Here are some of the main results of Chapter 5.

We know (or have no excuse for not knowing!) that if $I$ is an interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ is increasing, then

1. For every $x \in I$: If $x$ is an interior point of $I$, then
   \[ f(x^+) = \lim_{y \to x^+} f(y), \quad f(x^-) = \lim_{y \to x^-} f(y) \]
   exist; $f(x) \leq f(x^-) \leq f(x^+)$. One also has (and this could be a step in proving existence)
   \[ f(x^+) = \inf\{f(y) : y \in I, y > x\}, \quad f(x^-) = \sup\{f(y) : y \in I, y < x\}. \]

2. If $a$ is the left endpoint of $I$ (possibly $a = -\infty$, possibly $a \in I$ or $a \notin I$), then $f(a^+)$ exists.

3. If $b$ is the right endpoint of $I$ (possibly $b = \infty$, possibly $b \in I$ or $b \notin I$), then $f(b^-)$ exists.

So every discontinuity of an increasing function is a jump discontinuity. Because of this, one also sees that the set of
 discontinuities of an increasing function is countable (possibly finite, possibly empty).

One should not think because of this that increasing functions always have easily drawn graphs. In fact, one can prove:

**Lemma 1** Let $D$ be any countable subset of the interval $[a,b]$. There exists $f : [a,b] \to \mathbb{R}$ increasing discontinuous precisely at every $t \in D$.

The proof is not terribly difficult. To avoid distractions, I am including it in an appendix at the end. Feel free to ignore it. But because discontinuities, albeit countable, can form a dense subset of the domain of an increasing function, the first important theorem of Royden’s Chapter 5 is nothing but trivial. It states that an increasing function is differentiable almost everywhere.

Increasing functions are measurable; that should be obvious for many reasons. The most obvious one being, perhaps, that the set \{ $x : f(x) > a$\} has to be an interval. Suppose now that $f : [a,b] \to \mathbb{R}$ is measurable and $f'(x)$ exists for almost all $x \in [a,b]$; that is, $f'$ is differentiable a.e. Consider the functions $g_n$ defined for $n \in \mathbb{N}$ as follows. First, extend $f$ in some way to the right of $b$; for example by setting $f(x) = 0$ if $x > b$, or $f(x) = f(b)$ for $x > b$. The extended function should be measurable. Now define

\[ g_n(x) = n (f(x + \frac{1}{n}) - f(x)) \]
for $x \in [a,b]$. The extension was needed so $f(x + \frac{1}{n})$ makes sense if $x$ is close to $b$. Clearly each $g_n$ is measurable. At every $x \in (a,b)$ where $f$ is differentiable, one has $f'(x) = \lim_{n \to \infty} g_n(x)$, proving $f'$ is measurable.

Summarizing the narrative so far, increasing functions are a.e. differentiable, and a.e. derivatives of measurable functions are measurable. For an increasing function one also has $f' \geq 0$ (the proof given in calculus 1 textbooks for this fact applies also here), thus $\int_a^b f'$ makes perfect good sense. A natural question is whether one can recover $f$ from its derivative; in Newton’s immortal terms: can we recover the fluent from the fluxion? Specifically: Is \( \int_a^b f' = f(b) - f(a) \). As we shall see, the general answer is NO.

Another set of function friends found in this Chapter are the functions of bounded variation. The definition is:

Let $[a,b]$ be a closed and bounded interval in $\mathbb{R}$. A function $f : [a,b] \to \mathbb{R}$ is said to be of **bounded variation** ($f \in BV[a,b]$ for short) iff

\[ \sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| : a = x_0 < x < \cdots < x_n = b \right\} < \infty. \]
Here is the same definition in a more leisurely way. We denote, as we did when discussing integration, by $P_{a,b}$ the family of all partitions of $[a,b]$. If $P : x_0 < x_1 < \cdots < x_n = b$ is in $P_{a,b}$, define

$$V_{a,b}(f, P) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|.$$  

We say that $f$ is of bounded variation on $[a,b]$ iff

$$V_{a,b}(f) = \sup_{P \in P_{a,b}} V_{a,b}(f, P) < \infty.$$  

The number $V_{a,b}(f)$ is then called the total variation of $f$ in the interval $[a,b]$. The following types of functions are of bounded variation:

1. Increasing (and decreasing) functions. If $f : [a,b] \to \mathbb{R}$ is increasing, then (as is easily seen) $V_{a,b}(f, P) = f(b) - f(a)$ for all $P \in P_{a,b}$, thus also $V_{a,b}(f) = f(b) - f(a)$.

