Definition and notation. A partition of the interval \([a, b]\), where \(a, b \in \mathbb{R}\), \(a < b\) is a finite subset of \([a, b]\) that includes the points \(a, b\). If \(P\) is a partition of \([a, b]\) we’ll write \(P : x_0 < x_1 < \cdots < x_n = b\) to indicate that \(P\) is the set consisting of the points \(x_0, \ldots, x_n\) and that we have indexed the points of \(P\) beginning the indices at 0, with \(x_0 = a\), the point with the last index corresponding to \(b\), and the order of the points in the interval corresponds to the order of the indices. We denote by \(\mathcal{P}[a, b]\) the set of all partitions of \([a, b]\).

Definition and notation. Let \(P, Q\) be partitions of \([a, b]\). We say that \(Q\) is a refinement of \(P\), or that \(Q\) is finer than \(P\), iff \(P \subset Q\). We write \(P \leq Q\) or \(Q \geq P\) to indicate that \(Q\) is a refinement of \(P\).

We have:

Lemma 1 Any two partitions of \([a, b]\) have a common refinement: Let \(P, Q \in \mathcal{P}[a, b]\). There exists \(R \in \mathcal{P}[a, b]\) such that \(P \leq R, Q \leq R\).

Proof. \(R = P \cup Q\). \(\blacksquare\)

Notation. If \(f : [a, b] \to \mathbb{R}\) is bounded and \(P : x_0 = a < \cdots < x_n = b\) is a partition of \([a, b]\), we set

\[
M_i(f, P) = \text{l.u.b. } \{f(x) : x_{i-1} \leq x \leq x_i\}, \\
m_i(f, P) = \text{g.l.b. } \{f(x) : x_{i-1} \leq x \leq x_i\},
\]

In case \(P\) is understood (there is only one partition \(P\) in play) we write simply \(M_i(f), m_i(f)\) for \(M_i(f, P), m_i(f, P)\), respectively. If, in addition, \(f\) is also fixed, we just write \(M_i, m_i\), respectively. If \(f\) is fixed but \(P\) can vary, we write \(M_i(P), m_i(P)\), respectively. I think I covered all possible combinations.

Lemma 2 Let \(P, Q \in \mathcal{P}[a, b]\) with \(P \leq Q\). Then (for every bounded \(f : [a, b] \to \mathbb{R}\))

\[
U(f, P) \geq U(Q, P); \quad L(f, P) \leq L(f, Q).
\]

(When adding points to a partition, the lower sums increase, the upper sums decrease.)

Proof. It suffices to prove this result for the case in which \(Q\) has exactly one point more than \(P\), since one can obtain \(Q\) from \(P\) by adding one point at a time. The result is, incidentally, quite obvious if one draws a picture. Here is just an analytic proof.

Let \(P : x_0 = a < x_1 < \cdots < x_n = b\), \(Q : y_0 = a < y_1 < \cdots < y_{n+1} = b\) and all points of \(P\) are points of \(Q\). That means that there is \(j, 1 \leq j \leq n\) such that

\[
y_i = x_i \quad \text{for } 0 \leq i \leq j - 1, \\
x_{j-1} < y_j < x_j, \\
y_i = x_{i-1} \quad \text{for } j + 1 \leq i \leq n + 1.
\]

We’ll work with the upper sums, the proof for the lower sums is similar. Let

\[
M_i(P) = \text{l.u.b. } \{f(x) : x_{i-1} \leq x \leq x_i\}, \quad i = 1, \ldots, n, \\
M_i(Q) = \text{l.u.b. } \{f(x) : y_{i-1} \leq x \leq y_i\}, \quad i = 1, \ldots, n + 1.
\]
Then, by the relation between the points of \( P \) and \( Q \), we have

\[
M_i(Q) = M_i(P) \quad \text{for } i = 1, \ldots, j - 1,
\]

\[
M_i(Q) = M_{i-1}(P) \quad \text{for } i = j + 2, \ldots, n + 1,
\]

\[
M_j(Q) = \text{l.u.b.} \{ f(x) : y_{j-1} \leq x \leq y_j \} = \text{l.u.b.} \{ f(x) : x_{j-1} \leq x \leq x_j \} \leq M_j(P)
\]

(because \( \{ f(x) : y_{j-1} \leq x \leq y_j \} \subseteq \{ f(x) : x_{j-1} \leq x \leq x_j \} \)),

\[
M_{j+1}(Q) = \text{l.u.b.} \{ f(x) : y_j \leq x \leq y_{j+1} \} = \text{l.u.b.} \{ f(x) : x_j \leq x \leq x_{j+1} \} \leq M_j(P)
\]

(because \( \{ f(x) : y_j \leq x \leq y_{j+1} \} \subseteq \{ f(x) : x_j \leq x \leq x_{j+1} \} \)).

(If \( j = 1 \), then the set \( i = 1, \ldots, j - 1 \) is the empty set; a sum over the empty set is 0. Similarly, if \( j = n \), then the set \( i = j + 2, \ldots, n + 1 \) is empty.)

Thus

\[
U(f, Q) = \sum_{i=1}^{n+1} M_i(Q)(y_i - y_{i-1}) = \sum_{i=1}^{j-1} M_i(Q)(y_i - y_{i-1}) + M_j(Q)(y_j - y_{j-1}) + M_{j+1}(Q)(y_{j+1} - y_j) + \sum_{i=j+2}^{n+1} M_i(Q)(y_i - y_{i-1})
\]

\[
= \sum_{i=1}^{j-1} M_i(P)(x_i - x_{i-1}) + M_j(Q)(y_j - y_{j-1}) + M_{j+1}(Q)(x_{j+1} - y_j) + \sum_{i=j+2}^{n+1} M_i(P)(y_i - y_{i-1})
\]

\[
\leq \sum_{i=1}^{j-1} M_i(P)(x_i - x_{i-1}) + M_j(P)(x_j - x_{j-1}) + \sum_{i=j+1}^{n} M_i(P)(x_i - x_{i-1})
\]

\[
= \sum_{i=1}^{n} M_i(P)(x_i - x_{i-1}) = U(f, P).
\]

So here you have it; the detailed proof. I hope you agree it is essentially trivial. \( \blacksquare \)

As a consequence we get

**Lemma 3** Let \( f : [a, b] \to \mathbb{R} \) be bounded. Every lower sum is less than or equal every upper sum; that is, if \( P, Q \in \mathcal{P}[a, b] \), then \( L(f, P) \leq U(f, Q) \).

**Proof.** The result is obvious if \( P = Q \). In general, let \( R \) be a common refinement of \( P, Q \). Then, since \( L(f, R) \leq U(f, R) \), we have by Lemma 2

\[
L(f, P) \leq L(f, R) \leq U(f, R) \leq U(f, Q).
\]

\( \blacksquare \)

**Notation.** Let \( f : [a, b] \to \mathbb{R} \) be bounded. We associate with \( f \) two sets; the set of upper sums

\[
\mathcal{U} = \mathcal{U}(f) = \{ U(f, P) : P \in \mathcal{P}[a, b] \},
\]

and the set of lower sums

\[
\mathcal{L} = \mathcal{L}(f) = \{ L(f, P) : P \in \mathcal{P}[a, b] \}.
\]
Definition. Let \( f : [a, b] \rightarrow \mathbb{R} \) be bounded. We define the upper and the lower Riemann integrals of \( f \) over \([a, b]\) by
\[
\int_a^b f(x) \, dx = \text{g.l.b. } \mathcal{U},
\int_a^b f(x) \, dx = \text{l.u.b. } \mathcal{L},
\]
respectively.

Lemma 4  Let \( f : [a, b] \rightarrow \mathbb{R} \) be bounded. Then
\[
\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx.
\]
Proof. Let \( P \in \mathcal{P}[a, b] \). By Lemma 3, \( L(f, P) \) is a lower bound of the set \( \mathcal{U} \) of all upper sums, thus
\[
L(f, P) \leq \int_a^b f(x) \, dx.
\]
Since \( P \) is an arbitrary partition of \([a, b]\), this proves that \( \int_a^b f(x) \, dx \) is an upper bound of the set \( \mathcal{L} \) of all lower sums, thus
\[
\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx.
\]

Definition. Let \( f : [a, b] \rightarrow \mathbb{R} \) be bounded. We say \( f \) is Riemann integrable over \([a, b]\), and write \( f \in \mathcal{R}[a, b] \), iff
\[
\int_a^b f(x) \, dx = \int_a^b f(x) \, dx = \int_a^b f(x) \, dx.
\]
In this case we define the integral of \( f \) over \([a, b]\) to be this common value:
\[
\int_a^b f(x) \, dx = \int_a^b f(x) \, dx = \int_a^b f(x) \, dx.
\]

Theorem 5  Let \( f : [a, b] \rightarrow \mathbb{R} \) be bounded. Then \( f \) is integrable if and only if for every \( \epsilon > 0 \) there exists a partition \( P \) of \([a, b]\) such that \( U(f, P) - L(f, P) < \epsilon \).

