Last time, we saw how certain physical questions naturally give rise to Riemann sums, and hence to definite integrals. Specifically, if $F(x)$ is the force exerted upon an object at point $x$ over an interval $[a, b]$, then

$$W \approx \sum_{i=1}^{N} F(x_i^*) \Delta x, \quad N \to \infty \quad \int_{a}^{b} F(x) \, dx.$$ 

Therefore, the difficulty of solving a work problem lies in setting up the Riemann sum.

Today, we will look at some related constructions. First of all, we define the **average value** of a function $f(x)$ over an interval $[a, b]$. If we wish to take the arithmetic mean of a finite list of certain values of $f(x)$, say

$$\{ f(x_1^*), f(x_2^*), \ldots, f(x_n^*) \},$$

our formula is

$$f_{av} = \frac{1}{n} \left( f(x_1^*) + f(x_2^*) + \cdots + f(x_n^*) \right) = \frac{1}{(b-a) \Delta x} \sum_{i=1}^{n} f(x_i^*)$$

$$= \frac{1}{b-a} \sum_{i=1}^{n} f(x_i^*) \Delta x \quad \overset{N \to \infty}{\longrightarrow} \quad \frac{1}{b-a} \int_{a}^{b} f(x) \, dx.$$ 

That is: the average value of $f(x)$ on $[a, b]$ is defined to be

$$f_{av} = \frac{1}{b-a} \int_{a}^{b} f(x) \, dx$$
Example 46: Find the average value of $\sin x$ on the interval $[0, \pi]$.

Solution: We use the definition to compute

$$f_{av} = \frac{1}{\pi - 0} \int_0^\pi \sin x \, dx = \frac{1}{\pi} \int_0^\pi \sin x \, dx = \frac{1}{\pi} \left[ -\cos x \right]_0^\pi = \frac{2}{\pi},$$

and we are done.

\[\Box\]

Theorem 5.1 (The mean value theorem for integrals). Suppose $f(x)$ is continuous on $[a, b]$. Then there exists $c \in (a, b)$ such that

$$f(c) = \frac{1}{b - a} \int_a^b f(x) \, dx.$$

That is: $f(x)$ always attains its average value.

Proof. Let $F(x) = \int_a^x f(t) \, dt$. By the first fundamental theorem of calculus, we know that $F(x)$ is differentiable on $(a, b)$ and is continuous on $[a, b]$. Therefore, by the mean value theorem (for derivatives), there exists $c \in (a, b)$ such that

$$F'(c) = \frac{F(b) - F(a)}{b - a} = \frac{1}{b - a} \left( \int_a^b f(t) \, dt - \int_a^a f(t) \, dt \right) = \frac{1}{b - a} \int_a^b f(x) \, dx.$$

The same theorem implies that $F'(c) = f(c)$, and we are done. \[\Box\]

In fact, in our example, there are two points in the interval in which we have

$$\sin c_1 = \sin c_2 = \frac{2}{\pi} = f_{av}$$

(see left). The theorem doesn’t tell us this, however; it can’t even tell us where these points are.

Related to the concept of the average value is that of the centre of mass, which we now introduce.

It is a fact that, given two point masses $m_1$ and $m_2$ on a lever at respective distances $d_1$ and $d_2$ from the fulcrum, then the system is in equilibrium/fulcrum if and only if $m_1d_1 = m_2d_2$ (see the picture on the next page). In this case, the fulcrum is the centre of mass.
If we draw this picture on the $x$-axis with the centre of mass at $x = \bar{x}$, then we have the equation

$$m_1 \left( \bar{x} - x_1 \right) = m_2 \left( x_2 - \bar{x} \right)$$

$$\iff m_1 x_1 + m_2 x_2 = \bar{x}(m_1 + m_2),$$

where mass $m_i$ is at position $x_i$ (for $i = 1, 2$). Thus: the centre of mass of two objects is defined

$$\bar{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}.$$

By analogy, we define the centre of mass of any finite number of masses $m_1, \ldots, m_n$ at respective positions $x_1, \ldots, x_n$ to be

$$\bar{x} = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i}.$$

The quantity $\sum_{i=1}^n m_i x_i$ in the numerator is the sum of the individual moments $m_i x_i$ (with respect to the origin) of the masses $m_i$; their sum is called the moment of the system (about the origin).

We now generalize again, to two dimensions. This time, we have masses $m_1, \ldots, m_n$ at respective co-ordinates $(x_1, y_1), \ldots, (x_n, y_n)$, and we define two quantities: the moment of the system about the $y$-axis, defined

$$M_y = \sum_{i=1}^n m_i x_i,$$

and the moment of the system about the $x$-axis, defined

$$M_x = \sum_{i=1}^n m_i y_i.$$

Notice the subscripts – these are not typos!

