We’ve gotten pretty far with our “primitive” integral of functions of bounded variation. But not all continuous functions are BV (nor are all BV functions continuous), so we should define a more robust integral so that the set of integrable functions includes both cases.

The setup is pretty much the same as for the integral we already have: To calculate
\[ \int_{a}^{b} f(x) \, dx \]
we partition the interval \([a, b]\) with points
\[ x_0 = a < x_1 < x_2 < \cdots < x_n = b \]
– the new ingredient is how we choose to “sample” the function \(f\) in the interval \([x_{i-1}, x_i]\). When our functions were monotonic, we chose to evaluate \(f\) either the left or right endpoint, since that would give the maximum or minimum value of \(f\) on the subinterval.
Riemann vs Darboux

There are two approaches to sampling:

- **Riemann**: In the “Riemann integral” you choose a random point $x_i^*$ in the interval $[x_{i-1}, x_i]$ for each $i$, and create the *Riemann sum*

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

If the limit of $RS(f, P)$ as $P$ becomes finer and finer ($n \to \infty$ and $|P| \to 0$) exists, then that is the Riemann integral.

- **Darboux**: In the “Darboux integral” you make two sums: For the *upper Darboux sum* you sample $[x_{i-1}, x_i]$ by simply letting $f_i^+$ be the supremum of $f(x)$ on the interval $[x_{i-1}, x_i]$, and you make the *lower Darboux sum* by letting $f_i^-$ be the infimum of $f$ on the subinterval.
This is going to look familiar (compare with slide set 4). Let $f$ be any bounded function on $[a, b]$, and $P$ a partition of $[a, b]$. 

**Definition: (upper and lower sums)**

The *upper sum of $f$ corresponding to the partition $P$* is

$$U(f, P) = \sum_{i=1}^{n} f_i^+(x_i - x_{i-1})$$

and the *lower sum of $f$ corresponding to the partition $P$* is

$$L(f, P) = \sum_{i=1}^{n} f_i^-(x_i - x_{i-1})$$

Since $f_i^+ = \sup_{x \in [x_{i-1}, x_i]} f(x) \geq \inf_{x \in [x_{i-1}, x_i]} f(x) = f_i^-$ for each $i$, we have $U(f, P) \geq L(f, P)$.
Setup for the integral – definition

Since there’s always a common refinement $P$ for any partitions $P_1$ and $P_2$, we have

$$L(f, P_1) \leq L(f, P) \leq U(f, P) \leq U(f, P_2)$$

where $P$ is a common refinement for $P_1$ and $P_2$. So for any pair of partitions $P_1$ and $P_2$ we have

$$L(f, P_1) \leq U(f, P_2).$$

Therefore $\sup L(f, P)$ over all partitions $P$ (which is called the lower Darboux integral) is less than or equal to $\inf U(f, P)$ (the upper Darboux integral). If these are equal then we say the function $f$ is integrable and their common value is called the integral:

$$\int_a^b f(x) \, dx$$
To show that the integral exists, it is sufficient to find, for any \( \varepsilon > 0 \), a partition \( P \) such that

\[
U(f, P) < L(f, P) + \varepsilon.
\]

We did this before for monotonic functions, we can do this by using sufficiently fine regular partitions – these are partitions having the \( x_i \)'s evenly spaced (so \( x_i - x_{i-1} = \frac{b - a}{n} \) for all \( i = 1, \ldots, n \)).

**Theorem**

If \( f : [a, b] \to \mathbb{R} \) is monotonic, then \( \int_a^b f(x) \, dx \) exists.

Proof: Given \( \varepsilon > 0 \), choose \( n \) so large that the regular partition \( P \) with \( n \) steps, will have \( U(f, P) - L(f, P) = \frac{f(b) - f(a)}{n} < \varepsilon \).
Proposition (Greater generality)

Then we extended the integral to functions of bounded variation: If $f(x)$ can be written as the sum of two monotonic functions $p(x)$ and $q(x)$ with $p$ increasing and $q$ decreasing on $[a, b]$ (such a function is called a function of bounded variation), then

$$\int_{a}^{b} f(x) \, dx \text{ exists and is equal to } \int_{a}^{b} p(x) \, dx + \int_{a}^{b} q(x) \, dx.$$ 

Now, the goal is to extend the whole thing to continuous functions.
Because the upper and lower sums have these properties, it follows that the integral does:

1. **Linearity:**
   \[\int_a^b \alpha f(x) + \beta g(x) \, dx = \alpha \int_a^b f(x) \, dx + \beta \int_a^b g(x) \, dx \]
   for constants \(\alpha, \beta\) and \(f\) and \(g\) are integrable functions on \([a, b]\).

