Math 231E, Lecture 13.
Area & Riemann Sums

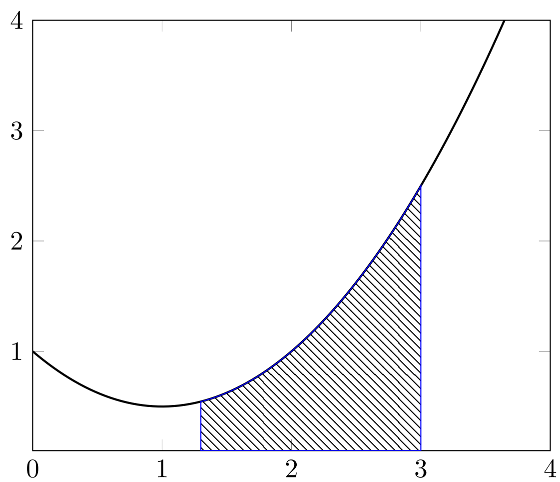
1 Motivation for Integrals

Question. What is an integral, and why do we care?

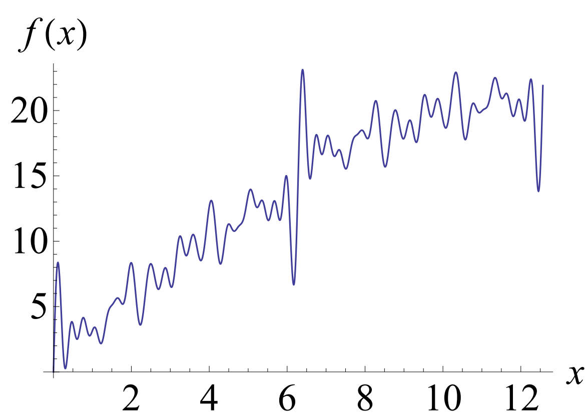
Answer 1. A tool to compute a complicated expression made up of smaller pieces.

Example 1.1. Some things that we use integrals to compute:

1. Let us say that we are given a function $y = f(x)$, and we want to know the area under the curve:



2. Let us say that we are given some crazy curve:



How long is the curve?

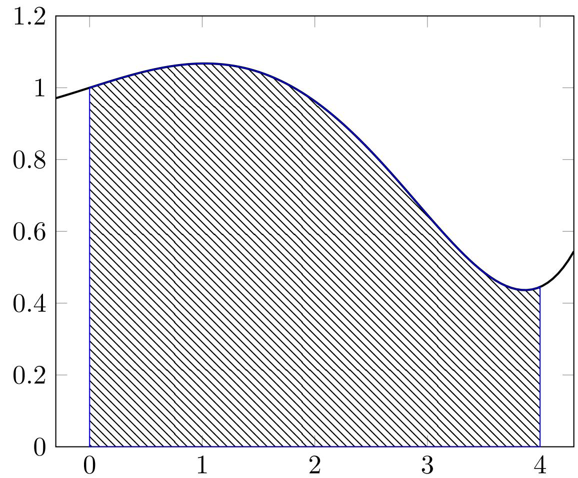
3. Given the dimensions of a tub filled with water, how much water does it hold? How much work does it take to move all of that water out?

4. How much energy does it take to leave a planet's gravity well? What's the escape velocity of Earth?

> The main idea behind all integration comes in three steps:
>
> 1. Break an object into n pieces.
> 2. Approximate those pieces by something we can compute, buy exploiting some kind of "smallness"
> 3. Add those pieces together. Take limit as $n \to \infty$.

2 Area under a curve

Let us say that we want to find the area under a curve:



Specifically, we are given some function $f(x)$, and we want the area under the curve from $x = 0$ to $x = 4$.