Proof. Assume first that \( f \) is integrable. Let \( \epsilon > 0 \) be given. Because \( \int_a^b f(x) \, dx \) is the greatest lower bound of \( \mathcal{U} \), there is \( U \in \mathcal{U} \) such that \( U < \int_a^b f(x) \, dx + \frac{\epsilon}{2} \); because of integrability \( \int_a^b f(x) \, dx = \int_a^b f(x) \, dx \) and elements of \( \mathcal{U} \) are determined by partitions, thus what we have just said is the same as saying that there exists a partition \( P_1 \) of \([a, b]\) such that
\[
U(f, P_1) < \int_a^b f(x) \, dx + \frac{\epsilon}{2}.
\]
Similarly, because $\int_a^b f(x) \, dx$ is the least upper bound of $L$, there exists a partition $P_2$ of $[a, b]$ such that
\[
L(f, P_2) > \int_a^b f(x) \, dx - \frac{\epsilon}{2}.
\]
Let $P = P_1 \cup P_2$ be a common refinement of $P_1, P_2$. By Lemma 2, $U(f, P) \leq U(F, P_1), L(f, P) \geq L(f, P_2)$, thus
\[
U(f, P) - L(f, P) \leq U(f, P_1) - L(f, P_2) < \left( \int_a^b f(x) \, dx + \frac{\epsilon}{2} \right) - \left( \int_a^b f(x) \, dx - \frac{\epsilon}{2} \right) = \epsilon.
\]

**Proposition 6** Let $f : [a, b] \to \mathbb{R}$ be increasing (or decreasing). Then $f$ is integrable.

**Proof.** For variety’s sake, we will assume that $f$ is decreasing. The proof for the increasing case is identical (and was done in class). Let $n \in \mathbb{N}$ and let $P_n \in \mathcal{P}[a, b]$ be given by
\[
P_n : x_0 = a < x_1 < \cdots < x_n = b, \quad \text{where } x_i = a + \frac{b - a}{n} \text{ for } i = 0, 1, \ldots, n.
\]
Because $f$ is decreasing, $M_i = f(x_{i-1}), m_i = f(x_i)$ for $i = 1, \ldots, n$ hence
\[
\begin{align*}
U(f, P_n) &= \sum_{i=1}^n f(x_{i-1}) \frac{b - a}{n} = (f(x_0) + f(x_1) + \cdots + f(x_{n-2}) + f(x_{n-1})) \frac{b - a}{n}, \\
L(f, P_n) &= \sum_{i=1}^n f(x_i) \frac{b - a}{n} = (f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + f(x_n)) \frac{b - a}{n}.
\end{align*}
\]
We see that
\[
U(f, P) - L(f, P) = (f(x_0) - f(x_n)) \frac{b - a}{n} = \frac{(b - a)(f(a) - f(b))}{n}.
\]
Select $n \in \mathbb{N}$ such that $\frac{(b - a)(f(a) - f(b))}{n} < \epsilon$. By Theorem 5, we are done. $\blacksquare$

What follows is a slightly more difficult result. But a useful one.

**Theorem 7** Assume $f : [a, b] \to \mathbb{R}$ is integrable. Let $I$ be a closed and bounded interval of $\mathbb{R}$ containing the range of $f$ and assume that $\varphi : I \to \mathbb{R}$ is continuous. Then $\varphi \circ f : [a, b] \to \mathbb{R}$ is integrable.

**Proof.** Because $I$ is compact, $\varphi$ is bounded. Let
\[
M = \text{l.u.b.}_{x \in I} |\varphi(x)| = \max_{x \in I} |\varphi(x)|.
\]
We assume $M > 0$; the case in which $M = 0$ is utterly trivial. Let $\epsilon > 0$ be given. Because $I$ is compact, $\varphi$ is uniformly continuous on $I$, hence there is $\delta > 0$ such that $|\varphi(x) - \varphi(y)| < \frac{\epsilon}{2(b - a)}$ whenever $x, y \in I, |x - y| < \delta$.

Because $f$ is integrable and $(\epsilon \delta)/(4M) > 0$, there exists (by Theorem 5) a partition
\[
P : x_0 = a < \cdots < x_m = b
\]
of \([a, b]\) such that
\[
U(f, P) - L(f, P) < \frac{\epsilon \delta}{4M}.
\]
We now partition the integers from 1 to \(m\) into two disjoint sets; namely
\[
A = \{ i \in \mathbb{N}, 1 \leq i \leq n, M_i - m_i < \delta \}, \\
B = \{ i \in \mathbb{N}, 1 \leq i \leq n, M_i - m_i \geq \delta \}.
\]
As usual, \(m_i = \text{g.l.b.} \sum_{i=1}^{x_i} f(x), M_i = \text{l.u.b.} \sum_{i=1}^{x_i} f(x)\). One of the sets \(A, B\) could be empty; a sum over the empty set is always interpreted as 0. Notice that \(f([x_{i-1}, x_i]) \subset [m_i, M_i]\).

Let \(i \in A\). If \(x, y \in [x_{i-1}, x_i]\), then \(f(x), f(y) \in [m_i, M_i]\), hence
\[
f(x) - f(y) \leq M_i - m_i < \delta.
\]
Similarly \(f(y) - f(x) \leq M_i - m_i < \delta\) thus \(|f(x) - f(y)| < \delta\).

Thus \(|\varphi(f(x)) - \varphi(f(y))| < \frac{\epsilon}{2(b - a)}\) for all \(x, y \in [x_{i-1}, x_i]\). This implies
\[
M_i(\varphi \circ f) - m_i(\varphi \circ f) < \frac{\epsilon}{2(b - a)} (\text{we can take } x \text{ so } \varphi(f(x)) \text{ is arbitrarily close to } M_i(\varphi \circ f), \text{ and } y \text{ so } \varphi(f(y)) \text{ is arbitrarily close to } m_i(\varphi \circ f)).
\]

\[
\sum_{i \in A} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1}) < \frac{\epsilon}{2(b - a)} \sum_{i \in A} (x_i - x_{i-1}) \leq \frac{\epsilon}{2(b - a)} \frac{m}{2} (b - a) = \frac{\epsilon}{2}.
\]

Recalling the choice of \(P\) and of \(B\) we have
\[
\frac{\epsilon \delta}{4M} > U(f, P) - L(f, P) = \sum_{i=1}^{m} (M_i - m_i)(x_i - x_{i-1}) \geq \sum_{i \in B} (M_i - m_i)(x_i - x_{i-1})
\]
\[
\geq \delta \sum_{i \in B} (x_i - x_{i-1}).
\]

It follows that
\[
\sum_{i \in B} (x_i - x_{i-1}) < \frac{\epsilon}{4M}.
\]

Using now that
\[
M_i(\varphi \circ f) - m_i(\varphi \circ f) \leq |M_i(\varphi \circ f)| + |m_i(\varphi \circ f)| \leq 2M,
\]
we get
\[
\sum_{i \in B} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1}) < \frac{\epsilon}{4M}(2M) = \frac{\epsilon}{2}.
\]

(\textbf{Note:} In class I thought one could get away with \(2M\) where I have \(4M\) here; a moment’s reflection shows that \(4M\) is needed; one can have \(M_i(f) = M, m_i(f) = -M\) for some \(i\)).

We proved:
\[
\sum_{i \in A} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1}) < \frac{\epsilon}{2},
\]
\[
\sum_{i \in B} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1}) < \frac{\epsilon}{2}.
\]
Thus

\[
U(\varphi \circ f, P) - L(\varphi \circ f, P) = \sum_{i=1}^{m} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1})
\]

\[
= \sum_{i \in A} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1}) + \sum_{i \in B} (M_i(\varphi \circ f) - m_i(\varphi \circ f))(x_i - x_{i-1})
\]

\[
< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\]

By Theorem 5 we are done.

\[\square\]

**Corollary 8** Continuous functions are integrable: If \(g : [a, b] \rightarrow \mathbb{R}\) is continuous, then \(g\) is integrable.

