Chapter 3

Differential and Integral Calculus

3.1 The derivative

3.1.1 Intuition

Figure 3.1: As we zoom in on the graph $y = f(x)$, it looks more and more like a linear function.

**Definition 3.1.1.** Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and let $x_0 \in (a, b)$. We say that $f$ is differentiable at $x_0$ iff the function

$$m : \mathcal{D} \to \mathbb{R}, \quad m(x) := \frac{f(x) - f(x_0)}{x - x_0},$$

(which is defined and continuous in the domain $\mathcal{D} = [a, b] \setminus \{x_0\}$) has a continuous extension $\tilde{m}$ to $[a, b]$. If this is the case, the derivative of $f$ at $x_0$ is defined as

$$f'(x_0) := \tilde{m}(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$
In other words, $f$ is differentiable at $x_0$ if there is a constant $k$ such that the function

$$m(x) = \begin{cases} \frac{f(x) - f(x_0)}{x - x_0} & x \neq x_0 \\ k & x = x_0 \end{cases}$$

is continuous at $x_0$. If this is the case we say

$$f'(x_0) := k = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

which is the number that makes the function $m$ above continuous at $x_0$.

**Important geometric interpretation:** Let $f : [a, b] \to \mathbb{R}$ be as above, and let $x_0$, $x \in (a, b)$ with $x \neq x_0$. These data determine two points, on the graph of $f$, namely the points:

$$P_0 = (x_0, f(x_0)) \quad \text{and} \quad P = (x, f(x)).$$

These two points are distinct (because $x_0 \neq x$), and so determine a straight line $\ell_{P_0P}$, which is called a *chord* of the graph:

![Figure 3.2: The chord of the graph of $f : x \mapsto 1 - 3x^2$ defined by the points $P_0 = (1/2, 1/4)$ (note that $f(1/2) = 1/4$) and $P = (x, 1 - 3x^2)$. The slope of the chord is $m(x) = \frac{1 - 3x^2 - 1/4}{x - 1/2} = \frac{-3x^2 + 3/4}{x - 1/2}$. The slope of that chord is precisely $m(x)$: $m(x) = \text{slope of chord } \ell_{P_0P} = \frac{f(x) - f(x_0)}{x - x_0}$. This function is well-defined in the domain of $f$ minus the point $x_0$.}
By definition, the function \( f \) is differentiable at \( P_0 \) if and only if the slope function \( m(x) \) has a continuous extension to \( x = x_0 \). Geometrically, as \( x \) varies the chord pivots around \( P_0 = (x_0, f(x_0)) \), and as \( x \) approaches \( x_0 \) the chord \( \ell_{P_0P} \) approaches the tangent to the graph at \( P_0 \), if it exists. One can take this as definition of the tangent line, and therefore this is why \( f'(x_0) \) is the slope of the tangent to the graph at \( P_0 = (x_0, f(x_0)) \).

*Problem 3.1.2. 1. Explain why the definition of derivative of \( f \) at \( x_0 \), if it exists, can also be expressed as:
\[
f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.
\]

Solution.

2. Make a sketch of the graph of some function \( f \) and indicate the geometric significance of the ratio \( \frac{f(x_0 + h) - f(x_0)}{h} \). Make sure you indicate the geometric significance of the parameter \( h \).

Solution.

*Problem 3.1.3. 1. Prove that the function \( g(x) = x^2 \) is differentiable at \( x_0 = 2 \), and find the value of \( f'(2) \).

Solution.

2. Prove that the function \( f(x) = |x| \) is continuous but not differentiable at \( x_0 = 0 \).

Solution.

3. In a previous problem you showed that the function
\[
f(x) = \begin{cases} 
    \frac{x^N - 3^N}{x - 3} & x \neq 3 \\
    N3^{N-1} & x = 3
\end{cases}
\]
is continuous everywhere. Deduce from this fact that the function \( w(x) = x^N \) is differentiable at \( x_0 = 3 \), and find the value of its derivative there.

Solution.

*Problem 3.1.4. 1. Let \( n \) be a positive integer, and \( c \in \mathbb{R} \) a constant. Use the definition to find the derivative of the function \( f(x) = cx^n \) at \( x_0 = a \).\(^1\)

\(^1\)Hint: Use the algebraic identity
\[
x^n - a^n = (x-a)(x^{n-1} + x^{n-2}a + x^{n-3}a^2 + \cdots + xa^{n-2} + a^{n-1}).
\]
Solution.