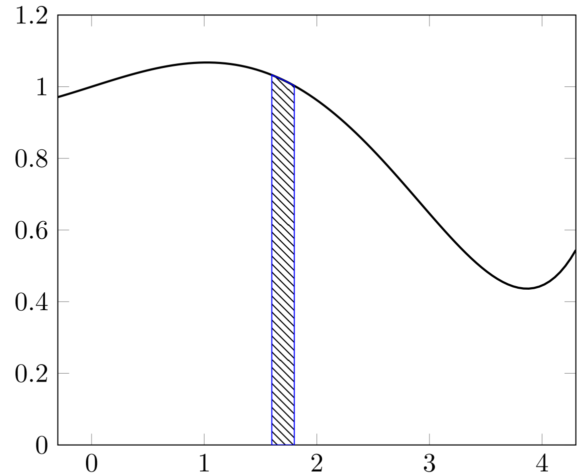
Step 1. Break it into n pieces. We will first break this area up into n vertical strips. If there are n of them, then they have to each be $\Delta x = 4/n$ wide, so we must consider the domains

$$[(k-1)\Delta x, k\Delta x], \quad k = 1, 2, 3, \ldots, n.$$

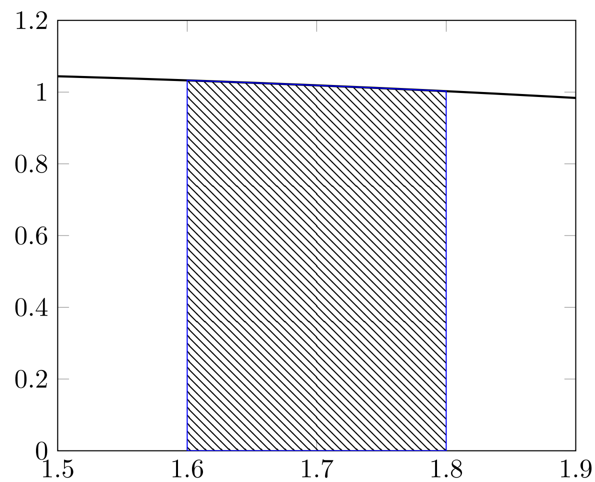
Notice that the first subdomain is $k = 0$ which is $[0, \Delta x] = [0, 4/n]$ and the last subdomain is

$$[(n-1)\Delta x, n\Delta x] = \left[(n-1)\frac{4}{n}, n\frac{4}{n}\right] = \left[4 - \frac{4}{n}, 4\right].$$

For example, if we choose $n = 20$, then one of these domains is $[8/5, 9/5)$, and we need to compute the area under the curve on this subdomain:

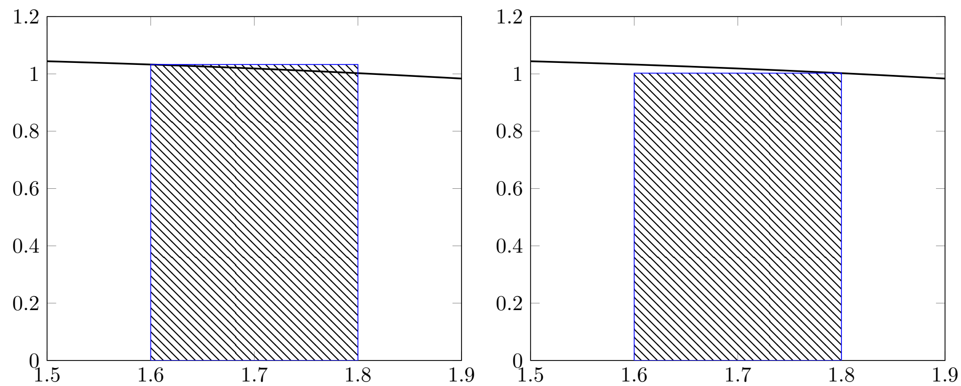


We can blow this picture up and concentrate on the region of interest:



Step 2. Approximate this area, exploiting a "smallness". Notice that if we can compute the area under the curve for the subdomain $[8/5, 9/5] = [1.6, 1.8]$ (and all the other areas), then we can get the total area by adding all of these up. However, *this problem isn't any simpler than the original problem.* We make it simpler by approximating.

Let us replace the area on the subdomain we considered above with a rectangle. The question arises, of course: which rectangle? We have several choices.



- In the left figure, we decided to use the height of the function on the left-hand end of the domain to define the rectangle, i.e. the height of the rectange is defined to be $f(1.6)$. This will, not surprisingly, be called the **left-hand rule**.