**Proof.** We can think of \(g\) as being \(\varphi \circ f\), where \(f\) is the identity function. \(f(x) = x\) for all \(x \in [a, b]\) (integrable by Proposition 6 since it is increasing) and \(\varphi = g\), thus continuous. Integrability follows by Theorem 7.

**Theorem 9** Let \(\mathcal{R} = \mathcal{R}[a, b]\) denote the set of all integrable functions over \([a, b]\).

Then \(\mathcal{R}\) is an algebra over \(\mathbb{R}\). Moreover, the map

\[
f \mapsto \int_{a}^{b} f(x) \, dx : \mathcal{R} \rightarrow \mathbb{R}
\]

is linear.

**Note:** An algebra over a field \(F\) is a vector space \(A\) over \(F\) in which there is defined a product \((u, v) \mapsto uv : A \times A \rightarrow A\) such that

1. \(u(vw) = (uv)w\) for all \(u, v, w \in A\),
2. \(u(v + w) = uv + uw\), \((u + v)w = uw + vw\) for all \(u, v, w \in A\),
3. \((cu)v = u(cv) = c(uv)\) for all \(u, v \in A\), \(c \in F\).

If, in addition, \(uv = vu\) for all \(u, v \in A\), we say \(A\) is commutative. If there exists \(e \in A\) such that \(cu = uc = u\) for all \(u \in A\), we say \(A\) is an algebra with an identity. Real valued functions on a set form an algebra under the usual operations, so when one says that a certain set of real valued functions is an algebra, one means with their usual operations. Thus stating that \(\mathcal{R}\) is an algebra boils down to:

1. \(f, g \in \mathcal{R}\) implies \(f + g, fg \in \mathcal{R}\).
2. \(f \in \mathcal{R}\), \(c \in \mathbb{R}\) implies \(cf \in \mathcal{R}\).

That the integral is linear means, of course, that

\[
\int_{a}^{b} (f + g)(x) \, dx = \int_{a}^{b} f(x) \, dx + \int_{a}^{b} g(x) \, dx, \quad \int_{a}^{b} cf(x) \, dx = c \int_{a}^{b} f(x) \, dx
\]

for all \(f, g \in \mathcal{R}\), \(c \in \mathbb{R}\).

**Proof.** Assume \(f, g \in \mathcal{R}\). If \(P : x_0 = a < x_1 < \cdots < x_n = b\) is a partition of \([a, b]\), we see that

\[
M_i(f + g) \leq M_i(f) + M_i(g), \quad m_i(f + g) \geq m_i(f) + m_i(g).
\]
In fact, \( M_t(f) + M_t(g) \) is clearly an upper bound for all the values \( f + g \) takes in the interval \([x_{i-1}, x_i]\), thus larger than the least upper bound \( M_t(f + g) \) of these values. Similarly for the lower bounds. We thus get

\[
(3) \quad U(f + g, P) \leq U(f, P) + U(g, P), \quad L(f + g, P) \geq L(f, P) + L(g, P)
\]

for all \( P \in \mathcal{P}[a, b] \). Let \( \epsilon > 0 \) be given. Since \( f \in \mathcal{R} \), the integral of \( f \) is both the l.u.b. of the lower sums and the g.l.b. of the upper sums. Because it is the former, there exists \( P_1 \in \mathcal{P}[a, b] \) such that

\[
\int_a^b f(x) \, dx - \frac{\epsilon}{2} < L(f, P_1);
\]

because it is the latter, there is \( P_2 \in \mathcal{P}[a, b] \) such that

\[
\int_a^b f(x) \, dx + \frac{\epsilon}{2} > U(f, P_2).
\]

Similarly, because \( g \in \mathcal{R} \), there exists \( P_3, P_4 \in \mathcal{P}[a, b] \) such that

\[
\int_a^b g(x) \, dx - \frac{\epsilon}{4} < L(g, P_3), \quad \int_a^b g(x) \, dx + \frac{\epsilon}{4} > U(g, P_4).
\]

Letting \( P = P_1 \cup P_2 \cup P_3 \cup P_4 \) be a common refinement of all these partitions, we get by lemma 2:

\[
\begin{align*}
\int_a^b f(x) \, dx - \frac{\epsilon}{2} &< L(f, P), \quad \int_a^b f(x) \, dx + \frac{\epsilon}{2} > U(f, P), \\
\int_a^b g(x) \, dx - \frac{\epsilon}{4} &< L(g, P), \quad \int_a^b g(x) \, dx + \frac{\epsilon}{4} > U(g, P).
\end{align*}
\]

By the definition of \( \int_a^b (f + g)(x) \, dx \), \( \int_a^b f(x) \, dx \), \( \int_a^b g(x) \, dx \), (3), and the choice of \( P \),

\[
\int_a^b (f + g)(x) \, dx \leq U(f + g, P) \leq U(f, P) + U(g, P)
\]

\[
< \int_a^b f(x) \, dx + \int_a^b g(x) \, dx + \frac{\epsilon}{2} < L(f, P) + L(g, P) + \epsilon \leq L(f + g, P) + \epsilon
\]

\[
\leq \int_a^b (f + g)(x) \, dx + \epsilon \leq \int_a^b (f + g)(x) \, dx + \epsilon
\]

the last inequality being due to the fact that the upper integral always dominates the lower one. Removing all reference to the one object that depends on \( \epsilon \), namely \( P \), we proved

\[
\int_a^b (f + g)(x) \, dx < \int_a^b f(x) \, dx + \int_a^b g(x) \, dx + \frac{\epsilon}{2} < \int_a^b (f + g)(x) \, dx + \epsilon \leq \int_a^b (f + g)(x) \, dx + \epsilon,
\]

since \( \epsilon > 0 \) is arbitrary, this implies that

\[
\int_a^b (f + g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx + \frac{\epsilon}{2} = \int_a^b (f + g)(x) \, dx
\]

from which we see that \( f + g \in \mathcal{R} \) and

\[
\int_a^b (f + g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx.
\]
That $cf \in \mathcal{R}$ if $f \in \mathcal{R}$ and $c \in \mathbb{R}$ is quite immediate. That is, it is immediate that

$$U(cf, P) = \begin{cases} cU(f, P) & \text{if } c > 0, \\ 0 & \text{if } c = 0, \\ cL(f, P) & \text{if } c < 0; \end{cases}$$

and

$$L(cf, P) = \begin{cases} cL(f, P) & \text{if } c > 0, \\ 0 & \text{if } c = 0, \\ cU(f, P) & \text{if } c < 0; \end{cases}$$

for all partitions $P \in \mathcal{P}[a, b]$. From this it is quite immediate that

$$\int_a^b cf(x) \, dx = c \int_a^b f(x) \, dx,$$

if $c \geq 0$ proving that if $\int_a^b f(x) \, dx = \int_a^b f(x) \, dx$; i.e., if $f \in \mathcal{R}$, then $cf \in \mathcal{R}$ and

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

If $c < 0$, then

$$\int_a^b cf(x) \, dx = c \int_a^b f(x) \, dx = c \int_a^b f(x) \, dx$$

and the result follows as well.

All that remains to be proved is that $f, g \in \mathcal{R}$, then $fg \in \mathcal{R}$. By Theorem 7, $f \in \mathcal{R}$ implies $f^2 \in \mathcal{R}$. Thus, if $f, g \in \mathcal{R}$, then

$$fg = \frac{1}{4} \left((f + g)^2 - (f - g)^2\right) \in \mathcal{R}.$$

**Theorem 10** Let $f : [a, b] \to \mathbb{R}$ be bounded and assume $c \in (a, b)$. Then $f$ is integrable over $[a, b]$ if and only if $f|_{[a, c]}$ is integrable over $[a, c]$ and $f|_{[c, b]}$ is integrable over $[c, b]$.

