Review. A function $y = f(x)$ has derivative $\frac{dy}{dx}|_{x=c} = f'(c)$ with four levels of meaning:

- Physical: If y is a quantity depending on x, the derivative $\frac{dy}{dx}|_{x=c}$ is the rate of change of y per tiny change in x, starting from $x = c$.

- Geometric: $f'(c)$ is the slope of the graph $y = f(x)$ near the point $(c, f(c))$, i.e. the slope of the tangent line at that point, $y = f(c) + f'(c)(x-c)$.

- Numerical: Approximate the derivative by the difference quotient: $f'(a) \approx \frac{\Delta f}{\Delta x} = \frac{f(x)-f(c)}{x-c}$ for x near c, equivalent to the linear approximation $f(x) \approx f(c)+f'(c)(x-c)$.

- Algebraic: Defining $f'(c) = \lim\limits_{x \to c} \frac{f(x)-f(c)}{x-c} = \lim\limits_{\Delta x \to 0} \frac{f(c+\Delta x)-f(c)}{\Delta x}$, we prove Basic Derivatives and Derivative Rules to find $f'(x)$ for any formula $f(x)$.

The physical meaning is the important one: derivatives are rates of change, so they are essential to analyze any dynamical (changing) system: for example velocity is the rate of change of position, $v = \frac{ds}{dt}$. The geometric meaning as slope is also useful to describe curved shapes. We formulate physical and geometric questions in terms of derivatives, then answer them by translating into numerical arithmetic or algebraic formulas, which are just technical tools. For example, to find maximum points of a physical function $f(x)$, or hill tops of a geometric curve $y = f(x)$, we consider that they must have horizontal (zero-slope) tangents, so we take the derivative $f'(x)$ and solve for the critical points $f'(x) = 0$ algebraically, or numerically with Newton's Method.

Accumulation of influence. In §3.9, we reversed the *algebraic* derivative operation: we could often recognize a given a function $f(x)$ as the derivative of some familiar function $F(x)$, obtaining an *algebraic antiderivative*.[*] In this chapter, we show the antiderivative has all four levels of meaning, giving tools to solve many new problems, which amazingly turn out to be essentially the same problem.[†]

To define antiderivatives *physically*, we reverse the intuition of a physical variable $y = f(x)$ being the rate of change of another variable $z = F(x)$, i.e. $y = f(x) = \frac{dz}{dx}$. Near $x = c$, this means $f(c) \approx \frac{\Delta z}{\Delta x}$, which reverses to:

$$\Delta z \approx f(c)\,\Delta x.$$

That is, starting from the base point $x = c$, a small input increment[‡] $\Delta x = x-c$ produces a proportional output increment $\Delta z = F(x)-F(c)$ controlled by the rate of change $f(c)$. Thus, $y = f(x)$ is an *influence* on the growth of $z = F(x)$ at each $x = c \in [a, b]$, and the antiderivative expresses the *accumulation* or *cumulative effect* of this influence.

- Position is the cumulative effect of velocity. (Velocity is rate of change of position.)

- Velocity is the cumulative effect of acceleration. (Accel is rate of change of velocity.)

Notes by Peter Magyar `magyarp@msu.edu`

[*]But many $f(x)$ are not the derivative of *any* formula $F(x)$, e.g. $f(x) = \sqrt{x^3+1}$, $\sin(x^2)$, or e^x/x.

[†]This was Newton's unparalleled spark of genius that ignited the scientific revolution in the 1700s. Coincidences like this keep occuring because mathematics reveals its intrinsic unity as we uncover deeper and deeper concepts: the apparent variety of questions and theories is only on the surface.

[‡]*Increment*: a small increase, a part added.

- Temperature is the cumulative effect of warming and cooling.
  (Warming is rate of change of temperature.)

- Electricity use is the cumulative effect of wattage. (Wattage is rate of energy use.)

- Total mass is the cumulative effect of density. (Density is rate of mass per volume.)

- Total force on a surface is the cumulative effect of pressure.
  (Pressure is rate of force per unit area.)

- Total population is the cumulative effect of birth and death rates.
  (Birth rate minus death rate is rate of population change.)

- Aggregate cost is the cumulative effect of unit prices (marginal cost).
  (Marginal cost is the price rate per extra unit.)

Distance integral. The physical intution of accumulation of $f(x)$ over $x \in [a, b]$ can be made clear and computable through a new operation called the *integral* $\int_a^b f(x)\, dx$, defined *numerically* as a limit of approximations.