2. Use the definition (or, equivalently, Problem 3.1.2) to find the derivative of \( g(x) = \sin(x) \) at an arbitrary point \( x \in \mathbb{R} \).

Solution.

3. Show that the derivative of \( h(x) = \cos(x) \) is \( h'(x) = -\sin(x) \) by using the interpretation of the derivative of \( h \) at \( a \) as the slope of the tangent line to \( y = h(x) \) at \( x = a \), the above result, and the identity \( \cos(x) = \sin(x + \frac{\pi}{2}) \).

Solution.

4. The number \( e \) is the real number defined so that \( \lim_{h \to 0} \frac{e^h - 1}{h} = 1 \). Interpret this definition of \( e \) geometrically (note that the expression on the left is a derivative...), and use it to find the derivative of the function \( k(x) = e^x \) at an arbitrary point \( x \in \mathbb{R} \).

Solution.

**Definition 3.1.5.** Let \( f : D \to \mathbb{R} \) be differentiable at \( x_0 \). The **linearization** of \( f \) at \( x_0 \) is the function

\[
L(x) = f(x_0) + (x - x_0)f'(x_0).
\]

**Example 3.1.6.** As an application, linearizations can be used to approximate values of functions, the idea being that the tangent line approximates the graph very well near the point of tangency.

As an example, let’s compute an approximate value of \( \sqrt{145} \). A first approximation is that \( \sqrt{145} \approx \sqrt{144} = 12 \). Can we do better? Yes, using the linearization of the square root function, \( f(x) = \sqrt{x} \), at the point \( x_0 = 144 \). The choice of this point is sensible, because (a) it’s not far from 145, and (b) we can compute both the function and its derivative at this point:

\[
f(144) = 12, \quad f'(x) = \frac{1}{2\sqrt{x}} \quad \Rightarrow \quad f'(144) = \frac{1}{24}.
\]

Therefore, the linearization is

\[
L(x) = 12 + \frac{1}{24}(x - 144) \quad \Rightarrow \quad L(145) = 12 + \frac{1}{24} = 12.04166.
\]

On the other hand, a calculator gives \( \sqrt{145} \approx 12.04159458 \).

\(^2\)Hint: Use the facts that

\begin{align*}
(a) \quad \sin(x + y) &= \sin(x) \cos(y) + \sin(y) \cos(x), \\
(b) \quad \lim_{h \to 0} \frac{\sin h}{h} &= 1, \text{ and} \\
(c) \quad -1 + \cos h &= -\frac{1 + \cos^2 h}{1 + \cos h} = \frac{\sin^2 h}{1 + \cos h}.
\end{align*}
3.2 The Riemann integral

The definition of the Riemann integral involves a fair amount of notation, and it will be good to develop some of it from our intuition of what we would like an integral to mean. Before we define the Riemann integral, consider the following problem:

*Problem 3.2.1.* The following table is a record of measurements of velocity vs. time for a moving object. This could be e.g. a Doppler effect measurement of an object moving straight out from your viewpoint, or a car set up to record its velocity versus time. The table shows data of velocity vs. time, for times $t = 0$ through $t = 2$, in steps of $\Delta t = .2$:

<table>
<thead>
<tr>
<th>t</th>
<th>v(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>.2</td>
<td>3</td>
</tr>
<tr>
<td>.4</td>
<td>6</td>
</tr>
<tr>
<td>.6</td>
<td>11</td>
</tr>
<tr>
<td>.8</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>23</td>
</tr>
<tr>
<td>1.4</td>
<td>29</td>
</tr>
<tr>
<td>1.6</td>
<td>27</td>
</tr>
<tr>
<td>1.8</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
</tr>
</tbody>
</table>

1. From these data, how would you go about estimating the distance the object has traveled between $t = 0$ and $t = 2$? Give your answer as an *algorithm*, meaning: How would you program a computer, e.g. an Excel worksheet, to estimate that distance? Find the numerical value that your method gives.

*Solution.*

2. If one were to repeat the experiment, what would you try to do to generate data to make your estimate more accurate?

*Solution.*

3. You are downloading a file from the internet. Suppose the table above now lists $v =$ the download speed (say in KB/sec) versus time. What is the interpretation of the result of applying your algorithm of part 1 now?
Solution.