2. Differentiable functions with a bounded derivative. More generally, Lipschitz functions: If $f : [a,b] \to \mathbb{R}$ satisfies that $|f(x) - f(y)| \leq M|x - y|$ for all $x, y \in [a,b]$, then

$$V_{a,b}(f, P) \leq M \sum_{i=1}^{n} (x_i - x_{i-1}) = M(b - a)$$  

for each partition $P = \{x_0, \ldots, x_n\}$.

3. If $f, g \in BV[a,b]$, then so are $f \pm g$, cf for all $c \in \mathbb{R}$.

A function of bounded variation does not need to be continuous. Increasing functions can have discontinuities, yet are of bounded variation.

A continuous function need not be of bounded variation. The classic example is the function defined by $f(x) = x \sin(1/x)$ for $x \neq 0$, $f(0) = 0$. It is continuous everywhere, yet not of BV in $[0,1]$. You should get a feeling that what $V_{a,b}(f)$ measures are the oscillations of $f$ in the interval $[a,b]$; the more oscillatory the function, the larger is $V_{a,b}(f)$.

A basic result about functions of bounded variation is

**Theorem 2** Let $f : [a,b] \to \mathbb{R}$. Then $f$ is of bounded variation if and only if $f = g - h$ where $g, h : [a,b] \to \mathbb{R}$ are increasing.

Important as this theorem is, it has a fairly simple proof. At least the idea is simple; one then needs to verify the details. Here is how the proof goes. It is sort of trivial that if $f \in BV[a,b]$, then $f \in BV[a,x]$ for all $x \in (a,b]$. one defines

$$g(x) = V_{a,x}(f)$$  

for $a \leq x \leq b$. It is quite easy to see that $g$ is increasing. It is also fairly easy to verify that $g - f$ is increasing and voilà, the theorem is proved.

As a consequence properties of increasing functions can be transmitted to functions of bounded variation. In particular, if $f : [a,b] \to \mathbb{R}$ is of bounded variation, then

1. The set of discontinuities of $f$ is countable, every discontinuity is a jump discontinuity.

2. $f$ is differentiable a.e.

We come back now to the question of whether we can recover a function that is differentiable a.e. from its derivative; i.e., is the fundamental theorem of calculus still valid?

Alas, no. Consider the Cantor ternary function $f$ defined in Problem 2.48 of Royden’s textbook. This function is increasing; $f(0) = 0$, $f(1) = 1$. It is constant on every open interval contained in the complement of the Cantor set; since every point not in the Cantor set is in such an interval, and being constant in an open interval implies zero derivative, one
sees that $f'(x) = 0$ for all $x \notin C$ ($C$ being the Cantor set). Since $m(C) = 0$, this implies that $f$ is differentiable almost everywhere (as expected) and $f'(x) = 0$ a.e. Thus

$$\int_0^1 f' = 0 < 1 = f(1) - f(0).$$

Notice that $f$ is continuous, so the question of what is needed so that a function can be recovered from its derivative does not have an incredibly easy answer.

But measure theory helps.

One has the following definition and theorem.

Let $f : [a,b] \to \mathbb{R}$. We say $f$ is absolutely continuous (a.c. for short) iff for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever $I_1, \ldots, I_m$ is a (finite) set of pairwise disjoint intervals in $a,b$, say $I_j = [x_j, y_j]$ for $j = 1, \ldots, m$, such that the sum of their lengths is $< \delta$, then

$$\sum_{j=1}^m |f(y_j) - f(x_j)| < \epsilon. \quad (1)$$

Here is the same definition with fewer words, more symbols. We say $f : [a,b] \to \mathbb{R}$ is absolutely continuous iff for each $\epsilon > 0$ there is $\delta > 0$ such that for every choice of points $x_1, y_1, x_2, y_2, \ldots, x_m, y_m$ in $[a,b]$ satisfying

$$a \leq x_1 < y_1 < x_2 < y_2 < \cdots < x_m < y_m \leq b, \quad \text{and} \quad \sum_{j=1}^m (y_j - x_j) < \delta$$

we have that $\sum_{j=1}^m |f(y_j) - f(x_j)| < \epsilon$.

