Example 2.29. If $\rho(x)$ is the linear density of a rod, then $\int_a^b \rho(x) \, dx = m(b) - m(a)$, where $m(x)$ is the mass of a length $b$ of the rod, measured from any arbitrary starting point.

Example 2.30. Suppose $g(x) = f(x) + c$, for a constant $c$. Then $g'(x) = f'(x)$. Thus

$$\int_a^b g'(x) \, dx = \int_a^b f'(x) \, dx$$

which implies by the Evaluation theorem that

$$g(b) - g(a) = f(b) - f(a)$$

which seems fishy until we check: $g(b) = f(b) + c, g(a) = f(a) + c$, so the difference $g(b) - g(a)$ is indeed equal to $f(b) - f(a)$. So the choice of antiderivative doesn’t matter.

So why is part 2 true? It relies on the following result, called the Mean Value Theorem (to be proven in MAT1325, but which you’ll agree is very reasonable):

**Theorem 2.31 (Mean Value Theorem).** Let $f$ be a differentiable function on an interval $[x_0, x_1]$. Then there is some point $x^* \in [x_0, x_1]$ such that

$$f'(x^*) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

that is, on any interval, your derivative function must attain the average value of the function on that interval.

This says that, for example, if your average velocity over a given time interval was $80 \text{ km/h}$ (the right hand side above), then necessarily there was at least one moment (possibly many!) during that interval when your instantaneous velocity was exactly $80 \text{ km/h}$.

How does this relate to our theorem? Let’s divide $[a, b]$ into $n$ subintervals and build a Riemann sum:

$$\sum_{i=1}^{n} f'(x^*_i) \Delta x$$

where $\Delta x = x_i - x_{i-1}$. Choose your sample point $x^*_i$ using the Mean Value theorem, so that we have

$$f'(x^*_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} \Rightarrow f(x_i) - f(x_{i-1}) = f'(x^*_i) \Delta x$$

so that our Riemann sum (writing it out in long notation) is just

$$\sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) = (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + (f(x_3) - f(x_2)) + \cdots + (f(x_n) - f(x_{n-1})).$$

This kind of sum — where almost everything cancels except in the first and last terms — is called a telescoping sum, and it simplifies to

$$f(x_n) - f(x_0) = f(b) - f(a)$$

as required.

### 2.4 Lecture 6: The rest of the fundamental theorem of Calculus

We have been discussing the definite integral for four lectures now.

Recall: We defined the definite integral of $f$ from $a$ to $b$ to be

$$\int_a^b f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x^*_i) \Delta x$$

(provided this limit exists) where
\[ \Delta x = (b - a)/n; \]

- for each \( i \in \{0, 1, \cdots, n\} \), \( x_i = a + i\Delta x; \)
- \( x^*_i \) is any choice of sample point in the interval \([x_{i-1}, x_i]\).

We say that \( f \) is integrable on \([a, b]\) when \( \int_a^b f(x) \, dx \) exists.

Last time, we saw the Evaluation Theorem (or: FTC Part 2). We can write it as

\[ \int_a^b f(t) \, dt = F(b) - F(a) \]

where \( F \) is any anti-derivative of \( f \), that is, \( F' = f \). The notation we use when solving problems is:

\[ \int_a^b f(t) \, dt = F(b) - F(a) \quad \text{where} \quad F' = f. \]

This theorem changes the problem of finding a limit of Riemann sums to the problem of finding an anti-derivative, which is often much easier.

**Example 2.32.** If \( f(x) = 2x \) then we saw \( F(x) = x^2 \) and so

\[ \int_2^7 2x \, dx = x^2 \bigg|_2^7 = 7^2 - 2^2 = 49 - 4 = 45. \]

Part 2 of the FTC said: the integral of the derivative of \( f \) is entirely determined by \( f \). Part 1 of the FTC will say: the derivative of the integral of \( f \) is \( f \). But for this to make any sense, we need to be able to think of the integral of \( f \) as a function, so that’s our first step.

### 2.4.1 The “integral so far” function

(The material from this section is in [S, Chapter 5.4].)

Let \( f \) be a function, and let \([a, b]\) be an interval in the domain of \( f \). How do we turn the integral into a function? It turns out, we can’t define a unique function, but only a family of functions.  