- In the right figure, we decided to use the height of the function on the right-hand end of the domain to define the rectangle, i.e. the height of the rectange is defined to be $f(1.8)$. This will, not surprisingly, be called the **right-hand rule**.

  There are other possibilities:

- We could choose the midpoint of the interval, and define the height to be $f(1.7)$. This is the **midpoint rule**.

- We could average the two values at the left- and right-hand endpoints, i.e. choose the height to be

$$\frac{1}{2}\left(f(1.6) + f(1.8)\right).$$

  This is called the **averaging rule** or the **trapezoid rule** (the motivation for the first name should be obvious, the second is also once one sees the correct picture. More on this below.)

- We can choose the largest value the function ever attains in the interval, i.e. we choose the height to be $\max_{x\in[1.6,1.8]} f(x)$. This is called the **upper rule**.

- We can choose the smallest value the function ever attains in the interval, i.e. we choose the height to be $\min_{x\in[1.6,1.8]} f(x)$. This is called the **lower rule**.

All of these rules will give slightly different answers. But they give an answer that we can compute. More generally, when computing the area associated to the domain $[k\Delta x, (k+1)\Delta x]$, we choose $x_k^* \in [k\Delta x, (k+1)\Delta x]$ by *some* rule, and then approximate the area by

$$\text{Area} = \text{width} \times \text{height} = \Delta x \times f(x_k^*).$$

Step 3. Add together; take limit.
The area of the kth rectange is $f(x_k^*)\Delta x$. Thus our approximation to the total area is

$$\sum_{k=1}^{n} f(x_k^*)\Delta x = f(x_1^*)\Delta x + f(x_2^*)\Delta x + \cdots + f(x_n^*)\Delta x.$$

Of course, this is an incorrect number! We have replaced each real small area with a rectangular area.

But notice that as we take more and more subdomains, the approximation error becomes smaller and smaller. So there is a hope that if we take the limit as the number of subdomains goes to infinity.

Thus we define

$$\int_0^4 f(x)\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^*)\Delta x,$$

if it exists.

There is another subtle problem to worry about. Notice that before we make some set of choices for the x_k^* to get the rectangle heights. Clearly, this will affect the computation, but if we are doing the right thing, then maybe it shouldn't matter?

One important thing to realize is this: For any Riemann sum (for a fixed value of n) the upper sum always gives the largest approximation to the integral and the lower sum always gives the smallest approximation to the integral. This is useful since it tells us that, in some sense, the upper sum and the lower sum are all that matter. Since all other Riemann sums lie in between those two if they tend to the same limit then (by the squeeze theorem) all of the other Riemann sums tend to the same limit as well.

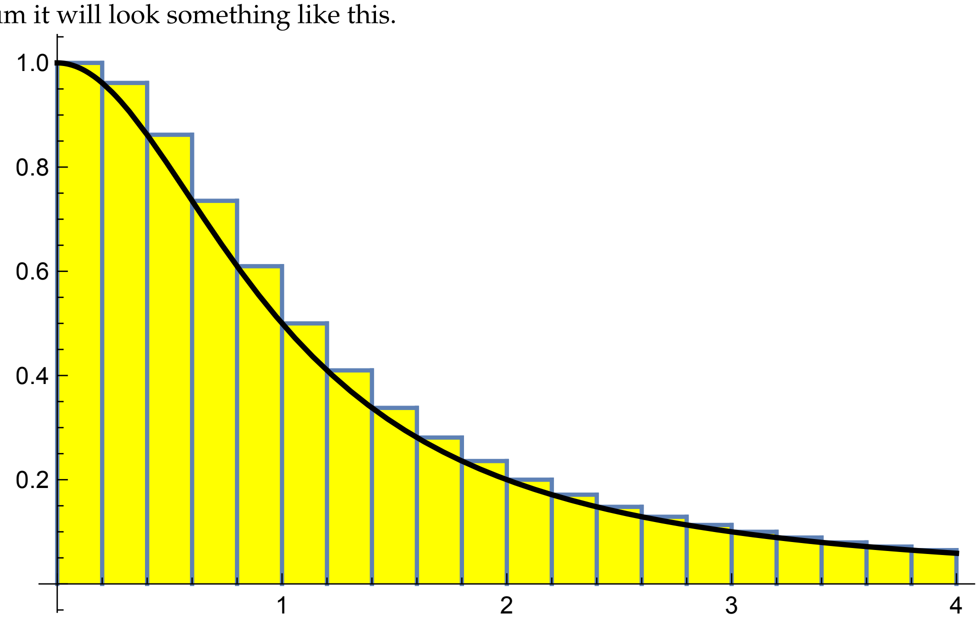
So let us make a definition:

Definition 2.1. A function is called **integrable** on the domain $[a, b]$ if the upper Riemann sum and the lower Riemann sum tend to the same limit as $n \to \infty$

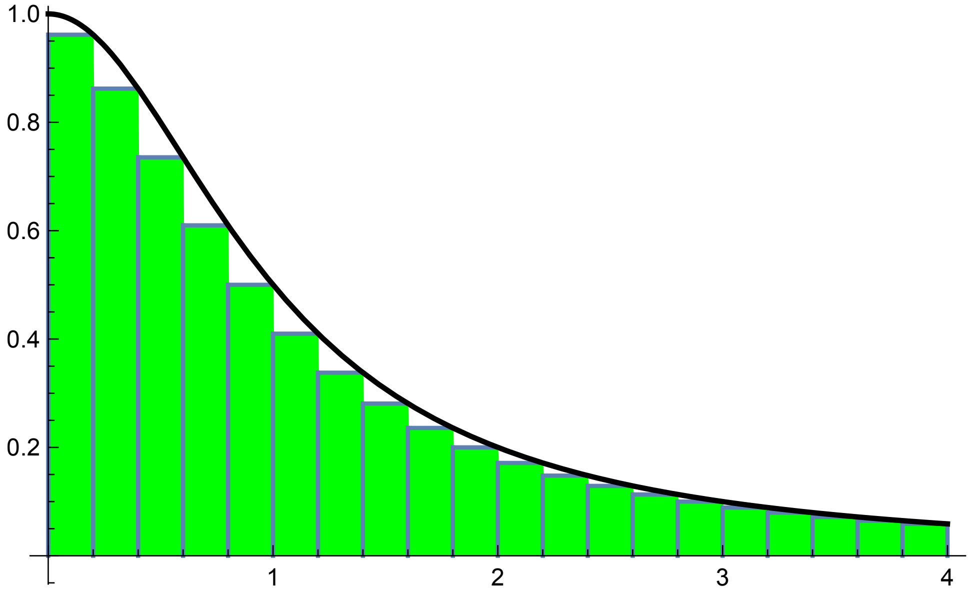
I will prove a first theorem and then state a somewhat stronger theorem.

Theorem 2.2. Every monotone function is integrable.

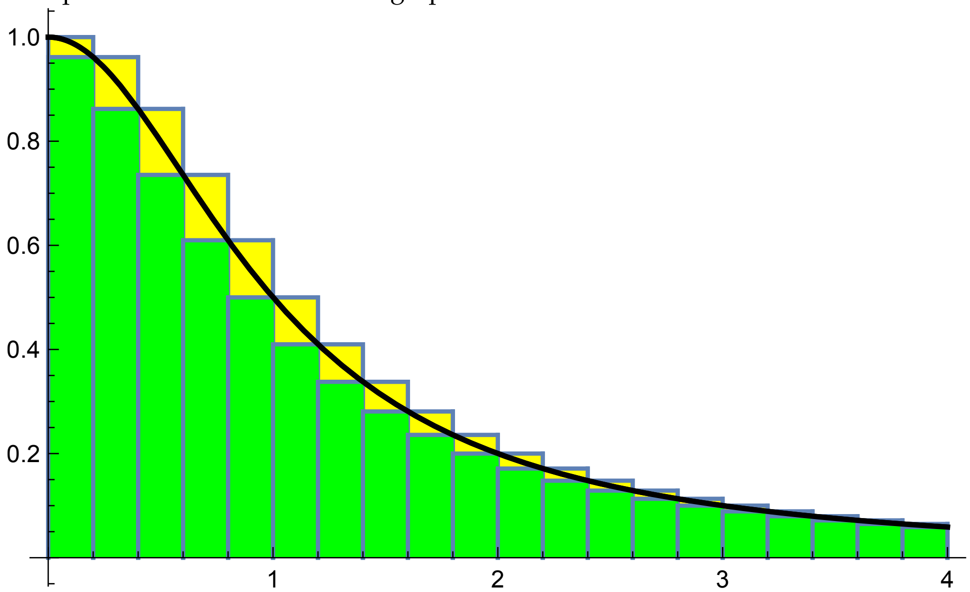
Most of the proof is by pictures. Lets take a monotone decreasing function. If I take an upper Riemann sum it will look something like this.