It is, or should be trivial that absolutely continuous functions are continuous. Lipschitz functions are absolutely continuous. The proof should be immediate, done in the blink of an eye. Absolutely continuous functions are of bounded variation. Proving this may take a bit more than a blink of an eye, but is still quite easy. Not every continuous function of bounded variation is absolutely continuous; the next theorem shows that Cantor’s ternary function (increasing and continuous) is not absolutely continuous.

We have the hierarchies going from smallest to largest:

- Lipschitz $\Rightarrow$ absolutely continuous $\Rightarrow$ bounded variation
- Absolutely continuous $\Rightarrow$ continuous
- Increasing, decreasing $\Rightarrow$ bounded variation

And also

- Bounded variation $\not\Rightarrow$ continuous
- Continuous $\not\Rightarrow$ bounded variation

Here is then the theorem we have all been waiting for.

**Theorem 3** A function $f : [a,b] \to \mathbb{R}$ is absolutely continuous if and only if there exists $g \in L^1([a,b])$ such that

$$f(x) = f(a) + \int_a^x g(t) \, dt.$$

In this case $g = f'$ a.e.
Maybe it is time to consider proving some of all of this. Maybe we should begin proving that if \( f \) is increasing, it is a.e. differentiable. To do this, we have to define first lateral \( \lim \inf \) and \( \lim \sup \) of a function.

Let \( a \in \mathbb{R} \) and let \( f \) be a function defined (at least) in some interval \((a, b)\). Roughly,

\[
\limsup_{x \to a^+} f(x)
\]

is a quantity (possibly equal to \( \infty \) or \( -\infty \)) with the property that as \( x \) approaches \( a \) from the right it holds that for every \( \epsilon > 0 \) eventually all values of \( f \) are below \( \limsup_{x \to a^+} f(x) + \epsilon \) but, however close \( x \) is to \( a \) one can always find some values above \( \limsup_{x \to a^+} f(x) - \epsilon \). (To understand this one should have clear the difference in meaning between “all” and “some.”)

A precise definition is similar to the one used for sequences. Suppose \( f \) is defined in \((a, b)\). For \( x \in (a, b) \) define \( g(x) \) by

\[
g(x) = \sup \{ f(y) : a < y < x \}.
\]

If \( a < x_1 < x_2 < b \) then \( \{ f(y) : a < y < x_1 \} \subset \{ f(y) : a < y < x_2 \} \) thus

\[
\sup \{ f(y) : a < y < x_1 \} \leq \{ f(y) : a < y < x_2 \};
\]

the function \( g \) is increasing, hence \( \lim_{x \to a^+} g(x) = \inf_{x > a} g(x) \) exists. By definition,

\[
\lim_{x \to a^+} \sup f(x) = g(x^+) = \inf_{a < x < b} \left( \sup_{a < y < x} f(y) \right).
\]

Similarly one defines

\[
\lim_{x \to a^+} \inf f(x) = \sup_{a < x < b} \left( \inf_{a < y < x} f(y) \right).
\]

All this is very nice; this definition has the great advantage that it shows that these right handed limits always exist. But the following characterizations are more useful.

**Lemma 4** Let \( f \) be defined (at least) in some interval \((a, b)\), \( a < b \), and let \( \bar{L} = \limsup_{x \to a^+} f(x) \), \( L = \liminf_{x \to a^+} f(x) \). Then \( L \leq \bar{L} \). Moreover, \( \bar{L} \) is the only element of \([-\infty, \infty]\) satisfying the following two properties:

1. If \( c > \bar{L} \), then there exists \( \delta > 0 \) such that \( f(x) < c \) for all \( x \in (a, a + \delta) \).
2. If \( c < \bar{L} \), then for every \( d > 0 \) there is \( x \in (a, a + \delta) \) such that \( f(x) > c \).

Similarly, \( \underline{L} = \liminf_{x \to a^+} f(x) \) is the unique (extended) number such that

1. If \( c < \underline{L} \), then there exists \( \delta > 0 \) such that \( f(x) > c \) for all \( x \in (a, a + \delta) \).
2. If \( c > \underline{L} \), then for every \( d > 0 \) there is \( x \in (a, a + \delta) \) such that \( f(x) < c \).

