Computational Integration

- symbolic integration.
- numerical integration or *quadrature*.
    - quadrature rules.
    - series approximation.
    - Monte-Carlo approximation.

Notation and Terminology

For function $f: \mathbf{R}^n \to \mathbf{R}$:

$$\int f.$$

If $f: \mathbf{R} \to \mathbf{R}$, then we may write something like

$$\int f(x)\, dx.$$

If f is multivariate, that is, a function of the form $f: \mathbf{R}^n \to \mathbf{R}$, then we can write

$$\int f(\mathbf{x})\, d\mathbf{x}.$$

If $\mathbf{f}$ is vector-valued, that is, a function of the form $\mathbf{f}: \mathbf{R}^n \to \mathbf{R}^m$, then we can write

$$\int \mathbf{f}(\mathbf{x})\, d\mathbf{x}.$$

Kinds of Integrals

- *definite* integral: domain of integration is given.
- *indefinite* integral: no domain of integration is given.
- *improper* integral: domain is infinite.
- *integral equation*: free variable in integrand or domain.

E.g., if $v(t)$ denotes velocity (in one dimension) then

$$\int_a^b v(t)\,dt$$

defines the distance travelled over the definite time interval $[a,b]$.

To define the distance travelled from time 0 up to any time $x > 0$, we can write

$$d(x) \quad = \quad \int_0^x v(t)\,dt.$$

This expression is called an *integral equation*.

If we allow the interval $[a,b]$ to grow by letting a approach $-\infty$ or b approach $+\infty$ (or both), then we have an *improper* integral of the form

$$\int_{-\infty}^{+\infty} f.$$

If a function f is continuous, then the *fundamental theorem of calculus* states that f has an (indefinite) integral F such that

$$\int_a^b f \quad = \quad F(b) - F(a).$$

Sometimes F is called an *antiderivative*, in that $dF/dx = f$.

Analytic or Closed-Form Solutions to Integrals

What is the indefinite integral of a polynomial in variable t of the form

$$p(t) \;=\; \sum_{i=0}^{n} a_i\, t^i \,?$$

If $f(x) = \sin x$, then

$$\int_0^{\pi} \sin x \;=\; 2$$

since $F(x) \;=\; -\cos x$ is such that $d(-\cos x)/dx = \sin x$.

Thus

$$\int_0^{\pi} \sin x \;=\; -\cos\pi \;+\; \cos 0 \;=\; 1 \;+\; 1 \;=\; 2.$$

The *sinc* function has the definition

$$\mathrm{sinc}(x) \;=\; \frac{\sin x}{x}\,.$$

Sometimes x is replaced by πt so that the zeros of $\mathrm{sinc}\,\pi t$ occur when t is integral.

It is not easy to show that

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x}\, dx \;=\; \pi\,.$$

Is the integral of $\mathrm{sinc}(\pi t)$ any different?

Maple is better than me at integration!

Maple easily finds a closed form for the integral of Runge's function:

```
> R := 1/(1+25*x^2);
                                   1
                      R := ---------
                                       2
                                1 + 25 x
> int(R,x=-1..1);
                      2/5 arctan(5)
> evalf(");
                        .5493603068
```

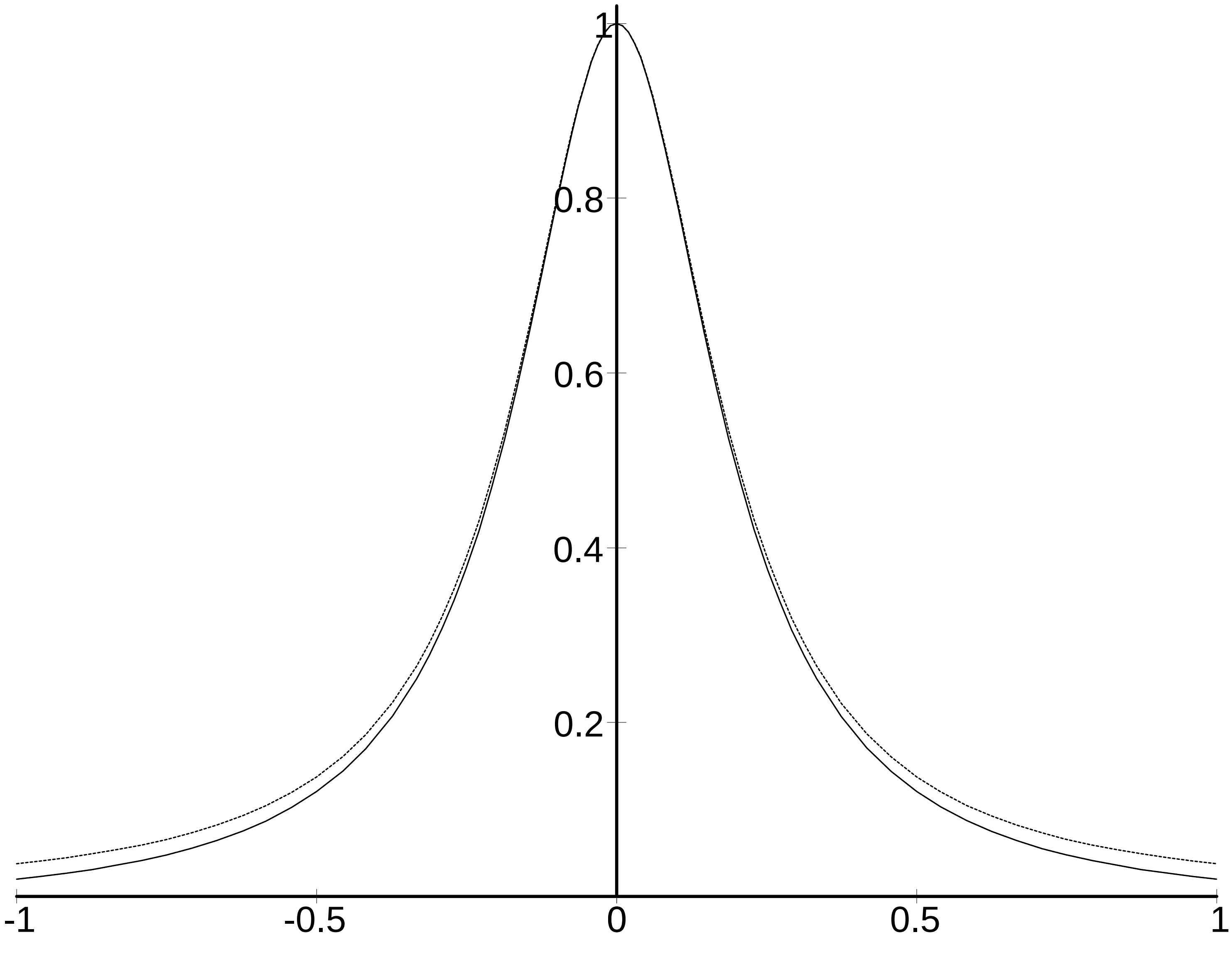
However, Maple is unable to find the closed-form solution for

