Abbreviated notes version for Test 3: just the definitions, theorem statements, proofs on the list. I may possibly have made a mistake, so check it. Also, this is not intended as a REPLACEMENT for your classnotes; the classnotes have lots of other things that you may need for your understanding, like worked examples.

6. The derivative

6.1. **Differentiation rules. Definition:** Let $f : (a,b) \to \mathbb{R}$ and $c \in (a,b)$. If the limit $\lim_{x \to c} \frac{f(x)-f(c)}{x-c}$ exists and is finite, then we say that f *is differentiable at* c, and we write this limit as $f'(c)$ or $\frac{df}{dx}(c)$. This is the derivative of f at c, and also obviously equals $\lim_{h \to 0} \frac{f(c+h)-f(c)}{h}$ by setting $x = c + h$ or $h = x - c$. If f is differentiable at every point in (a,b), then we say that f is differentiable on (a,b).

Theorem 6.1. *If $f : (a,b) \to \mathbb{R}$ is differentiable at at a point $c \in (a,b)$, then f is continuous at c.*

Proof. If f is differentiable at c then

$$f(x) = \frac{f(x) - f(c)}{x - c}(x - c) + f(c) \to f'(c)0 + f(c) = f(c),$$

as $x \to c$. So f is continuous at c. □

Theorem 6.2. (Calculus I differentiation laws) *If $f, g : (a,b) \to \mathbb{R}$ is differentiable at a point $c \in (a,b)$, then*

(1) *$f(x) + g(x)$ is differentiable at c and $(f + g)'(c) = f'(c) + g'(c)$.*
(2) *$f(x) - g(x)$ is differentiable at c and $(f - g)'(c) = f'(c) - g'(c)$.*
(3) *$Kf(x)$ is differentiable at c if K is a constant, and $(Kf)'(c) = Kf'(c)$.*
(4) (Product rule) *$f(x)g(x)$ is differentiable at c, and $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.*
(5) (Quotient rule) *$\frac{f(x)}{g(x)}$ is differentiable at c if $g(c) \neq 0$, and $(\frac{f}{g})'(c) = \frac{f'(c)g(c)-f(c)g'(c)}{g(c)^2}$.*

Proof. In the last step of many of the proofs below we will be silently using the definition of the derivative, and the 'limit laws' from Theorem 5.2.

(1) As $x \to c$ we have

$$\frac{f(x) + g(x) - (f(c) + g(c))}{x - c} = \frac{f(x) - f(c)}{x - c} + \frac{g(x) - g(c)}{x - c} \to f'(c) + g'(c).$$

(4) We have

$$\frac{f(x)g(x) - f(c)g(c)}{x - c} = \frac{f(x)g(x) - f(x)g(c)}{x - c} + \frac{f(x)g(c) - f(c)g(c)}{x - c} = f(x)\,\frac{g(x) - g(c)}{x - c} + g(c)\,\frac{f(x) - f(c)}{x - c}.$$

By Theorem 6.1, f is continuous at c, that is, $\lim_{x \to c} f(x) = f(c)$. So

$$\lim_{x \to c} \frac{f(x)g(x) - f(c)g(c)}{x - c} = \lim_{x \to c} f(x)\,\frac{g(x) - g(c)}{x - c} + g(c)\,\frac{f(x) - f(c)}{x - c} = f(c)g'(c) + g(c)f'(c).$$

(3) Set $g(x) = K$ in (4), to get $(Kf)'(c) = Kf'(c) + 0f(c) = Kf'(c)$.

(2) By (1) and (3), $(f + (-g))' = f' + (-g)' = f' - g'$.

(5) Since $g(c) \neq 0$, so that $|g(c)| > 0$, by Proposition 5.7 $|g(x)|$, and hence also $g(x)$, is nonzero on a neighborhood of c. So division by $g(x)$ in what follows is justified. We have

$$\frac{\frac{f(x)}{g(x)} - \frac{f(c)}{g(c)}}{x - c} = \frac{f(x)g(c) - g(x)f(c)}{g(x)g(c)(x - c)} = \frac{f(x)g(c) - f(c)g(c) + f(c)g(c) - f(c)g(x)}{g(x)g(c)(x - c)}.$$