Your answer to part 1 of the previous problem was probably something of this sort: You were given \( N + 1 \) data points \( \{t_j\}_{j=0}^{N} \) at which the velocity was measured, which divide the overall interval into \( N \) time segments. The time interval between measurements was constant, although it might not have been. We’ll denote it by \( \Delta t \). To estimate the total distance, you probably formed a sum such as:

\[
L = \sum_{j=0}^{N-1} v(t_j) \Delta t
\]

or maybe

\[
R = \sum_{j=1}^{N} v(t_j) \Delta t = \sum_{j=0}^{N-1} v(t_{j+1}) \Delta t
\]

or maybe

\[
M = \sum_{j=0}^{N-1} \frac{v(t_{j+1}) + v(t_j)}{2} \Delta t,
\]

and it turns out that \( M = \frac{L + R}{2} \). (From previous courses you might recognize \( L \) as the left-hand Riemann sum, \( R \) as the right-hand Riemann sum.)

More generally, we define:

**Definition 3.2.2.** Let \( f : [a, b] \to \mathbb{R} \). A Riemann sum for \( f \) over \([a, b]\) is a sum of the form

\[
\mathcal{R} = \sum_{j=1}^{j=N} f(x_j^*) (x_j - x_{j-1})
\]

where:

- \( a = x_0 < x_1 < x_2 < \cdots < x_N = b \) are \((N + 1)\) points subdividing \([a, b]\) into \( N \) intervals

\[
[x_j, x_{j-1}], \quad j = 1, 2, \ldots N
\]

and

- \( \forall j = 1, 2, \ldots N \quad x_j^* \in [x_j, x_{j-1}] \).

There are some particular instances of the above definition that we will be interested in:

1. If \( \forall j \quad x_j^* = x_{j-1} \), \( \mathcal{R} \) is called a left-hand Riemann sum.

2. If \( \forall j \quad x_j^* = x_j \), \( \mathcal{R} \) is called a right-hand Riemann sum.
Below is a graphical representation of a Riemann sum:

In the velocity example above, if we were given the function $v(t)$, $t \in [a, b]$ and were asked to compute exactly the distance traveled between the times $t = a$ and $t = b$, we’d want to take limits of Riemann sums as the subdivisions get finer and finer. This is the intuitive idea of the definite integral $\int_{a}^{b} v(t) \, dt$.

### 3.2.1 Integrals of monotone functions

It turns out that it’s much easier to develop the theory of integration of functions that are piece-wise monotone (that is, a function that is monotone over intervals). Most “reasonable” functions are piece-wise monotone, so that’s all we’ll consider. In fact, in this subsection we only consider *monotone* functions.

**Definition 3.2.3.** A function $f$ is *monotone increasing* over a domain $\mathcal{D}$ if $\forall s, t \in \mathcal{D}$,

$$s < t \Rightarrow f(s) \leq f(t).$$

*Problem 3.2.4.* In this problem $f : [a, b] \to \mathbb{R}$ denotes a *monotone increasing* function. For each positive integer $N$, consider the subdivision $a = x_0 < x_1 < \cdots < x_N = b$ of $[a, b]$ into $N$ subintervals. As before, let $\mathcal{L}_N$ and $\mathcal{R}_N$ denote the left and right-hand Riemann sums associated to $f$ and this subdivision.

Show that:
1. For each $N$ one has the 3 inequalities

$$(b - a) f(a) \leq \mathcal{L}_N \leq \mathcal{R}_N \leq (b - a) f(b).$$

Solution.

2. What can you say in the spirit of part (1) if the function $f$ is monotone decreasing?

Solution.

3. What can you say in the spirit of part (1) if the function $f$ is not monotone?

Solution.

Now let’s fix a standard way of subdividing the interval $[a, b]$. Let $\mathcal{L}_N$ (resp. $\mathcal{R}_N$), $N \geq 1$, be the left-hand (resp. right-hand) Riemann sum associated to $f$ and the partition of $[a, b]$ into $2^N$ subintervals of equal length $\Delta x = x_j - x_{j-1} = \frac{b-a}{2^N}$. Let $\ell_N(x)$ (resp. $r_N(x)$) be the corresponding step function defined on $[a, b]$ (i.e. the one such that $\mathcal{L}_N$ (resp. $\mathcal{R}_N$) calculates the area under its graph). Specifically,

$$r_N(x) = f(x_{j+1}) \quad \text{if} \quad x \in [x_j, x_{j+1}],$$

and

$$\ell_N(x) = f(x_j) \quad \text{if} \quad x \in [x_j, x_{j+1}].$$

*Problem 3.2.5.* In this problem, we will define the integral $\int_a^b f(x)\,dx$ using the Axiom of Completeness.