For concreteness, imagine a toy car on a track adjusting its speed by the simple velocity function $f(t) = v(t) = t^2$ over time $t \in [0, 2]$; the antiderivative is the car's position $F(t) = s(t)$. Assuming initial position $s(0) = 0$, we compute the total distance traveled $s(2) = \int_0^2 v(t)\, dt$, the cumulative effect of velocity $v(t)$. Since distance = velocity × time, or $\Delta s = v(t)\, \Delta t$, we can say very roughly:

$$s(2) \;=\; \Delta s \;\approx\; v(2)\, \Delta t \;=\; 4\,(2{-}0) \;=\; 8.$$

This would be exact if there were a constant velocity $v(t) = v(2) = 4$ for the whole time, but in fact the car accelerates from $v(0) = 0$, so $s(2) \approx 8$ is a gross overestimate.

For a good approximation, we split the time interval $[0, 2]$ into $n = 20$ increments of size $\Delta t = 0.1$, with dividing points:

$$0.0 < 0.1 < 0.2 < \cdots < 1.8 < 1.9 < 2.0\,.$$

We approximate a distance increment during each time increment, and add these up to get the total distance traveled; this is called a *Riemann sum*:

$$\begin{aligned}
s(2) \;&\approx\; v(0.1)\Delta t + v(0.2)\Delta t + \cdots + v(1.9)\, \Delta t + v(2.0)\, \Delta t \\
&=\; (0.1)^2\,(0.1) + (0.2)^2\,(0.1) + \cdots + (1.9)^2\,(0.1) + (2.0)^2\,(0.1) \\
&\approx\; 2.87\,.
\end{aligned}$$

Here we sample the velocity $v(t)$ at the end of each increment: the first sample point is $t_1 = 0.1$, the right endpoint of $t \in [0.0, 0.1]$, and similarly for $t_2 = 0.2$, $t_3 = 0.3$, ..., $t_{20} = 2.0$. This is still an *overestimate* (upper Riemann sum), since the velocity is slightly less at the beginning of each increment than the sample $v(t_i)$ at the end.

To get an *underestimate* (lower Riemann sum), we should instead sample velocity at the beginning of each increment, where it is smallest, i.e. $t_1 = 0.0$, ..., $t_{20} = 1.9$:

$$\begin{aligned}
s(2) \;&\approx\; v(0.0)\Delta t + v(0.1)\Delta t + \cdots + v(1.8)\, \Delta t + v(1.9)\, \Delta t \\
&=\; (0.0)^2\,(0.1) + (0.1)^2\,(0.1) + \cdots + (1.8)^2\,(0.1) + (1.9)^2\,(0.1) \\
&\approx\; 2.47.
\end{aligned}$$

Thus the position $s(2)$ lies between 2.47 and 2.87. To get more accurate estimates, we take more and more increments getting smaller and smaller, $n \to \infty$ and $\Delta t = \frac{2}{n} \to 0$. The estimates converge to a limiting value called the *integral*, giving the exact position:

$$s(2) = \int_0^2 v(t)\, dt \stackrel{\text{def}}{=} \lim_{n \to \infty} v(t_1)\, \Delta t + v(t_2)\, \Delta t + \cdots + v(t_n)\, \Delta t.$$

The word "integral" is related to "entire": it computes the *whole* distance $s(2)$ accumulated from the n small increments. The $\int$ is an elongated S meaning *sum*; $v(t)$ stands for all sample values $v(t_1), \ldots, v(t_n)$; and dt suggests very small Δt splitting more and more increments.

For this simple function $v(t) = t^2$, we can compare the numerical answers with the known algebraic solution: $s(t) = \frac{1}{3}t^3$ is the unique antiderivative with $s(0) = 0$, so:

$$s(2) = \tfrac{1}{3}(2^3) = \tfrac{8}{3} \approx 2.667.$$

which is indeed between the lower and upper estimates above. In fact, the average of the two estimates is $\frac{2.47+2.87}{2} = 2.67$, which is the exact answer rounded to 2 decimal places. For this simple $v(t)$, the approximations are not actually needed, since we luckily know an algebraic antiderivative $s(t)$; but the numerical limit formula is valid for *any* $v(t)$.