**Proof.** Assume first that $f|_{[a, c]}$ is integrable over $[a, c]$ and $f|_{[c, b]}$ is integrable over $[c, b]$. However, this is somewhat awkward notation, so we’ll say simply: $f$ is integrable over $[a, c]$ and over $[c, b]$. Let $\epsilon > 0$ be given. Because the integral is the l.u.b. of the set of lower sums, there exist partitions $P_1 \in \mathcal{P}[a, c]$, $P_2 \in \mathcal{P}[c, b]$ such that

$$\int_a^c f(x) \, dx - \frac{\epsilon}{2} < L(f, P_1), \quad \text{i.e.,} \quad \int_a^c f(x) \, dx < L(f, P_1) + \frac{\epsilon}{2};$$

$$\int_c^b f(x) \, dx - \frac{\epsilon}{2} < L(f, P_2), \quad \text{i.e.,} \quad \int_c^b f(x) \, dx < L(f, P_2) + \frac{\epsilon}{2};$$

On the other hand, because the integral is the greatest lower bound of the upper sums, there exist partitions $P_3 \in \mathcal{P}[a, c]$, $P_4 \in \mathcal{P}[c, b]$ such that

$$\int_a^c f(x) \, dx + \frac{\epsilon}{2} > U(f, P_3), \quad \text{i.e.,} \quad \int_a^c f(x) \, dx > U(f, P_3) - \frac{\epsilon}{2};$$

$$\int_c^b f(x) \, dx + \frac{\epsilon}{2} > U(f, P_4), \quad \text{i.e.,} \quad \int_c^b f(x) \, dx < U(f, P_2) - \frac{\epsilon}{2};$$

Taking common refinements, $P = P_1 \cup P_3 \in \mathcal{P}[a, c]$; $Q = P_2 \cup P_4 \in \mathcal{P}[c, b]$, lower sums get larger, upper sums get smaller; thus

$$U(f, P) - \frac{\epsilon}{2} < \int_a^c f(x) \, dx < L(f, P) + \frac{\epsilon}{2},$$

$$U(f, Q) - \frac{\epsilon}{2} < \int_c^b f(x) \, dx < L(f, Q) + \frac{\epsilon}{2}.$$
Now let \( R = P \cup Q \), which is a partition of \([a, b]\). It is (or should be) clear that
\[
L(f, R) = L(f, P) + L(f, Q), \quad U(f, R) = U(f, P) + U(f, Q),
\]
thus adding (4) and (5) we get
\[
U(f, R) - \epsilon < \int_a^c f(x) \, dx + \int_c^b f(x) \, dx < L(f, R) + \epsilon.
\]
The upper integral is smaller than all upper sums, the lower integral bounds all lower sums from above, thus our last inequalities imply
\[
\int_a^b f(x) \, dx - \epsilon < \int_a^c f(x) \, dx + \int_c^b f(x) \, dx < \int_a^b f(x) \, dx + \epsilon.
\]
Since \( \epsilon > 0 \) was arbitrary, the last inequalities can only hold if
\[
(6) \quad \int_a^b f(x) \, dx \leq \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \leq \int_a^b f(x) \, dx.
\]
But \( \int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx \); this last inequality together with (6) implies
\[
\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx.
\]
In particular, \( \int_a^b f(x) \, dx = \int_a^b f(x) \, dx \), hence \( f \) is integrable over \([a, b]\) and
\[
\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.
\]
Conversely, assume that \( f \in \mathcal{R}[a, b] \). Let \( \epsilon > 0 \). By Theorem 5, there exists a partition \( P \) of \([a, b]\) such that \( U(f, P) - L(f, P) < \epsilon \). Since this inequality is preserved under refinements, we may assume \( c \in P \). In this case \( P = P_1 \cup P_2 \), where \( P_1 \in \mathcal{P}[a, c] \), \( P_2 \in \mathcal{P}[c, b] \). We now have
\[
U(f, P_1) - L(f, P_1) \leq U(f, P_1) - L(f, P_1) + U(f, P_2) - L(f, P_2) = U(f, P) - L(f, P) < \epsilon
\]
hence \((\epsilon > 0 \text{ being arbitrary}) f \) is integrable over \([a, c]\). Similarly
\[
U(f, P_2) - L(f, P_2) \leq U(f, P_1) - L(f, P_1) + U(f, P_2) - L(f, P_2) = U(f, P) - L(f, P) < \epsilon
\]
proving \( f \) integrable over \([c, b]\).

**Theorem 11** Let \( f : [a, b] \to \mathbb{R} \) be integrable and assume that \( f(x) \geq 0 \) for all \( x \in [a, b] \). Then
\[
\int_a^b f(x) \, dx \geq 0.
\]
More generally, if \( f, g : [a, b] \to \mathbb{R} \) are integrable and \( f(x) \leq g(x) \) for all \( x \in [a, b] \), then
\[
\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx
\]
Proof. Assume \( f \) is integrable and non-negative at all points of \([a, b]\). It is then obvious that \( L(f, P) \geq 0 \) for all partitions \( P \) of \([a, b]\). It is enough that it be \( \geq 0 \) for a single partition to be able to say that \( \int_a^b f(x) \, dx \geq 0 \), hence
\[
\int_a^b f(x) \, dx = \int_a^b f(x) \, dx \geq 0.
\]
If now \( f(x) \leq g(x) \) for all \( x \in [a, b] \), then \((f - g)(x) \geq 0 \) for all \( x \in [a, b] \) and
\[
\int_a^b g(x) \, dx - \int_a^b f(x) \, dx = \int_a^b (f - g)(x) \, dx \geq 0.
\]

Theorem 12 Let \( f : [a, b] \to \mathbb{R} \) be integrable. Then \(|f| \) (the function defined by \(|f|(x) = |f(x)|\)) is integrable and
\[
\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx.
\]
Proof. Since \( x \mapsto |x| \) is continuous, \(|f|\) is integrable by Theorem 7. Since
\[-|f(x)| \leq f(x) \leq |f(x)|\]
for all \( x \in [a, b] \), Theorem 11 implies
\[-\int_a^b |f(x)| \, dx \leq \int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx;
\]
i.e.,
\[
\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx.
\]

When a function is integrable, Theorem 5 states that given any \( \epsilon > 0 \) there is a partition for which the upper and lower sums differ by less than \( \epsilon \). An immediate consequence that we have used actually several times is: If \( f \) is integrable over \([a, b]\), there exists a partition \( P \) such that
\[
\int_a^b f(x) \, dx \leq L(f, P) + \epsilon, \quad \int_a^b f(x) \, dx > U(f, P) - \epsilon.
\]
In fact, if \( P \) is the partition such that \( U(f, P) - L(f, P) < \epsilon \), since \( \int_a^b f(x) \, dx \leq U(f, P) \), we get
\[
\int_a^b f(x) \, dx - L(f, P) < \epsilon; \quad \text{i.e.,} \quad \int_a^b f(x) \, dx < L(f, P) + \epsilon.
\]
Since \( \int_a^b f(x) \, dx \geq L(f, P) \) one gets \( \int_a^b f(x) \, dx > U(f, P) - \epsilon \). Because lower sums increase, upper sums decrease, if one refines one has:

Lemma 13 Assume \( f : [a, b] \to \mathbb{R} \) is integrable. Let \( \epsilon > 0 \). There exists \( P \in \mathcal{P}[a, b] \) such that for every \( Q \in \mathcal{P}[a, b] \) such that \( Q \geq P \) we have
\[
\int_a^b f(x) \, dx \leq L(f, Q) + \epsilon, \quad \int_a^b f(x) \, dx > U(f, Q) + \epsilon.
\]
We may frequently use this lemma without giving a specific reference.

**Proposition 14** Assume \( f : [a, b] \to \mathbb{R} \) is integrable and let \( g : [a, b] \to \mathbb{R} \) be bounded. Assume that \( g \) differs from \( f \) (at most) at a finite number of points of \([a, b]\). Then \( g \) is integrable and

\[
\int_a^b g(x) \, dx = \int_a^b f(x) \, dx.
\]

**Proof.** Let \( S = \{ x \in [a, b] : g(x) \neq f(x) \} \cup \{ a, b \} \). By adding \( a, b \) to the set (if they are not already there) we can list the elements of \( S \) as we do for partitions; in fact, \( S \) is a partition. We cab write: \( S : x_0 = a < x_1 < \cdots < x_m = b \). For each subinterval of this partition select an interior point; that is, let \( c_i \in (x_{i-1}, x_i) \) for \( i = 1, \ldots, m \). For example, one could take \( c_i = (x_i + x_{i-1})/2 \) for \( i = 1, \ldots, m \).

