Learning Integration with Maple

Attila Kovács
Eötvös Loránd University, Budapest, Hungary
attila@compalg.inf.elte.hu

Submitted to HEJ. Manuscript no.: ANM-010201-A

This version is for internal use only!

Abstract

The aim of the present paper is to introduce the notion and basic properties of integrals as well as to show their importance. According to Buchberger's didactic principle (see [1]) the computer algebra software Maple assists making easier the way of learning. The worksheet was written during my stay in Nijmegen under the supervision of A. H. M. Levelt. The procedures are free and can be obtained via e-mail from the author or by downloading from this site. I hope that this introduction will be useful not only for calculus students but for everyone who wants to know more about this important part of mathematics. Throughout this paper we used MapleV Release 5, which is a registered trademark of Waterloo Maple Software. This work was supported by TEMPUS JEP 09269-95.
AMS indices: 00A35, 26A42

1 Introduction

Problem: Let f be a positive valued (bounded) function on a closed (finite) interval $[a, b]$ of the real numbers. Compute the area of the region in the x, y plane bounded by the graph of f, the x-axis and the straight lines $x = a,\ x = b$.
EXAMPLE 1.1.

```
> f:=x->x*sin(x)^2;
```

$$f := x \to x \sin(x)^2$$

```
> plot(f(x),x=1..3,y=0..2,scaling=CONSTRAINED);
```

For "general" functions this looks like a hard problem. However, for some types of special functions (e.g. linear functions) the problem is easy. Let us look at *stepfunctions*. A stepfunction g on $[a, b]$ is given by a *subdivision* of $[a, b]$, i.e. a finite sequence of numbers $a = a_0 \le a_1 \le \ldots \le a_n = b$ and function values $v_1, \ldots, v_n$ such that $g(x) = v_i$ for all x such that $a_{i-1} < x \le a_i$ for all i, whereas $g(x) = 0$ for all $x \le a$ and $x > b$.

Since stepfunctions play an important role we introduce the procedure **stepfunction** to produce them. It has two arguments, a subdivision S and a list of function values V. With the notations above $S = [a_0, \ldots, a_n]$ and $V = [v_1, \ldots, v_n]$.
EXAMPLE 1.2.

```
> S:=[3,3.7,5,6.2,7.1,8]: V:=[2.5,-1,4.2,3,-2]:
> g:=stepfunction(S,V):
> plot(g,x=0..10,scaling=CONSTRAINED);
```

For positive stepfunctions the solution to the *area problem* is immediate and given by the formula

$$\sum_{i=1}^{n} (a_i - a_{i-1})\, v_i.$$

Note that this formula makes sense for general stepfunctions: the resulting number represents the area above the x-axis minus the area below the x-axis. The procedure **int_step** yields the above sum for given S and V.
EXAMPLE 1.3.