General integral. Take any given function $f(x)$ which we consider as the rate of change of an unknown antiderivative $F(x)$ for $x \in [a, b]$, i.e. $f(x) = F'(x)$. Then we compute the *total change* $F(b) - F(a)$ as the integral of the rate of change $f(x)$ from $x = a$ to b:

$$F(b) - F(a) = \int_a^b f(x)\, dx \stackrel{\text{def}}{=} \lim_{n \to \infty} f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x.$$

Here we split the interval $[a, b]$ into n increments of size $\Delta x = \frac{b-a}{n}$, and choose a sample point in each increment: $x_1, x_2, \ldots, x_n$ can be the left or right endpoints, or anywhere between. The terms are the approximate incremental changes in $F(x)$, computed as the rate of change $f(x_i)$ times the length of the increment Δx. Finally, we take the limit as $n \to \infty$ and $\Delta x \to 0$, making the approximations converge to the exact integral value.

Area integral. Now we come to one of the most surprising results in mathematics: the *geometric* interpretation of the integral. Suppose we have a function with $f(x) \geq 0$ for $x \in [a, b]$, and we wish to determine the area under the graph $y = f(x)$ and above the interval $[a, b]$ on the x-axis. For example, let us again take $f(x) = x^2$ over $x \in [0, 2]$.
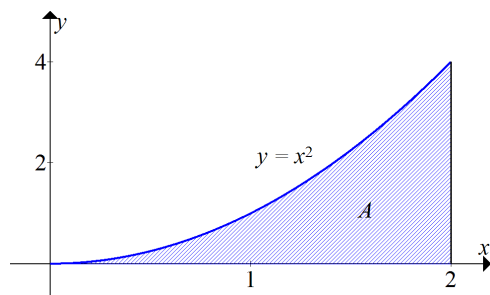
To approximate area A, we cover it by 20 thin rectangles of width $\Delta x = 0.1$ (below left).
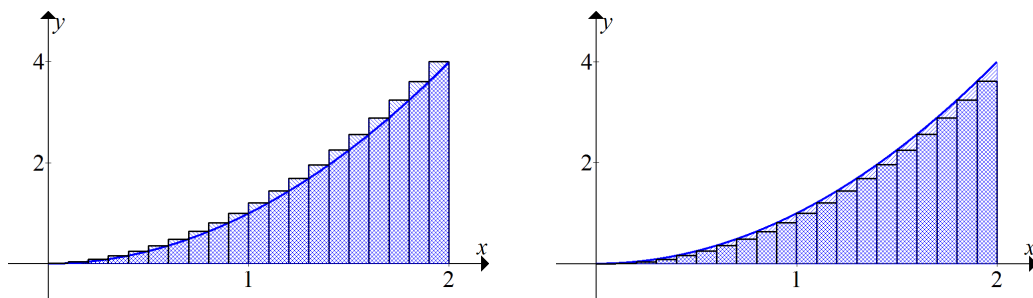
The dividing points are again $0.0 < 0.1 < \cdots < 1.9 < 2.0$, and each rectangle reaches up to the graph at the right endpoint of each increment, giving heights $f(0.1), f(0.2), \ldots, f(2.0)$. The area A under the curve is close to the total area of the rectangles; adding up (height)×(width) for each rectangle gives:

$$
\begin{aligned}
A &\approx f(0.1)\,\Delta x + f(0.2)\,\Delta x + \cdots + f(2.0)\,\Delta x \\
&= (0.1)^2\,(0.1) + (0.2)^2\,(0.1) + \cdots + (2.0)^2\,(0.1) \approx 2.87\,.
\end{aligned}
$$

This is an overestimate since the rectangles slightly overshoot the curve. For an underestimate, we take heights at the left endpoint of each increment, fitting rectangles under the graph (above right):

$$
A \approx f(0.0)\,\Delta x + f(0.1)\,\Delta x + \cdots + f(1.9)\,\Delta x \approx 2.47\,.
$$

Clearly, this is the same computation we did before, upper and lower Riemann sums approximating the integral $s(2) = \int_0^2 v(t)\,dt = \int_0^2 t^2\,dt$. Taking the limit of thinner and thinner rectangles, the exact area is equal to the exact integral:

$$
A = \int_0^2 x^2\,dx.
$$

So the amazing coincidence is:

> The area under the curve $y = x^2$ above $x \in [0,2]$ is the *same* as the distance traveled with velocity $v(t) = t^2$ during $t \in [0,2]$.