To avoid distractions, the simple proof is relegated to the appendix. A few trivial consequences to notice are:

1. \( \bar{L} = \infty \) is equivalent to saying that for every \( c \in \mathbb{R} \), every \( \delta > 0 \), there is \( x \in (a, a + \delta) \) such that \( f(x) > c \). In other words, \( f \) is not bounded above to the right of \( a \).
2. \( \bar{L} = -\infty \) is equivalent to saying that for every \( c \in \mathbb{R} \) there exists \( \delta > 0 \) such that \( f(x) < c \) for all \( x \in (a, a + \delta) \); i.e., \( \lim_{x \to a^+} = -\infty \).
3. \( \underline{L} = \infty \) is equivalent to saying that for every \( c \in \mathbb{R} \) there exists \( \delta > 0 \) such that \( f(x) > c \) for all \( x \in (a, a + \delta) \); i.e., \( \lim_{x \to a^+} = \infty \).
4. \( \underline{L} = -\infty \) is equivalent to saying that for every \( c \in \mathbb{R} \), every \( \delta > 0 \), there is \( x \in (a, a + \delta) \) such that \( f(x) < c \). In other words, \( f \) is not bounded below to the right of \( a \).
5. \( \bar{L} = \underline{L} \) if and only if \( \lim_{x \to a^+} f(x) \) exists, in which case

\[
\lim_{x \to a^+} f(x) = \bar{L} = \underline{L}.
\]
There being no right without left, one also defines \( \limsup_{x \to a^-} f(x) \), \( \liminf_{x \to a^-} f(x) \) by

\[
\limsup_{x \to a^-} f(x) = \inf_{x < a} \left( \sup_{x < y < a} f(y) \right), \quad \liminf_{x \to a^-} f(x) = \sup_{x < a} \left( \inf_{x < y < a} f(y) \right).
\]

The counterpart to Lemma 4 is, of course,

**Lemma 5** Let \( f \) be defined (at least) in some interval \((a, b)\), \( a < b \), and let \( L = \limsup_{x \to b^-} f(x) \), \( \underline{L} = \liminf_{x \to b^-} f(x) \). Then \( \underline{L} \leq L \). Moreover, \( \underline{L} \) is the only element of \([-\infty, \infty]\) satisfying the following two properties:

1. If \( c > \underline{L} \), then there exists \( \delta > 0 \) such that \( f(x) < c \) for all \( x \in (b - \delta, b) \).
2. If \( c < \underline{L} \), then for every \( d > 0 \) there is \( x \in (b - \delta, b) \) such that \( f(x) > c \).

Similarly, \( \overline{L} = \liminf_{x \to b^-} f(x) \) is the unique (extended) number such that

1. If \( c < \overline{L} \), then there exists \( \delta > 0 \) such that \( f(x) > c \) for all \( x \in (b - \delta, b) \).
2. If \( c > \overline{L} \), then for every \( d > 0 \) there is \( x \in (b - \delta, b) \) such that \( f(x) < c \).

Having defined these new acquaintances (will they become friends?), we try to prove that if \( f : (a, b) \to \mathbb{R} \) is increasing, then \( f \) is differentiable a.e. For this we define four new functions on \((a, b)\) by

\[
D^+ f(x) = \limsup_{y \to x^+} \frac{f(y) - f(x)}{y - x},
\]

\[
D_+ f(x) = \liminf_{y \to x^+} \frac{f(y) - f(x)}{y - x},
\]

\[
D^- f(x) = \limsup_{y \to x^-} \frac{f(y) - f(x)}{y - x},
\]

\[
D_- f(x) = \liminf_{y \to x^-} \frac{f(y) - f(x)}{y - x},
\]

if \( x \in (a, b) \). A first observation is that because \( f \) is increasing, all of these quantities are \( \geq 0 \) (possibly equal to \( \infty \)). One also has \( D^+ f(x) \geq D_+ f(x), D^- f(x) \geq D_- f(x) \) for all \( x \), but there is no reason to assume any other relation. The proof of a.e. differentiability consists in proving

1. The set where any two of these four quantities differ is a null set. Thus

\[
g'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}
\]

is defined a.e. But it could, in principle, be infinity on more than a null set.