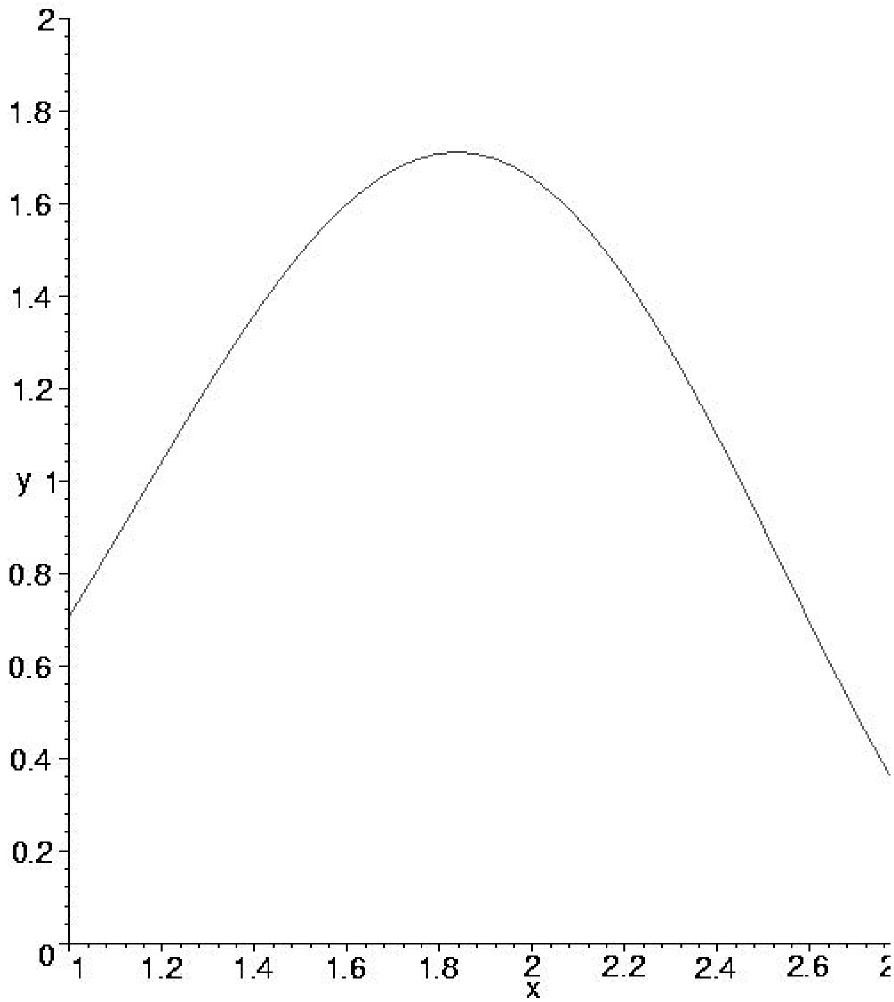


Figure 1:

For the stepfunction g in Example 1.2. the area under the graph of g is

```
>  int_step(S,V);
```

$$6.39$$

Now back to the problem for general functions. The idea is to approximate such a function f by stepfunctions. E.g., a continuous function can be well approximated with constant on sufficiently small intervals. Let us look at the function f of Example 1.1.

EXAMPLE 1.4.

Take a *regular* subdivision of $[a, b]$, i.e. a subdivision into n subintervals of equal length.

```
>  f:=x->x*sin(x)^2:
>  n:=10:
>  S:=[seq(1.+2*'i'/n,'i'=0..n)];
```

$$S := [1., 1.200000000, 1.400000000, 1.600000000, 1.800000000, 2., 2.200000000, 2.400000000, 2.600000000, 2.800000000, 3.]$$

For the value of the stepfunction on $[a_{i-1}, a_i]$ we choose $f(a_i)$. Hence

```
>  V:=[seq(f(S['i']),'i'=2..n+1)];
```

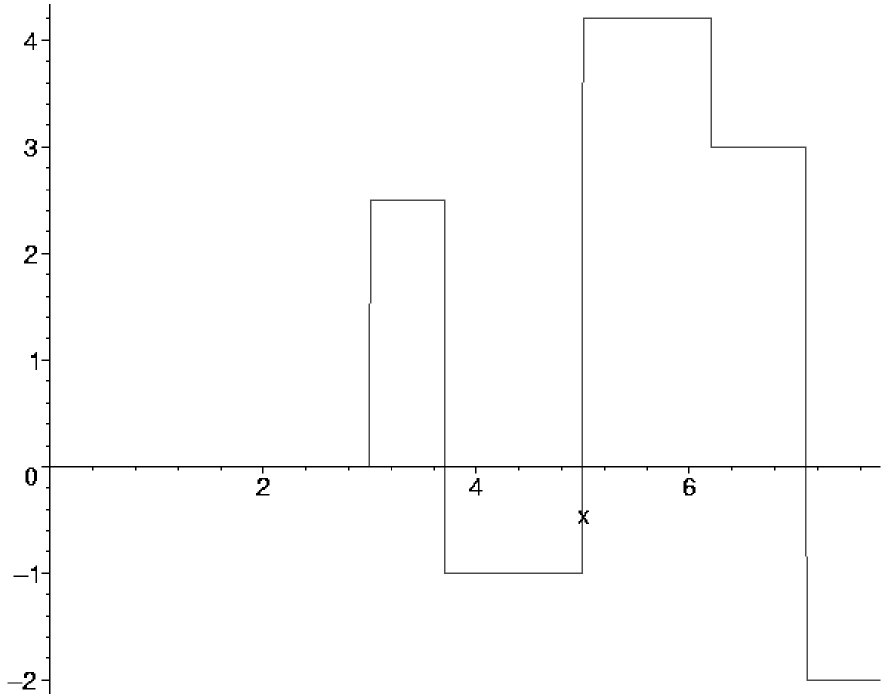


Figure 2:

$$V := [1.042436229, 1.359555639, 1.598635820, 1.707082575, 1.653643621,$$
$$1.438066157, 1.095001220, .6909283272, .3142077702, .05974457007]$$

```
> g:=stepfunction(S,V):
> plot(\{f(x),g(x)\},x=1..3);
```

This plot shows the function f and its approximating stepfunction g. Some daemon (in fact Maple's procedure **int**) tells us that the area under the graph of f equals to 2.264846340 whereas the area under the graph of g is given by

```
> int_step(S,V);
```

$$2.191860385$$

which can be rounded to 2.192. Neither very good nor totally bad as an approximation. We can get a better approximation by increasing the number of subintervals:

```
> n:=50: S:=[seq(1.+2*'i'/n,'i'=0..n)]:
> V:=[seq(f(S['i']),'i'=2..n+1)]:
> g:=stepfunction(S,V):
> plot(\{f(x),g(x)\},x=1..3,numpoints=500); g:='g':
> int_step(S,V);
```

$$2.251554949$$

rounded to 2.252, a much better approximation.