$$\int \frac{\cos x}{1+25x^2}\, dx,$$

despite the fact that this function looks very similar to, and is bounded above by, Runge's function on $[-1,1]$.

I couldn't find a closed-form solution either but. Maple reports that its definite integral on $[-1,1]$ is roughly 0.5213413.

What is and isn't symbolically integrable?



Simple functions do not necessarily have closed-form integrals.

Arc Length

Let $\mathbf{x}(t) = (x(t), y(t), z(t))$ be a parametric space curve.

How long is the trajectory made by $\mathbf{x}$ on an interval $t \in [a,b]$?

Let $\mathbf{x}'(t) = (x'(t), y'(t), z'(t))$ be the velocity vector of the curve at time t.

This corresponds to the componentwise derivative of $\mathbf{x}$ with respect to t.

The function $\mathbf{x}'$ defines the instantaneous *direction* of travel and the magnitude of this vector,

$$\|\mathbf{x}'(t)\| = \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2},$$

defines its *speed*.

The *arc length* of the curve $\mathbf{x}$ between points $\mathbf{x}(a)$ and $\mathbf{x}(b)$ is

$$\int_a^b \|\mathbf{x}'(t)\| \, dt.$$

If the function is in explicit form, namely of the form

$$y = f(x),$$

then its arc length is

$$\int_a^b \sqrt{1 + (f')^2} \, dx.$$

Examples of Arc Length Computations

Use Maple to derive the arc length of arbitrary cubic curve segments $\mathbf{x}(t) = (x(t), y(t), z(t))$ over the interval [0,1].

Use the parametric definition of a circle to show that the circumference (i.e., the arc length) of a circle of radius r has length $2\pi r$.

Can this be easily extended to the arc length for ellipses? A class of functions called the *elliptic* integrals arises from this.

Some very simple functions do not have closed-form solutions. Let us compute the arc length of the curve

$$y = \sin x$$

over $x \in [0,\pi]$.

The arc length is

$$L = \int_0^\pi \sqrt{1 + (y')^2}\, dx$$

$$= \int_0^\pi \sqrt{1 + \cos^2 x}\, dx.$$

Maple can’t solve it analytically. Neither can I. But it is easy to compute numerically.

Numerical Quadrature

The idea is the same for every technique: to integrate f on an interval $[a,b]$, compute a set of points on (or near) the curve, multiply each by a weight, and sum the result.

Where do the weights come from?

The number of points used per interval, their spacing, and the nature of interpolation defines a specific rule.

A *quadrature rule* looks like this:

$$\int_a^b f(x)\,dx \approx \sum_{i=0}^{n} w_i f_i,$$

where the w_i are weights and the f_i are points on the curve.

If you think this looks like filtering, you are among the lucky few!

In general, a huge variety of approaches exist. Our approach will be based on interpolation.

Quadrature rules from basis functions

In our case, the weights will derive from interpolation (as an approximation). I.e., if f is approximated by a polynomial $p(t)$, then we know that $p(t)$ has an analytic integral.