This is because area under the graph is the cumulative effect of the height $f(x) = x^2$ above the horizontal interval $x \in [0,2]$,[§] while distance traveled is the cumulative effect of the velocity $v(t) = t^2$ during $t \in [0,2]$. These are both given by the same integral:

$$
\int_0^2 x^2\,dx = \int_0^2 t^2\,dt,
$$

where the name of the variable x or t is mathematically irrelevant.

We have seen the distance problem before in §2.7, when we related the graphs of acceleration $a(t)$, velocity $v(t)$, and distance traveled $s(t)$ in terms of slopes. Now we can visualize $s(t)$ more directly at a particular $t = b$ as the area under the $v(t)$ graph above the interval $t \in [0,b]$, and similarly for $v(t)$ as the area under $a(t)$.

[§]Reverse: rate of change of area up to a moving endpoint is equal to the height at that endpoint.

Approximating an integral. We compute the area A under one arch of the graph $y = \sin^2(x)$ numerically to an accuracy of one decimal place (error less than ± 0.05).
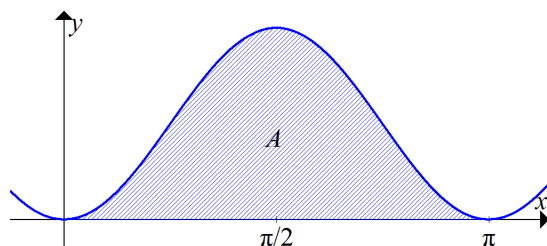
That is, we approximate:

$$A = \int_0^\pi \sin^2(x)\, dx \approx \sin^2(x_1)\, \Delta x + \sin^2(x_2)\, \Delta x + \cdots + \sin^2(x_n)\, \Delta x,$$

where n is a large enough number of rectangles, the corresponding horizontal increment is $\Delta x = \frac{b-a}{n} = \frac{\pi}{n}$, and $x_1, \ldots, x_n$ are appropriate sample points in each increment.

To make sure of the required accuracy, we compute an overestimate and an underestimate (upper and lower Riemann sums). For an *overestimate*, we take $x_1, \ldots, x_n$ so that $\sin^2(x)$ is *largest* within each increment. These are not always the right endpoints, because the function is decreasing on the second half of the interval. Rather, for the increments within $[0, \frac{\pi}{2}]$, we take the right endpoints, and for the increments within $[\frac{\pi}{2}, \pi]$ we take the left endpoints. For an *underestimate*, we do the opposite, taking sample points where $\sin^2(x)$ is *smallest* within each increment.

With a spreadsheet or algebra software, we can easily take $n = 100$ increments of size $\Delta x = \frac{\pi - 0}{100} = 0.01\pi$. The upper estimate is:

$$\sin^2(0.01\pi)\,(0.01\pi) + \sin^2(0.02\pi)\,(0.01\pi) + \cdots + \sin^2(0.50\pi)\,(0.01\pi)$$

$$+ \sin^2(0.50\pi)\,(0.01\pi) + \sin^2(0.51\pi)\,(0.01\pi) + \cdots + \sin^2(0.99\pi)\,(0.01\pi) \approx 1.60\,.$$

The lower estimate is:

$$\sin^2(0.00\pi)\,(0.01\pi) + \sin^2(0.01\pi)\,(0.01\pi) + \cdots + \sin^2(0.49\pi)\,(0.01\pi)$$

$$+ \sin^2(0.51\pi)\,(0.01\pi) + \sin^2(0.52\pi)\,(0.01\pi) + \cdots + \sin^2(1.00\pi)\,(0.01\pi) \approx 1.54\,.$$

Averaging the upper and lower bounds gives our best estimate for the area A:

$$1.54 < A < 1.60 \quad \Longleftrightarrow \quad A = 1.57 \pm 0.03\,.$$

As we have seen, the integral $F(b) = \int_0^b f(x)\, dx$ always defines an antiderivative function numerically, whether or not we can find an algebraic antiderivative. But in §3.9 we did find a tricky algebraic antiderivative: $F(x) = \frac{1}{2}x - \frac{1}{4}\sin(2x)$, satisfying $F(0) = 0$. Since there can be only one such antiderivative, we find that:

$$\int_0^\pi \sin^2(x)\, dx = F(\pi) = \frac{1}{2}\pi - \frac{1}{4}\sin(2\pi) = \frac{\pi}{2} \approx 1.571\,,$$

so in fact our numerical approximation was accurate to 2 decimal places.