We then define the centre of mass $(\bar{x}, \bar{y})$ to be

$$\bar{x} = \frac{M_y}{m}, \quad \bar{y} = \frac{M_x}{m},$$
where

\[ m = \sum_{i=1}^{n} m_i \]

is the total mass of the system.

**Example 47:** Find the centre of mass of the system consisting of three objects at positions \((-1, 1), (2, -1), \text{ and } (3, 2)\), of respective masses 3, 4, and 8.

**Solution:** Putting these values in terms of our variables gives us

\[ m_1 = 3, \quad (x_1, y_1) = (-1, 1); \quad m_2 = 3, \quad (x_2, y_2) = (2, -1); \quad m_3 = 8, \quad (x_3, y_3) = (3, 2). \]

Thus

\[ m = m_1 + m_2 + m_3 = 3 + 4 + 8 = 15, \]

and similarly

\[ M_y = m_1 x_1 + m_2 x_2 + m_3 x_3 = (3)(-1) + (4)(2) + (8)(3) = 29, \]

and

\[ M_y = m_1 y_1 + m_2 y_2 + m_3 y_3 = (3)(1) + (4)(-1) + (8)(2) = 15. \]

Hence

\[ \bar{x} = \frac{M_y}{m} = \frac{29}{15}, \quad \bar{y} = \frac{M_x}{m} = \frac{15}{15}, \]

and it follows that the centre of mass is at \((\bar{x}, \bar{y}) = \left(\frac{29}{15}, 1\right)\).

These formulas naturally give rise to Riemann sums: if \(R\) is a region in the plane bounded below by \(y = 0\), above by \(y = f(x)\), to the left by \(x = a\), and to the right by \(x = b\), and which moreover has uniform density \(\rho = \rho \text{ kg/m}^2\), we have for any collection of sample points \((x^*_i, f(x^*_i))\) that

\[
\bar{x} = \frac{M_y}{m} \approx \frac{1}{\rho \int_a^b f(x) \, dx} \sum_{i=1}^{N} \left[ \rho f(x^*_i) \Delta x_i \right] x^*_i \xrightarrow{N \to \infty} \frac{1}{A} \int_a^b x f(x) \, dx,
\]

where

\[ A = \int_a^b f(x) \, dx \]

is the area under the graph of \(f(x)\) between \(x = a\) and \(b\). Similarly,

\[
\bar{y} = \frac{M_x}{m} \approx \frac{1}{\rho \int_a^b f(x) \, dx} \sum_{i=1}^{N} \left[ \rho f(x^*_i) \Delta x_i \right] \left( \frac{1}{2} f(x^*_i) \right) \xrightarrow{N \to \infty} \frac{1}{A} \int_a^b \frac{f(x)^2}{2} \, dx;
\]
the factor \( \frac{1}{2} \) appears because the centre of mass of a uniform rectangle is “halfway up.”

A region \( R \) like the one I have just described is called a lamina, and its centre of mass – namely,

\[ \bar{x}, \bar{y} = \left( \frac{1}{A} \int_a^b x f(x) \, dx, \frac{1}{A} \int_a^b \frac{1}{2} [f(x)]^2 \, dx \right) \]

is called its centroid.

**Example 48:** Find the centroid of a semicircular disc of radius \( R \) and uniform density \( \rho \).

**Solution:** We consider the function \( f(x) = \sqrt{R^2 - x^2} \), whose graph is the desired lamina. Since our lamina is uniformly dense, we have by our formulas above that

\[ \bar{x} = \frac{1}{A} \int_{-R}^{R} x \sqrt{R^2 - x^2} \, dx \]

and

\[ \bar{y} = \frac{1}{A} \int_{-R}^{R} \frac{1}{2} \left[ \sqrt{R^2 - x^2} \right]^2 \, dx. \]

We know a priori that \( A = \frac{1}{2} \pi R^2 \); moreover, by symmetry (or using the fact that the integrand is odd), we know that \( \bar{x} = 0 \), and it remains only to calculate \( \bar{y} \). We have

\[ \bar{y} = \frac{2}{\pi R^2} \int_{-R}^{R} \frac{1}{2} (R^2 - x^2) \, dx = \frac{1}{\pi R^2} \int_{-R}^{R} (R^2 - x^2) \, dx. \]

We can now use the fact that this integrand is even to write

\[ \bar{y} = \frac{2}{\pi R^2} \int_{0}^{R} (R^2 - x^2) \, dx = \frac{2}{\pi R^2} \left[ R^2 x - \frac{1}{3} x^3 \right]_{x=0}^{R} = \frac{4R^3}{3\pi}, \]

from which we see that our centroid is at \( (0, \frac{4R^3}{3\pi}) \), regardless of the value of \( \rho \).