So, if

$$p(t) \;=\; \sum_{i=0}^{n} b_i(t)\, f_i,$$

where the b_i are basis functions, then

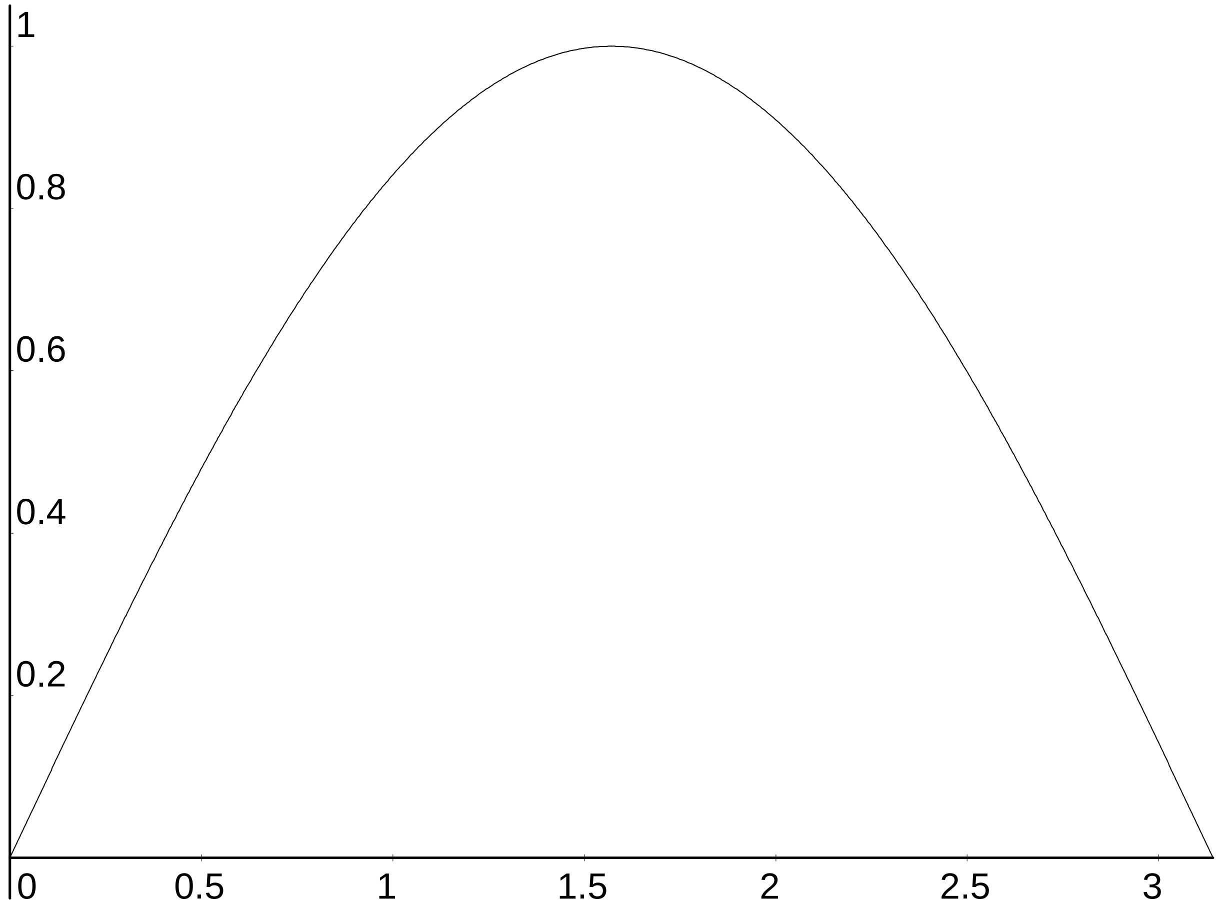
$$\begin{aligned}
\int_a^b f(t)\,dt &\approx \int_a^b p(t)\,dt \\
&= \int_a^b \Big(\sum_{i=0}^{n} b_i(t)\, f_i \Big)\, dt \\
&= \sum_{i=0}^{n} \Big(\int_a^b b_i(t)\, f_i\, dt \Big) \\
&= \sum_{i=0}^{n} f_i \Big(\int_a^b b_i(t)\, dt \Big) \\
&= \sum_{i=0}^{n} f_i\, w_i.
\end{aligned}$$

Nifty!

Riemann Sums

Suppose we want to compute numerically

$$\int_0^{\pi} \sin x \; dx \quad = \quad 2.$$



Enclose it by one box

Let us start with one rectangle of width $b - a = \pi$ and of height

$$h_0 = \sin\left(\frac{b+a}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1.$$

The area under the rectangle is simply its width times its height, which is $\pi \times 1 = \pi$.

Not good enough.

So try two boxes

Split our interval $[0,\pi]$ into two subintervals $[0,\pi/2]$ and $[\pi/2,\pi]$, and compute two new rectangles.

One of the rectangles will have height

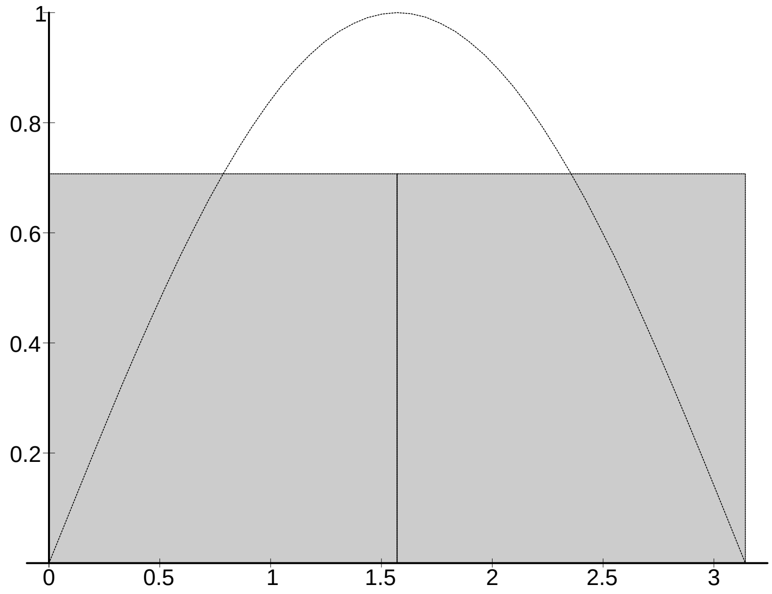
$$h_{left} = \sin\left(\frac{\pi/2+0}{2}\right) = \sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}.$$

