5. INTEGRALS

In this chapter we will begin with an overview of the problem of finding areas—we will discuss what the term “area” means, and we will outline two approaches to defining and calculating areas. Following this overview, we will discuss the Fundamental Theorem of Calculus, which is the theorem that relates the problems of finding tangent lines and areas, and we will discuss techniques for calculating areas.

5.1 Areas and Distances

The Area Problem: Given a function \( f \) that is continuous and nonnegative on an interval \([a, b]\), find the area between the graph of \( f \) and the interval \([a, b]\) on the \( x \)-axis.
Activity 1: Use rectangles to estimate the area under the parabola $y = x^2$ from 0 to 1.

(a) Using left endpoints

(b) Using right endpoints

Note: You can use the applet in http://mathworld.wolfram.com/RiemannSum.html to verify your results and experiment by increasing the number of rectangles.
We can repeat this procedure with a larger number of strips. The below figure shows what happens when we divide the region \( S \) into eight strips of equal width.

By computing the sum of the areas of the smaller rectangles \( L_8 \) and the sum of the areas of the larger rectangles \( R_8 \), we obtain better lower and upper estimates for \( A \): 

\[
0.2734375 < A < 0.3984375.
\]

The area \( A \) is the number that is smaller than all upper sums and larger than all lower sums.
From the above figures, we see that if we increase the number of rectangles we get better approximation to the area. Here is a table of various values of right and left sums.

<table>
<thead>
<tr>
<th>$n$</th>
<th>4</th>
<th>8</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_n$</td>
<td>0.21875</td>
<td>0.27344</td>
<td>0.28500</td>
<td>0.30875</td>
<td>0.31685</td>
<td>0.32340</td>
<td>0.32835</td>
<td>0.33283</td>
</tr>
<tr>
<td>$R_n$</td>
<td>0.46875</td>
<td>0.39844</td>
<td>0.38500</td>
<td>0.35875</td>
<td>0.35018</td>
<td>0.34340</td>
<td>0.33835</td>
<td>0.33383</td>
</tr>
</tbody>
</table>

We see from the table that both $L_n$ and $R_n$ approach $\frac{1}{3}$. Let us verify this:

**Activity 2:** Show that $\lim_{n \to \infty} R_n = \lim_{n \to \infty} L_n = \frac{1}{3}$. 

![Graph of $y = x^2$ with rectangles approximating the area under the curve]
\[ R_n = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x \]

**Definition:** The area \( A \) of the region \( S \) that lies under the graph of the continuous function \( f \) is the limit of the sum of the areas of approximating rectangles:

\[ A = \lim_{n \to \infty} R_n = \lim_{n \to \infty} [f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x] \]

**Note:** It can be proved that the limit in this definition always exists, since we are assuming that \( f \) is continuous. It can also be shown that we will get the same value if we use the left end points:

\[ A = \lim_{n \to \infty} L_n = \lim_{n \to \infty} [f(x_0)\Delta x + f(x_1)\Delta x \cdots + f(x_{n-1})\Delta x] \]

In fact, instead of using the right or left endpoints of each interval to compute the height of each rectangle, we could choose any number \( x_i^* \) inside the interval \([x_{i-1}, x_i]\). These \( x_i^* \) are called **sample points**.

For example, if we take the **midpoint** of each interval as sample points, the above limit will still give the same value.

We will usually use the **sigma notation** to write sums with many terms more compactly. For example

\[ \sum_{i=1}^{n} f(x_i)\Delta x = f(x_1)\Delta x + f(x_2)\Delta x \cdots + f(x_n)\Delta x \]

**Note:** If you need practice with sigma notation, you can look at the examples and try some of the exercises in Appendix F of the textbook.
Activity 3: Let \( A \) be the area of the region that lies under the graph of \( f(x) = e^{-x} \) between \( x = 0 \) and \( x = 2 \).

a) Using right endpoints, find an expression for \( A \) as a limit. Do not evaluate the limit.

b) Estimate the area by taking sample points to be midpoints and using four subintervals.
The Distance Problem: Find the distance traveled by an object during a certain time period if the velocity of the object is known at all times.

Preview Activity: Suppose that a person is taking a walk along a long straight path and walks at a constant rate of 4 miles per hour.

(a) On the left-hand axes provided below, sketch a graph of the velocity function \( v(t) = 4 \) mph.

The right-hand axes will be used in question (d).

(b) How far did the person travel during the two hours? How is this distance related to the area of a certain region under the graph of \( y = v(t) \)?

(c) Find an algebraic formula, \( s(t) \), for the position of the person at time \( t \), assuming that \( s(0) = 0 \). Explain your thinking.

(d) On the right-hand axes provided in part (a), sketch a labeled graph of the position function \( y = s(t) \).

(e) For what values of \( t \) is the position function \( s \) increasing? Explain why this is the case using relevant information about the velocity function \( v \).
But if the velocity varies, it’s not so easy to find the distance traveled. We investigate the problem in the following activity.

**Activity 4:** Suppose that a person is walking in such a way that her velocity varies slightly according to the information given in the table and the graph below.