1. Draw the graph of a continuous monotone increasing function $f$ on an interval $[a, b]$. Draw in the functions $\ell_N(x)$ and $r_N(x)$ (as defined above) for a couple $N$ values.

Solution.

2. Compute the difference $\mathcal{R}_N - \mathcal{L}_N$ (it should depend on $N$). Explain why $|\mathcal{R}_N - \mathcal{L}_N| \to 0$ as $N \to \infty$.

Solution.

3. Explain why there is exactly one real number $c$ in the intervals $[\mathcal{L}_N, \mathcal{R}_N]$ for all $N$ (you will need to argue, for instance, that these intervals are nested...). Then we can define

$$\int_a^b f(x)\,dx := c = \lim_{N \to \infty} \mathcal{L}_N = \lim_{N \to \infty} \mathcal{R}_N.$$
**Solution.**

Given the results above and the Axiom of Completeness, the following definition is valid:

**Definition 3.2.6.** Let \( f : [a, b] \to \mathbb{R} \) be a monotone increasing function. For each positive integer \( N = 1, 2, \ldots \) let \( L_N \) and \( R_N \) denote the left and right Riemann sums associated to the even subdivision of \([a, b]\) into \( 2^N \) intervals. Then

\[
\int_a^b f(x) \, dx = \lim_{N \to \infty} R_N = \lim_{N \to \infty} L_N.
\]

**Problem 3.2.7.** Let \( f(x) = x^2 \). Compute the integral \( \int_0^1 x^2 \, dx \) using the definition via the following steps:

1. Write down the expression for the right-hand Riemann sum \( R_N \) (as defined above).
2. Compute the limit \( \lim_{N \to \infty} R_N \).

**Solution.**

3.2.2 Integrals of piece-wise monotone functions

We can take integrals of non-monotone functions, as long as we can chop up the domain into segments on which the function is monotone. This is the case for most functions you are used to writing down (e.g. many polynomials are not monotone, but all polynomials are piecewise monotone).

**Definition 3.2.8.** A piecewise monotone function on an interval \([a, b]\) is a function \( f : [a, b] \to \mathbb{R} \) such that there is a (fixed! finite!) subdivision of the interval \([a, b]\) such that on each subinterval the function is monotone.

**Definition 3.2.9.** If \( f : [a, b] \to \mathbb{R} \) is piecewise monotone, then \( \int_a^b f(x) \, dx \) is the sum of the integrals of the monotone pieces of \( f \).

We’ll accept without proof the following properties of the integral:

\[\text{Hint: You may find the identity } \sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6} \text{ helpful.}\]
Proposition 3.2.10. Let \( f \) and \( g \) be piecewise monotone functions defined on \([a, b]\). Then:

1. If \( \forall x \in [a, b] \ f(x) \leq g(x) \) then \( \int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx \).

2. \( \int_a^b (f + g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \).

3. If \( c \in (a, b) \), then \( \int_a^c f(x) \, dx + \int_c^b f(x) \, dx = \int_a^b f(x) \, dx \).

4. If \( k \in \mathbb{R} \), then \( \int_a^b kf(x) \, dx = k \int_a^b f(x) \, dx \).

3.3 The Mean Value Theorem

We begin with the following theorem, which will be used in the proof of the Mean Value Theorem (in fact it is a special case of it).

Theorem 3.3.1. (Rolle’s theorem) Let \( f \) be a function continuous on \([a, b]\) and differentiable on \((a, b)\). Suppose that \( f(a) = f(b) \).

Then there is a point \( c \in (a, b) \) such that \( f'(c) = 0 \).

Proof. Since \( f \) is continuous it attains a maximum and a minimum value on \([a, b]\). If \( f \) attains either of those at a point in \((a, b)\), let \( c \) be that point. (By an argument in the next section, \( f'(c) = 0 \) at such a point.) If both max and min are attained at points in \((a, b)\), pick \( c \) to be one such point.

Otherwise, \( f \) attains both its max and min at the endpoints, and since \( f(a) = f(b) \) the function has to be constant, so any \( c \in (a, b) \) will do.

The Mean Value Theorem is the following:

| Mean Value Theorem. | Let \( f \) be a function continuous on \([a, b]\) and differentiable on \((a, b)\). Then there is a point \( c \in (a, b) \) such that \( f(b) - f(a) \) \( b - a \) = \( f'(c) \). |

Proof. The chord of the graph of \( f \) defined by the points \((a, f(a))\) and \((b, f(b))\) is the graph of the function

\[
\chi(x) = (x - a) \cdot \frac{f(b) - f(a)}{b - a} + f(a).
\]