The other rectangle has the same height, since

$$h_{right} = \sin\left(\frac{\pi + \pi/2}{2}\right) = \sin\left(\frac{3\pi}{4}\right) = \frac{\sqrt{2}}{2}.$$

Both rectangles are of width $\pi/2$, and therefore the area of the two rectangles is

$$\frac{\pi}{2}h_{left} + \frac{\pi}{2}h_{right} = \frac{\pi\sqrt{2}}{2} \approx 2.22144.$$



Riemann sums in general

Break our original interval $[a,b]$ into many subintervals with x components $a = x_0, x_1, \cdots, x_n = b$.

Form n rectangles $R_0, \cdots, R_{n-1}$.

Each rectangle R_i has midpoint

$$m_i = \frac{x_{i+1} + x_i}{2},$$

height $\sin m_i$, and width $x_{i+1} - x_i$.

Thus the area of all the rectangles is

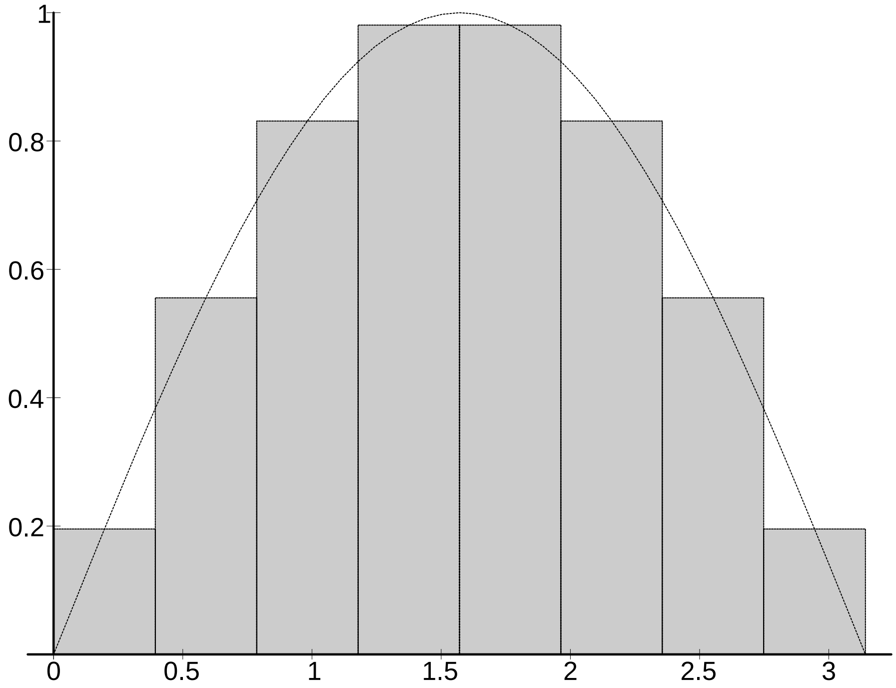
$$A_n = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \sin m_i.$$

For a general function f, the approximation of the definite integral over n subintervals is given by the following rule.

> ***Riemann Sums Quadrature Rule.***
>
> For each subinterval $[x_i, x_{i+1}]$, compute a single sample value $f(m_i)$, where $m_i = (x_{i+1} + x_i)/2$. If there are n subintervals, then the overall area is
>
> $$A_n = \sum_{i=0}^{n-1} (x_{i+1} - x_i)\, f(m_i).$$

Riemann sums on eight subintervals

$$\int_0^{\pi} \sin x \, dx \;\approx\; A_8 \;\approx\; 2.0129.$$



n	Area (A_n)
1	3.14159
2	2.22144
4	2.05234
8	2.01291
16	2.00321
32	2.00080
64	2.00020
128	2.00005
256	2.00001
512	2.00000

