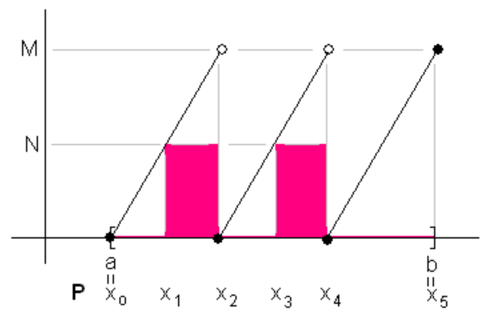
Example. On the interval $[a, b]$ with partition P as shown, consider the function f. We have

$$U(f, P) = N(x_1 - x_0) + M(x_2 - x_1) + N(x_3 - x_2) + M(x_4 - x_3) + M(x_5 - x_4)$$
$$L(f, P) = 0(x_1 - x_0) + N(x_2 - x_1) + 0(x_3 - x_2) + N(x_4 - x_3) + 0(x_5 - x_4)$$

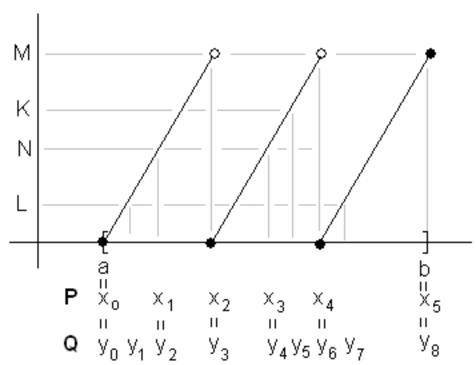
The blue region shows the area representing the upper sum, and the pink region the lower sum.

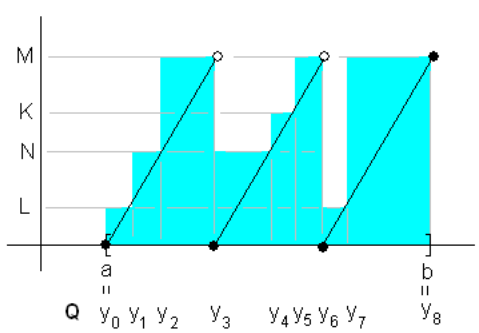


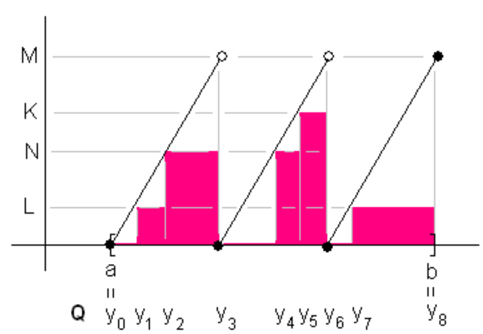




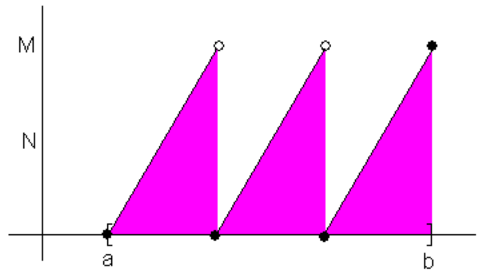
With a partition Q that refines P, the upper sum decreases and the lower sum increases:







Replacing the partitions with ever finer ones, we see that the upper sums and lower sums approach each other, with the area of the three triangles as the common limit: $U(f) = L(f) = \int_a^b f$.