**Definition 2.33.** Let \( f \) be a function with domain \( A \subseteq \mathbb{R} \) and range \( \mathbb{R} \). Let \( a \in A \) and let \( B \) be the set of all points \( x \) in the domain \( A \) such that \( f \) is integrable on \([a, x]\). (If \( x < a \), then we mean instead that \( f \) is integrable on \([x, a]\).) Define a new function \( F \) with domain \( B \) by setting, for each \( x \in B \),

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

Since \( F \) depends on the choice of \( a \), we sometimes write \( F_a \) to be more clear. We call it the “area so far” function, since it measures the area from \( a \) as far as \( x \).

**Remark 2.34.** Remember that the \( t \) inside the integrand is just a “dummy variable”; we could have called it anything. But it wouldn’t be nice to call it \( x \), because we’re already using \( x \) for the variable representing the upper limit of integration, and that’s the independent variable of this function.

**Example 2.35.** Suppose \( f = g' \). Then

\[ \int_a^x f(t) \, dt = \int_a^x g'(t) \, dt = g(x) - g(a). \]

So the integral so far function in this case is just an antiderivative of the integrand. Typically, it’s not the one you started with \((g(x))\) but some vertical shift of this \((g(x) - g(a))\); note that \( g(a) \) is just a constant.

---

1What? Well, you can’t talk about your position without reference to a starting point. You can talk about velocity without an absolute reference, but if you talk about displacement, you have to identify a starting point, or starting time, or something.
A more interesting example is \( f(x) = e^{-x^2} \). The graph of this function is a bell-curve, and this function is used everywhere in statistics. However, (something we’ll show later!), there is no simple function \( g(x) \) such that \( f(x) = g'(x) \). Nevertheless, we can still define a function

\[
F(x) = \int_a^x e^{-t^2} dt
\]

but in this case we have no other formula for this function — this expression is as good as it gets. OK, fine, it’s a function — the question is, is this function still an anti-derivative of \( f \)?

Yes; this is the Fundamental Theorem, part 1.

2.4.2 The Fundamental Theorem of Calculus, part 1

We can finally state the rest of the theorem:

**Theorem 2.36** (Fundamental Theorem of Calculus, Part 1). Let \( f \) be an integrable function on \([a, b]\). Then

\[
\frac{d}{dx} \int_a^x f(t) \, dt = f(x)
\]

for every \( a < x < b \). That is, if \( F(x) = \int_a^x f(t) \, dt \), then \( F'(x) = f(x) \); \( F \) is an anti-derivative of \( f \).

Notice the symmetry with part 2 (evaluation theorem): to calculate an integral, we use an anti-derivative; to find an anti-derivative, we use an integral.

In any given situation, you only use one half of the FTC: either you find an anti-derivative somehow, and apply Part 2 to evaluate an integral; or else you find the integral-so-far function somehow, and use that to produce the anti-derivative.

2.4.3 What does part 1 mean? Why is it true? What does it give us?

So part 1 says: the derivative of the area-so-far function is just your function back again. The rate of change of the area under your curve is precisely equal to the \( y \)-value at that point on the curve.

**Example 2.37.** Let \( f(x) = 3 \). Then for any \( a \), \( F_a(x) = 3(x - a) = 3x - 3a \). Then \( F'(x) = 3 = f(x) \), by what we calculated before.

It is also helpful to think about what the “integral so far” function means, to recognize why the theorem is true.

**Example 2.38.** If \( v(t) \) represents the instantaneous velocity of a particle at time \( t \), then since each rectangle of the Riemann sum is velocity times time, giving displacement, we see that \( V(t) \) represents the displacement of the object since an initial time \( t = a \).

**Example 2.39.** If \( f(t) = |v(t)| \) represents the instantaneous speed of a particle at time \( t \), then \( F_a(t) \) is just the distance travelled by the particle since time \( t = a \).

**Example 2.40.** If \( f(x) \) represents the linear density of a rod, then each rectangle of the Riemann sum corresponds to density times length, which is mass. So in the limit, \( F_a(x) \) is the total mass of a piece of rod starting at the point \( a \) and ending at \( x \).