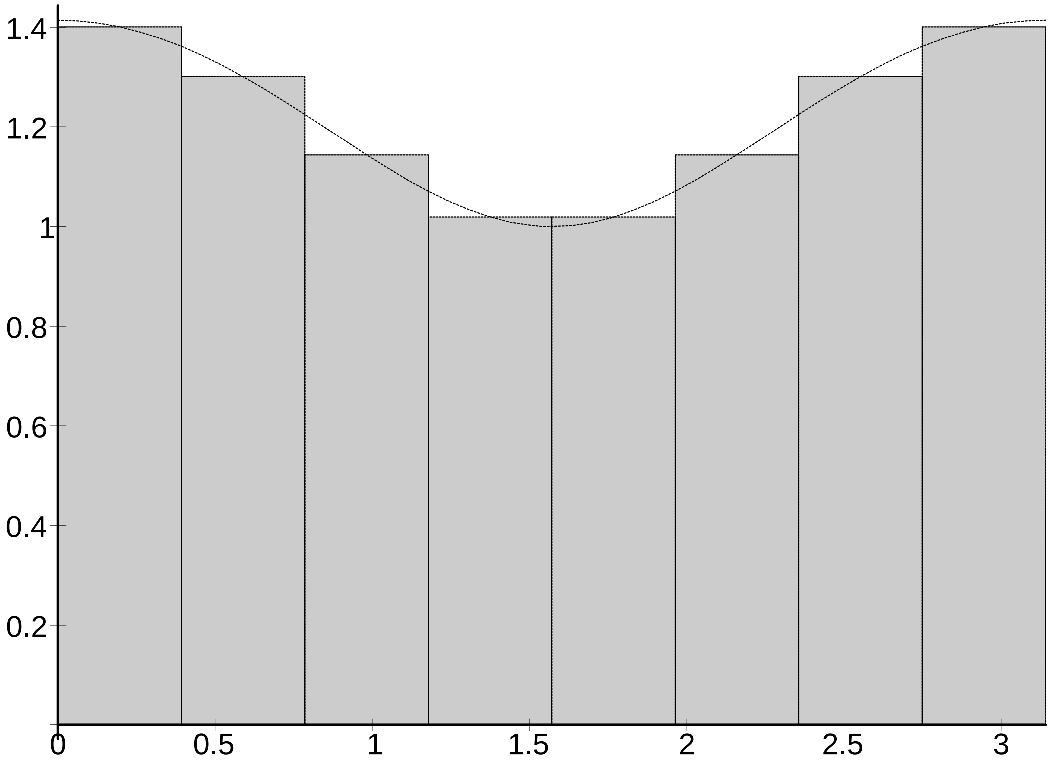
Arc Length Computation

Approximate the arc length of $y = \sin x$.

Using only 32 rectangles, we find that

$$\int_0^{\pi} \sqrt{1 + \cos^2 x}\, dx \;\approx\; 3.820197792,$$

which is correct to eight decimal places.



Trapezoid Rule

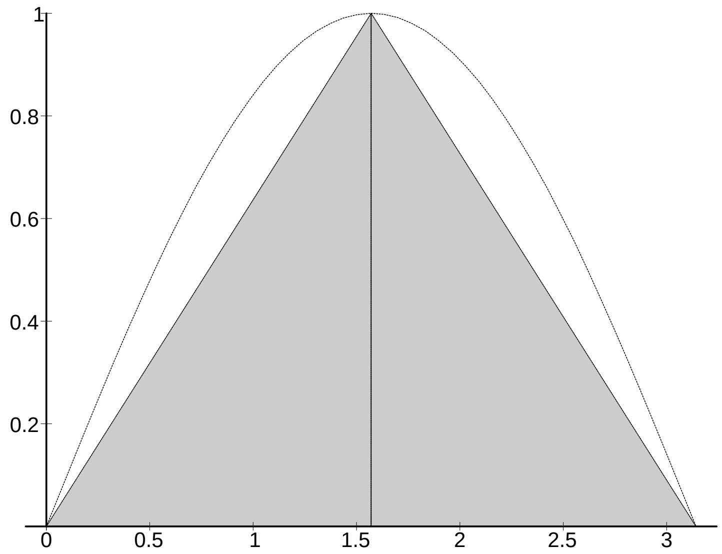
Based on piecewise linear interpolation.

We once again break our interval $[a,b]$ up into pieces with x components $a = x_0, x_1, \cdots, x_n = b$.

In this case, we have a set of samples

$$(x_0, f(x_0)), (x_1, f(x_1)), \cdots (x_n, f(x_n)).$$

We join adjacent points by a line segment, forming trapezoids rather than rectangles.



The area of a given trapezoid T_i is

$$A(T_i) \quad = \quad \frac{f(x_{i+1}) \; + \; f(x_i)}{2} \, (x_{i+1} - x_i).$$

In General

By summing the area of the trapezoids, we have another approximation scheme, given by the following summary.

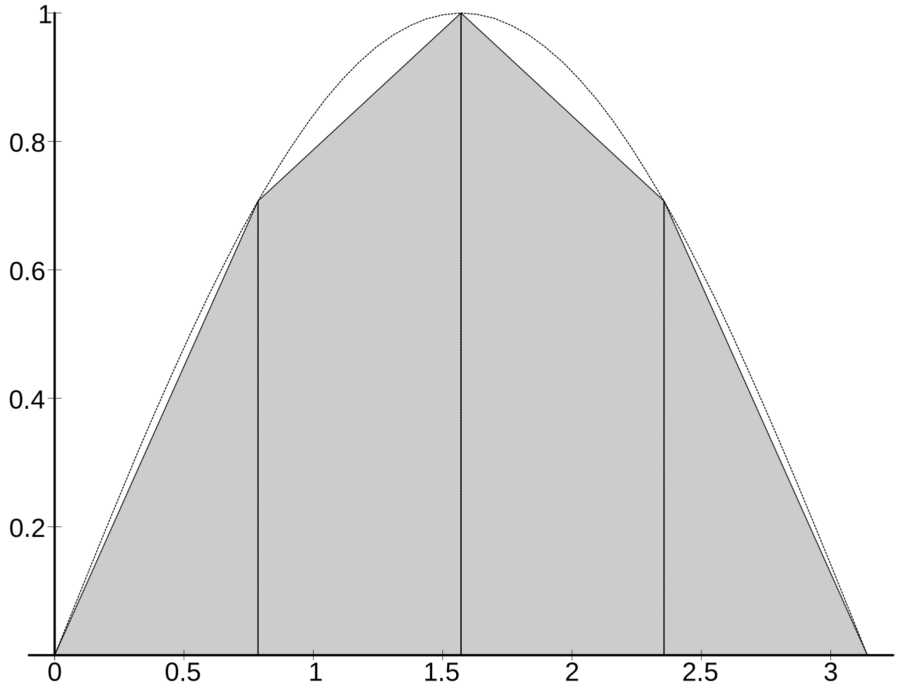
Trapezoid Quadrature Rule.