2. \( g' \in L^1(a, b) \), hence finite a.e.

To see all four “quasi-derivatives” are equal a.e. it suffices to show that for every two of them, the set where one is larger than the other one is a null set. Suppose we have two measurable functions \( \phi, \psi \) defined on an interval \((a, b)\). To show the set \( \{ \phi < \psi \} \) is a null set one can use (and this is what will be used here) the rational numbers in between trick. Let us define

\[
G(\phi, \psi) = \{ x \in (a, b) : \phi(x) > \psi(x) \}.
\]

Then

\[
G(\phi, \psi) = \bigcup_{(u, v) \in \mathbb{Q} \times \mathbb{Q}, u > v} \{ x \in (a, b) : \phi(x) > u > v > \psi(x) \}.
\]

Because the rationals are countable, it is a breeze to see that \( m(G(\phi, \psi)) = 0 \) if and only if

\[
m(\{ x \in (a, b) : \phi(x) > u > v > \psi(x) \}) = 0 \quad \text{for all } u, v \in \mathbb{Q}, u > v.
\]
And now things get a bit painful. We have to see that if \( \phi, \psi \) are any two from the set \( \{ D^+, D_+, D_-^f, D_-f \} \), if \( u, v \in \mathbb{Q} \), \( u > v \), then the set
\[
G(\phi, \psi, u, v) = \{ x \in (a, b) : \phi(x) > u > v > \psi(x) \}
\]
is a null set. Considering that if \( \phi = D_-f, \psi = D_+ f \) or if \( f = D_-, \psi = D_-^f \), the sets \( G(\phi, \psi, u, v) \) are empty (so we don’t need to consider these cases), we are left with “only” 14 cases, if I counted right. Royden proves that \( G(D^+ f, D_- f, u, v) \) is a null set for all \( u, v \in \mathbb{Q}, u > v \). Let’s try to see here that \( G(D^+ f, D_+ f, u, v) \) is a null set if \( u > v \); in other words (in more words) let us prove that if \( u > v \) then the set of all \( x \in (a, b) \) such that
\[
(2) \quad \limsup_{y \to x^+} \frac{f(y) - f(x)}{y - x} > u > v > \liminf_{y \to x^+} \frac{f(y) - f(x)}{y - x}
\]
is a null set. I’ll call this set \( E \), for short.

We want to see \( m^*(E) = 0 \) (we could argue that \( E \) is measurable and that we can use \( m \) instead of \( m^* \), but why bother?). As is usual, we prove that a non-negative number equals 0 by proving it is \( < \epsilon \) for any given \( \epsilon > 0 \), so let us give \( \epsilon > 0 \). Imitating Royden, there is some open set \( U, E \subset U \subset (a, b) \) such that \( m(U) < m^*(E) + \epsilon \). Incidentally, since \( E \subset (a, b) \) and we may assume (why?) that \( a, b \) are finite, there is no question that \( m^*(E) < \infty \).

Suppose \( x \in E \) so that \( 2 \) holds. Looking at the third (last) inequality it tells us that a liminf is \( < v \), so by Lemma 4 for every \( \eta > 0 \) there is \( y_{x, \eta} \in (x, x + \eta) \) such that
\[
\frac{f(y_{x, \eta}) - f(x)}{y_{x, \eta} - x} < v, \quad \text{or} \quad f(y_{x, \eta}) - f(x) < v(y_{x, \eta} - x).
\]
We may assume that \( y_{x, \eta} \) is taken close enough to \( x \) so that \( [x, y_{x, \eta}] \subset U \).

