There is a class of functions on which one ‘knows’ what the integral should be. Namely, if \( f : [a, b] \to \mathbb{R} \) is a constant function, i.e. \( f(t) = \beta \) for some \( \beta \in \mathbb{R} \), then one should have
\[
\int_a^b f = \beta(b - a)
\]
since \( b - a \) is the length of the interval. More generally, one should have that the integral is additive for \( a < b < c \):
\[
\int_a^c f = \int_a^b f + \int_b^c f.
\]
This gives that one also knows what the integral must be for piecewise constant functions: if \( f : [a, b] \to \mathbb{R} \) and \( a = t_0 < t_1 < \ldots < t_n = b \), with \( f(t) = \beta_i \), \( t \in (t_{i-1}, t_i) \), with \( \beta_i \in \mathbb{R} \), then one should have
\[
\int_a^b f = \sum_{i=1}^n \beta_i(t_i - t_{i-1}).
\]
(Note that we do not care about the value of \( f \) at the points \( t_0, t_1 \), etc., the integral should not depend on the values of \( f \) on a finite number of points.)

Now, it should not be completely ‘obvious’ that there can be a notion of integral that satisfies this. The issue if that \( f \) is piecewise constant, there may be many ways of choosing the \( t_i \); for instance, if one such partition works, one could always refine it by adding more division points. One then has to show that the result is independent of the particular choice of division points (subject to \( f \) being affine on each subinterval). This is quite easy though (good exercise); if one has two such partitions, one can take their common refinement (consisting of all the division points in either), and then show that the integrals defined using the original two partitions are equal to that defined using their common refinement, and thus to each other.

Once one knows that the integral is well-defined on the set \( D = D([a, b]) \) of piecewise constant functions, one gets the basic properties of the integral very easily for functions \( f \in D \):

1. linearity: \( \int_a^b (c_1 f_1 + c_2 f_2) = c_1 \int_a^b f_1 + c_2 \int_a^b f_2 \), \( c_j \in \mathbb{R} \), \( f_j \in D \),
2. positivity: if \( f \in D \), \( f \geq 0 \) (i.e. \( f(t) \geq 0 \) for all \( t \in [a, b] \)) then \( \int_a^b f \geq 0 \),
3. boundedness: if \( f \in D \) then \( \int_a^b f \leq \int_a^b |f| \leq (b - a)\|f\| \), where \( \|f\| \) is the norm on continuous functions: \( \|f\| = \sup \{|f(t)| : t \in [a, b] \} \),
4. additivity: if \( a < b < c \) and \( f \in D([a, c]) \) then \( \int_a^b f + \int_b^c f = \int_a^c f \).

Note that all properties but the last refer to a single interval \([a, b]\), which is why we are using \( D \) simply there in the notation.

Notice that positivity plus linearity imply that if \( f, g \in D \), \( f \geq g \) (i.e. \( f(t) \geq g(t) \) for all \( t \in [a, b] \)) then \( \int_a^b f \geq \int_a^b g \). Indeed, \( \int_a^b f - \int_a^b g = \int_a^b (f - g) \geq 0 \) where the first equality is linearity of \( \int_a^b \) and the second is positivity using \( f - g \geq 0 \). In particular, positivity plus linearity imply boundedness.

We now consider integral for more general functions. The idea is as follows. Given a bounded function \( f : [a, b] \to \mathbb{R} \), and a partition \( P \) given by \( a = t_0 < t_1 < \ldots < t_n = b \), we can construct two piecewise constant functions
\[
f_{L,P}(t) := \inf_{s \in (t_{i-1}, t_i]} f(s) \text{ if } t \in (t_i, t_{i+1}), i = 0, 1, \ldots, n - 1
\]
and
\[
f_{U,P}(t) := \sup_{s \in (t_{i-1}, t_i]} f(s) \text{ if } t \in (t_i, t_{i+1}), i = 0, 1, \ldots, n - 1.
\]
We then define the upper integral and the lower integral by
\[ U(f) = \inf_{\mathcal{P}} \int_{a}^{b} f_{U,\mathcal{P}}, \quad L(f) = \sup_{\mathcal{P}} \int_{a}^{b} f_{L,\mathcal{P}}. \]

The following is easy to see.

**Lemma 1** For any bounded function \( f : [a, b] \to \mathbb{R} \), \( U(f) \) and \( L(f) \) are well-defined and \( L(f) \leq U(f) \).

**Proof.** Well-definedness just follows from the fact that the supremum and infimum is well-defined for any bounded set of real numbers.