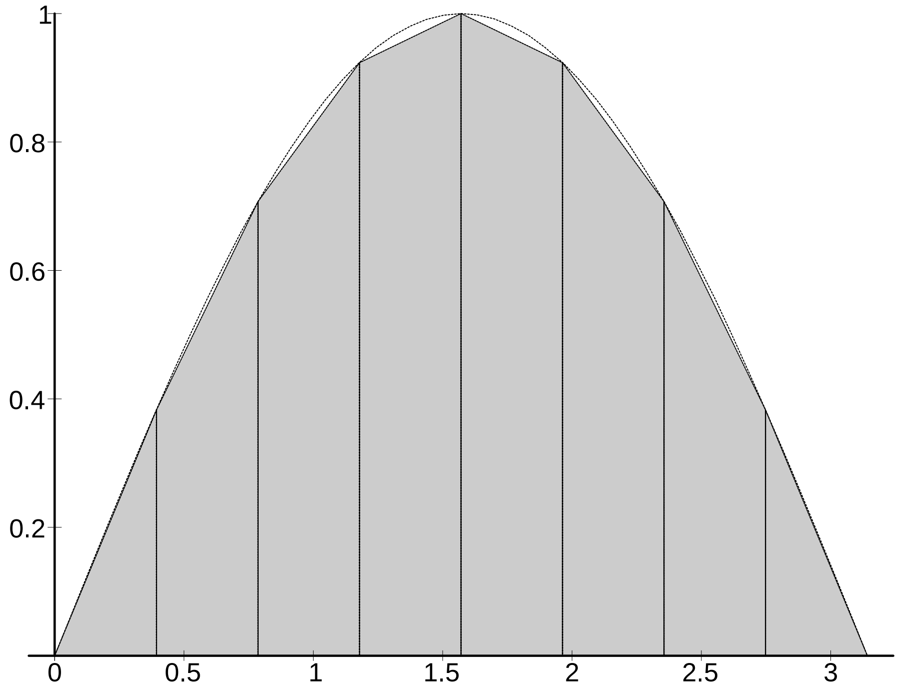
Over each subinterval $[x_i, x_{i+1}]$, evaluate $f(x_i)$, and $f(x_{i+1})$. Each point has weight 1/2 in computing the area under each subinterval. If there are n subintervals, the overall area is

$$A_n = \sum_{i=0}^{n-1} A(T_i) = \sum_{i=0}^{n-1} \frac{f(x_{i+1}) + f(x_i)}{2}(x_{i+1} - x_i).$$

n	Area (A_n)
2	1.57079
4	1.89611
8	1.97423
16	1.99357
32	1.99839
64	1.99959
128	1.99989
256	1.99997
512	2.00000