So why is the theorem true? The left hand side is

\[
\frac{d}{dx} F_a(x) = \lim_{h \to 0} \frac{F_a(x + h) - F_a(x)}{h}.
\]

So let’s work out this difference quotient:

\[
\frac{1}{h}(F_a(x + h) - F_a(x)) = \frac{1}{h} \left( \int_a^{x+h} f(t) \, dt - \int_a^x f(t) \, dt \right)
= \frac{1}{h} \int_x^{x+h} f(t) \, dt
\]
(using the properties of the integrals). Now what is this term? It is \( \frac{1}{h} \) times the area under \( f \) between \( x \) and \( x + h \). The area of this slice is approximately \( f(x)((x + h) - x) = f(x)h \) (and in the limit as the interval gets infinitesimally small, this estimate gets more and more accurate), so say

\[
P'_a(x) = \lim_{h \to 0} \frac{F_a(x + h) - F_a(x)}{h} = \lim_{h \to 0} \frac{f(x)h}{h} = f(x)
\]

which is what we wanted.

Now what does it give us? It tells us how to produce a function whose derivative is \( f(x) \), that is, to antidifferentiate. Unfortunately, as a formula, it’s not so wonderful, since it’s so hard to compute. Think of it as a theorem of last resort: you weren’t able to find an anti-derivative, but you really need one because it will solve all your problems, so you grab some software and produce this area-so-far function.\(^2\)

### 2.4.4 The indefinite integral

A consequence of the Fundamental Theorem is the following definition.

**Definition 2.41.** Let \( f \) be a function. The **indefinite integral** of \( f \) is the general anti-derivative of \( f \), and is written

\[
\int f(x) \, dx.
\]

That is, if \( F \) is any particular anti-derivative of \( f \), then

\[
\int f(x) \, dx = F(x) + c
\]

where \( c \) is an arbitrary constant.\(^3\)

**Example 2.42.** \( \int 2x \, dx = x^2 + c \) where \( c \) is an arbitrary constant

**Example 2.43.** For any constant \( c \), \( \int c \, dx = cx + a \), where \( a \) is an arbitrary constant (Note that we had to change the name of the constant since \( c \) was already being used.)

So the indefinite integral is a family of functions. Part 1 of the FTC says that the members of this family are all of the different “area-so-far” functions; Part 2 says that

\[
\int_a^b f(x) \, dx = \left[ F(x) \right]_a^b
\]

(and we can just forget the arbitrary constant in this case, since it will disappear in the difference on the right).

### 2.4.5 (Extra) food for thought about the “integral-so-far” function

*We didn’t do this in class; it’s simpler to use the FTC Part 2 to generate examples than to work them out by hand, so the value in this section is the practice in thinking about how to integrate and what it means.*

We can calculate this function (area-so-far) explicitly in some cases:

**Example 2.44.** Suppose \( f(x) = 3 \), the constant function. Then for any \( a \), we have

\[
F_a(x) = \int_a^x 3 \, dt = 3(x - a)
\]

by the properties of the integral last time. (Note that if \( x < a \) we get a negative answer, as expected for the “backwards integral”.)

---

\(^2\)There are nice examples in Physics and in Probability where you need an antiderivative and the best description that exists is as an area-so-far function. For example, the general normal distribution (bell curve) is the function \( e^{-(x-\mu)^2/\sigma^2} \) and the probability that a sample lies in the interval \([a,b]\) is the integral of this function from \( a \) to \( b \) — and the only way to compute this integral is numerical approximation. The only anti-derivative you can ever write down is \( \int e^{-(t-\mu)^2/\sigma^2} \, dt \).

\(^3\)We have seen that adding a constant doesn’t change the derivative; we’ll need to show later that that is the only transformation of \( F \) which doesn’t change the derivative.
Example 2.45. Suppose \( f(x) = 2x \). Suppose we choose \( a = 0 \). Then for \( x \geq 0 \) we have

\[
F(x) = \int_0^x f(t) \, dt = \int_0^x 2t \, dt
\]

which is the area of a triangle with base \( x \) and height \( 2x \) so area \( x^2 \). And if \( x < 0 \), then

\[
F(x) = \int_0^x f(t) \, dt = -\int_x^0 2t \, dt.
\]

The region described by this integral is a triangle below the \( x \)-axis, which makes the integral negative; but we are then multiplying by \(-1\), so our answer should be positive. The area is again \( x^2 \), a positive value.

We conclude that for any \( x \),

\[
F(x) = \int_0^x f(t) \, dt = x^2
\]

What happens if we change the lower limit of integration \( a \)?