This equals

$$\frac{g(c)\,\frac{f(x)-f(c)}{x-c} - f(c)\,\frac{g(x)-g(c)}{x-c}}{g(x)g(c)} \to \frac{g(c)f'(c) - f(c)g'(c)}{(g(c))^2}$$

as $x \to c$, since by Theorem 6.1, $\lim_{x\to c} g(x) = g(c)$. □

Theorem 6.3. (Calculus I chain rule) *If $f : (a, b) \to \mathbb{R}$ is differentiable at a point $c \in (a, b)$, and if $g : I \to \mathbb{R}$ is differentiable at $f(c)$, where I is an open interval containing $f((a, b))$, then the composition $g \circ f$ is differentiable at c and $(g \circ f)'(c) = g'(f(c))f'(c)$.*

As in Calculus, a point c is called a local minimum (resp. local maximum) point for a function f if $\exists$ a neighborhood U of c s.t. $f(c) \leq f(x)$ (resp. $f(c) \geq f(x)$) $\forall x \in V$. In this case we say that $f(c)$ is a local minimum (resp. maximum) value of f. As in Calculus I, the word 'extreme' means either 'minimum' or 'maximum'. So we have *local extreme points* and *local extreme values*, just as in Calculus I.

Recall from Calculus that a *critical point* is a point c s.t. $f'(c)$ does not exist or $f'(c) = 0$.

Theorem 6.4. (First Derivative Test) *If $f : (a, b) \to \mathbb{R}$ is differentiable at a local extremum point $c \in (a, b)$, then $f'(c) = 0$ (and so c is a critical point).*

Proof. Suppose that c is a local maximum point of f, so $f(x) \leq f(c)$ for all x in a neighborhood of c. Then $\lim_{x\to c+} \underbrace{\frac{f(x) - f(c)}{x - c}}_{\leq 0} \leq 0$ (by the bullet just before Proposition 5.4 with $g = 0$). Similarly, $\lim_{x\to c-} \underbrace{\frac{f(x) - f(c)}{x - c}}_{\geq 0} \geq 0$. Since f is differentiable $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$ exists, and so these two one-sided limits must be equal by Proposition 5.4, and hence must be 0. That is, $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} = 0$. The local minimum case is similar. □

6.2. **The mean value theorem.**

Lemma 6.5. (Rolle's theorem) *If $f : [a, b] \to \mathbb{R}$ is continuous, and f is differentiable on (a, b), and $f(a) = f(b) = 0$, then there exists $c \in (a, b)$ with $f'(c) = 0$.*

Proof. To prove this, note that it is clearly true if f is constant. By the MAX-MIN theorem f has a maximum and a minimum value on $[a, b]$. At least one of these extreme values must be nonzero if f is not constant, and hence this value is achieved at a point c which is not a or b. By the first derivative test (Theorem 6.4), $f'(c) = 0$. □

Theorem 6.6. (The mean value theorem) *If $f : [a, b] \to \mathbb{R}$ is continuous, and f is differentiable on (a, b), then there exists $c \in (a, b)$ with $f'(c) = (f(b) - f(a))/(b - a)$. Equivalently, $f(b) - f(a) = f'(c)(b - a)$.*

Proof. Let $g(x) = \frac{f(b)-f(a)}{b-a}(x - a) + f(a)$. Clearly $g' = \frac{f(b)-f(a)}{b-a}$. Then $h = f - g$ is continuous on $[a, b]$ and differentiable on (a, b). Also it is easy to check that $h(a) = h(b) = 0$, so by Rolle's theorem $\exists c \in (a, b)$ s.t. $h'(c) = f'(c) - g'(c) = 0$. Hence $f'(c) = g'(c) = \frac{f(b)-f(a)}{b-a}$. □

Corollary 6.7. *If $f : (a, b) \to \mathbb{R}$ has $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on (a, b).*

Proof. If $a < x < y < b$ then by the MVT there exists c such that $f(y) - f(x) = f'(c)(y - x) = 0$. Thus $f(x) = f(y)$. That is, f is constant. □

Corollary 6.8. *If $f : (a, b) \to \mathbb{R}$ has $f'(x) > 0$ for all $x \in (a, b)$ then $f(x)$ is strictly increasing on (a, b). Similarly, if $f'(x) \geq 0$ (resp. $f'(x) < 0, f'(x) \leq 0$) for all $x \in (a, b)$ then $f(x)$ is increasing (resp. strictly decreasing, decreasing) on (a, b).*

Proof. We just prove one, the others are similar. Suppose that $f'(x) > 0$ for all $x \in (a, b)$. If $a < x < y < b$ then by the MVT there exists c such that $f(y) - f(x) = f'(c)(y - x) > 0$, since $f'(c) > 0$. Thus $f(x) < f(y)$. So f is strictly increasing on (a, b). □