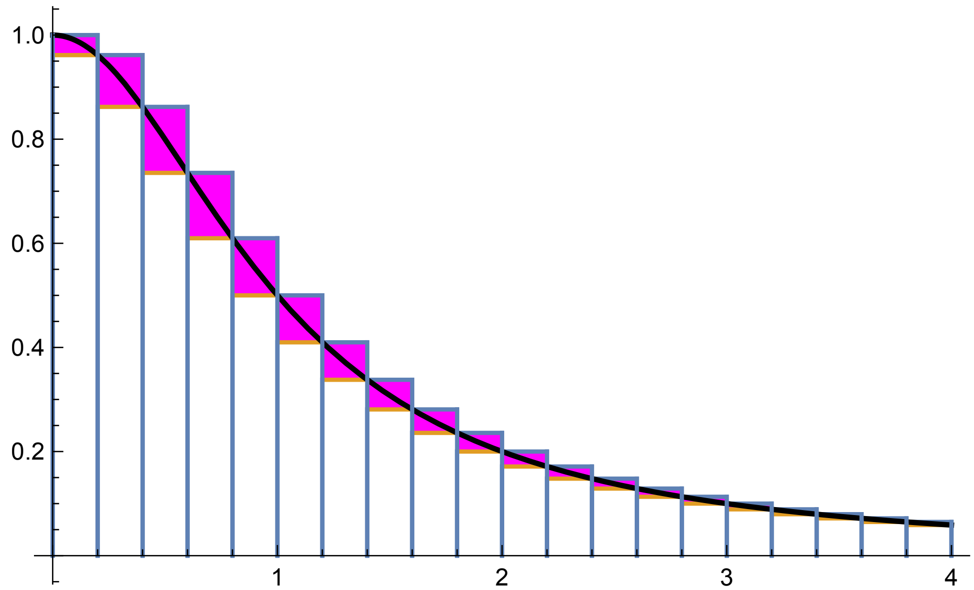
The lower Riemann sum will look similar but will lie entirely below the curve, like this:



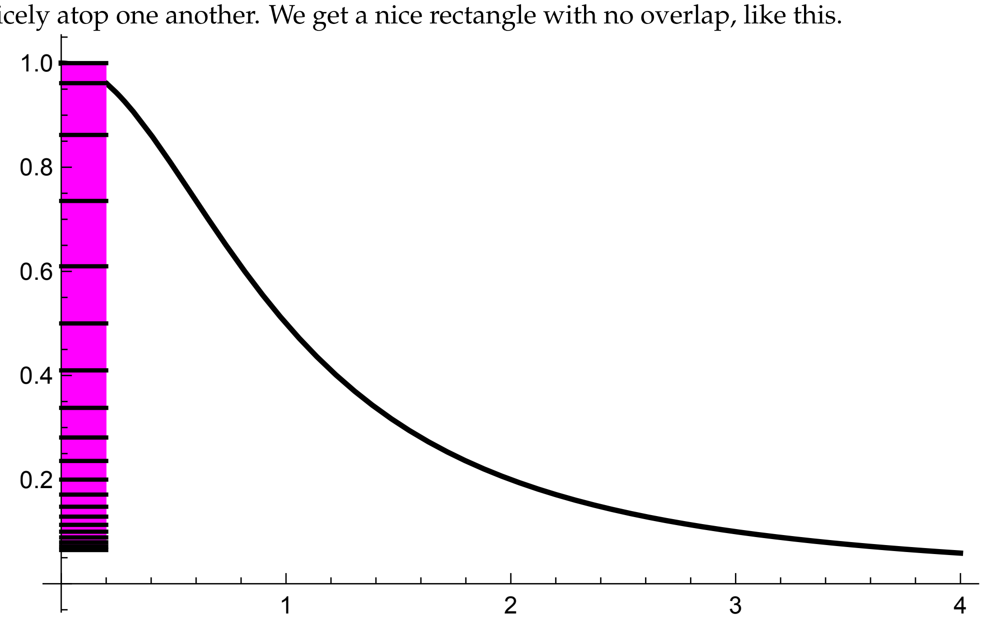
Lets plot them both on the same graph.