The collection of intervals \( \{ [x, x + y_{x, \eta}] : x \in E, \eta > 0 \} \) constitute a Vitali covering of \( E \), thus there exists a finite number of these intervals, say \( \{ I_n \}_{1 \leq n \leq N} \), \( I_n = [x_n, y_n] \) (\( y_n = y_{x_n, \eta_n} \)) for \( n = 1, 2, \ldots, N \) such that \( I_n \cap I_m = \emptyset \) if \( n \neq m \) and
\[
m^*(E \setminus \bigcup_{n=1}^{N} I_n) < \epsilon.
\]

Let
\[
A = E \cap \bigcup_{n=1}^{N} (x_n, y_n).
\]
Now \( E \setminus A \) differs from \( E \setminus \bigcup_{n=1}^{N} I_n \) by a finite set, hence a by a null set, so \( m^*(E \setminus A) < \epsilon \). From \( E = A \cup (E \setminus A) \) we get \( m^*(E) \leq m^*(A) + m^*(E \setminus A) < m^*(A) + \epsilon \) so that \( m^*(A) > m^*(E) - \epsilon \). On the other hand, because \( A \subset U \) (all intervals \( [x_n, y_n] \) are contained in \( U \)), \( m^*(A) \leq m^*(U) < m^*(E) + \epsilon \). The point is that \( A \) is sort of close to \( E \). By the definition of \( y_n = y_{x_n, \eta_n} \) we also have
\[
(3) \quad \sum_{n=1}^{N} (f(y_n) - f(x_n)) < v \sum_{n=1}^{N} (y_n - x_n) \leq v m(U) < v(m^*(E) + \epsilon).
\]
The reason \( \sum_{n=1}^{N} (y_n - x_n) \leq m(U) \) is that the intervals \( I_n \) are pairwise disjoint and contained in \( U \) so that \( \bigcup_{n=1}^{N} I_n \subset U \) and
\[
\sum_{n=1}^{N} (y_n - x_n) = \sum_{n=1}^{N} m(I_n) = m \left( \bigcup_{n=1}^{N} I_n \right) \leq m(U).
\]
Now \( A \subset E \), so \( 2 \) holds if \( x \in A \). But we should change the notation to avoid overloading, so let \( z \in A \) and consider now the first inequality in \( 2 \); that is, (to rewrite it replacing \( x \) by \( z \), etc.)
\[
\liminf_{w \to z^+} \frac{f(w) - f(z)}{w - z} > u.
\]
It states that a limsup is larger than \( u \), hence for every \( \delta > 0 \) there will be \( w_{z, \delta} \in (z, z + \delta) \) such that
\[
f(w_{z, \delta}) - f(z) > u, \quad \text{or} \quad f(w_{z, \delta}) - f(z) > u(w_{z, \delta} - z).
\]
Because \( z \in A \) and \( A \) is included in the union of the interiors of the \( I_n \), we have \( z \in (x_n, y_n) \) for some \( n \), then we may assume that \([z, w_{z,\delta}] \subset (x_n, y_n)\) by taking \( w_{z,\delta} \) sufficiently close to \( z \). The intervals \([z, w_{z,\delta}]\) for \( z \in A \), \( \delta > 0 \) are thus a Vitali covering of \( A \), hence there is a pairwise disjoint finite number of them covering \( A \) up to \( \epsilon \); that is, there exist \( z_1, \ldots, z_M \in A \), \( \delta_1, \ldots, \delta_M > 0 \) such that setting \( w_m = w_{z_m,\delta_m}, J_m = [z_m, w_m] \) for \( m = 1, \ldots, M \), \( J_m \cap J_r = \emptyset \) if \( m \neq r \), and

\[
m^* \left( A \setminus \bigcup_{m=1}^{M} J_m \right) < \epsilon.
\]

Recalling the construction of \( A \) and how we saw that \( m^*(A) > m^*(E) - \epsilon \), we can now introduce

\[
B = A \cap \bigcap_{m=1}^{M} J_m
\]

(we don’t need to work with the interiors here), then \( B \) is a subset of \( A \) covered by the \( J_m \)'s and \( m^*(B) > m^*(A) - \epsilon > m^*(E) - 2\epsilon \). Also, because \( f(w_m) - f(z_m) > u(w_m - z_m) \),

\[
\sum_{m=1}^{M} (f(w_m) - f(z_m)) > u \sum_{m=1}^{M} (w_m - z_m) = m \left( \bigcup_{j=1}^{M} J_m \right) \geq um^*(B) > u(m^*(E) - 2\epsilon).
\]

Here we could ask an interesting question. Where have we used that \( f \) is increasing? The answer is: So far, nowhere. What was done, so far, could be done with any function. Now comes the one and only point where the hypothesis of \( f \) being increasing is used.