2. **Monotonicity:** If \(f(x) \geq g(x)\) for all \(x \in [a, b]\) then
   \[\int_a^b f(x) \, dx \geq \int_a^b g(x) \, dx.\]

3. **Absolute Value:** If \(f(x)\) is integrable on \([a, b]\), then so is \(|f(x)|\) and
   \[\left|\int_a^b f(x) \, dx\right| \leq \int_a^b |f(x)| \, dx.\]
More basic properties

4 If $a < b < c$ then

$$\int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx.$$ 

5 If $m \leq \inf\{f(x) \mid x \in [a, b]\}$ and $M \geq \sup\{f(x) \mid x \in [a, b]\}$ then

$$(b - a)m \leq \int_a^b f(x) \, dx \leq (b - a)M.$$ 

6 Mean value theorem for integrals: If $f$ is continuous and integrable (soon we can remove this as an assumption) on $[a, b]$ then there is a $c$ with $a < c < b$ such that

$$f(c) = \frac{1}{b - a} \int_a^b f(x) \, dx.$$
Let $F(x)$ be a differentiable function on $[a, b]$ with derivative $F'(x)$, and suppose $F'$ is an integrable function on $[a, b]$. Then

$$\int_{a}^{b} F'(x) \, dx = F(b) - F(a).$$

For any interval $(x_{i-1}, x_i)$ in the partition $P$ we have

$$F'_i^- \leq \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}} \leq F'_i^+$$

by the mean value theorem (for derivatives), since the middle quantity is $F'(x)$ for some $x$ in the interval.
But then

\[ L(F', P) = \sum_{i=1}^{n} F_i^+(x_i - x_{i-1}) \]

\[ \leq \sum_{i=1}^{n} F(x_i) - F(x_{i-1}) \]

\[ \leq \sum_{i=1}^{n} F_i^+(x_i - x_{i-1}) = U(F', P). \]

But the middle sum telescopes to \( F(x_n) - F(x_0) = F(b) - F(a) \).

Now \( F(b) - F(a) \) is trapped between \( \sup_{P} L(F', P) \) and \( \inf_{P} U(F', P) \), both of which are equal to the integral \( \int_{a}^{b} F'(x) \, dx \) because \( F' \) is integrable.
Fundamental theorem of calculus II: Derivatives of integrals

Let $f(x)$ be a continuous (and integrable) function on $[a, b]$. Define the function $F(x)$ via

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

Then $F$ is differentiable and $F'(x) = f(x)$.

First, if $h > 0$ we have

$$F(x + h) - F(x) = \int_x^{x+h} f(t) \, dt = hf(c)$$

for some $c$ between $x$ and $x + h$ by properties 4 and 6 (mean value theorem for integrals).
Therefore the difference quotient

\[
\frac{F(x + h) - F(x)}{h} = f(c) \quad \text{where } x < c < x + h
\]

But since \( f \) is continuous, \( f(c) \to f(x) \) as \( h \to 0^+ \).

For \( h < 0 \) we use that \( F(x + h) - F(x) = -\int_{x+h}^{x} f(t) \, dt = hf(c) \)

for some \( c \) between \( x + h \) and \( x \), and the proof goes through as for the \( h > 0 \) case.
Uniform continuity

To show that continuous functions on closed intervals are integrable, we’re going to define a slightly stronger form of continuity:

**Definition (uniform continuity):**

A function $f(x)$ is *uniformly continuous* on the domain $D$ if for every $\varepsilon > 0$ there is a $\delta > 0$ that depends only on $\varepsilon$ and not on $x \in D$ such that for every $x, y \in D$ with $|x - y| < \delta$, it is the case that $|f(x) - f(y)| < \varepsilon$.