Think about what the difference between the upper and lower sums will be: it will be the area of a bunch of little rectangles. Lets plot just the difference between the two, in a new color.



Now all of those little rectangles have the same width, so think about sliding them over so that they sit nicely atop one another. We get a nice rectangle with no overlap, like this.



The area of this rectangle is the width times the height, or $(f_{max} - f_{min})\Delta x$. But $\Delta x \propto \frac{1}{n}$ and so tends to 0 as $n \to \infty$. Thus we have that, for monotone functions $\lim_{n\to\infty} U_n - L_n = 0$, where U_n is the upper sum and L_n is the lower sum. By the Squeeze theorem this means that for monotone functions all Riemann sums converge to the same limit.

This covers a lot of the common functions: exponential, log, powers. Polynomials and sines and cosines are not monotone but they can be broken up into monotone pieces, so you can just apply the monotone argument to each piece. A more general theorem is

Theorem 2.3. Every continuous function, or even piecewise continuous function, is integrable.

There are non-integrable functions but they are pretty weird – there can be no interval on which they are monotone.

This theorem is very useful, since it tells us that we can use whatever rectangle rule we find convenient for a given problem, and this won't mess up the answer.

Let us work out an example. We want to compute

$$\int_0^1 x^2\,dx.$$

We first do the case where $n = 4$ to warm up, then compute the general n formula and then the limit. If $n = 4$, $\Delta x = 1/4$, and our four subdomains are

$$[0, 1/4], \quad [1/4, 1/2], \quad [1/2, 3/4], \quad [3/4, 1].$$

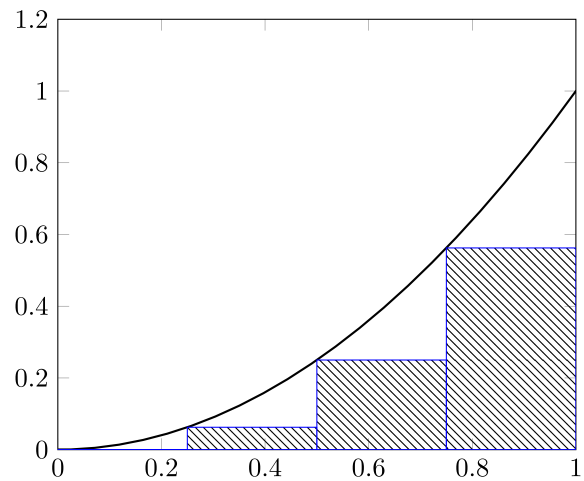
The **left-hand rule** says that the four rectangle heights we will use are

$$f(0) = 0, \quad f(1/4) = 1/16, \quad f(1/2) = 1/4, \quad f(3/4) = 9/16.$$

So our sum is

$$0 \cdot \frac{1}{4} + \frac{1}{16} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{9}{16} \cdot \frac{1}{4} = \frac{1}{4}\left(0 + \frac{1}{16} + \frac{1}{4} + \frac{9}{16}\right) = \frac{1}{4} \cdot \frac{14}{16} = \frac{7}{32}.$$

See picture:



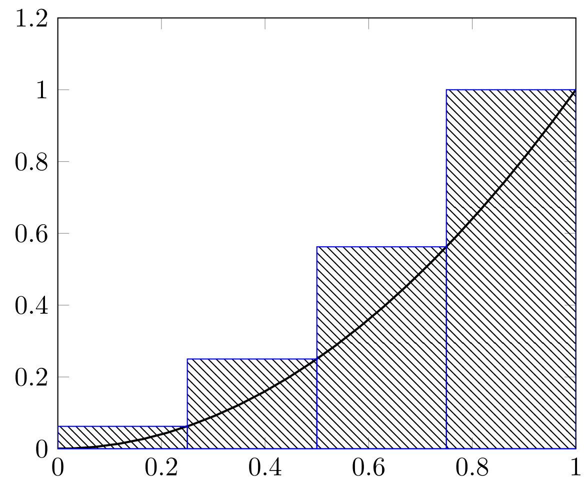
The **right-hand rule** says that the four rectangle heights we will use are

$$f(1/4) = 1/16, \quad f(1/2) = 1/4, \quad f(3/4) = 9/16, \quad f(1) = 1.$$

So our sum is

$$\frac{1}{16} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{9}{16} \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{1}{4}\left(\frac{1}{16} + \frac{1}{4} + \frac{9}{16} + 1\right) = \frac{1}{4} \cdot \frac{30}{16} = \frac{15}{32}.$$

See picture:



Note that although neither of these calculations are exact, we have determined upper and lower bounds for the integral: clearly the first number is smaller, and the second number larger, than the area under the curve. So we have established that

$$\frac{7}{32} \leq \int_0^1 x^2\,dx \leq \frac{15}{32}.$$

Let us now consider the general n formula. In this case, we will have the subdomains $[(k-1)\Delta x, k\Delta x]$, with $k = 1, 2, \ldots, n$. In the left-hand rule, we will evaluate on the left-hand endpoint, so we will have

$$\sum_{k=1}^{n} f((k-1)\Delta x)\Delta x = \sum_{k=1}^{n} ((k-1)\Delta x)^2 \Delta x = \sum_{k=1}^{n} \frac{(k-1)^2}{n^3} = \frac{1}{n^3}\sum_{k=1}^{n}(k-1)^2.$$

The right-hand rule will have us using the right-hand endpoint, so we have

$$\sum_{k=1}^{n} f(k\Delta x)\Delta x = \sum_{k=1}^{n} (k\Delta x)^2 \Delta x = \sum_{k=1}^{n} \frac{k^2}{n^3} = \frac{1}{n^3}\sum_{k=1}^{n} k^2.$$

Now, let us use the fact[1] that

$$\boxed{\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.}$$

Then the right-hand rule becomes, for any n,

$$\frac{1}{n^3}\frac{n(n+1)(2n+1)}{6}.$$

Now take the limit as $n \to \infty$, and we have

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n^3}\frac{n(n+1)(2n+1)}{6} &= \frac{1}{6}\lim_{n\to\infty}\frac{n(n+1)(2n+1)}{n^3} \\ &= \frac{1}{6}\lim_{n\to\infty}\frac{n}{n}\lim_{n\to\infty}\frac{n+1}{n}\lim_{n\to\infty}\frac{2n+1}{n} = \frac{1}{6}\cdot 1\cdot 1\cdot 2 = \boxed{\frac{1}{3}}. \end{aligned}$$

[1]If this fact seems a bit of a *deus ex machina* to you at this point, then rest assured, firstly, that yes it is, and secondly, that you'll know how to derive these sorts of rules yourself before too long.

For the left-hand rule, we need to work out a few substeps. If we accept the formula given above works for any n, then it will also work for $n-1$, so we have

$$\sum_{k=1}^{n-1} k^2 = \frac{(n-1)((n-1)+1)(2(n-1)+1)}{6} = \frac{(n-1)n(2n-1)}{6}.$$

But then we also have

$$\sum_{k=1}^{n}(k-1)^2 = \sum_{j=0}^{n-1} j^2 = 0 + \sum_{j=1}^{n-1} j^2 = \frac{(n-1)n(2n-1)}{6}.$$

Thus the left-hand rule gives

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n^3}\frac{(n-1)n(2n-1)}{6} &= \frac{1}{6}\lim_{n\to\infty}\frac{(n-1)n(2n-1)}{n^3} \\ &= \frac{1}{6}\lim_{n\to\infty}\frac{n-1}{n}\lim_{n\to\infty}\frac{n}{n}\lim_{n\to\infty}\frac{2n-1}{n} = \frac{1}{6}\cdot 1\cdot 1\cdot 2 = \boxed{\frac{1}{3}}.\end{aligned}$$