It suffices to prove:

\( g \) is integrable over \([x_{i-1}, c_i]\) and over \([c_i, x_i]\) for \( i = 1, \ldots, m \), and

\[
\int_{x_{i-1}}^{x_i} g(x) \, dx = \int_{x_{i-1}}^{x_i} f(x) \, dx,
\]

In fact, assuming this done, Theorem 10 kicks in proving \( g \) is integrable over \([x_{i-1}, x_i]\) and

\[
\int_{x_{i-1}}^{x_i} g(x) \, dx = \int_{x_{i-1}}^{x_i} f(x) \, dx
\]

for \( i = 1, \ldots, n \). But now we can keep on applying Theorem 10 (the proper mathematical term for this procedure is induction) to get from \( g \) integrable over \([x_0, x_1]\), \([x_1, x_2]\) and

\[
\int_{x_0}^{x_1} g(x) \, dx = \int_{x_0}^{x_1} f(x) \, dx,
\]

that \( g \) is integrable over \([x_0, x_2]\) and

\[
\int_{x_0}^{x_2} g(x) \, dx = \int_{x_0}^{x_2} f(x) \, dx.
\]

Since \( g \) is also integrable over \([x_2, x_3]\), and its integral over that interval coincides with that of \( f \), we then get that \( g \) is integrable over \([x_0, x_3]\) with the same integral as \( f \). And so forth.

So it all reduces to proving that if \( f \) is integrable over an interval, which we may as well assume to be again \([a, b]\), and \( g \) differs from \( f \) only at one of the endpoints, then \( g \) is integrable and

\[
\int_a^b g(x) \, dx = \int_a^b f(x) \, dx.
\]

I’ll assume that \( g(x) = f(x) \) if \( a < x \leq b \); the case in which \( g(x) = f(x) \) for \( a \leq x < b \) is quite similar. Since \( f, g \) are bounded, we let \( M > 0 \) be such that \(|f(x)| \leq M, |g(x)| \leq M\) for all \( x \in [a, b] \). Let \( \epsilon > 0 \) be given. Let \( P \in [a, b] \), be such that

\[
\int_a^b f(x) \, dx \leq L(f, P) + \epsilon, \quad \int_a^b f(x) \, dx > U(f, P) - \epsilon.
\]

Such a \( P \) exists by Lemma 13; since we may refine, we may assume that \( P : x_0 = a < x_1 < \cdots < x_n = b \) where we select \( x_1 \) so that \( 0 < x_1 - x_0 \leq \epsilon/(2M) \). Since \( g = f \) on \([x_1, b]\), we will have

\[
|U(g, P) - U(f, P)| = |(M_1(g) - M_1(f))(x_1 - x_0)| \leq 2M \frac{\epsilon}{2M} = \epsilon.
\]
From this we get
\[ \int_a^b g(x) \, dx \leq U(g, P) \leq U(f, P) + \epsilon < \int_a^b f(x) \, dx + 2\epsilon; \]

thus
\[ \int_a^b g(x) \, dx > \int_a^b f(x) \, dx + 2\epsilon; \]

this being true for all \( \epsilon > 0 \), we conclude \( \int_a^b g(x) \, dx \leq \int_a^b f(x) \, dx \). Similarly,
\[ |L(g, P) - L(f, P)| = |(m_1(g) - m_1(f))(x_1 - x_0)| \leq 2M \frac{\epsilon}{2M} = \epsilon, \]
hence
\[ \int_a^b g(x) \, dx \geq L(g, P) \geq L(f, P) - \epsilon > \int_a^b f(x) \, dx - 2\epsilon; \]

thus
\[ \int_a^b g(x) \, dx > \int_a^b f(x) \, dx - 2\epsilon; \]

this holding for all \( \epsilon > 0 \), we conclude \( \int_a^b g(x) \, dx \geq \int_a^b f(x) \, dx \). We proved
\[ \int_a^b g(x) \, dx \leq \int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx \leq \int_a^b g(x) \, dx, \]
the last inequality holding always. Thus
\[ \int_a^b g(x) \, dx = \int_a^b f(x) \, dx = \int_a^b g(x) \, dx \]
proving the proposition.

**Examples.** Here are some integrals computed using the definition. Doing a lot of these exercises can make one appreciate more the fundamental theorem of calculus.

But first an exercise.

**Exercise.** For \( k = 0, 1, 2, \ldots, n = 1, 2, 3, \ldots \) let
\[ S_n(k) = 1^k + 2^k + \cdots + n^k = \sum_{j=1}^{n} j^k. \]

Prove: There exists a constant \( C_k \) depending on \( k \) but not on \( n \) such that
\[ \left| S_n(k) - \frac{(n+1)^{k+1}}{k+1} \right| \leq C_k n^k. \] 

For example, to make sure the thing to prove is understood (and also to start an inductive proof), let’s see what happens for \( k = 0 \) and for \( k = 1 \). If \( k = 0 \), then
\[ S_n(0) = 1 + \cdots + 1 = n; \quad \frac{(n+1)^{0+1}}{0+1} = n + 1, \]
so \(|S_n(0) - \frac{(n+1)^{1+1}}{0+1}| = 1\); we can take \(C_0 = 1\).

If \(k = 1\) then
\[
S_n(1) = \frac{n(n+1)}{2}; \quad \frac{(n+1)^{1+1}}{1+1} = \frac{(n+1)^2}{2},
\]
so
\[
|S_n(1) - \frac{(n+1)^{1+1}}{1+1}| = \frac{n+1}{2} \leq n,
\]
we can take \(C_1 = 1\). (\(C_k\) is not always going to be 1!)

Here are some hints on how to proceed that almost constitute a proof. The idea is to use induction on \(k\). A first thing to notice (many times one realizes what the first thing to notice is once one is far along in a proof and suddenly one hits a snag) is that assuming one has proved (8) for some \(k\), then
\[
S_n(k) \leq D_k n^{k+1}
\]
where \(D_k\) is a constant depending only on \(k\); for example
\[
D_k = \frac{2^{k+1}}{k+1} + C_k.
\]
In fact, by (8), since \(n+1 \leq 2n\), \(n^k \leq n^{k+1}\),
\[
S_n(k) \leq \left| S_n(k) - \frac{(n+1)^{k+1}}{k+1} \right| + \frac{(n+1)^{k+1}}{k+1} \leq \frac{(n+1)^{k+1}}{k+1} + C_k n^k
\]
\[
\leq \left( \frac{2^{k+1}}{k+1} + C_k \right) n^{k+1}.
\]

Now we are ready to prove (8) by induction. In fact, the approach I suggest does much more, it provides a recurrence formula for \(S_n(k)\) that allows you to get \(S_n(k)\) once you know \(S_n(i)\) for \(0 \leq i \leq k-1\).

We verified (8) for \(k = 0, 1\), in other words, for all cases less than or equal to \(k-1\) when \(k = 2\). Assume now the estimate proved up to and including \(k-1\) for some \(k \geq 2\); that is, assume proved for some constants \(C_0, \ldots, C_{k-1}\) independent of \(n\) that
\[
|S_n(i) - \frac{(n+1)^{i+1}}{i+1}| \leq C_i n^k
\]
for all \(n \in \mathbb{N}\). Then it also holds that
\[
S_n(i) \leq D_i n^{i+1}
\]
for all \(n \in \mathbb{N}, i = 0, 1, \ldots, k-1\), where \(D_0, \ldots, D_{k-1}\) do not depend on \(n\).

The interesting trick now is to look at \(S_n(k+1)\) and \(S_{n+1}(k+1)\) to get an expression for \(S_n(k)\). In other words we go higher than apparently needed. But it works. Following all the steps leading to the recurrence formula could be a good exercise in working with summations, assuming you have never done any such thing before. The main trick could be contained in the next line.
\[
S_{n+1}(k+1) = \sum_{j=1}^{n+1} j^{k+1} = \sum_{j=0}^{n} (j+1)^k = \sum_{j=0}^{n} \sum_{i=0}^{k+1} \binom{k+1}{i} j^i.
\]
The last equality is due to the binomial formula. But we have to be a bit careful. One of the terms, when both \( j \) and \( i \) are 0 is 0, and one can wonder how this is to be interpreted. It turns out that in the usual way, as 1. But just to be on the safe side, let us separate the case \( j = 0 \) from the rest, and redo the computation.

\[
S_{n+1}(k+1) = \sum_{j=1}^{n+1} j^{k+1} = \sum_{j=0}^{n} (j+1)^k = 1 + \sum_{j=1}^{n} (j+1)^k
\]

\[
= 1 + \sum_{j=1}^{n} \sum_{i=0}^{k+1} \binom{k+1}{i} j^i
\]

We can now exchange the order of summation and continue the equalities:

\[
S_{n+1}(k+1) = 1 + \sum_{j=1}^{n} \sum_{i=0}^{k+1} \binom{k+1}{i} j^i = 1 + \sum_{j=0}^{n} \left( \sum_{i=0}^{k+1} \binom{k+1}{i} j^i \right) S_n(i)
\]