Exercise 1.1.

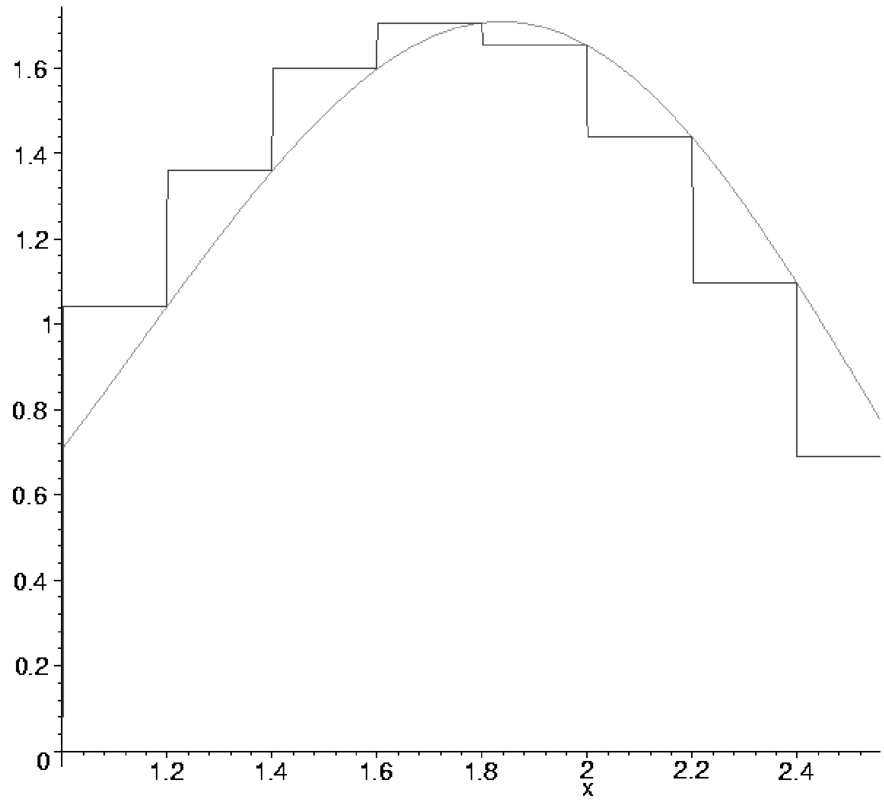


Figure 3:

Consider the function

$$f\ :\ x \to \sqrt{1-x^2}.$$

The region under the graph of f between $x = 0,\ x = 1$ is a 90 degree circle sector. Hence the area equals to $\pi/4$. Use the above method of approximation by regular stepfunctions to find an approximation of π. How many subintervals are needed to approximate π with an error less then or equal to $1/50$?

The Maple package "student" contains the functions **leftsum** and **rightsum**, which are in case of regular subdivision the same as the function **int_step**. I.e., the function **leftsum** (**rightsum**) computes a numerical approximation to a definite integral using rectangles. The height of each rectangle is determined by the value of the function at the left (right) side of each subinterval. A graph of the approximation can be obtained by the Maple procedure **leftbox** (**rightbox**).

Exercise 1.2.
Approximate the function

$$f(x) = \sqrt{1+x^3}$$

on the intarval $[4, 6]$ by regular stepfunctions using the procedure

- **leftsum**,
- **rightsum**

with an error less than 0.01. How many subintervals were needed in each case?