A uniformly continuous function is necessarily continuous, but on non-compact sets (i.e., sets that are not closed and bounded) a continuous function need not be uniformly so.
Not uniformly continuous

To help understand the import of uniform continuity, we’ll reverse the definition:

Definition (not uniformly continuous):

A function $f(x)$ is not uniformly continuous on $D$ if there is some $\varepsilon > 0$ such that for every $\delta > 0$, no matter how small, it is possible to find $x, y \in D$ with $|x - y| < \delta$ but $|f(x) - f(y)| > \varepsilon$.

For instance $f(x) = x^2$ is not u.c. on the set $[0, \infty)$ because we can choose $\varepsilon = 1$ and then for any $\delta > 0$, we have $(x + \delta)^2 - x^2 = 2x\delta + \delta^2$ and we can choose $x > 1/(2\delta)$ so that $2x\delta > 1$. So there is no $\delta$ that works for every $x$ in the infinite interval.

Likewise, $g(x) = 1/x$ is not u.c. on $(0, 1]$ because, again using $\varepsilon = 1$, for any $\delta > 0$ we’ll pick an integer $n > 0$ so that $1/n < \delta$.
Then let $x = 1/n$ and $y = 1/(n+1)$ and $|x - y| < \delta$ but $|1/x - 1/y| = 1$. 
Integrating continuous functions

Our goal is:

**Theorem**

If $f(x)$ is a continuous function on the closed, bounded interval $[a, b]$, then $f$ is integrable on $[a, b]$.

We’ll accomplish this in two jumps:

**Lemma 1**

If $f(x)$ is a **uniformly** continuous function on the closed, bounded interval $[a, b]$, then $f$ is integrable on $[a, b]$.

**Lemma 2**

If $f(x)$ is continuous on the closed, bounded interval $[a, b]$ then $f$ is uniformly continuous on $[a, b]$.

It’s easy to see how the theorem follows from the lemmas.
Lemma 1

If \( f(x) \) is a \textbf{uniformly} continuous function on the closed, bounded interval \([a, b]\), then \( f \) is integrable on \([a, b]\).

Proof: Given \( \varepsilon > 0 \), let \( \delta > 0 \) be such that
\[ |f(x) - f(y)| < \varepsilon/(b - a) \]
whenever \( x, y \in [a, b] \) and \( |x - y| < \delta \).

Now choose \( n \) so that \((b - a)/n < \delta\), and let \( P \) be the regular partition of \([a, b]\) into \( n \) subintervals. We’ll have \( x_i - x_{i-1} < \delta \) for all \( i \), so \( f_i^+ - f_i^- < \varepsilon/(b - a) \) for all \( i \). Then

\[
U(f, P) - L(f, P) = \sum_{i=1}^{n} (f_i^+ - f_i^-) \frac{b - a}{n} < \sum_{i=1}^{n} \frac{\varepsilon}{b - a} \frac{b - a}{n} = \varepsilon.
\]

so \( U(f, P) < L(f, P) + \varepsilon \), proving the integrability of \( f \).
Lemma 2

If \( f(x) \) is continuous on the closed, bounded interval \([a, b]\) then \( f \) is uniformly continuous on \([a, b]\).

We (almost!) did the proof of this in class using Bolzano-Weierstrass, so here’s the whole thing.

Proof by contrapositive: Assume that \( f \) is not uniformly continuous on \([a, b]\), and we’ll find a point of discontinuity of \( f \).
To do it, recall that since \( f \) is not u.c., there is an \( \varepsilon > 0 \) such that for any \( \delta > 0 \), no matter how small, there is a pair of points \( p, q \in [a, b] \) with \( |p - q| < \delta \) but \( |f(p) - f(q)| > \varepsilon \).
Construct two sequences \( \{p_n\} \) and \( \{q_n\} \) where all of the points of each sequence are in \([a, b]\), and such that \( |p_n - q_n| < 1/n \) (so we’re making \( \delta \) smaller and smaller) but \( |f(p) - f(q)| > \varepsilon \). Since \( \{p_n\} \) is contained in \([a, b]\) it is bounded and by Bolzano Weierstrass it has a convergent subsequence \( \{p_{n_i}\} \), with limit \( L \in [a, b] \).