We now get a nice recurrence formula for \( S_n(k) \), if we solve for \( S_n(k) \) noticing that

\[
S_{n+1}(k+1) - S_n(k) = (n+1)^{k+1};
\]

namely,

\[
(11) \quad S_n(k) = \frac{(n+1)^{k+1} - 1}{k+1} - \frac{1}{k+1} \sum_{i=0}^{k-1} \binom{k+1}{i} S_n(i).
\]

**Digression.** I think this is a really nice formula. From it we can get step by step formulas for all these sums. For example

\[
S_n(2) = \frac{(n+1)^3 - 1}{3} - \frac{1}{3} \left( \binom{3}{0} S_n(0) + \binom{3}{1} S_n(1) \right)
\]

\[
= \frac{(n+1)^3 - 1}{3} - \frac{n(n+1)}{2} = \frac{(2n+1)(n+1)n}{6};
\]

\[
S_n(3) = \frac{(n+1)^4 - 1}{4} - \frac{1}{4} \left( \binom{4}{0} S_n(0) + \binom{4}{1} S_n(1) + \binom{4}{2} S_n(2) \right)
\]

\[
= \frac{(n+1)^4 - 1}{4} - \frac{n(n+1)}{2} = \frac{3(2n+1)(n+1)n}{6} = \frac{n^2(n+1)^2}{4};
\]

and so forth.

**End of digression.**

Returning to our exercise, we proved

\[
\left| S_n(k) - \frac{(n+1)^{k+1}}{k+1} \right| \leq \frac{1}{k+1} + \frac{1}{k+1} \sum_{i=0}^{k-1} \binom{k+1}{i} S_n(i).
\]

Since \( S_n(i) \leq D_n i + 1 \leq D_n k \) if \( i \leq k - 1 \), (8) follows. What I wrote is not really almost a proof; it is a proof. We will need the following immediate consequence of (8):

\[
(12) \quad \lim_{n \to \infty} \frac{1}{n^k} S_n(k) = \frac{1}{k+1}, \quad \text{for } k = 0, 1, 2, \ldots.
\]
Using this result we can show that the function \( x \mapsto x^k \) \((k \in \mathbb{N} \cup \{0\})\) is integrable over every interval \([a, b]\) and compute its integral using only the definition. Of course, we know already it is integrable because it is continuous, but we don’t know its integral. However, to avoid too bad computations (and because as I’ll show it suffices to do this case), I will assume \( a = 0 \). So let \( b \in \mathbb{R} \). \( 0 < b \) and consider \( f : [0, b] \to \mathbb{R} \) defined by \( f(x) = x^k \), where \( k \) is a non-negative integer. For each \( n \in \mathbb{N} \) we let \( P_n \in \mathcal{P}[0, b] \) be a partition of \([0, b]\) into \( n \) equal parts; that is
\[
P_n : x_0 = 0 < \cdots < x_n = b, \quad x_i = \frac{b}{n} \text{ for } i = 0, 1, \ldots, n.
\]
Because \( f \) is increasing,
\[
m_i = f(x_{i-1}) = \frac{(i - 1)b^k}{n^k}, \quad M_i = f(x_i) = \frac{ib^k}{n^k}.
\]
It follows that
\[
L(f, P_n) = \sum_{i=1}^{n} (i - 1)b^k \frac{b}{n^k} = \frac{b^{k+1}}{nk+1} S_{n-1}(k),
\]
\[
U(f, P_n) = \sum_{i=1}^{n} ib^k \frac{b}{n^k} = \frac{b^{k+1}}{nk+1} S_n(k).
\]
This proves that for every \( n \in \mathbb{N} \),
\[
\frac{b^{k+1}}{nk+1} S_{n-1}(k) \leq \int_0^b f(x) \, dx \leq \int_0^b f(x) \, dx \leq \frac{b^{k+1}}{nk+1} S_n(k).
\]
This comprises the lower and upper integral between the terms of two sequences both converging to the same limit; by (12)
\[
\lim_{n \to \infty} \frac{b^{k+1}}{nk+1} S_{n-1}(k) = \lim_{n \to \infty} \frac{b^{k+1}}{nk+1} S_n(k) = \frac{b^{k+1}}{k+1}.
\]
This proves the \( f \) is integrable and, in fact, that
\[
\int_0^b x^k \, dx = \frac{b^{k+1}}{k+1}, \quad k = 0, 1, \ldots
\]
A totally identical computation would prove that if \( b < 0 \) then
\[
\int_b^0 x^k \, dx = -\frac{b^{k+1}}{k+1}, \quad k = 0, 1, \ldots
\]
Assuming now \( a < b \) it is easy to get that
\[
\int_a^b x^k \, dx = \frac{1}{k+1} (b^{k+1} - a^{k+1}), \quad k = 0, 1, \ldots
\]
dividing into cases and using the equality in Theorem 10.

The next example was done in class the day we introduced the concept of Riemann integral. Let \( f \) be the Dirichlet function on the interval \([0, 1]\); that is, let \( f(x) = 0 \) if \( 0 \leq x \leq 1 \) and \( x \) is irrational; \( f(x) = 1 \) if \( 0 \leq x \leq 1 \) and \( x \) is rational. Given any partition \( P : x_0 = 0 < x_1 < \cdots < x_n = 1 \), every interval \([x_{i-1}, x_i]\) contains both rational and irrational points, thus points where \( f \) is
0 and points where \( f \) is 1. Thus \( m_i = 0, M_i = 1 \) for \( i = 1, \ldots, n \), hence 
\[ U(f, P) = 1, \quad L(f, P) = 0 \]
and it since \( P \) was an arbitrary partition we see that all upper sums equal 1, all lower sums equal 0, thus 
\[ U = \{1\}, \quad L = \{0\} \] and 
\[ \int_0^1 f(x) \, dx = 0 < 1 = \text{uintopenf}01. \]
The Dirichlet function is not integrable.

**The fundamental theorem of calculus is almost here!** As a first step, we need to extend the equality of Theorem 10. And make some definitions.

**Definition.** Let \( b < a \) and assume that \( f \) is integrable over \([b, a]\). We define:
\[ \int_a^b f(x) \, dx = -\int_b^a f(x) \, dx. \]
In addition, for any point \( a \) in the domain of a function \( f \), we define
\[ \int_a^a f(x) \, dx = 0. \]
The consequence of these definitions is that if \( f \) is defined on an interval \( I \) and is integrable over every closed and bounded subinterval of \( I \) (this is equivalent to saying it is integrable over \( I \) if \( I \) happens to be closed and bounded), then 
\[ \int_a^b f(x) \, dx \] is defined for all \( a, b \in I \). We have

**Lemma 15** Let \( I \) be an interval in \( \mathbb{R} \), let \( f : I \to \mathbb{R} \) and assume \( f \) is integrable over every closed and bounded subinterval of \( I \). Then
\[ \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \]
for all \( a, b, c \in I \).

**Proof.** If \( a = b \) the formula reduces to
\[ 0 = \int_a^c f(x) \, dx + \int_c^a f(x) \, dx \]
which is the definition of \( \int_c^a f(x) \, dx \) if \( a < c \), or of \( \int_c^a f(x) \, dx \) if \( c < a \), or the equality \( 0 = 0 + 0 \) if \( c = a \). Similar considerations hold if \( a = c \) or if \( c = b \).
We may thus assume that \( a, b, c \) are distinct and then we have six possibilities:
\( a < b < c \), \( a < c < b \), \( b < a < c \), \( b < c < a \), \( c < a < b \), \( c < b < a \). The second case \( a < c < b \) is covered by Theorem 10. And now one can just wade through the others. If \( a < b < c \), then Theorem 10 implies
\[ \int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx = \int_a^b f(x) \, dx - \int_c^b f(x) \, dx \]
from which the result follows. Similarly for the other cases.

\[ \blacksquare \]

Given that now we have the possibility of the lower limit of integration being larger than the upper limit, we have to be careful with the signs. First of all, we
state a corollary of Theorem 11, a rather immediate one. To prove it we need to now that if \( f(x) = c \) for all \( x \in [a, b] \) (constant function) then
\[
\int_a^b f(x) \, dx = \int_a^b c \, dx = c(b - a)
\]
This was done in class assuming \( a < b \); the case \( b \leq a \) is an immediate consequence. The proof in case \( a < b \) is rather obvious since one sees at once that \( U(f, P) = L(f, P) = c(b - a) \) for all partitions \( P \) of \([a, b]\).