The other situation we consider is when the region \( R \) in question is bounded above and below by curves. In this case, the formula is mostly the same, but slightly different.
Let $R$ be the region bounded above by $y = f(x)$ and below by $y = g(x)$, to the left by the line $x = a$ and to the right by the line $x = b$. Then the centroid of $R$ is given by the formula

$$(\bar{x}, \bar{y}) = \left( \frac{1}{A} \int_{a}^{b} x[f(x) - g(x)] \, dx, \frac{1}{A} \int_{a}^{b} \frac{1}{2} [(f(x))^2 - (g(x))^2] \, dx \right)$$

where $A$ is as before (we omit the derivation of these formulas here).

**Example 49:** Find the centre of mass of the region bounded by the line $y = x$ and the parabola $y = x^2$.

**Solution:** First, we find the points of intersection to obtain the limits of integration; in example 42 in the last lecture, we found these points to be at $x = 0$ and $x = 1$. When $x$ is between 0 and 1, then it is a fact that $x^2 \leq x$ (check this!), and we can now use the formulas we just introduced: we have

$$A = \int_{0}^{1} (x - x^2) \, dx = \left[ \frac{1}{2}x^2 - \frac{1}{3}x^3 \right]_{x=0}^{1} = \frac{1}{6},$$

and thus

$$\bar{x} = 6 \int_{0}^{1} x(x - x^2) \, dx = 6 \int_{0}^{1} x^2 - x^3 \, dx = 6 \left[ \frac{1}{3}x^3 - \frac{1}{4}x^4 \right]_{x=0}^{1} = 6 \left( \frac{1}{12} \right) = \frac{1}{2},$$

$$\bar{y} = 6 \int_{0}^{1} \frac{1}{2} ((x)^2 - (x^2)^2) \, dx = 3 \left[ \frac{1}{3}x^3 - \frac{1}{5}x^5 \right]_{x=0}^{1} = \frac{2}{5}.$$

Therefore, the centroid of our region is $(\bar{x}, \bar{y}) = \left( \frac{1}{2}, \frac{2}{5} \right)$. ♣

We leave for now the discussion of “practical calculations” and “real-world scenarios” to introduce the most powerful tool we have in our integral calculus utility belt: integration by parts.

### 6 Integration by parts and trigonometric integrals

Recall from the differential calculus the **product rule**, which states that

$$\frac{d}{dx} (f(x)g(x)) = f'(x)g(x) + f(x)g'(x).$$
We can rewrite this equation as
\[ f(x) \frac{d}{dx} g(x) = \frac{d}{dx} (f(x)g(x)) - \left( \frac{d}{dx} f(x) \right) g(x), \]
and by again forgetting that \( \frac{d}{dx} \) is DEFINITELY NOT A FRACTION, we “clear denominators” to get
\[ f(x) \frac{d}{dx} g(x) = \frac{d}{dx} (f(x)g(x)) - g(x) \frac{d}{dx} f(x). \]
More compactly, with \( u = f(x), v = g(x) \), we have
\[ u \frac{dv}{dx} = \frac{d(uv)}{dx} - v \frac{du}{dx}; \]
finally, we integrate, and we obtain the integration by parts formula:
\[ \int u \frac{dv}{dx} = uv - \int v \frac{du}{dx}. \]
Much like \( u \)-substitution, integration by parts requires critical thought only for one single step, and is otherwise totally mechanical.

**Example 50:** Find \( \int x \sin x \, dx \).

**Solution:** We have to make a choice of what is to be \( u \). Just like with \( u \)-substitution, we have to guess, and this is the only difficult part. Once again, once we choose \( u \), the rest of the integrand must be \( dv \) (it’s been decided for us!) So, if we choose \( u = x \), then \( dv \) must be \( \sin x \, dx \).

**Note the inclusion of \( dx \) in the expression for \( dv \)!**

Because \( u = x \), we know that \( \frac{du}{dx} = 1 \), and so \( du = dx \). Similarly,
\[ v = \int dv = \int \sin x \, dx = - \cos x; \]
we will save our added constants until the last step. Thus, by our formula,
\[ \int x \sin x \, dx \bigg|_{u \frac{dv}{dx}}^{uv} = x \cos x \bigg|_u^v - \int \cos x \, dx \bigg|_{v \frac{du}{dx}}^{uv} = -x \cos x + \sin x + C. \]
Because this is our final answer, we now include the \( +C \) term.

Some remarks on this example:
1. As with $u$-substitution, making the right choice for $u$ is crucial to make the technique useful.

2. Although we “pick up” an added constant when we integrate $dv$, we will never include it in the formula for $v$ (that is, we will take the added constant to be zero.) We will, however, still need the added constant when we integrate for the last time.

3. For a “geometric” interpretation of integration by parts, see https://en.wikipedia.org/wiki/Integration_by_parts#Visualization.