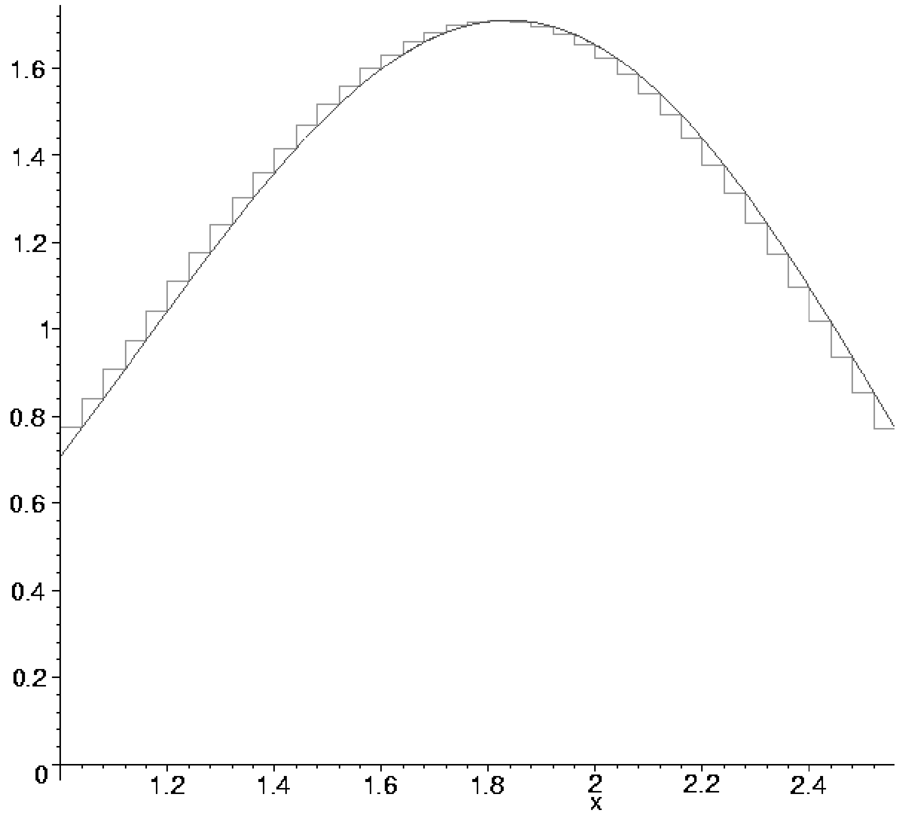


Figure 4:

2 The Riemann integral

For a bounded function f on a (finite) interval $[a, b]$ and a subdivision S of $[a, b]$ one can define two special stepfunctions: U and L (upper and lower stepfunction) as follows.

Let $S := [a_0, \ldots, a_n]$. Then for all i $(1 \leq i \leq n)$ and x satisfying $a_{i-1} < x \leq a_i$ one puts

$$U(x) = \sup(f(y) : a_{i-1} < y \leq a_i)$$

and

$$L(x) = \inf(f(y) : a_{i-1} < y \leq a_i),$$

respectively.

EXAMPLE 2.1. (cf. EXAMPLE 1.4.)

```
> f:=x->x*sin(x)^2: n:=10: S:=[seq(1.+2*'i'/n,'i'=0..n)]:
> integral_plot(f,S,'riemann');
```

5 By definition the graph of f lies below that of U and above the one of L. Hence the area I below the graph

of f must satisfy

$$\textbf{lower}(f, S) \leq I \leq \textbf{upper}(f, S).$$

Here **lower** (resp. **upper**) is the area under the graph of L (resp. U). In our example it means the values

```
> lower(f,S);
```