**Proposition 16** Assume \( a < b \) and \( f : [a, b] \to \mathbb{R} \) is integrable. Assume \( n, M \) are such that \( m \leq f(x) \leq M \) for all \( x \in [a, b] \). Then
\[
m(b - a) \leq \int_a^b f(x) \, dx \leq M(b - a).
\]

**Proof.** This is an immediate consequence of Theorem 11.

We also want a result that doesn’t require us to know whether \( a < b \). The following one will do.

**Proposition 17** Let \( f \) be integrable over the interval \( J \), where \( J = [a, b] \) if \( a \leq b \), \( J = [b, a] \) if \( b < a \). Assume \( |f(x)| \leq M \) for all \( x \in J \). Then
\[
\left| \int_a^b f(x) \, dx \right| \leq M|b - a|.
\]

**Proof.** If \( a < b \), this is an immediate consequence of Theorem 12 and Theorem 11:
\[
\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx \leq \int_a^b M \, dx = M(b - a) = M|b - a|.
\]

Assume now \( b < a \). Then, also by Theorem 12 and Theorem 11,
\[
\left| \int_a^b f(x) \, dx \right| = \left| \int_b^a f(x) \, dx \right| \leq \int_b^a |f(x)| \, dx \leq \int_b^a M \, dx = M(a - b) = M|b - a|.
\]

**Theorem 18** **Fundamental Theorem of Calculus, Version 1.** Let \( I \) be an interval in \( \mathbb{R} \) and let \( f : I \to \mathbb{R} \) be integrable over all closed and bounded subintervals of \( I \). Let \( a \in I \) and define \( F : I \to \mathbb{R} \) by
\[
F(x) = \int_a^x f(t) \, dt
\]
for \( x \in I \). Then

1. \( F \) is continuous on \( I \).
2. \( F \) is differentiable at all \( x \in I \) at which \( f \) is continuous and \( F'(x) = f(x) \) at such points.
Proposition 17 implies over \( [a, b] \) if \( f \) is Lipschitz in \([a, b]\). Once this is done, I'll explain why this proves continuity, just in case. So assume \([a, b]\) is a closed and bounded sub-interval of \( I \). Then \( f \) is bounded on \([a, b]\) (because integrability in our sense implies boundedness), say \(|f(x)| \leq M\) for all \( x \in [a, b] \). Then, if \( x, y \in [a, b] \),

\[
|F(y) - F(x)| = \left| \int_x^y f(t) \, dt \right| \leq M|y - x|
\]

by Proposition 17, which is Lipschitz. In particular, of course, the restriction of \( F \) to all closed and bounded subintervals of \( I \) is continuous. If \( x \in I \) then either \( x \) is an interior point of \( I \) or it is an endpoint of \( I \). If it is an interior point, then exist \( a, b, a < x < b \) and \([a, b] \subset I \). Continuity of \( F|_{[a, b]} \) at \( x \) implies continuity of \( F \) at \( x \). If \( x \) is an endpoint, say \( x \) is the left endpoint of \( I \). Then let \( b > x \) be such that \([x, b] \subset I \). Continuity of \( F|_{[x, b]} \) at \( x \) implies continuity of \( F \) at \( x \). Similarly if \( x \) is the right endpoint of \( I \).

1. **\( F \) is continuous.** The proof given here is probably a bit better than the one given in class. I will show: If \( \int_a^b f(t) \, dt = F(b) - F(a) \) and \( f \) is continuous at \( x \), then the limit in question will be one-sided.

**Proof.** Notice that if \( x, y \in I \), then \( F(y) - F(x) = \int_x^y f(t) \, dt \). In fact,

\[
F(y) - F(x) = \int_a^y f(t) \, dt - \int_a^x f(t) \, dt = \int_x^y f(t) \, dt + \int_x^y f(t) \, dt - \int_a^x f(t) \, dt = \int_x^y f(t) \, dt.
\]

1. **\( F \) is differentiable where \( f \) is continuous.** Let \( x \in I \) and assume \( f \) is continuous at \( x \). What we prove here is that if \( x \) is an interior point, then \( F \) is differentiable at \( x \); if \( x \) happens to be an endpoint then the one-sided derivative exists. In other words, we prove

\[
\lim_{y \to x} \frac{F(y) - F(x)}{y - x} = f(x).
\]

The domain of the function whose limit we take at \( x \) is \( I \setminus \{x\} \), a set of which \( x \) is a cluster point. If \( x \) happens to be an endpoint of \( I \), then the limit in question will be one-sided.

Let \( \epsilon > 0 \) be given. Since \( f \) is continuous at \( x \), there is \( \delta > 0 \) such that if \( y \in I \) and \(|y - x| < \delta\), then \(|f(y) - f(x)| < \epsilon\). Now we use:

\[
\frac{1}{y - x} \int_x^y f(x) \, dt = \frac{f(x)}{y - x} \int_x^y dt = f(x)
\]

(the variable of integration is \( t \! \! \! \) ) to get

\[
\left| \frac{F(y) - F(x)}{y - x} - f(x) \right| = \left| \frac{1}{y - x} \int_x^y f(t) \, dt - \frac{1}{y - x} \int_x^y f(x) \right| = \left| \frac{1}{y - x} \int_x^y (f(t) - f(x)) \, dt \right|.
\]

Assume now \(|y - x| < \delta\). The last integral is over the interval \([x, y]\) if \( x < y \) or over \([y, x]\) if \( y < x \). In either case every \( t \) in the interval of integration will be at least \( \delta \) from \( x \) thus \(|f(t) - f(x)| < \epsilon \) for all \( t \) in the interval of integration. Proposition 17 implies

\[
\left| \frac{F(y) - F(x)}{y - x} - f(x) \right| = \frac{1}{|y - x|} \int_x^y (f(t) - f(x)) \, dt \leq \frac{1}{|y - x|} \epsilon|y - x| = \epsilon.
\]
Since $\epsilon > 0$ was arbitrary, differentiability is proved.

**Example/Exercise.** Let $f : [0, 3] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ x, & 1 < x \leq 2 \\ 1, & 2 < x \leq 3. \end{cases}$$

Then, if we define $F(x) = \int_0^x f(t) \, dt$,

$$F(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ \frac{1}{2} + \frac{1}{2}x^2, & 1 < x \leq 2 \\ x - \frac{1}{2}, & 2 < x \leq 3 \end{cases}$$

Prove that the expression(s) given for $F$ are correct, using only results developed in these notes. Then prove that $F$ is continuous and differentiable everywhere except at $x = 2$. The only problems for continuity and differentiability would be at $x = 1$ and $x = 2$; see that neither is a problem for continuity, only $x = 2$, which happens to be the only point where $f$ is discontinuous, is a problem for differentiability.

An immediate corollary is

**Corollary 19** Let $I$ be an interval in $\mathbb{R}$ and let $f : I \to \mathbb{R}$ be continuous. Then $f$ has an antiderivative defined in $I$; that is, there exists a differentiable $F : I \to \mathbb{R}$ such that $F'(x) = f(x)$ for all $x \in I$.

**Proof.** By Corollary 8, $f$ is integrable over all closed and bounded subintervals of $I$. Select $a \in I$ and define

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

Then $F'(x) = f(x)$ for all $x \in I$ by Theorem 18.

We now come to the more common Calculus 1 (or 2, soon to be Calculus 3 or 4 if current trends continue) version of the theorem.

**Theorem 20** Fundamental Theorem of Calculus, Version 2. Let $f : [a, b] \to \mathbb{R}$ be integrable and assume there exists $F : [a, b] \to \mathbb{R}$ differentiable such that $F'(x) = f(x)$ for all $x \in [a, b]$. Then

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

Before we go into a proof, some remarks might be in order. First of all, if $f$ is continuous, then the theorem is a simple consequence of the first version, Theorem 18. In fact, in the first place such an $F$ always exists and we should rephrase the statement by saying “if $F' = f$” instead “assume there exists . . .”, etc. The proof then goes as follows: define $G : [a, b] \to \mathbb{R}$ by

$$G(x) = \int_a^x f(t) \, dt;$$
By Theorem 18, \( G'(x) = f(x) = F(x) \) for all \( x \in [a, b] \). By the consequence of the mean value theorem of differential calculus that states that two functions having the same derivative in an interval differ by a constant, there is \( C \in \mathbb{R} \) such that \( G(x) = F(x) + C \) at all \( x \in [a, b] \). Evaluating at \( a \), \( G(a) = \int_a^a f(t) \, dt = 0 \), thus \( 0 = F(a) + C \), so \( C = -F(a) \). Then

\[
\int_a^b f(x) \, dx = G(b) = F(b) + C = F(b) - F(a).
\]

The Theorem in Rosenlicht assumes \( f \) is continuous and proves it as I just did.