By construction, each \( J_m \) is contained in some \( I_n \). For \( n \in \{1, \ldots, N\} \), consider

\[
\sum_{m, J_m \subset I_n} (f(w_m) - f(z_m)).
\]

The intervals \( J_m \) are pairwise disjoint and because \( f \) is increasing it is easy to see that this sum is dominated by \( f(y_n) - f(x_n) \). What we are doing is adding up the rise of the function over each of the intervals \( J_m \), and the sum of all these rises cannot exceed the total rise of the function in the containing interval \( I_n \), which is \( f(y_n) - f(x_n) \). Combining this fact with (3) and (4), we get

\[
u(m^*(E) - 2\epsilon) < \sum_{m=1}^{M} (f(w_m) - f(z_m)) = \sum_{n=1}^{N} \sum_{m, J_m \subset I_n} (f(w_m) - f(z_m)) \leq \sum_{n=1}^{N} (f(y_n) - f(x_n)) < v(m^*(E) + \epsilon).
\]

The conclusion is:

\[
u(m^*(E) - 2\epsilon) < v(m^*(E) + \epsilon).
\]

This inequality has to hold for all \( \epsilon > 0 \), hence we proved \( um^*(E) \leq vm^*(E) \). This last inequality can only hold if \( m^*(E) = 0 \), since \( u > v \).

And so it goes. Two down, twelve to go.

The proof of Theorem 3 is a bit less exhausting. In one direction it is even fairly easy, if you did your exercises. Let \( f : [a, b] \to \mathbb{R} \) and suppose there exists \( g \in L^1([a, b]) \), \( \gamma \in \mathbb{R} \), such that

\[
f(x) = \gamma + \int_{a}^{x} g(t) \, dt.
\]

(I am not assuming the constant in front is \( f(a) \); as we shall see, it has to be \( f(a) \)).
A perhaps useful comment. One defines $\int_a^b f(x)\,dx$ to be the Lebesgue integral of $f$ over the interval $[a,b]$ or, equivalently, over $(a,b)$. The value is the same whether one uses $[a,b]$ or $(a,b)$ because by Proposition 18 of the Concordance notes we would have

$$\int_{[a,b]} f \, dm = \int_{(a,b)} f \, dm + \int_{\{a\}} f \, dm + \int_{\{b\}} f \, dm$$

and the two last integrals are zero because $\{a\}, \{b\}$ are null sets.

We want to prove $f$ is absolutely continuous and $g = f'$ a.e. This has to be divided up into several components:

1. $f(y) - f(x) = \int_{[x,y]} f \, dm$ for all $x, y \in [a,b], x \leq y$.

2. $f$ is a.c.

3. $f' = g$ a.e.

For the first of these properties, we invoke again Proposition 18 of the Concordance notes. If $a \leq x \leq y \leq b$,

$$f(y) = f(y) + \int_{[a,y]} g \, dm = f(y) + \int_{[a,x]} g \, dm + \int_{[x,y]} g \, dm = f(y) + \int_{[a,x]} g \, dm + \int_{[x,y]} g \, dm = f(x) + \int_{[x,y]} g \, dm$$

and the first result follows.

To see $f$ is absolutely continuous, let $\epsilon > 0$. Recall Proposition 13 from the Concordance notes. Because $g$ is non-negative, integrable, there is $\delta > 0$ such that if $E \subset [a,b]$ and $m(E) < \delta$, then

(5)

Assume now we have any (finite) number of pairwise disjoint subintervals $I_j = [x_j, y_j]$ of $[a,b]$, say for $j = 1, \ldots, m$, such that the sum of their lengths satisfies $\sum_{j=1}^m (y_j - x_j) < \delta$. Let $E = \bigcup_{j=1}^m I_j$. Then $m(E) < \delta$ and (5) holds. By Proposition 12 of the concordance notes we then have

$$\sum_{j=1}^m |f(y_j) - f(x_j)| = \sum_{j=1}^m \int_{x_j}^{y_j} |g(t)| \, dt \leq \sum_{j=1}^m \int_{x_j}^{y_j} g(t) \, dt = \sum_{j=1}^m \int_{I_j} g \, dm = \int_E |g| \, dm < \epsilon$$

and absolute continuity follows. Because $f$ is continuous, it is clear that $\gamma = f(a)$.