More trapezoids



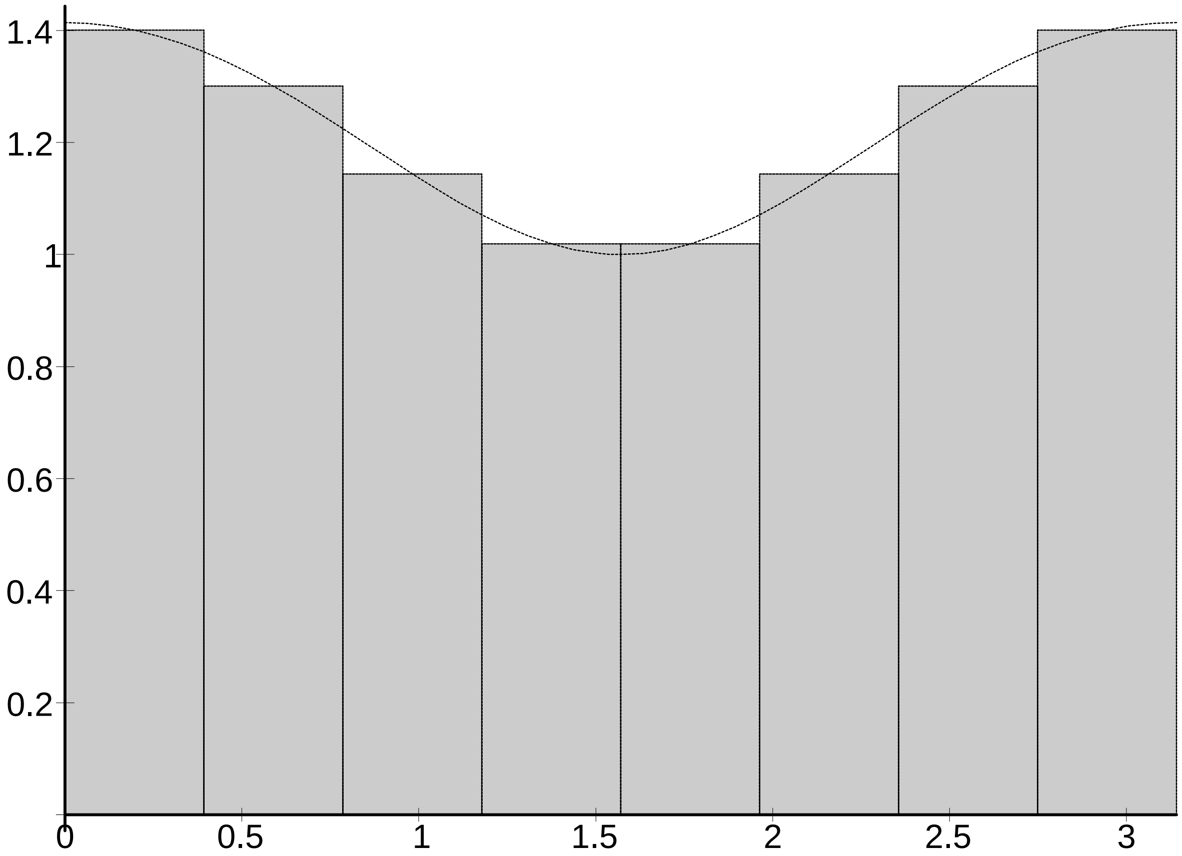


Arc Length, Revisited

Once again, approximate the arc length of $y = \sin x$.

Using ten trapezoids the result is again correct to eight decimal places:

$$\int_0^{\pi} \sqrt{1 + \cos^2 x}\; dx \quad = \quad 3.820197718.$$



Simpson's Rule

When we work with piecewise quadratic interpolation, we get *Simpson's rule*

Let's map a parabola through $f(a)$, $f(m)$, and $f(b)$, where $m = \frac{a+b}{2}$.

We can define a basis to do this using Maple.

```
> B := matrix([[a^2,a,1], [m^2, m,1], [b^2,b,1]]);

> M := inverse(B);

> T := [t^2,t,1];

> basis := seq(dotprod(T,col(M,i)),i=1..3);

> m := (a+b)/2;

> BASIS := seq(int(basis[i],t=a..b), i=1..3);

> BASIS := seq(simplify(BASIS[i]),i=1..3);

   BASIS := - 1/6 a + 1/6 b, - 2/3 a + 2/3 b,
                  - 1/6 a + 1/6 b
```

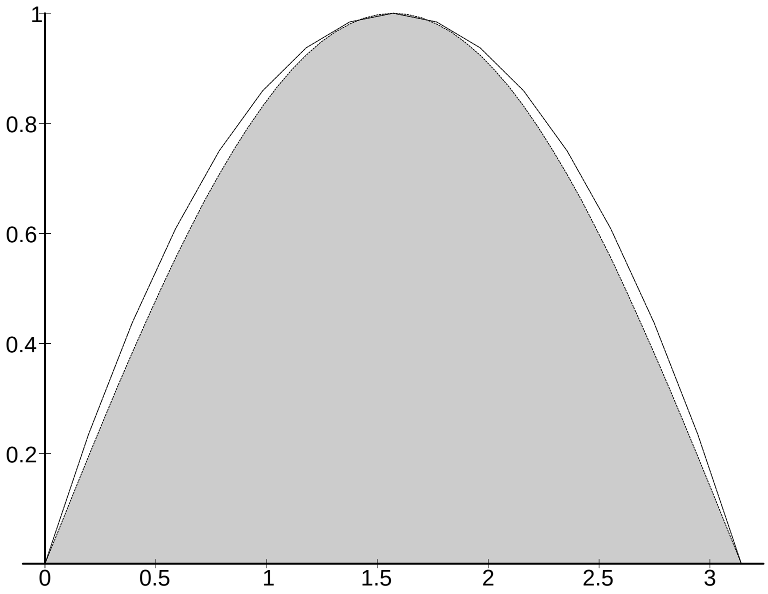
This gives us the weights for our quadrature rule:

$$\int_a^b Q(x)\,dx \quad = \quad \frac{b-a}{6}\,f(a) \;+\; \frac{b-a}{6}\,f(b) \;+\; \frac{2(b-a)}{3}\,f\Big(\frac{a+b}{2}\Big).$$

***Simpson's Quadrature Rule*.**

On each subinterval $[x_i, x_{i+1}]$, evaluate $f(x_i)$, $f(x_{i+1})$, and $f((x_{i+1}+x_i)/2)$. Each function value has weight 1/6, 1/6, and 2/3, respectively, in computing the area under each subinterval. If there are n subintervals, the overall area is

$$A_n = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \left(\frac{1}{6} f(x_i) + \frac{1}{6} f(x_{i+1}) + \frac{2}{3} f\left(\frac{x_i + x_{i+1}}{2} \right) \right).$$



n	Area (A_n)
1	2.094395
2	2.004559
4	2.000269
8	2.000017
16	2.000001
32	1.999999

A Cubic Quadrature Rule

Analogously, we can work with the cubic Catmull-Rom basis to get another quadrature rule. The cubic Catmull-Rom basis functions are:

$$
\begin{aligned}
b_0(t) &= -\frac{t^3}{2} + t^2 - \frac{t}{2}, \\
b_1(t) &= \frac{3t^3}{2} - \frac{5t^2}{2} + 1, \\
b_2(t) &= -\frac{3t^3}{2} + 2t^2 + \frac{t}{2}, \\
b_3(t) &= \frac{t^3}{2} - \frac{t^2}{2}.
\end{aligned}
$$

Computing the integral of each b_j is very easy. If

$$b_j(t) = \sum_{i=0}^{3} a_i t^i,$$

then

$$B_j(t) = \int_0^1 b_j(t)\, dt = \sum_{i=0}^{3} \frac{a_i}{i+1}.$$

Therefore

$$B_0 = \frac{-1}{24}, \quad B_1 = \frac{13}{24}, \quad B_2 = \frac{13}{24}, \quad B_3 = \frac{-1}{24}.$$

Gathering everything up, our quadrature rule looks like this...