Lemma 2.46. Suppose \( f \) is integrable on its domain \( A \) and \( a < b \in A \). Define

\[
F_a(x) = \int_a^x f(t) \, dt \quad \text{and} \quad F_b(x) = \int_b^x f(t) \, dt.
\]

Then for every \( x \in A \),

\[
F_a(x) = F_b(x) + C
\]

where \( C = \int_a^b f(t) \, dt \) is a constant.

Proof. We need to show that

\[
\int_a^x f(t) \, dt = \int_a^b f(t) \, dt + \int_b^x f(t) \, dt
\]

for any \( a, b, x \in A \). If \( a < b < x \), then this is precisely one of the properties of the integral we saw last time.

If \( a < x < b \), we similarly have

\[
\int_a^x f(t) \, dt + \int_x^b f(t) \, dt = \int_a^b f(t) \, dt
\]

which we can rewrite as

\[
\int_a^x f(t) \, dt - \int_b^x f(t) \, dt = \int_a^b f(t) \, dt
\]

so that we again have (after rearranging):

\[
F_a(x) = c + F_b(x).
\]

If \( x < a < b \) (the last case) then

\[
\int_x^a f(t) \, dt + \int_a^b f(t) \, dt = \int_x^b f(t) \, dt
\]

gives

\[
-F_a(x) + c = -F_b(x)
\]

or

\[
F_a(x) = c + F_b(x).
\]

Thus in fact: we see that \( F_a \) and \( F_b \) have exactly the same graph, just separated by the addition of a constant. \( \square \)
Remark 2.47. Our proof showed that in fact the property
\[ \int_a^c f(t) \, dt + \int_c^b f(t) \, dt = \int_a^b f(t) \, dt \]
holds for any \(a, b, c\), regardless of their order. This is very handy.

Example 2.48. Suppose \(f(x) = 2x\) and we choose a different starting point \(a\). Then the relation
\[ F_0(x) = \int_0^a 2t \, dt + F_a(x) \]
and the fact that \(\int_0^a f(t) \, dt = F_0(a) = a^2\) allows us to conclude that
\[ F_a(x) = x^2 - a^2. \]

We could also compute this directly, by working out the area under the curve as above; the difference is just that the regions are trapezoids rather than triangles.

2.4.6 Now what?

We see that to solve the area problem, we need to be able to compute integrals. To compute integrals using limits of Riemann sums is very hard; but the FTC gives us a shortcut, using anti-derivatives.

The thing to recall: so Part 1 is the only known method for computing anti-derivatives. But it’s not useful! So what we’ll do to take advantage of Part 2 is just learn all known derivatives really well, plus a few general techniques, so that we can identify anti-derivatives.

Therefore, we need to differentiate all the functions we can think of, so that we can start to recognize anti-derivatives.
Chapter 3

Lots of differentiation and derivatives

Our plan for the next few classes:

- discuss the finer points of differentiation, and what it’s useful for, and when it fails;
- find the derivatives of all of the basic functions, and learn different techniques of differentiation that make differentiation easier
- use these to build a table of known anti-derivatives, and to compute definite integrals.

3.1 Derivatives: tangent lines and where functions aren’t differentiable

The material from this section is found in [S, Chapter 2.7].

We saw that the derivative of a function \(f\) at a point \(x\) in the domain of \(f\) is given by

\[
f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h},
\]

when this exists. If it exists, then \(f'\) is said to be differentiable at \(x\). Alternate notation, if \(y = f(x)\):

\[
f' = \frac{dy}{dx} = \frac{df}{dx} = \frac{d}{dx}f,
\]

sometimes with the variable in parentheses \((f'(x), \frac{d}{dx}f(x), \text{etc})\). If we evaluate the derivative at a specific point \(x = a\), then we might write

\[
f'(a) = \left. \frac{dy}{dx} \right|_{x=a} = \left. \frac{df}{dx} \right|_{x=a}
\]

Geometrically, we know that \(f'(a)\) is the slope of the tangent line to the curve \(y = f(x)\) at \((a, f(a))\).

Let’s consider some examples.

Example 3.1. Find the equation of the tangent line to the curve \(y = x^2\) at \(x = 5\).