Our version is more general and not a consequence of version 1. But how much more general is it in practice? That is, how many functions \( f \) are there with the property that \( f \) is the derivative at EVERY point of an interval of a differentiable function, yet \( f \) is not continuous? The answer is that it isn’t so easy to find such functions. But anyway, here goes the proof of the stronger, more flavorful, improved version of the Fundamental Theorem of Calculus. As you can see, the proof is not much harder than the one assuming continuity, the one i did above, the one in Rosenlicht.

**Proof.** Let \( P : x_0 = a < x_1 < \cdots < x_n = b \) be a partition of \([a, b]\). By the mean value theorem of differential calculus, for \( i = 1, \ldots, n \) there exists \( c_i \in (x_{i-1}, x_i) \) such that

\[
F(x_i) - F(x_{i-1}) = f'(c_i)(x_i - x_{i-1}) = f(c_i)(x_i - x_{i-1}).
\]

Thus,

\[
m_i(f)(x_i - x_{i-1}) \leq f'(c_i)(x_i - x_{i-1}) = F(x_i) - F(x_{i-1}) \leq M_i(f)(x_i - x_{i-1})
\]

and adding up we get

\[
L(f, P) \leq \sum_{i=1}^n (F(x_i) - F(x_{i-1})) \leq U(f, P);
\]

that is, \( L(f, P) \leq F(b) - F(a) \leq U(f, P) \) since \((F(x_i) - F(x_{i-1})) = F(b) - F(a)\). But \( P \) was an arbitrary partition of \([a, b]\) and \( f \) was integrable, thus the only number that can be above all lower sums and below all upper sums is the integral of \( f \); i.e.,

\[
F(b) - F(a) = \int_a^b f(x) \, dx.
\]

\[ \blacksquare \]

And this pretty much does it. Except perhaps to show that this integral is the same as the one Rosenlicht brings in. I won’t do any of this in class! To prove the equivalence of our definition with Rosenlicht, we have to bring in the concept of a Riemann sum. I’ll be a bit more informal here than usual. Given a partition \( P : x_0 = a < x_1 < \cdots < x_n = b \) of \([a, b]\) and a function \( f \) defined on \([a, b]\), a Riemann sum corresponding to this partition is any number of the form

\[
S(f, P) = S(f, P, c) = \sum_{i=1}^n f(c_i)(x_i - x_{i-1})
\]

where \( c = (c_1, \ldots, c_n) \) is an \( n \)-tuple of points in \([a, b]\) such that \( x_{i-1} \leq c_i \leq x_i \) for \( i = 1, \ldots, n \). To abbreviate, I’ll say \( c \subseteq P \) to indicate that \( c \) is an \( n \)-tuple of points, the \( i \)-th entry in \( c \) being in the \( i \)-th interval of \( P \).
We denote the width of \( P \) by \(|P|\); specifically if \( P \) is as above, then
\[ |P| = \max_{1 \leq i \leq n} (x_i - x_{i-1}). \]
To distinguish Rosenlicht’s notion of integrability from ours, until equivalence is established, I’ll say that a bounded function \( f \) is R-integrable (R for Rosenlicht rather than Riemann) iff there exists a number \( A \) such that for every \( \epsilon > 0 \) there exists \( \delta > 0 \) with the property that if \( P \in \mathcal{P}[a,b] \), \(|P| < \delta\), then
\[ |S(f, P, c) - A| < \epsilon \]
for all \( c \in P \).

Assume first that \( f \) is R-integrable, so that this number \( A \) of the definition exists. Let \( \epsilon > 0 \) be given. Let \( \delta > 0 \) correspond to \( \epsilon \) as in the definition of R-integrable. Let \( P \in \mathcal{P}[a,b] \), \(|P| < \delta\). If \( P : x_0 = a < \cdots < x_n = b \), we select \( c = (c_1, \ldots, c_n) \) as follows. Let \( c_i \in [x_{i-1}, x_i] \) be such that
\[ M_i(f) - \frac{\epsilon}{n(b-a)} < f(c_i). \]
Then
\[ U(f, P) = \sum_{i=1}^{n} M_i(f)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} \left( M_i(f) + \frac{\epsilon}{n(b-a)} \right)(x_i - x_{i-1}) = S(f, P, c) + \epsilon < A + 2\epsilon. \]
This implies that \( \int_{a}^{b} f(x) \, dx < A + 2\epsilon \). Since this is true for all \( \epsilon > 0 \), this proves that \( \int_{a}^{b} f(x) \, dx \leq A \). Similarly one proves that \( \int_{a}^{b} f(x) \, dx \geq A \). Since \( \int_{a}^{b} f(x) \, dx \) both have to be equal and equal to \( A \); that is, \( f \) is integrable and \( A = \int_{a}^{b} f(x) \, dx \).

The converse is much harder. Assume \( f \) is integrable (our sense). Then \( f \) is bounded, let \( M \geq 0 \) be such that \(|f(x)| \leq M \) for all \( x \in [a,b] \). Let \( \epsilon > 0 \) be given. There is \( P : x_0 = a < x_1 < \ldots < x_n = b \in \mathcal{P}[a,b] \) such that \( U(f, P) - L(f, P) < \epsilon \). Let
\[ \rho = \min_{1 \leq i \leq n} x_i - x_{i-1}; \quad \delta = \min(\rho, \frac{\epsilon}{(M+1)n}) \]
where \( n \) is the number of subintervals of \( P \), and consider a partition \( Q : y_0 = a < y_1 < \ldots y_m = b \) such that \(|Q| < \delta \). We will divide the indices \( 1, 2, \ldots, m \) into pairwise disjoint groups:
\[ A_i = \{ \ell \in \{1, 2, \ldots, m\} \mid [y_{\ell-1}, y_\ell] \subset [x_{i-1}, x_i] \}, \quad i = 1, 2, \ldots, n \]
\[ B = \{ \ell \in \{1, 2, \ldots, m\} : \ell \notin \bigcup_{i=1}^{n} A_i \}. \]
Assume now \( c = (c_1, \ldots, c_m) \) is a choice of points such that \( c \subset Q \). If \( \ell \in A_i \) then the fact that the interval \([y_{\ell-1}, y_\ell]\) is contained in the interval \([x_{i-1}, x_i]\) implies that \( m_i(f, P) \leq f(c_\ell) \leq M_i(f, P) \). It should be clear that
\[ m_i(f, P)(x_i - x_{i-1}) \leq \sum_{\ell \in A_i} f(c_\ell)(y_\ell - y_{\ell-1}) \leq M_i(f, P)(x_i - x_{i-1}), \]
then setting \( A = \bigcup_{i=1}^{n} A_i \),
\[ (13) \quad L(f, P) \leq \sum_{\ell \in A} f(c_\ell)(y_\ell - y_{\ell-1}) \leq U(f, P). \]
How far away is $\sum_{\ell \in A} f(c_{\ell})(y_{\ell} - y_{\ell-1})$ from the full Riemann sum $S(f, P, c)$? The answer is that not very far. Suppose $\ell \in B$; that is, suppose that $[y_{\ell-1}, y_{\ell}]$ is not contained in any subinterval of $P$. Because $y_{\ell} - y_{\ell-1} \leq |Q| < \delta < \rho$, the only way this can happen is if for some $i \in \{1, \ldots, n-1\}$ we have $y_{\ell-1} < x_i < y_{\ell}$. There are thus no more than $n - 1$ indices in $B$, hence

$$\left| \sum_{\ell \in B} f(c_{\ell})(y_{\ell} - y_{\ell-1}) \right| \leq M \sum_{\ell \in B} (y_{\ell} - y_{\ell-1}) \leq M(n - 1)\delta < \epsilon.$$

Combining this with (13) and using the fact that $\sum_{m=1}^{m} = \sum_{\ell \in A} + \sum_{\ell \in B}$ we get

$$L(f, P) - \epsilon < S(f, Q, c) < U(f, P) + \epsilon.$$

Since the choice of $P$ is such that the integral of $f$ is within $\epsilon$ of $U(f, P)$ and of $L(f, P)$ we proved that $|Q| < \delta$ implies that

$$\left| S(f, Q, c) - \int_a^b f(x) \, dx \right| < 2\epsilon$$

for all $c \subset Q$. That is, $f$ is R-integrable.

And so it goes.