***Cubic Catmull-Rom Quadrature Rule*.**

Over each subinterval i simulate the effect of integrating the cubic Catmull-Rom curve passing through the points $f(x_i)$ and $f(x_{i+1})$. Doing so requires the values $f(x_{i-1})$, $f(x_i)$, $f(x_{i+1})$, and $f(x_{i+2})$. Each function value has weight –1/24, 13/24, 13/24, –1/24, respectively, in computing the area under each subinterval. If there are n subintervals, then the overall area is

$$A_n = \sum_{i=0}^{n-1} (x_i - x_{i-1}) \left(\frac{-1}{24} f(x_{i-1}) + \frac{13}{24} f(x_i) + \frac{13}{24} f(x_{i+1}) - \frac{1}{24} f(x_{i+2}) \right).$$

n	Area (A_n)
3	1.964949
6	1.997730
12	1.999857
24	1.999991
48	1.999999

Compound Quadrature Rules

We notice that some points are shared across segments, so we *compound* those values into a single result.

For Riemann sums, the compound rule is the same as the quadrature rule, because no points are shared between segments.

For Trapezoid rule, lots of points are shared:

$$A_n = \sum_{i=0}^{n-1} \frac{f(x_{i+1}) + f(x_i)}{2} (x_{i+1} - x_i),$$

the righthand endpoint on subinterval I_i becomes the left-hand endpoint of interval I_{i+1}. If we assume uniformly spaced x, we can rewrite this equation as

$$\begin{aligned} A_n &= \Delta x \sum_{i=0}^{n-1} \frac{f(x_i) + f(x_{i+1})}{2} \\ &= \Delta x \left(\frac{f(x_0)}{2} + \frac{f(x_1)}{2} + \frac{f(x_1)}{2} + \cdots + \frac{f(x_{n-1})}{2} + \frac{f(x_{n-1})}{2} + \frac{f(x_n)}{2} \right) \\ &= \Delta x \left(\frac{f(a)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(b)}{2} \right). \end{aligned}$$

Compound Trapezoid Rule.

If $x_0, x_1, \cdots, x_n$ are uniformly spaced points on the interval $[a,b]$ with $x_0 = a$ and $x_n = b$, then the compound trapezoid rule on these points is

$$A_n = \Delta x \left(\frac{f(a)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(b)}{2} \right).$$

Compound Simpson's Rule

To derive a compound Simpson's rule, we once again assume that the x_i are uniformly spaced:

$$A_n = \sum_{i=0}^{n-1} \frac{x_{i+1}-x_i}{6} f(x_i) + \frac{x_{i+1}-x_i}{6} f(x_{i+1}) + \frac{2(x_{i+1}-x_i)}{3} f\left(\frac{x_i + x_{i+1}}{2} \right)$$

$$= \frac{\Delta x}{6} \sum_{i=0}^{n-1} f(x_i) + f(x_{i+1}) + 4f\left(\frac{x_i + x_{i+1}}{2} \right)$$

$$= \frac{\Delta x}{6} \left(f(a) + 4f(m_0) + 2f(x_1) + 4f(m_1) + \cdots + 4f(m_{n-1}) + f(b) \right)$$

where $m_i = (x_i + x_{i+1})/2$. Note, however, that the sequence of x values

$$x_0,\ m_0,\ x_1,\ m_1,\ \cdots,\ x_{n-1},\ m_{n-1},\ x_n$$

contains $2n+1$ points with an equal spacing of $\Delta x/2$. We relabel these values as

$$\bar{x}_0,\ \bar{x}_1,\ \cdots,\ \bar{x}_{2n},$$

where $\bar{x}_{2i} = x_i$ for $i = 0, 1, \cdots, n$. and $\bar{x}_{2i+1} = m_i$, for $i = 0, 1, \cdots, n-1$.

We also let

$$\Delta\bar{x} = \frac{(x_{i+1}-x_i)}{2} = \frac{\Delta x}{2}.$$

With this we can rewrite the above equation for compound Simpson's rule as

$$A_n = \frac{\Delta\bar{x}}{3} \left(f(a) + 4f(\bar{x}_1) + 2f(\bar{x}_2) + 4f(\bar{x}_3) + \cdots + 4f(\bar{x}_{2n-1}) + f(b) \right),$$

since $\Delta x/6 = \Delta\bar{x}/3$.

We shall drop the $\overline{x}$ notation in the following rule.

***Compound Simpson's Rule*.**

The result of approximating a function f piecewise on n intervals, using n piecewise continuous quadratic curves and integrating the resulting approximation is

$$A_n = \frac{\Delta x}{3}\Big(f(a) + 4f(x_1) + 2f(x_2) + \cdots + 4f(x_{2n-1}) + f(b)\Big),$$

where $\Delta x = \dfrac{b-a}{2n}$ and $x_i = a + i\Delta x,\ i = 0, 1, \cdots, 2n.$