Corollary 6.9. *If $f'(x) = g'(x)$ for all $x \in (a, b)$ then there exists a constant C such that $f(x) = g(x) + C$ for all $x \in (a, b)$.*

6.3. **Taylors theorem and the second derivative test.** Here is a version of the Calculus II Taylor's theorem:

Theorem 6.10. *Suppose that $n \in \mathbb{N}$ (or $n = 0$). If $f : (a, b) \to \mathbb{R}$ is $n + 1$ times differentiable, and the first n of these derivatives are continuous on (a, b), and if $x_0 \in (a, b)$, then for every $x \in (a, b)$ with $x \neq x_0$ there is a number c between x_0 and x such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!} (x - x_0)^{n+1}.$$

Remark. Note that the MVT 6.6 is essentially the case $n = 0$ of the last theorem.

A number c is a *strict local minimum (resp. local maximum) point* for a function f if $\exists$ a deleted neighborhood U of c s.t. $f(c) < f(x)$ (resp. $f(c) > f(x)$) $\forall x \in U$.

Theorem 6.11. (The second derivative test (Calculus I)) *Suppose that f''' is continuous on a neighborhood of c. If $f'(c) = 0$ and $f''(c) > 0$ then f has a strict local minimum at $x = c$. If $f'(c) = 0$ and $f''(c) < 0$ then f has a strict local maximum at $x = c$.*

Proof. Suppose $f''(c) > 0$. Setting $n = 2$ and $x = c + h, x_0 = c$ in Taylors theorem above, we have

$$f(c+h) - f(c) = f'(c)\,h + \frac{f''(c)}{2}\,h^2 + \frac{f'''(d)}{6}\,h^3,$$

where d is a number between c and $c + h$. As $h \to 0$ we have $d \to c$, so that $f'''(d) \to f'''(c)$ and $f'''(d)h \to f'''(c) \cdot 0 = 0$. So since $f'(c) = 0$,

$$\frac{f(c+h) - f(c)}{h^2} = \frac{f''(c)}{2} + \frac{f'''(d)}{6}\,h \;\to\; \frac{f''(c)}{2} > 0,$$

as $h \to 0$. By Proposition 5.3, there exists $\delta > 0$, such that $\frac{f(c+h)-f(c)}{h^2} > 0$ for $h \in (-\delta, \delta)$. So $f(c + h) - f(c) > 0$, or $f(c + h) > f(c)$, for $h \in (-\delta, \delta)$. This says that f has a strict local minimum at $x = c$.

The case $f''(c) < 0$ is similar, except

$$\frac{f(c+h) - f(c)}{h^2} = \frac{f''(c)}{2} + \frac{f'''(d)}{6}\,h \;\to\; \frac{f''(c)}{2} < 0,$$

as $h \to 0$. This implies that there exists $\delta > 0$, such that $\frac{f(c+h)-f(c)}{h^2} < 0$ for $h \in (-\delta, \delta)$, and so as above $f(c + h) < f(c)$, for $h \in (-\delta, \delta)$. This says that f has a strict local maximum at $x = c$. $\square$

6.4. **The open mapping/Inverse function theorems.**

Theorem 6.12. *Suppose that $f : (a, b) \to \mathbb{R}$ is a one-to-one continuous function. Then f is either strictly increasing on (a, b), or it is strictly decreasing on (a, b).*

Recall that the notation $f(E)$ means $\{f(x) : x \in E\}$.

Proposition 6.13. *Suppose that $f : (a, b) \to \mathbb{R}$ is a one-to-one continuous function. Then $f((a, b))$ is an open interval.*

Theorem 6.14. *Suppose that $f : (a, b) \to \mathbb{R}$ is a one-to-one continuous function. Then f^{-1} is continuous.*

Theorem 6.15. (The inverse function theorem) *Suppose that $f : (a, b) \to \mathbb{R}$ is a differentiable function with $f'(x) \neq 0$ for all $x \in (a, b)$. Then f is one-to-one, its range $f((a, b))$ is an open interval (c, d), and for any $y \in (c, d)$, we have*

$$(f^{-1})'(y) \ = \ \frac{1}{f'(x)} \,, \qquad f(x) = y.$$

7. Integration–the Riemann integral

Throughout this chapter $f : [a, b] \to \mathbb{R}$ is a bounded function. That is, there are two constants m and M such that $m \leq f(x) \leq M$ for all $x \in [a, b]$. Equivalently, there is a constant K such that $|f(x)| \leq K$ for all $x \in [a, b]$, or equivalently, Range(f) is a bounded set.