We see now why we restricted to bounded functions: So that the sup and inf of f on each subinterval exists (and is finite). And similarly we restricted to bounded intervals so that we do not have infinitely long subintervals or infinitely many subintervals, making the upper and lower sums much harder or impossible to interpret. (We restricted to closed intervals so that the function has values at the endpoints.) Later in a calculus course integrals of unbounded functions or over unbounded or unclosed intervals might be allowed, but such things are always called "improper integrals" and interpreted as limits of proper integrals, limits that may or may not exist:

Example. Improper integrals:

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{\varepsilon\to 0^+}\int_\varepsilon^1 \frac{1}{\sqrt{x}}\,dx = \lim_{\varepsilon\to 0^+}\left[2x^{1/2}\right]_\varepsilon^1 = \lim_{\varepsilon\to 0^+} 2(1-\sqrt{\varepsilon}) = 2$$

$$\int_1^\infty \frac{1}{\sqrt{x}}\,dx = \lim_{M\to\infty}\int_1^M \frac{1}{\sqrt{x}}\,dx = \lim_{M\to\infty}\left[2x^{1/2}\right]_1^M = \lim_{M\to\infty} 2(\sqrt{M}-1) = \infty$$

$$\int_{-\infty}^\infty \sin x\,dx = ?\lim_{M\to\infty}\int_{-M}^M \sin x\,dx = \lim_{M\to\infty}\left[-\cos x\right]_{-M}^M = \lim_{M\to\infty} 0 = 0$$

The correct version of the last one is

$$\int_{-\infty}^\infty \sin x\,dx = \lim_{N\to-\infty}\lim_{M\to\infty}\int_N^M \sin x\,dx = \lim_{N\to-\infty}\lim_{M\to\infty}\left[-\cos x\right]_N^M \text{ does not exist}$$

The limit where the left and right endpoint approach $-\infty$ and ∞ at the same rate is the called the "principal value of the improper integral."

Example. A function that is not integrable: The Dirichlet function $\chi_\mathbb{Q}$ on $[0,1]$. Every subinterval in every partition contains rational numbers, so the supremum of the $\chi_\mathbb{Q}$-values on the subinterval is 1, so the upper sum for every partition is 1, so the upper integral is 1. But every subinterval in every partition also contains irrational numbers, so the infimum of the $\chi_\mathbb{Q}$-values on the subinterval is 0, so the lower sum for every partition is 0, so the lower integral is 0.

Lemma. *For any bounded function f on the compact interval $[a,b]$:*

(a) For any partition P of $[a,b]$, $L(f,P)\le U(f,P)$.

(b) For any partitions P,Q of $[a,b]$, if $P\subseteq Q$, then $L(f,P)\le L(f,Q)$ and $U(f,P)\ge U(f,Q)$.

(c) For any partition P of $[a,b]$, $L(f,P)\le U(f)$ Therefore, $L(f)\le U(f)$.

Proof. (a) On any subinterval $[x_{i-1},x_i]$, $\inf\{f(x): x_{i-1}\le x\le x_i\}\le \sup\{f(x): x_{i-1}\le x\le x_i\}$.

(b) It is enough to assume that Q has only one more element than P, say $P=\{x_0,x_1,\dots,x_n\}$ and Q has the additional division point x^* between x_{j-1} and x_j. Then $p=\inf\{f(x): x_{j-1}\le x\le$

$x_j\}$ is less than or equal to each of $q_1 = \inf\{f(x) : x_{j-1} \le x \le x^*\}$ and $q_2 = \inf\{f(x) : x^* \le x \le x_j\}$ — it is equal to at least one, but no larger than either. So the term $p(x_j - x_{j-1})$ in $L(f,P)$ that corresponds to $[x_{j-1}, x_j]$ is at most the sum $q_1(x^* - x_{j-1}) + q_2(x_j - x^*)$ of the two terms in $L(f,Q)$ that correspond to $[x_{j-1}, x^*]$ and $[x^*, x_j]$ — because $x_j - x_{j-1} = (x^* - x_{j-1}) + (x_j - x^*)$. Adding up all the terms for all the subintervals in each partition, we see that $L(f,P) \le L(f,Q)$. Similarly, $U(f,P) \ge U(f,Q)$.