$$1.992058554$$

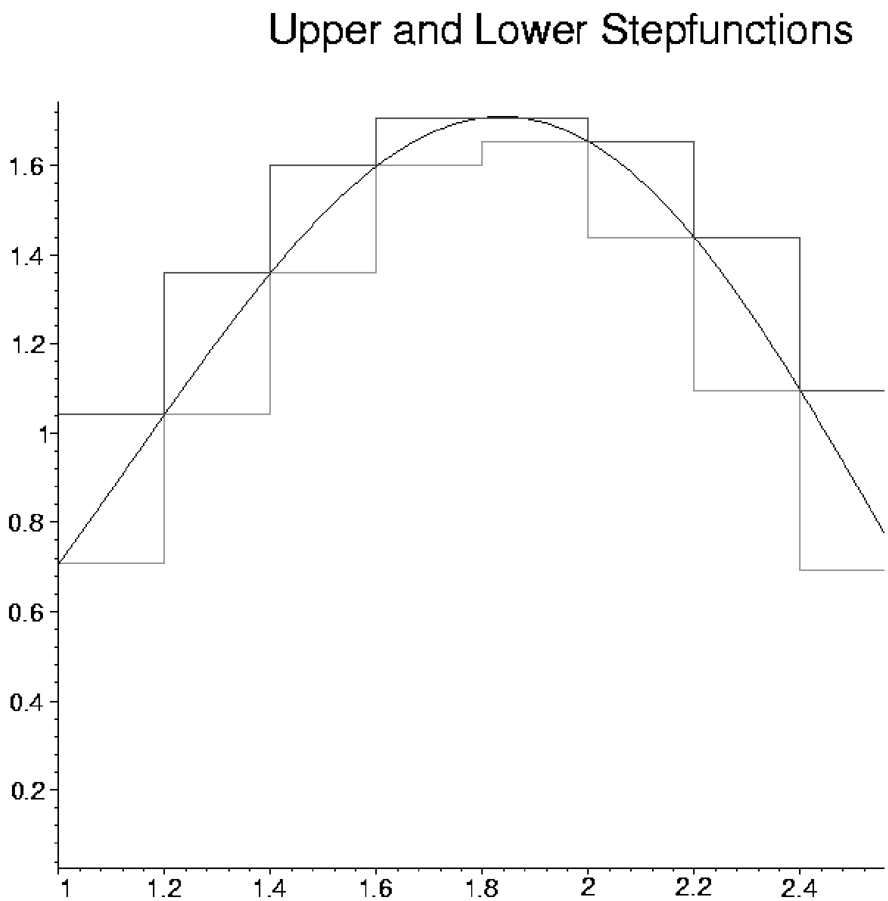


Figure 5:

```
> upper(f,S);
```

$$2.521327986$$

If we make the subdivision finer and finer the approximation of I by **upper** and **lower** should be better and better.

```
> n:=25: S:=[seq(1.+2*`i`/n,`i`=0..n)]:
> integral_plot(f,S,'riemann');
> lower(f,S);
```

$$2.157692545$$

```
> upper(f,S);
```

$$2.369400318$$

Let us denote the subdivision S devided into n subintervals by S_n. Now we must realize that “the area under the graph of f” has never been properly defined (though it is a very intuitive notion). The notions introduced here, however, give a good opportunity.

Definition: f is called (Riemann-) integrable if there exists a real number I such that

$$I = \lim_{n\to\infty} \mathbf{lower}(f, S_n) = \lim_{n\to\infty} \mathbf{upper}(f, S_n).$$

The number I is called the “integral of f over $[a, b]$”.

Notations: $\int_a^b f(x)\,dx$ commonly or sometimes in text form
int$(f(x), x = a..b)$.

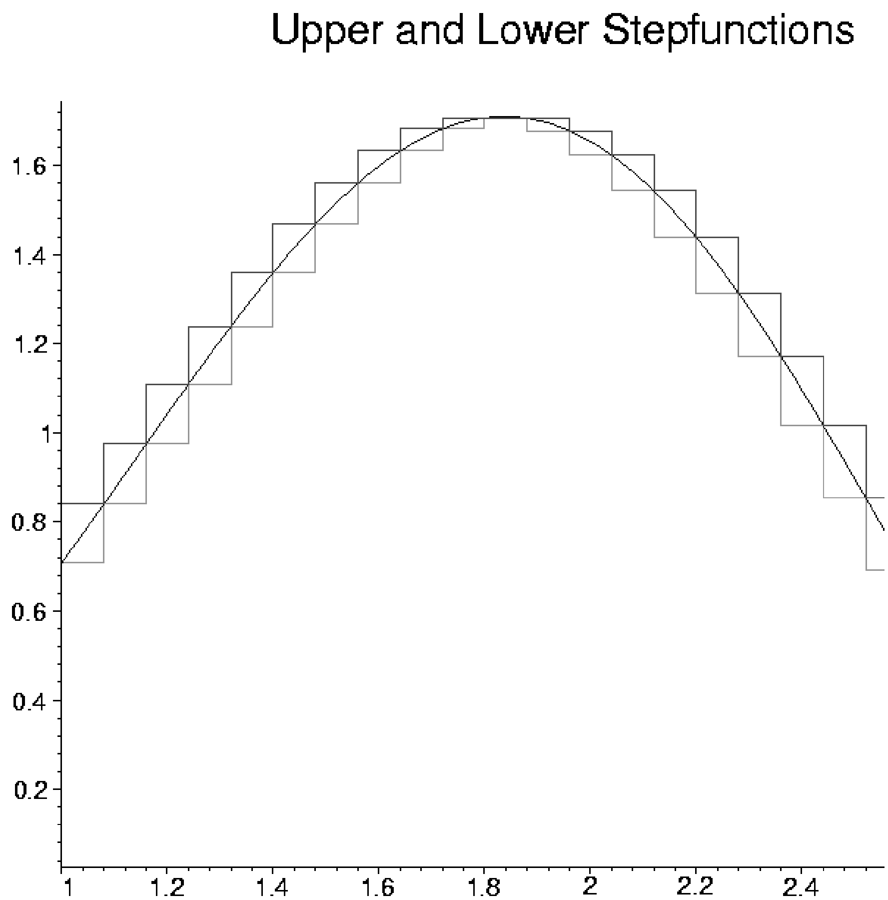


Figure 6:

2.1 Basic properties

The proofs can be found in any textbook on calculus.

(a) f is integrable over $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a subdivision S of $[a, b]$ such that **upper**$(f, S)-$**lower**$(f, S) < \varepsilon$. (The existence of a real number I with the property **lower**$(f, S) \leq I \leq$ **upper**(f, S) is a consequence.)