Because \( |p_n - q_n| < 1/n \), the subsequence \( \{q_{n_i}\} \) must converge to \( L \) as well, since

\[
|q_{n_i} - L| < |q_{n_i} - p_{n_i}| + |p_{n_i} - L|
\]

and the first of these terms can be made arbitrarily small since it is less than \( 1/n_i \) and the second can be made small because \( p_{n_i} \to L \).
Now we’ll show that $f$ is not continuous at $L$. If it were, then for any $E > 0$ there would be a $D > 0$ such that if $x \in [a, b]$ with $|x - L| < D$, then $|f(x) - f(L)| < E$. Let $E = \varepsilon/2$, so we have the corresponding $D > 0$. We know that for $i$ sufficiently large we will have

$$|p_{n_i} - L| < D \text{ and } |q_{n_i} - L| < D$$

because both (sub)sequences converge to $L$. But then

$$|f(p_{n_i}) - f(q_{n_i})| \leq |f(p_{n_i}) - f(L)| + |f(q_{n_i}) - f(L)| < 2E < \varepsilon$$

contradicting the fact that $|f(p_n) - f(q_n)| > \varepsilon$ for all $n$. 
Integration by substitution

We still have the substitution rule:

Integration by substitution

Suppose $f : [a, b] \to \mathbb{R}$ and $g : [c, d] \to \mathbb{R}$ are integrable, and that the derivative of $g$ exists and is bounded and continuous on $(c, d)$. Further, assume that the image of $g : (c, d) \to \mathbb{R}$ is contained in the interval $(a, b)$. Then

$$\int_c^d f(g(x))g'(x) \, dx = \int_{g(c)}^{g(d)} f(x) \, dx.$$ 

Proof: Apply FTC to the function

$$H(x) = \int_c^x f(g(t)) \, dt - \int_{g(c)}^{g(x)} f(t) \, dt.$$
Integration by parts

Suppose \( f \) and \( g \) are differentiable functions on \([a, b]\), with \( f' \) and \( g' \) continuous as well. Then

\[
\int_a^b f(x)g'(x) \, dx = f(x)g(x) \bigg|_{x=a}^{x=b} - \int_a^b f'(x)g(x) \, dx.
\]

To prove this, consider the function

\[
H(x) = \int_a^x f(t)g'(t) \, dt - f(t)g(t) \bigg|_{t=a}^{t=x} + \int_a^x f'(t)g(t) \, dt.
\]

Using the fundamental theorem and the product rule for derivatives, we have

\[
H'(x) = f(x)g'(x) - \left( f(x)g'(x) + f'(x)g(x) \right) + f'(x)g(x) = 0
\]

so \( H(x) \) is constant. But \( H(a) = 0 \), so the constant is zero. The fact that \( H(b) = 0 \) is the formula in the box.
Trapezoidal rule

You probably recall the trapezoidal rule for estimating integrals from elementary calculus:

Trapezoidal rule

If $f$ is a twice-differentiable function on $[a, b]$, then

$$\int_a^b f(x) \, dx \approx \frac{h}{2} \left( f_0 + 2f_1 + 2f_2 + \cdots + 2f_{n-1} + f_n \right),$$

where $f_i = f(x_i)$, and $x_i = a + hi$ for $i = 0, \ldots, n$ and $h = (b - a)/n$

We’ll derive the trapezoidal rule and then find an estimate for the error (difference between the approximation on the right side and the actual integral on the left).
Derivation of the trapezoidal rule