Now note that for any partition \( \mathcal{P} \), \( f_{L,\mathcal{P}}(t) \leq f_{U,\mathcal{P}}(t) \) for all \( t \) (except for the finitely many points \( t_0, t_1, \) etc.). Now given two different partitions \( \mathcal{P}' \) and \( \mathcal{P}'' \), by considering a common refinement \( \mathcal{P} \), we have
\[ f_{L,\mathcal{P}'}(t) \leq f_{L,\mathcal{P}}(t) \leq f_{U,\mathcal{P}'}(t) \leq f_{U,\mathcal{P}''}(t). \]

Hence \( L(f) \leq U(f) \). \( \square \)

There is in general no reason that \( U(f) = L(f) \), but if this happens, this is a reasonable definition for the integral.

**Definition 1** We say that a bounded function \( f : [a, b] \to \mathbb{R} \) is Riemann integrable if \( U(f) = L(f) \). In this case we write
\[ \int_{a}^{b} f = U(f) = L(f). \]

There are many Riemann integrable functions. In particular, all continuous functions are Riemann integrable.

**Proposition 1** If \( f : [a, b] \to \mathbb{R} \) is continuous, then \( f \) is Riemann integrable.

**Proof.** Since we know \( L(f) \leq U(f) \) by Lemma 1, it suffices to show the opposite inequality \( L(f) \geq U(f) \). It in turn suffices to show that for every \( \varepsilon > 0 \),
\[ L(f) \geq U(f) - \varepsilon. \]

It in turn suffices to find a partition \( \mathcal{P} \) such that
\[ \int_{a}^{b} f_{L,\mathcal{P}} \geq \int_{a}^{b} f_{U,\mathcal{P}} - \varepsilon. \]

From now on fix \( \varepsilon > 0 \).

Recall from the problem set that every continuous function on the closed interval \([a,b]\) is in fact uniformly continuous. Therefore, given \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that \( |f(x) - f(y)| < \frac{\varepsilon}{b-a} \) whenever \( x, y \in [a, b], |x - y| < \delta \). Fix a partition \( a = t_0 < t_1 < \ldots < t_n = b \) such that \( t_{i+1} - t_i < \delta \).

Then for any \( x, y \in [t_i, t_{i+1}] \), it follows that \( |f(x) - f(y)| < \frac{\varepsilon}{b-a} \).

Therefore, \( f_{U,\mathcal{P}}(t) - f_{L,\mathcal{P}}(t) \leq \frac{\varepsilon}{b-a} \) for every \( t \in (t_i, t_{i+1}), i \in \{0, \ldots n - 1\} \), and hence
\[ \int_{a}^{b} f_{U,\mathcal{P}} - \int_{a}^{b} f_{L,\mathcal{P}} \leq \frac{\varepsilon}{b-a} \cdot (b - a) = \varepsilon, \]
which is what we wanted to show. \( \square \)

We now prove that the properties of integration that hold for piecewise constant functions in fact hold for all Riemann integrable functions:
Theorem 1 The following properties of integration hold:

1. linearity: if $f_j$ are Riemann integrable on $[a, b]$ and $c_j \in \mathbb{R}$, then $c_1 f_1 + c_2 f_2$ is Riemann integrable on $[a, b]$ and
   \[ \int_a^b (c_1 f_1 + c_2 f_2) = c_1 \int_a^b f_1 + c_2 \int_a^b f_2, \]
2. positivity: if $f$ is Riemann integrable on $[a, b]$ and $f \geq 0$ (i.e. $f(t) \geq 0$ for all $t \in [a, b]$) then
   \[ \int_a^b f \geq 0, \]
3. boundedness: if $f$ is Riemann integrable on $[a, b]$ then $|\int_a^b f| \leq \int_a^b |f| \leq (b - a)\|f\|$, where $\|f\|$ is the norm on continuous functions: $\|f\| = \sup\{|f(t)| : t \in [a, b]\}$,
4. additivity: if $a < b < c$ and $f$ is Riemann integrable on $[a, c]$, then $f$ is Riemann integrable on $[a, b]$ and $[b, c]$ and
   \[ \int_a^b f + \int_b^c f = \int_a^c f. \]

Proof. Let us first consider linearity. Clearly, for any partition $P$,

\[ c_1 \inf_{s \in (t_i, t_{i+1})} f_1(s) + c_2 \inf_{s' \in (t_i, t_{i+1})} f_2(s') \leq \inf_{s \in (t_i, t_{i+1})} (c_1 f_1 + c_2 f_2)(s). \]

Hence

\[ c_1 (f_1)_{L, P}(t) + c_2 (f_2)_{L, P}(t) \leq (c_1 f_1 + c_2 f_2)_{L, P}(t) \]