(b) In the definition of "integrable" the notions *subdivision* and *regular subdivision* are equivalent. To be more precise the subintervals can have different length, provided only that as n increases the length of the longest subinterval tends to 0.

(c) If f is continuous then f is integrable. (It is a consequence of uniform continuity of f on $[a, b]$.)

(d) Every monotonic function is integrable. (**upper**$(f, S)-$**lower**(f, S) can easily be estimated.)

(e) If f and g are integrable over $[a, b]$ then $f + g$ is integrable and

$$\int_a^b (f(x) + g(x))\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

(f) If f is integrable over $[a, b]$ and c is a real number then $c \cdot f$ is integrable and

$$\int_a^b c\, f(x)\, dx = c \int_a^b f(x)\, dx.$$

(g) Let f be integrable over $[a,b]$ and $\varepsilon > 0$. Then there exists $\delta > 0$ with the following property: for any subdivision $S = [a_0, \ldots, a_n]$ into subintervals of length smaller than δ and any $b_1, \ldots, b_n$ satisfying $a_{i-1} \leq b_i \leq a_i$ the difference of the "Riemann sum" from the integral of f is less than epsilon, i.e.

$$\left| \sum_{i=1}^{n} (a_i - a_{i-1})\, f(b_i) - \int_a^b f(x)\, dx \right| < \varepsilon.$$

(h) (Immediate consequence of g.) Let f be integrable over $[a,b]$ and $\varepsilon > 0$. Then there exists a positive integer n such that for all $b_1, \ldots, b_n$ satisfying $b_i \in [a_{i-1}, a_i]$

$$\left| (\sum_{i=1}^{n} f(b_i))/n - \int_a^b f(x)\, dx \right| < \varepsilon.$$

(i) (another important consequence of g.) Let f be continuous on $[a,b]$. Then there exists a $c \in [a,b]$ such that

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

It is often called the "mean value theorem" of the integral calculus.

EXAMPLE 2.2.
(c) shows that in EXAMPLE 1.4 the function

$$f \;:\; x \to x \sin(x)^2$$

is integrable, whereas (h) implies that **int_step**(S, V) goes to $\int_1^3 x \sin(x)^2\, dx$ for $n \to \infty$. Hence

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{(1 + \frac{2\,i}{n}) \sin(1 + \frac{2\,i}{n})^2}{n} = \int_1^3 x \sin(x)^2\, dx.$$

Exercise 2.1.
Prove that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{\sqrt{n^2 - i^2}}{n^2} = \pi/4.$$

Hint: Look at $\int_0^1 \sqrt{1 - x^2}\, dx$.

3 Integration in closed form

Let us recall the so-called *fundamental theorem of calculus*.

Theorem 1. Let f be a continuous function on the interval $[a,b]$. For any $c \in [a,b]$ define $g(c) = \int_a^c f(x)\, dx$. Then $c \to g(c)$ is a differentiable function on $[a,b]$ and $g'(c) = f(c)$ for all $c \in [a,b]$.

This theorem is extremely useful. Because if we have to compute the value $\int_a^b f(x)\, dx$ and we are so fortunate to find a function g such that $g' = f$, then

$$\int_a^b f(x)\, dx = g(b) - g(a).$$

Such a function g is called a *primitive of* f (or *anti-derivative of* f, *indefinite integral of* f). g is determined up to an additive constant. The usual notation for $g(x)$ is $\int f(x)\, dx$. Now we can compute immediately lots of integrals:
EXAMPLE 3.1.

```
> Int(3*x^2-5*x^7,x=3..5);
```

$$\int_3^5 3\,x^2 - 5\,x^7\, dx$$

In order to compute this integral we must find a function g such that

```
> diff(g(x),x)=3*x^2-5*x^7;
```

$$\frac{\partial}{\partial x}\,\mathrm{g}(x) = 3\,x^2 - 5\,x^7$$

One sees immediately that

```
> g:=x->x^3-5/8*x^8;
```

$$g := x \to x^3 - \frac{5}{8}\,x^8$$

is such a function. And so

```
> Int(3*x^2-5*x^7,x=3..5)=g(5)-g(3); g:='g':
```

$$\int_3^5 3\,x^2 - 5\,x^7\,dx = -239942$$

EXAMPLE 3.2.

```
> Int(2*x*exp(x^2),x=1..2);
```

$$\int_1^2 2\,x\,e^{(x^2)}\,dx$$

We must find a function g such that

```
> diff(g(x),x)=2*x*exp(x^2);
```

$$\frac{\partial}{\partial x}\,\mathrm{g}(x) = 2\,x\,e^{(x^2)}$$

Here we are “fortunate” again, because one sees at once that

```
> g:=x->exp(x^2);
```