Main definitions for this Chapter (from Calculus 1): [Pictures drawn in class.]

- A *partition* P of $[a, b]$ is an ordered set $\{x_0, x_1, \cdots, x_n\}$ where $a = x_0 < x_1 < x_2 < \cdots < x_n = b$.
- Let $\mathcal{P}$ be the set of all partitions P of $[a, b]$.
- We define $\Delta x_k = x_k - x_{k-1}$, for each $k = 1, 2, \cdots, n$.
- We define $M_k = \sup\{f(t) : x_{k-1} \leq t \leq x_k\}$ and $m_k = \inf\{f(t) : x_{k-1} \leq t \leq x_k\}$, for each $k = 1, 2, \cdots, n$.
- We define the *upper sum* $U(f, P) = \sum_{k=1}^{n} M_k \, \Delta x_k$. This is the sum of the areas of the red rectangles in picture.
- We define the *lower sum* $L(f, P) = \sum_{k=1}^{n} m_k \, \Delta x_k$. This is the sum of the areas of the green rectangles in picture.
- We define the *upper integral* $U(f) = \inf\{U(f, P) : P \in \mathcal{P}\}$.
- We define the *lower integral* $L(f) = \sup\{L(f, P) : P \in \mathcal{P}\}$.
- We say that a bounded function $f : [a, b] \to \mathbb{R}$ is *integrable* if $L(f) = U(f)$. In this case we write $\int_a^b f \, dx$ for the number $L(f) = U(f)$.

Some observations:

Observation 1: If $m \leq f(x) \leq M$ for all $x \in [a, b]$, then $m(b - a) \leq L(f, P) \leq U(f, P) \leq M(b - a)$, for any partition P of $[a, b]$.

Proof. Since $m_k \leq M_k$ clearly for each $k = 1, 2, \cdots, n$, we have

$$L(f, P) = \sum_{k=1}^{n} m_k \, \Delta x_k \leq \sum_{k=1}^{n} M_k \, \Delta x_k = U(f, P).$$

Since $m_k \geq m$ clearly for each $k = 1, 2, \cdots, n$, we have

$$L(f, P) = \sum_{k=1}^{n} m_k \, \Delta x_k \geq \sum_{k=1}^{n} m \, \Delta x_k = m(b - a).$$

Similarly, $U(f, P) = \sum_{k=1}^{n} M_k \, \Delta x_k \leq \sum_{k=1}^{n} M \, \Delta x_k = m(b - a)$. □

Definition. If P, Q are two partitions of $[a, b]$, then we say that P *refines* Q, or that P is *finer* than Q, if $Q \subseteq P$. That is, P consists of Q with some additional points added.

Observation 2: If P refines Q then $L(f, Q) \leq L(f, P) \leq U(f, P) \leq U(f, Q)$.

Observation 3: If P, Q are two partitions of $[a, b]$ then $L(f, P) \leq U(f, Q)$.

Observation 4: $L(f) \leq U(f)$.

Proof. If P, Q are two partitions of $[a, b]$ then by Observation 3, $L(f, P) \leq U(f, Q)$. So for fixed Q, $U(f, Q)$ is an upper bound for $\{L(f, P) : P \in \mathcal{P}\}$. Thus $L(f) \leq U(f, Q)$, by definition of $L(f)$. Hence $L(f)$ is a lower bound for $\{U(f, Q) : Q \in \mathcal{P}\}$. Thus $L(f) \leq U(f)$ by definition of $U(f)$. □

Observation 5: $L(f, P) \leq L(f) \leq U(f) \leq U(f, P)$ for all partitions P of $[a, b]$. In particular, if f is integrable, then $L(f, P) \leq \int_a^b f\,dx \leq U(f, P)$.

Theorem 7.1. (Riemann condition) *A bounded function $f : [a, b] \to \mathbb{R}$ is integrable if and only if $\forall\, \epsilon > 0,\ \exists$ a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \epsilon$.*

Proof. By definition, f is integrable if and only if $U(f) = L(f)$.