(for all $t$ except for the end-points of the intervals). Now since $f_1$ and $f_2$ are Riemann integrable, for every $\varepsilon > 0$, there exists a partition $P$ such that

\[ c_1 \int_a^b f_1 + c_2 \int_a^b f_2 - \varepsilon \leq c_1 \int_a^b (f_1)_{L, P}(t) + c_2 \int_a^b (f_2)_{L, P}(t) \leq c_1 \int_a^b f_1 + c_2 \int_a^b f_2. \]

Combining the above inequalities, and using linearity for piecewise constant functions, we obtain that

\[ L(c_1 f_1 + c_2 f_2) \geq \int_a^b (c_1 f_1 + c_2 f_2)_{L, P} = c_1 \int_a^b (f_1)_{L, P}(t) + c_2 \int_a^b (f_2)_{L, P}(t) \geq c_1 \int_a^b f_1 + c_2 \int_a^b f_2 - \varepsilon. \]

Since $\varepsilon > 0$ is arbitrary, this implies

\[ L(c_1 f_1 + c_2 f_2) \geq c_1 \int_a^b f_1 + c_2 \int_a^b f_2. \]

A similar argument applies to the upper integrals, except with the opposite inequality, i.e. we have

\[ U(c_1 f_1 + c_2 f_2) \leq c_1 \int_a^b f_1 + c_2 \int_a^b f_2. \]

We also know that $L(c_1 f_1 + c_2 f_2) \leq U(c_1 f_1 + c_2 f_2)$ (from Lemma [1]). Hence,

\[ L(c_1 f_1 + c_2 f_2) = U(c_1 f_1 + c_2 f_2) = c_1 \int_a^b f_1 + c_2 \int_a^b f_2. \]

This implies that $c_1 f_1 + c_2 f_2$ is Riemann integrable, and

\[ f_a^b (c_1 f_1 + c_2 f_2) = c_1 f_a^b f_1 + c_2 f_a^b f_2. \]

We then consider positivity. This is obvious since by definition, if $f \geq 0$, then both the upper and lower integrals are $\geq 0$.

For boundedness, as we argued above, it follows from linearity and positivity.

Finally, we prove additivity. Let us denote the upper and lower integrals on the interval $[a, b]$ by $U_{[a,b]}(f)$ and $L_{[a,b]}(f)$, and similarly for other intervals.
Let \( \varepsilon > 0 \), there exists a partition \( \mathcal{P} \) of \([a, c]\) such that
\[
\int_a^c f_L,\mathcal{P} \leq L^{[a,c]}(f) \leq \int_a^c f_U,\mathcal{P} + \varepsilon.
\]

Notice that after refining a partition, \( \int_a^b f_L,\mathcal{P} \) increases. We can therefore assume without loss of generality that the partition \( \mathcal{P} \) takes the form \( a = t_0 < \cdots < t_k = b = s_0 < s_1 < \cdots < s_\ell = c \).

Now by the additivity of the integral for piecewise constant functions, we have
\[
\int_a^c f_L,\mathcal{P} = \int_a^b f_L,\mathcal{P} + \int_b^c f_L,\mathcal{P}.
\]

Therefore,
\[
L^{[a,b]}(f) + L^{[b,c]}(f) \geq \int_a^b f_L,\mathcal{P} + \int_b^c f_L,\mathcal{P} = \int_a^c f_L,\mathcal{P} \geq L^{[a,c]}(f) - \varepsilon.
\]

By the arbitrariness of \( \varepsilon > 0 \), it follows that
\[
L^{[a,b]}(f) + L^{[b,c]}(f) \geq L^{[a,c]}(f).
\]

Now we can repeat the above argument but for the upper integrals. The inequalities will be reversed and we have
\[
U^{[a,b]}(f) + U^{[b,c]}(f) \leq U^{[a,c]}(f).
\]

Since \( f \) is Riemann integrable on \([a, c]\), it follows that \( L^{[a,c]}(f) = U^{[a,c]}(f) \). Thus the chain of inequalities
\[
L^{[a,c]}(f) \leq L^{[a,b]}(f) + L^{[b,c]}(f) \leq U^{[a,b]}(f) + U^{[b,c]}(f) \leq U^{[a,c]}(f)
\]
are in fact all equalities, giving the desired result. \( \square \)

This is already a nice achievement, but let us now prove the fundamental theorem of calculus. For this purpose it is convenient to define \( \int_a^a f = 0 \), with which definition all properties, including the additivity, still hold.

**Theorem 2** (Fundamental theorem of calculus, part I) Suppose \( f : [a, b] \to \mathbb{R} \) is continuous, and let \( F(t) = \int_a^t f \), \( t \in [a, b] \). Then \( F \) is \( C^1 \), and \( F' = f \).