(c) Assume BWOC that there is a partition P for which $L(f,P) > U(f)$. Then $L(f,P)$ is not a lower bound for the set of all upper sums, so there is a partition Q for which $L(f,P) > U(f,Q)$. But then $P \cup Q$ is a refinement of both P and Q, so by (b) and (a), we have

$$L(f,P) \le L(f, P\cup Q) \le U(f, P\cup Q) \le U(f,Q) \ ,$$

a contradiction. Thus, $U(f)$ is an upper bound on the set of all lower sums, it is at least as large as the least upper bound, and we have the "Therefore" sentence. □

Here is another way to say "integrable", in which, instead of finding one partition P that makes $L(f,P)$ almost as large as possible and another, Q, that makes $U(f,P)$ almost as small as possible, we say that it is enough to find a single partition that makes the upper and lower sums close to each other:

Proposition. *A bounded function $f : [a,b] \to \mathbb{R}$ is integrable iff, $\forall\, \varepsilon > 0$, $\exists\, P$ partition of $[a,b]$ s.t. $U(f,P) - L(f,P) < \varepsilon$.*

Proof. ($\Leftarrow$) Assume BWOC that $U(f) \ne L(f)$. Then then $U(f) - L(f) = \varepsilon > 0$. By hypothesis, $\exists\, P$ s.t. $U(f,P) - L(f,P) < \varepsilon$. But because $L(f,P) \le L(f)$ and $U(f,P) \ge U(f)$, we have $U(f) - L(f) \le U(f,P) - L(f,P) < \varepsilon = U(f) - L(f)$, $\rightarrow\!\leftarrow$.

($\Rightarrow$) Let $\varepsilon > 0$ be given, and pick P, Q s.t. $U(f,P) - \int_a^b f < \varepsilon/2$ and $\int_a^b f - L(f,Q) < \varepsilon/2$ Then because $P \cup Q$ is a common refinement of P, Q, we have

$$L(f,Q) \le L(f, P\cup Q) \le U(f, P\cup Q) \le U(f,P) \ ,$$

so

$$\begin{aligned} U(f, P\cup Q) - L(f, P\cup Q) &\le U(f,P) - L(f,Q) \\ &= \left(U(f,P) - \int_a^b f\right) + \left(\int_a^b f - L(f,Q)\right) < \varepsilon \ . \end{aligned}$$

□

Now we show that a continuous function (on a compact interval) is integrable. Because not all continuous functions are differentiable, we see that it is harder for a function to be differentiable than for it to be integrable.

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

Proof. Let $\varepsilon > 0$ be given. Because f is uniformly continuous, there is a $\delta > 0$ for which $|x_i - x_j| < \delta$ implies $|f(x_i) - f(x_j)| < \varepsilon/(b-a)$. Pick a partition P for which all the subintervals $[x_{i-1}, x_i]$ have length less than δ. (For example, we might choose n in $\mathbb{N}$ so large that $(b-a)/n < \delta$ and set $P = \{x_i = a + i(b-a)/n : i = 0, 1, \ldots, n\}$.) Then because f attains its inf and sup on each $[x_{i-1}, x_i]$, say at t_i and u_i respectively, we have $|t_i - u_i| \le |x_i - x_{i-1}| < \delta$, so $|f(t_i) - f(u_i)| < \varepsilon/(b-a)$, so

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (f(t_i) - f(u_i))(x_i - x_{i-1}) < \sum_{i=1}^{n} \frac{\varepsilon}{b-a}(x_i - x_{i-1}) = \frac{\varepsilon}{b-a}(x_n - x_0) = \varepsilon \ ,$$

the second last equality because the sum

$$(x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1})$$

"telescopes", i.e., all the terms cancel except the second and second last. $\square$

But there are many discontinuous integrable functions; our first example (the "three triangles") was discontinuous at two points but still integrable. We have seen that the Dirichlet function on $[0,1]$ is not integrable. But we claim that the Thomae function <u>is</u> integrable on this interval:

Example. Recall that the Thomae function is given by

$$t(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ \frac{1}{n} & \text{if } x = \frac{m}{n} \text{ in lowest terms} \end{cases}$$