<table>
<thead>
<tr>
<th>t</th>
<th>0.00</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>1.25</th>
<th>1.50</th>
<th>1.75</th>
<th>2.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>v(t)</td>
<td>1.5000</td>
<td>1.7891</td>
<td>1.9375</td>
<td>1.9922</td>
<td>2.0000</td>
<td>2.0078</td>
<td>2.0625</td>
<td>2.2109</td>
<td>2.5000</td>
</tr>
</tbody>
</table>

(a) Using the grid, graph, and given data appropriately, estimate the distance traveled by the walker during the two hour interval from $t = 0$ to $t = 2$. You should use time intervals of width $\Delta t = 0.5$, choosing a way to use the function consistently to determine the height of each rectangle in order to approximate distance traveled.
(b) How could you get a better approximation of the distance traveled on \([0, 2]\)? Explain, and then find this new estimate.

(c) Now suppose that you know that \(v\) is given by \(v(t) = \frac{1}{2}(t^3 - 3t^2 + 3t + 3)\). Remember that \(v\) is the derivative of the walker’s position function \(s\). Find a formula for \(s\) so that \(s' = v\).

(d) Based on your work in (c), what is the value of \(s(2) - s(0)\)? What is the meaning of this quantity?
Activity 5: Suppose that an object moving along a straight line path has its velocity $v$ (in meters per second) at time $t$ (in seconds) given by the piecewise linear function whose graph is given below. We view movement to the right as being in the positive direction (with positive velocity), while movement to the left is in the negative direction. Suppose further that the object’s initial position at time $t = 0$ is $s(0) = 1$.

(a) Determine the total distance traveled and the total change in position on the time interval $0 \leq t \leq 2$. What is the object’s position at $t = 2$?

(b) On what time intervals is the moving object’s position function increasing? Why? On what intervals is the object’s position function decreasing? Why?
(c) What is the object’s position at $t = 8$? How many total meters has it traveled to get to this point (including distance in both directions)? Is this different from the object’s total change in position on $t = 0$ to $t = 8$?

(d) Find the exact position of the object at $t = 1, 2, 3, ..., 8$ and use this data to sketch an accurate graph of $y = s(t)$. How can you use the provided information about $y = v(t)$ to determine the concavity of $s$ on each relevant interval?
5.2 The Definite Integral

In the previous section we saw that a limit of the form

\[ \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i^*) \Delta x = \lim_{n \to \infty} \left[ f(x_1^*) \Delta x + f(x_2^*) \Delta x + \ldots + f(x_n^*) \Delta x \right] \]

arises when we compute an area. We have a special terminology for such limits:

**Definition:** If \( f \) is a function defined for \( a \leq x \leq b \), we divide the interval \([a,b]\) into \( n \) subintervals of equal width \( \Delta x = (b-a)/n \). We let \( x_0 (=a), x_1, \ldots, x_n (=b) \) be the endpoints of these sub-intervals and we let \( x_1^*, x_2^*, \ldots, x_n^* \) be any sample points in these intervals. Then the definite integral of \( f \) from \( a \) to \( b \) is

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

provided that this limit exists. If it does exists, we say that \( f \) is integrable on \([a,b]\).

**Note 1:** The symbol \( \int \) is called an integral sign \(^1\). In the notation \( \int_a^b f(x) \, dx \), \( f(x) \) is called the integrand and and \( a \) and \( b \) are called the limits of integration; \( a \) is the lower limit and \( b \) is the upper limit. For now, the symbol \( dx \) has no meaning by itself; \( \int_a^b f(x) \, dx \) is all one symbol. The \( dx \) simply indicates that the independent variable is \( x \). The procedure of calculating an integral is called integration.

**Note 2:** The sum \( \sum_{i=1}^{n} f(x_i^*) \Delta x \) that occurs in the definition is called a Riemann sum.

---

\(^1\)This notation was devised by Leibniz. It is an elongated \( S \) and was chosen because an integral is a limit of sums. This notation is so useful and so powerful that its development by Leibniz must be regarded as a major milestone in the history of mathematics and science.
We know from previous section that if \( f \) is positive then the Riemann sum can be interpreted as a sum of areas of approximating rectangles (see Figure 1). The integral however corresponds to the area under the curve \( y = f(x) \) from \( a \) to \( b \). (See Figure 2).

If the function \( f \) takes both positive and negative values on \([a, b]\) then the integral corresponds to the net area. That is, a difference of areas:

\[
\int_a^b f(x) \, dx = A_1 - A_2
\]

where \( A_1 \) is the area of the region above the \( x \)-axis and below the graph of \( f \), and \( A_2 \) is the area of the region below the \( x \)-axis and above the graph of \( f \).

**Theorem:** If \( f \) is continuous on \([a, b]\), or if \( f \) has only finitely many jumps on \([a, b]\), then \( f \) is integrable on \([a, b]\) (i.e. the definite integral \( \int_a^b f(x) \, dx \) exists.)
If $f$ is integrable on $[a, b]$ then the definition of the integral gives the same value no matter how we choose the sample points $x_i^*$. To simplify the situation we often take the sample points to be the right endpoints. Then $x_i^* = x_i$ and the definition of the integral simplifies as follows:

**Theorem:** If $f$ is integrable on $[a, b]$, then

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

where $\Delta x = \frac{b-a}{n}$ and $x_i = a + i\Delta x$.

**Activity 1:** Express $\lim_{n \to \infty} \sum_{i=1}^{n} (x_i^2 + 3\cos x_i)\Delta x$ as an integral on the interval $[0, 1]$.

**Activity 2:** a) Evaluate the Riemann sum for $f(x) = x^2 - x$ by taking the sample points to be right endpoints and $a = 0$, $b = 2$ and $n = 4$. 
b) Evaluate \( \int_{0}^{2} (x^2 - x) \, dx \)