$$g := x \to e^{(x^2)}$$

will do. So we find

```
> Int(x*exp(x^2),x=1..2)=g(2)-g(1);
```

$$\int_1^2 x\,e^{(x^2)}\,dx = e^4 - e$$

Apply **evalf** if you want a numerical result,

```
> evalf(g(2)-g(1)); g:='g':
```

$$51.87986820$$

EXAMPLE 3.3.

```
> Int(ln(x)^2,x);
```

$$\int \ln(x)^2\,dx$$

Now we should find g such that

```
> diff(g(x),x)=ln(x)^2;
```

$$\frac{\partial}{\partial x}\,\mathrm{g}(x) = \ln(x)^2$$

This is a bit harder. However, we can proceed as follows. We can find a g which is *almost* OK:

```
> g:=x->x*ln(x)^2;
```

$$g := x \to x\ln(x)^2$$

Then

```
> diff(g(x),x);
```

$$\ln(x)^2 + 2\ln(x)$$

In this expression we should get rid of the term $2\ln(x)$. So if we can find a function h satisfying $h'(x) = 2\ln(x)$, we are done, because then we can replace g by $g - h$. So we are reduced to a similar, but simpler problem. A candidate for h is given by $h(x) = 2x\ln(x)$. Then $h'(x) = 2\ln(x) + 2$. Now we must get rid of the term 2. This is easy: replace $h(x)$ by $h(x) - 2x$:

```
> h:=x->2*x*ln(x)-2*x;
```

$$h := x \to 2\,x\ln(x) - 2\,x$$

```
> diff(h(x),x);
```

$$2\ln(x)$$

This is what we were looking for. Hence our g becomes

```
> g:=x->(x*ln(x)^2-(2*x*ln(x)-2*x));
```

$$g := x \to x\ln(x)^2 - 2\,x\ln(x) + 2\,x$$

Check:

```
> diff(g(x),x);
```

$$\ln(x)^2$$

The trick applied twice in the latter example is the so-called *integration by parts* which can be expressed by the following formula

$$\int f(x)\,(\frac{\partial}{\partial x}\,g(x))\,dx = f(x) * g(x) - \int (\frac{\partial}{\partial x}\,f(x))\,g(x)\,dx.$$

This is valid for continuously differentiable functions f and g. The proof is by differentiation.

Exercise 3.1.
Compute the primitives of the following functions. Check the results by differentiating:

- $\int \frac{1}{(x+1)^2}\,dx$
- $\int x\sin(x)\,dx$
- $\int x^2\cos(x)\,dx$

Exercise 3.2. (cf. EXAMPLE 1.4.)
Compute $\int_1^3 x\sin(x)^2\,dx$.
Hint: use the relation $\cos(2x) = 1 - 2\sin(x)^2$.
EXAMPLE 3.4.

```
> Int(1/(x*ln(x)),x);
```

$$\int \frac{1}{x\ln(x)}\,dx$$

(restricted to $x > 1$). Here integration by parts seems powerless. In textbooks you may find:

$$\int \frac{1}{x\ln(x)}\,dx = \int \frac{1}{\ln(x)}\,d\ln(x) = \ln(\ln(x)).$$

The first relation makes sense when one applies the thumb rule: $dg(x) = g'(x)dx$; the second one by putting $\ln(x) = y$, namely

$$\int \frac{1}{y}\,dy = \ln(y)$$

and substituting by $y = \ln(x)$. This juggling can be turned into a precise theorem:

Theorem 2. ("Substitution rule") Let g be a continuously differentiable function on $[a, b]$ and f a continuous function defined for all $g(c)$ where $c \in [a, b]$. Then

$$\int_{g(a)}^{g(y)} f(x)\, dx = \int_a^y f(g(y))\,(\frac{\partial}{\partial y}\, g(y))\, dy$$

for all $y \in [a, b]$.

The proof of this useful theorem is very simple: differentiate both sides of the formula with respect to y. Then one gets an equality. So both members differ by a constant (i.e. independent of y). Finally, substitute $y = a$. Often the substitution rule is written in a less careful way:

$$\int f(x)\, dx = \int f(g(y))\,(\frac{\partial}{\partial y}\, g(y))\, dy.$$

Example 3.4. now goes as follows: let

$$f(x) = 1/(x \ln(x))$$

and

$$g(y) = \exp(y).$$

Then

$$f(g(y)) = 1/(y \exp(y))$$

and

$$g'(y) = \exp(y).$$

Hence we find

$$\int f(x)\, dx = \int \frac{1}{y}\, dy = \ln(y) = \ln(\ln(x)).$$

EXAMPLE 3.5.

```
>  Int(1/sqrt(1-x^2),x);
```

$$\int \frac{1}{\sqrt{1 - x^2}}\, dx$$

Here the substitution $x = \sin(y)$ is helpful. To be a bit more precise the function f$(x) = \frac{1}{\sqrt{1-x^2}}$ is defined and it is greater than 0 for $-1 < x < 1$ and $g(y) = \sin(y)$ is restricted to $-\pi < y < \pi$. Then