Because every subinterval of every partition of $[0,1]$ contains irrational numbers, the lower sum of t with respect to every partition is 0, so the lower integral of t is 0. Thus, to see that t is integrable, we need to show that the upper integral is 0, i.e., that, given $\varepsilon > 0$, we can find a partition P with respect to which $U(t,P) < \varepsilon$. We can do that: There are only finitely many rationals x in $[0,1]$ for which $f(x) > \varepsilon/2$. Pick a partition P of $[0,1]$ so that these finitely many rationals are the centers (or, for $x = 0$ and $x = 1$, the ends) of subintervals with total length $\varepsilon/2$. Then in $U(f,P)$, the terms corresponding to those subintervals add up to at most 1 times the total length of these subintervals and hence less than $\varepsilon/2$. In the other subintervals the value of the function is at most $\varepsilon/2$, so those subintervals contribute to $U(f,P)$ a total of less than $(\varepsilon/2)\cdot 1$. It follows that $U(f,P) < \varepsilon$.

In fact, it is shown in the text that a bounded function on a compact interval is integrable iff its set of discontinuities has "measure zero." We won't try to explain this term, because the right way to do that is to take a totally different approach to the course, in terms of "Lebesgue integration." So we will only say that all countable sets, including finite sets, and the Cantor set have "measure zero," so that functions that are discontinuous only on these sets are integrable. ("Measure" is a concept that extends the idea of length, so a set of length ℓ has measure ℓ; but it is hard to assign a "length" to a set like the rationals, which has measure 0.) Remember that the Dirichlet function is discontinuous everywhere, so its set of discontinuities in $[0,1]$ has measure one; but the Thomae function is discontinuous only on the rationals, which have measure 0. And sure enough, the Dirichlet function is not integrable, while the Thomae function is.

`[At this point students are ready to do the eleventh problem set.]`

Proposition. (Properties of the definite integral) *Suppose f, g are integrable on $[a,b]$ and k is a constant. Then:*

(0) If $a < c < b$, then $\int_a^b f = \int_a^c f + \int_c^b f$.

(1) $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

(2) $\int_a^b kf = k\int_a^b f$.

(3) If $f \le g$ on $[a,b]$, then $\int_a^b f \le \int_a^b g$.

Corollary. *(of (3)): (3a) If $m \le f \le M$ on $[a,b]$, then $m(b-a) \le \int_a^b f \le M(b-a)$.*
(3b) $|f|$ is integrable and $|\int_a^b f| \le \int_a^b |f|$.

The fact that $|f|$ is integrable is new information — the proof is an exercise. To do it, you should verify that, if m_i, M_i are the inf and sup of f on $[x_{i-1}, x_i]$ and m_i', M_i' are the same for $|f|$, then $M_i - m_i \geq M_i' - m_i'$. To get the second assertion in (3b), use $-|f| \leq f \leq |f|$ on $[a.b]$.

Remarks on the proposition:

(0) Part of what is to be proved here is that f is integrable on $[a, b]$ iff it is integrable on both $[a, c]$ and $[c, b]$. But if we define $\int_a^a f = 0$ and $\int_b^a f = -\int_a^b f$, then we don't need to add the "If" part of this statement as long as f is integrable on the intervals determined by a, b, c.

(1) The proof is an exercise.

(2) Here is a proof of the $k < 0$ case: Let $\varepsilon > 0$ be given, and pick a partition P of $[a, b]$ for which $U(f, P) - \int_a^b f < \varepsilon/(2|k|)$ and $\int_a^b f - L(f, P) < \varepsilon/(2|k|)$. Then because, on each subinterval $[x_{i-1}, x_i]$ in P we have

$$\sup\{kf(x) : x_{i-1} \leq x \leq x_i\} = k \cdot \inf\{f(x) : x_{i-1} \leq x \leq x_i\}$$

and similarly with sup and inf reversed, it follows that

$$\begin{aligned} U(kf, P) - L(kf, P) &= kL(f, P) - kU(f, P) = |k|(U(f, P) - L(f, P)) \\ &\leq |k| \left(U(f, P) - \int_a^b f + \int_a^b f - L(f, P) \right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \ , \end{aligned}$$

so kf is integrable, and

$$U(kf, P) - k \int_a^b f = k \left(L(f, P) - \int_a^b f \right) = |k| \left(\int_a^b f - L(f, P) \right) < \frac{\varepsilon}{2} < \varepsilon \ ,$$

so $k \int_a^b f = \int_a^b kf$.//

(3) Because $f \leq g$, for any subinterval $[x_{i-1}, x_i]$ of any partition P of $[a, b]$ we have

$$\sup\{f(x) : x_{i-1} \leq x \leq x_i\} \leq \sup\{g(x) : x_{i-1} \leq x \leq x_i\}$$

so $U(f, P) \leq U(g, P)$, so $\int_a^b f = \inf_P U(f, P) \leq \inf_P U(g, P) = \int_a^b g$.