($\Leftarrow$) Suppose that $\forall\, \epsilon > 0,\ \exists$ a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \epsilon$. Using this, and the definition of $L(f)$ and $U(f)$,

$$U(f) \leq U(f, P) < L(f, P) + \epsilon \leq L(f) + \epsilon.$$

Since this is true for every $\epsilon > 0$, we have $U(f) \leq L(f)$, by Theorem 3.6. But $L(f) \leq U(f)$ by Observation 4, so $L(f) = U(f)$.

($\Rightarrow$) If $U(f) = L(f)$, and if $\epsilon > 0$ is given, choose (by definition of $L(f)$ and $U(f)$, and the 'principles in terms of ϵ' on page 18 of these notes), partitions Q and R such that $L(f, Q) > L(f) - \frac{\epsilon}{2}$, and $U(f, R) < U(f) + \frac{\epsilon}{2}$. Let $P = Q \cup R$, the refinement of both Q and R obtained by taking their union. Then by the last equations, and Observation 2, we have

$$U(f, P) \leq U(f, R) < U(f) + \frac{\epsilon}{2} = L(f) + \frac{\epsilon}{2} < L(f, Q) + \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq L(f, P) + \epsilon.$$

Looking at the left and right side of the last line, we see that $U(f, P) - L(f, P) < \epsilon$, which is what was required. □

Definition. A function $f : D \to \mathbb{R}$ is called *uniformly continuous* if $\forall\, \epsilon > 0$, $\exists\, \delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $x, y \in D$, and $|x - y| < \delta$.

Any uniformly continuous function is clearly continuous.

Theorem 7.2. *If D is compact, and $f : D \to \mathbb{R}$ is continuous, then f is uniformly continuous.*

Theorem 7.3. *If $f : [a, b] \to \mathbb{R}$ is continuous, then f is integrable.*

Proof. By the last theorem, f is uniformly continuous. Thus given $\epsilon > 0$, there is a number $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{b-a}$ whenever $x, y \in [a, b]$ and $|x - y| < \delta$. Choose a partition $P = \{x_0, x_1, \cdots, x_n\}$ of $[a, b]$ such that $\Delta x_k = x_k - x_{k-1} < \delta$ for every $k = 1, 2, \cdots, n$. Consider the interval $[x_{k-1}, x_k]$. By the Min-Max theorem 5.8, f has a maximum value M_k and a minimum value m_k on this interval; so there are numbers s and t in $[x_{k-1}, x_k]$ with $f(s) = M_k, f(t) = m_k$. Since $|s-t| \leq \Delta x_k < \delta$, we conclude that

$$M_k - m_k \; = \; |f(s) - f(t)| \; < \; \frac{\epsilon}{b-a} \; .$$

Now

$$U(f, P) - L(f, P) \; = \; \sum_{k=1}^{n} \Delta x_k \, M_k \; - \; \sum_{k=1}^{n} \Delta x_k \, m_k \; = \; \sum_{k=1}^{n} \Delta x_k \, (M_k - m_k),$$

and so

$$U(f, P) - L(f, P) \; < \; \sum_{k=1}^{n} \Delta x_k \, \frac{\epsilon}{b-a} \; = \; (b-a) \, \frac{\epsilon}{b-a} \; = \; \epsilon.$$

Thus f satisfies the 'Riemann condition' 7.1, and so f is integrable. ☐

Fact I1: If f is integrable on $[a, b]$, and if K is a constant, then $\int_a^b Kf \, dx \; = \; K \int_a^b f \, dx$.

Fact I2: If $f : [a, b] \to \mathbb{R}$ is integrable, and if $m \leq f(x) \leq M$ for all $x \in [a, b]$, then

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

From this it is easy to deduce that if $f : [a, b] \to \mathbb{R}$ is integrable, and if $f(x) \geq 0$ for all $x \in [a, b]$, then $\int_a^b f \, dx \geq 0$. Indeed, simply take $m = 0$ in Fact I2.

Fact I3: $\int_a^b K \, dx \; = \; K(b-a)$, if K is a constant.

In the following we will use a simple fact about supremums: if A and B are two sets of numbers, and if for every element $x \in A$ there exists some element y in B with $x \leq y$, then $\sup A \leq \sup B$.

Fact I4: If f and g are integrable on $[a, b]$, and if $f(x) \leq g(x)$ for all $x \in [a, b]$, then $\int_a^b f \, dx \leq \int_a^b g \, dx$.