To derive the trapezoidal rule, we’ll look at a single “panel” from $x_{i-1}$ to $x_i$. Via the change of variables $x = x_{i-1} + t$ we can write

$$
\int_{x_{i-1}}^{x_i} f(x) \, dx = \int_0^h f(x_{i-1} + t) \, dt.
$$

Now let

$$f(x_{i-1} + t) = c_0 + c_1 t + c_2 t^2 + \cdots$$

be the beginning of the Taylor expansion of $f$ around the point $t = 0$ (which is $x = x_{i-1}$). Then $f(x_{i-1}) = c_0$, $f(x_i) = c_0 + c_1 h + c_2 h^2 + \cdots$, and

$$
\int_0^h f(x_{i-1} + t) \, dt = c_0 h + c_1 \frac{h^2}{2} + \cdots.
$$

Therefore,

Both $\frac{h}{2}(f(x_{i-1}) + f(x_i))$ and $\int_0^h f(x_{i-1} + t) \, dt = c_0 h + c_1 \frac{h^2}{2} + \cdots$
So we have

\[ \int_{0}^{h} f(x_{i-1} + t) \, dt - \frac{h}{2} (f(x_{i-1}) + f(x_{i})) = \cdots \]

where \( \cdots \) hides terms that have powers of \( h \) beginning with \( h^3 \) and derivatives of \( f \) beginning with \( f'' \). So we would expect that the magnitude of this difference can be expressed in terms of these.

Also, when the trapezoidal rule is implemented with many \( n \) “panels” or steps, the right endpoint of each panel is the left endpoint of the next. That’s where all the 2’s come from in the rule. In the worst case, the errors in each panel will reinforce each other rather than cancel, so the error will be \( n \) times something that involves \( h^3 \) and \( f'' \).
Single-step error

We use integration by parts to estimate the error incurred in a single panel \((u = f, \ dv = dt\), say \(v = t + A\) for a tbd constant \(A\)):

\[
\int_0^h f(x_{i-1} + t) \, dt = (t + A)f(x_{i-1} + t) \bigg|_{t=0}^h - \int_0^h (t+A)f'(x_{i-1}+t) \, dt
\]

First concentrate on

\[
(t + A)f(x_{i-1} + t) \bigg|_{t=0}^h = (h + A)f(x_i) - Af(x_{i-1})
\]

We choose \(A = -h/2\) so that this is the trapezoidal formula. So we have

\[
\int_0^h f(x_{i-1} + t) \, dt - \frac{h}{2}(f(x_{i-1}) + f(x_i)) = - \int_0^h \left( t - \frac{h}{2} \right) f'(x_{i-1}+t) \, dt
\]
Single-step error

So far, our expression for the single-step error is

$$E_T = - \int_0^h \left( t - \frac{h}{2} \right) f'(x_{i-1} + t) \, dt,$$

so we integrate by parts again:

$$E_T = - \left. \left( \frac{t^2}{2} - \frac{h}{2} t + B \right) f'(x_{i-1} + t) \right|_0^h$$

$$+ \int_0^h \left( \frac{t^2}{2} - \frac{h}{2} t + B \right) f''(x_{i-1} + t) \, dt$$

We want to choose $B$ so that the term on the first line is zero. That term, expanded, is $B(f'(x_{i-1}) - f'(x_i))$ so we should choose $B = 0$. And so now:

$$E_T = \int_0^h \frac{1}{2} (t(t - h)) f''(x_{i-1} + t) \, dt.$$
So far, \( E_T = \int_0^h \frac{1}{2} (t(t - h)) f''(x_i - 1 + t) \, dt \). If we were to try and integrate by parts again, we would find that we cannot choose the integration constant for \( dv \) (i.e., the “C”) to make the first part zero again. So this expression is the best we can do.

Turn this into an estimate by letting \( M = \max_{x \in [x_{i-1}, x_i]} |f''(x)| \), so

\[
|E_T| \leq \int_0^h \left| \frac{1}{2} (t(t - h)) \right| f''(x_i - 1 + t) \, dt \leq \frac{M}{2} \int_0^h t(t-h) \, dt = \frac{Mh^3}{12}
\]

This is the standard formula for the single-step error in the trapezoidal rule. Multiplying this by \( n \) for the global estimate, and recalling that \( h = (b - a)/n \) gives the standard global error:

\[
E_T(f, [a, b]) \leq \frac{(b - a)^3 M}{12n^2} \quad \text{where} \quad M = \max_{x \in [a,b]} |f''(x)|.
\]