`[At this point students are ready to do the twelfth problem set.]`

Theorem. (Fundamental Theorem of Calculus) *(1) If f is the derivative of F on $[a, b]$ and f is integrable, then $\int_a^b f = F(b) - F(a)$.*

(2) Let $g : [a, b] \to \mathbb{R}$ be integrable, and define $G(t) = \int_a^t g(x)\, dx$. Then $G : [a, b] \to \mathbb{R}$ is continuous. If g is continuous at c in $[a, b]$, then G is differentiable at c and $G'(c) = g(c)$.

Proof. (1) For any subinterval $[x_{i-1}, x_i]$ in any partition P of $[a, b]$, by the Mean Value Theorem $\exists\, x_i^* \in [x_{i-1}, x_i]$ s.t. $F(x_i) - F(x_{i-1}) = f(x_i^*)(x_i - x_{i-1})$, so $\sum_{i=1}^n (F(x_i) - F(x_{i-1}))$ is between $L(f, P)$ and $U(f, P)$. But the sum is telescoping, to $F(b) - F(a)$, independent of the partition P, so $F(b) - F(a)$ is between $L(f)$ and $U(f)$, which are both equal to $\int_a^b f$.

(2) In fact, we can show that G is Lipschitz, which is stronger than continuous: Let M be an upper bound on $|g|$ in $[a,b]$. Then for all $x_1, x_2 \in [a,b]$, we have

$$|G(x_1) - G(x_2)| = \left| \int_a^{x_1} g - \int_a^{x_2} g \right| = \left| \int_{x_2}^{x_1} g \right| \le M|x_1 - x_2| \; .$$

Now suppose g is continuous at c. Then

$$G'(c) = \lim_{x \to c} \frac{G(x) - G(c)}{x - c} = \lim_{x \to c} \frac{\int_c^x g}{x - c} \; ,$$

so we want to show, given $\varepsilon > 0$, $\exists\, \delta > 0$ s.t. $0 < |x - c| < \delta \implies |\int_c^x g/(x-c) - g(c)| < \varepsilon$. Now there <u>is</u> a $\delta > 0$ s.t., if $|x - c| < \delta$, then $|g(x) - g(c)| < \varepsilon$, and for <u>that</u> δ, if $0 < |x - c| < \delta$, then

$$\begin{aligned} \left| \frac{\int_c^x g}{x - c} - g(c) \right| &= \left| \frac{1}{x - c} \left(\int_c^x (g(t) - g(c)) dt \right) \right| \\ &\le \frac{1}{x - c} \int_c^x |g(t) - g(c)| dt \le \frac{1}{x - c} \int_c^x \varepsilon \, dt = \varepsilon \; . \end{aligned}$$

(In the middle of this, note that, if $x - c < 0$, then any $\int_c^x$ of a nonnegative function is also ≤ 0.) $\square$

Here are some examples of functions g and the functions G defined by their integrals, $G(x) = \int_c^x g(t)\, dt$.

Example. In these two examples, $c = -2$, so that, no matter what g is, $G(-2) = \int_{-2}^{-2} g = 0$. We are interested especially in the points at which the definitions of g change from one formula to another. First:

$$g(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x \ge 0 \end{cases} \; ,$$

$$G(x) = \int_{-2}^x g(t)\, dt = \begin{cases} \int_{-2}^x -1\, dt = -x - 2 & \text{if } x < 0 \\ G(0) + \int_0^x 1\, dt = -2 + x & \text{if } x \ge 0 \end{cases}$$