Proof: If we take a partition P of $[a, b]$, and if m_k are the infimums used in the definition of $L(f, P)$, and if m'_k are the infimums used in the definition of $L(g, P)$, then by the fact mentioned above I4,

$$m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\} \leq \inf\{g(x) : x \in [x_{k-1}, x_k]\} = m'_k.$$

Thus

$$L(f,P) = \sum_{k=1}^{n} m_k \, \Delta x_k \leq \sum_{k=1}^{n} m'_k \, \Delta x_k = L(g,P) \leq L(g) = \int_a^b g \, dx.$$

Taking the supremum over all partitions P we deduce that $L(f) \leq \int_a^b g \, dx$, which is what we need since $L(f) = \int_a^b f \, dx$. □

Fact I5: $\int_a^b (f+g) \, dx = \int_a^b f \, dx + \int_a^b g \, dx$, if f and g are integrable on $[a,b]$.

Fact I6: If f is integrable on $[a,b]$, then so is $|f(x)|$, and $|\int_a^b f \, dx| \leq \int_a^b |f| \, dx$.

Fact I7: $\int_a^b f \, dx = \int_a^c f \, dx + \int_c^b f \, dx$ if $a \leq c \leq b$, and if f is integrable on $[a,c]$ and on $[c,b]$.

Fact I8: If f is monotone on $[a,b]$, then f is integrable on $[a,b]$.

Fact I9: (The first fundamental theorem of Calculus) If f is integrable on $[a,b]$, define $F(x) = \int_a^x f(t) \, dt$, for $x \in [a,b]$. If f is continuous at a point $c \in (a,b)$, then $F'(c) = f(c)$.

We mention *Riemann sums* briefly. Suppose that $P = \{x_0, x_1, \cdots, x_n\}$ is a partition of $[a,b]$, and that $t_k \in [x_{k-1}, x_k]$ for every $k = 1, 2, \cdots, n$. Then $\sum_{k=1}^n f(t_k) \, \Delta x_k$ is called a Riemann sum for f, and is sometimes written as $R(f,P)$. Note that because $m_k \leq f(t_k) \leq M_k$ for every $k = 1, 2, \cdots, n$ (by definition of m_k and M_k), we have

$$L(f,P) = \sum_{k=1}^{n} m_k \, \Delta x_k \leq \sum_{k=1}^{n} f(t_k) \, \Delta x_k \leq \sum_{k=1}^{n} M_k \, \Delta x_k \leq U(f,P).$$

That is, any Riemann sum $R(f,P)$ lies between $L(f,P)$ and $U(f,P)$.

Fact I10: (The second fundamental theorem of Calculus) If $f : [a,b] \to \mathbb{R}$ is continuous, and is differentiable on (a,b), and if f' is integrable on $[a,b]$ (set $f'(a) = f'(b) = 0$ if they are not already defined), then $\int_a^b f' \, dx = f(b) - f(a)$.

Proof: Suppose that $P = \{x_0, x_1, \cdots, x_n\}$ is a partition of $[a,b]$. By the MVT on $[x_{k-1}, x_k]$ there is a number $t_k \in (x_{k-1}, x_k)$ such that $f(x_k) - f(x_{k-1}) = f'(t_k)(x_k - x_{k-1})$. Thus

$$f(b) - f(a) = \sum_{k=1}^{n} (f(x_k) - f(x_{k-1})) = \sum_{k=1}^{n} f'(t_k)(x_k - x_{k-1}).$$

On the other hand, we have by the fact above the theorem we are proving,

$$L(f',P) \leq \sum_{k=1}^{n} f'(t_k)(x_k - x_{k-1}) \leq U(f',P),$$

and so

$$L(f',P) \leq f(b) - f(a) \leq U(f',P).$$

Taking the supremum over partitions P we get

$$\int_a^b f'\,dx = L(f') = \sup\{L(f',P) : \text{partitions } P\} \leq f(b) - f(a).$$

Similarly, taking the infimum over partitions P we get

$$f(b) - f(a) \leq U(f') = \int_a^b f'\,dx.$$

Thus $\int_a^b f'\,dx = f(b) - f(a)$. $\square$

Fact I11: (Integration by parts) If f' and g' are continuous on an open interval containing $[a,b]$ then $\int_a^b f\,g'\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'\,g\,dx$.

Fact I12: ('change of variable'/'substitution') If g is a differentiable function defined on an open interval containing numbers $c < d$, with g' integrable on $[c,d]$, and if f is a continuous function on an open interval I containing the range of g, then $\int_c^d f(g(x))\,g'(x)\,dx = \int_{g(c)}^{g(d)} f(x)\,dx$.