Solution: We have previously seen that \(y' = 2x\). Therefore at \(x = 5\), the tangent line has slope \(2(5) = 10\). The point on the curve at \(x = 5\) is \((5, 25)\), so our tangent line is given by

\[
m = \frac{y - y_0}{x - x_0} \Rightarrow 10 = \frac{y - 25}{x - 5} \Rightarrow y = 10(x - 5) + 25 = 10x - 25.
\]

Example 3.2. Find the derivative of \(y = \frac{1}{\sqrt{x}}\) and use this to find the equation of the tangent line to the curve at \(x = 4\).
Solution: Set \( f(x) = \frac{1}{\sqrt{x}} \) and compute the difference quotient:

\[
\frac{\Delta y}{\Delta x} = \frac{f(x + h) - f(x)}{h}
\]

\[
= \frac{1}{\sqrt{x+h}} - \frac{1}{\sqrt{x}}
\]

\[
= \frac{h}{\sqrt{x} - \sqrt{x+h}}
\]

The standard thing to do to simplify an expression with a difference of radicals is to rationalize, so we try that.

Multiply by \( \frac{\sqrt{x+h}}{\sqrt{x+h}} \):

\[
\frac{\Delta y}{\Delta x} = \frac{1}{h} \cdot \frac{x - (x+h)}{\sqrt{x} \cdot \sqrt{x+h} \cdot (\sqrt{x} + \sqrt{x+h})}
\]

\[
= \frac{1}{h} \cdot \frac{-h}{\sqrt{x} \cdot \sqrt{x+h} \cdot (\sqrt{x} + \sqrt{x+h})}
\]

\[
= \frac{-1}{\sqrt{x} \cdot \sqrt{x+h} \cdot (\sqrt{x} + \sqrt{x+h})}
\]

and we see that this function is continuous at \( h = 0 \), so the limit exists:

\[
f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}
\]

\[
= \lim_{h \to 0} \frac{-1}{\sqrt{x} \cdot \sqrt{x+h} \cdot (\sqrt{x} + \sqrt{x+h})}
\]

\[
= -\frac{1}{2} x^{-3/2}
\]

Now, at \( x = 4 \) the derivative is

\[
f'(4) = -\frac{1}{2} (4)^{-3/2} = -\frac{1}{2} \left( \frac{1}{8} \right) = -\frac{1}{16}
\]

which is the slope of the tangent line to \( y = \frac{1}{\sqrt{x}} \) at \( (4, \frac{1}{2}) \). Thus the equation of the tangent line is

\[
y = -\frac{1}{16} x + b
\]

where \( b \) is obtained by plugging in the point \( (4, \frac{1}{2}) \): \( b = \frac{1}{2} + \frac{1}{16} \cdot 4 = \frac{3}{4} \); giving final answer

\[
y = -\frac{1}{16} x + \frac{3}{4}.
\]

**Example 3.3.** Find the derivative of the following function at \( x = 1 \), if it exists.

\[
f(x) = \begin{cases} x^2 & x \leq 1 \\ x & x > 1 \end{cases}
\]

In this case, the difference quotient will be different at \( x = 1 \), depending on whether \( h > 0 \) or \( h < 0 \).

Case \( h < 0 \):

\[
\frac{f(1+h) - f(1)}{h} = \frac{(1+h)^2 - 1}{h} = \frac{2h + h^2}{h} = 2 + h
\]
which is continuous at $h = 0$ so we conclude

$$\lim_{h \to 0^-} \frac{f(1 + h) - f(1)}{h} = \lim_{h \to 0^-} (2 + h) = 2$$

where this notation means we’ve only considered the case that $h$ is to the left of 0.

Case $h > 0$:

$$\frac{f(1 + h) - f(1)}{h} = \frac{(1 + h) - 1}{h} = \frac{h}{h} = 1$$

which is a constant, so we conclude

$$\lim_{h \to 0^+} \frac{f(1 + h) - f(1)}{h} = 1$$

where the notation means we’ve only considered the case that $h$ is to the right of 0.

So

$$\lim_{h \to 0^-} \frac{f(1 + h) - f(1)}{h} \neq \lim_{h \to 0^+} \frac{f(1 + h) - f(1)}{h}$$

that is, the slopes don’t match up at the splicing point $x = 1$; this function is not differentiable, since the limit of a function doesn’t exist if the left and right limits aren’t equal; our reasoning is that since the slope at $x = 1$ depends on the direction in which you approach it, there is no well-defined tangent line.

We see that this is the only point at which such a difficulty can occur; this function is differentiable everywhere except at $x = 1$. 

34