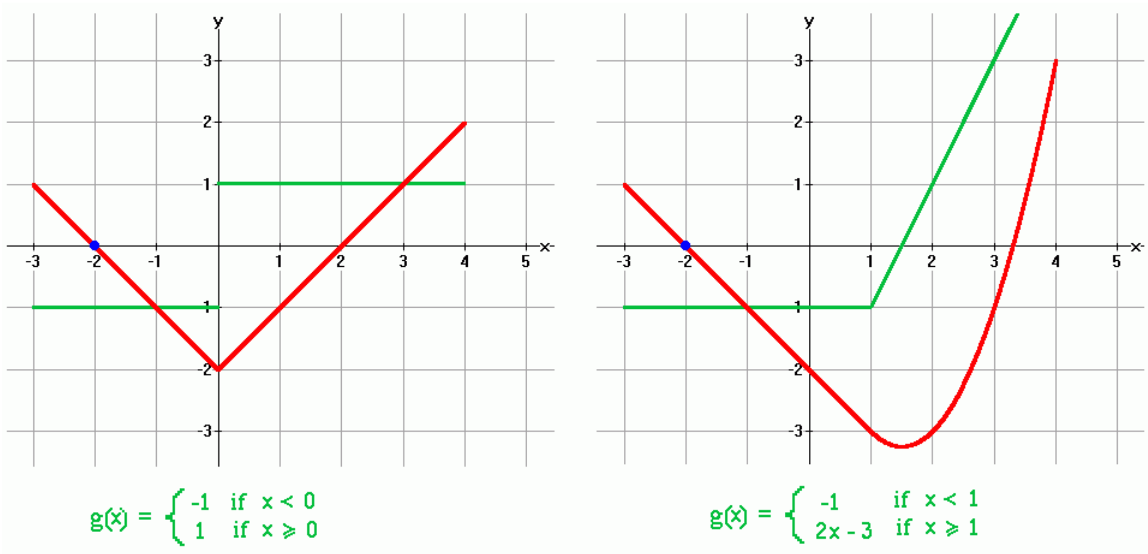
Here, G is differentiable except at the discontinuity $x = 0$ of g. Second:

$$g(x) = \begin{cases} -1 & \text{if } x \le 1 \\ 2x - 3 & \text{if } x > 1 \end{cases} \; ,$$

$$G(x) = \int_{-2}^x g(t)\, dt = \begin{cases} -x - 2 & \text{if } x \le 1 \\ G(1) + \int_1^x (2t - 3) dt = -3 + (x^2 - 3x) - (-2) = x^2 - 3x - 1 & \text{if } x > 1 \end{cases}$$

Because this g is continuous at the point where its definition changes, the corresponding G is differentiable there.

Here are the corresponding graphs, with the g's in green and the G's in red:



Though our text doesn't seem to include them, I have seen in other texts examples of functions defined by integrals where the upper, and maybe also the lower, limit(s) of integration are themselves functions. In order to find their derivatives, we only need to throw in the Chain Rule: Suppose we are given $H(x) = \int_c^{f(x)} g(t)\,dt$. We can isolate an intervening function: $G(u) = \int_c^u g(t)\,dt$. Then $H(x) = G(f(x))$, so $(d/dx)(H(x)) = G'(f(x)) \cdot f'(x) = g(f(x))f'(x)$.

Example.

$$\begin{aligned}\frac{d}{dx}\int_{2x}^{5x^2+3} \exp(-t^2)\,dt &= \frac{d}{dx}\left(\int_0^{5x^2+3} \exp(-t^2)\,dt - \int_0^{2x} \exp(-t^2)\,dt\right)\\ &= \exp(-(5x^2+3)^2)\cdot(10x) - \exp(-(2x)^2)\cdot 2\\ &= 10x\exp(-(5x^2+3)^2) - 2\exp(-(2x)^2)\ .\end{aligned}$$

Here, the intervening function is $G(u) = \int_0^u \exp(-t^2)\,dt$, so that $G'(u) = \exp(-u^2)$. There is nothing special about 0; because $\exp(-t^2)$ is continuous everywhere, any constant would have worked.

`[At this point students are ready to do the thirteenth problem set.]`