**Proof:** If we show \( F \) is differentiable with \( F' = f \), continuity follows from that of \( f \). So it suffices to show that \( F \) is differentiable with \( F' = f \).

In order to do so, let \( t \in [a, b] \) and consider \( h > 0 \); the case of \( t \in (a, b) \) and \( h < 0 \) is similar. Then (for \( h < b - t \))
\[
F(t + h) - F(h) = \int_a^{t+h} f - \int_a^t f = \int_t^{t+h} f
\]
by the additivity, and
\[
\int_t^{t+h} f = \int_t^{t+h} f(t) + \int_t^{t+h} (f - f(t)) = hf(t) + \int_t^{t+h} (f - f(t)),
\]
where the first equality is the linearity of the integral and the second is the integral of a constant function directly from the definition of \( D \). Thus,
\[
|F(t + h) - F(t) - hf(t)| = \left| \int_t^{t+h} (f - f(t)) \right| \leq h \sup\{ |f(s) - f(t)| : s \in [t, t+h] \},
\]
where the inequality is the boundedness statement for the integral on \([t, t + h]\). By the continuity of \(f\) at \(t\), given \(\varepsilon > 0\) there exists \(\delta > 0\) such that \(|s - t| < \delta\) implies \(|f(s) - f(t)| < \varepsilon\); thus for \(0 < h < \min(\delta, b - t)\) we have
\[
|F(t + h) - F(t) - hf(t)| \leq \varepsilon h.
\]
Together with the analogous argument for \(h < 0\), this is exactly the definition of differentiability at \(t\) with derivative \(f\), proving the theorem. \(\square\)

In order to do the second half of the fundamental theorem of calculus, we just need an observation:

**Proposition 2** If \(f \in C([a, b])\) is differentiable on \((a, b)\) and \(f'(t) = 0\) for all \(t \in (a, b)\) then \(f\) is constant, i.e. \(f(t) = f(a)\) for all \(t \in [a, b]\).

**Proof:** Let \(t_1 < t_2, t_1, t_2 \in [a, b]\). By the mean value theorem on \([t_1, t_2]\), there is \(c \in (t_1, t_2)\) such that \(f'(c) = \frac{f(t_2) - f(t_1)}{t_2 - t_1}\); by the assumption \(f'(c) = 0\), so \(f(t_1) = f(t_2)\). Thus, for all \(t \in [a, b]\), \(f(t) = f(a)\), so \(f\) is a constant function. \(\square\)

**Theorem 3** (Fundamental theorem of calculus, part II) If \(f \in C^1([a, b])\) then \(f(b) - f(a) = \int_a^b f'\).

This is the theorem that is actually used to evaluate integrals explicitly: given a function \(g \in C([a, b])\) find \(f \in C^1([a, b])\) such that \(f' = g\), and then \(f(b) - f(a) = \int_a^b g\). Note that part I of the fundamental theorem says that the indefinite integral of \(g\) is such a function, but that does not give an explicit evaluation, i.e. you need a different method of finding the antiderivative for the explicit calculation. For instance, if \(g(t) = t^n\), then you check that \(f(t) = \frac{1}{n+1}t^{n+1}\) satisfies \(f'(t) = g(t)\) using the product rule for differentiation, and then apply the second part of the fundamental theorem of calculus to find \(\int_a^b g\).

**Proof:** Let \(F(t) = \int_a^t f'\). By the first part of the fundamental theorem of calculus proven above, \(F\) is continuously differentiable with \(F' = f'\). Thus, the function \(g = f - F\) satisfies \(g \in C^1([a, b])\) and \(g' = 0\). By the previous proposition, \(g(b) = g(a)\), i.e. \(f(b) - F(b) = f(a) - F(a)\), i.e.
\[
f(b) - f(a) = F(b) - F(a) = \int_a^b f',
\]
completing the proof. \(\square\)

Note that integration by parts is a simple consequence of the fundamental theorem of calculus and the product rule for differentiation: \((fg)' = f'g + fg'\); indeed, for \(f, g \in C^1([a, b])\),
\[
f(b)g(b) - f(a)g(a) = (fg)(b) - (fg)(a) = \int_a^b (fg)' = \int_a^b (f'g + fg') = \int_a^b f'g + \int_a^b fg',
\]
and now a simple rearrangement gives the formula:
\[
\int_a^b f'g = fg|^b_a - \int_a^b fg', \quad fg|^b_a = f(b)g(b) - f(a)g(a).
\]
Finally, we remark that we can define integration for bounded functions \(f : [a, b] \to \mathbb{R}^n\) by considering the componentwise integration; see the homework problem.