For the third point, one might want to try to imitate the proof used for the fundamental theorem of calculus, but there one proves $f' = g$ at all points of continuity of $g$. In the current environment, these points could be scarce. However, one can approximate $g$ by continuous functions (Lusin’s Theorem) and one can get a proof from this. Or one can use the result established with so much effort, that increasing functions, hence also functions of bounded variation, are differentiable a.e. (as Royden does in Lemma 9, Theorem 10 of Chapter 5). Since absolute continuity implies bounded variation, $f$ is differentiable a.e.

I’ll take the easy way out, and skip the proof. You probably shouldn’t.

The converse requires once more Vitali to prove first: If $f$ is a.c. and $f' = 0$ a.e., then $f$ is constant. After this result, the converse is almost trivial; all one needs to do is use the fact that if $f$ is a.c., then $f'$ is integrable, set up a new function $h$ by $h(x) = f(a) + \int_a^x f'(t) \, dt$. By what was proved (or left unproved), $h$ is a.c and $h' = f'$ a.e.; i.e., the absolutely continuous function $f - h$ satisfies $(f - h)' = 0$ a.e.. Thus $f = h$ a.e. and, by continuity, $f = h$ everywhere.

Appendix

Proof of Lemma 1 Order the elements of $D$ as a sequence, $D = \{t_1, t_2, \ldots\}$. let $c_n$ be a positive real number for each $n \in \mathbb{N}$, such that $\sum_{n=1}^\infty c_n < \infty$. For example $c_n = 1/n^2$ for each $n$. Or $c_n = 2^{-n}$. Or the terms of your favorite convergent series of positive terms. If $x \in [a,b]$ let $N_x = \{n \in \mathbb{N} : t_n < x\}$. For a given $x \in [a,b]$, the set $N_x$ can be finite or infinite. For
example, if $D$ is the set of all rational numbers in $[a, b]$, then $N_x$ is infinite for all $x \in (a, b)$. In general, $N_a = \emptyset$ and $N_b$ is either all of $\mathbb{N}$ or $\mathbb{N}$ minus one integer, depending on whether $b$ does not belong, or belongs to $D$. Define $f$ by

$$f(x) = \sum_{n \in N_x} c_n.$$ 

So $f(a) = 0$ (because $N_a = \emptyset$, if $b \notin D$ then $f(b) = \sum_{n=1}^{\infty} c_n$. It is quite easy to see that $f$ is increasing. It is a bit harder, but not by much, to see that $f(t-) = f(t) < f(t+) = f(t) + c_m$ if $t = t_m \in D$. 

**Proof of Lemma** That $\limsup_{x \to a^+} f(x) \geq \liminf_{x \to a^+} f(x)$ is an easy consequence of the obvious inequality $\sup_{a < y < x} f(y) \geq \inf_{a < y < x} f(y)$, valid for all $x > a$. This implies at once that

$$\limsup_{x \to a^+} f(x) = \lim_{x \to a^+} \sup_{a < y < x} f(y) \geq \lim_{x \to a^+} \inf_{a < y < x} f(y) = \liminf_{x \to a^+} f(x).$$

Let $c > \limsup_{x \to a^+} f(x)$. Going to the definition,

$$c > \inf_{y > x} f(y);$$

to be larger than the inf of a set means not to be a lower bound of the set, thus there is $x > a$ such that $c > \sup_{a < y < x} f(y)$, hence $c > f(y)$ for all $y \in (a, x)$. If we set $\delta = x - a$ we see $\delta > 0$ and $f(y) < c$ for all $y \in (a, a + \delta)$. Now let $c < \limsup_{x \to a^+} f(x)$. To be less than the inf of a set means the same as being smaller than everything in the set, thus $c < \sup_{a < y < x} f(y)$ for all $x > a$; thus $c$ is not an upper bound of $\{f(y) : a < y < x\}$ for all $x > a$, which is to say that $c$ is not an upper bound of $\{f(y) : a < y < a + \delta\}$ for all $\delta > 0$. Thus for every $\delta > 0$ there is $y \in (a, a + \delta)$ such that $f(y) > c$. This proves that $\limsup_{x \to a^+} f(x)$ satisfies the required properties. The uniqueness of a number satisfying these two properties is left as an (easy) exercise. The proof for the liminf is essentially the same, and will be omitted.