$$f(g(y)) = \frac{1}{\sqrt{1 - \sin(y)^2}} = \frac{1}{\sqrt{\cos(y)^2}} = \frac{1}{\cos(y)},$$

because $\cos(y) > 0$ for the values of y under consideration. Since $g'(y) = \cos(y)$ the substitution formula yields $\int f(x)\, dx = y$. Obviously, we want y as a function of x. This is easy: $y \to x$ is the inverse function arcsin of sin. Hence

$$\int \frac{1}{\sqrt{1 - x^2}}\, dx = \arcsin(x).$$

Exercise 3.3.
Compute $\int \frac{1}{1+\sin(x)}\, dx$ and check your result by differentiating.
Hint: substitute $x = 2 \arctan(t)$. First show that

- $\sin(x) = \frac{2\,t}{1+t^2}$,
- $\cos(x) = \frac{1-t^2}{1+t^2}$,
- $\frac{\partial}{\partial t}\, 2 \arctan(t) = \frac{2}{1+t^2}$.

The substitution rule now shows that our integral is replaced by $\int g(t)\, dt$ where g is a rational function of t.
Remark: This method works for any integral of the form $\int f(\sin(x),\, \cos(x))\, dx$, where f is rational function of two variables.
Exercise 3.4.
Compute $\int \frac{1}{\sin(x)+\cos(x)}\, dx$.

4 How the computer works?

The method applied in the two preceding exercises is not obvious. There are many more methods for computing primitives of special classes of functions. Until recently the computation of primitives of *elementary functions* (an elementary function is traditionally meant to be an element of the set of functions that can be built up recursively by starting from $\mathbb{Q}(x)$ and taking sums, products, quotients, logarithms, exponentials, algebraic functions and compositions of them) came down to skillfully applying a whole "bag of tricks". This has been so for centuries and many great mathematicians have been occupied with this problem. A practical approach consists of making *integral tables*, large collections of functions and their primitives. Only since the fundamental work of Robert Risch we dispose of an algorithm which decides whether the primitive of an elementary function is again an elementary function, and if it is, it will be computed. His algorithm is (partly) implemented in the bigger computer algebra systems (like Maple). It is too complicated to explain here. However, because of the Risch algorithm there is less need of exercises with the bag of tricks. So we shall not dive into integration methods for special classes of functions. Instead some examples of Maple's **int** procedure will be shown.

EXAMPLE 4.1.

```
> Int(1/sqrt(3-2*x-x^2),x);
```

$$\int \frac{1}{\sqrt{3-2\,x-x^2}}\,dx$$

```
> value(%);
```

$$\arcsin(\frac{1}{2}+\frac{1}{2}\,x)$$

Check by differentiating!

EXAMPLE 4.2.

```
> Int(sin(x)^2*cos(x)^5,x);
```

$$\int \sin(x)^2\cos(x)^5\,dx$$

```
> value(%);
```

$$-\frac{1}{7}\sin(x)\cos(x)^6+\frac{1}{35}\cos(x)^4\sin(x)+\frac{4}{105}\cos(x)^2\sin(x)+\frac{8}{105}\sin(x)$$

EXAMPLE 4.3.

```
> Int((2*x^3-2*x^2-1)/(x-1)^2*exp(x^2),x);
```

$$\int \frac{(2\,x^3-2\,x^2-1)\,e^{(x^2)}}{(x-1)^2}\,dx$$

```
> value(%);
```

$$\frac{x\,e^{(x^2)}}{x-1}$$

Now let us see how Maple is working when computing such integrals.

```
> infolevel[int]:=2:
> Int((-exp(x)-x+ln(x)*x+ln(x)*x*exp(x))/(x*(exp(x)+x)^2),x);
```

$$\int \frac{-e^x-x+\ln(x)\,x+\ln(x)\,x\,e^x}{x\,(e^x+x)^2}\,dx$$

```
> value(%);
> infolevel[int]:=0:
```

```
int/indef:  first-stage indefinite integration
int/indef2:  second-stage indefinite integration
int/indef2:  trying integration by parts
int/rischnorm:  enter Risch-Norman integrator
int/risch:  enter Risch integration
int/risch/algebraic1:  RootOfs should be algebraic numbers and functions
int/risch:  the field extensions are
```

$$[\_X,\, e^{-\_X},\, \ln(\_X)]$$

`int/risch:  Introduce the namings:`

$$\{\_th_1 = e^{-\_X},\, \_th_2 = \ln(\_X)\}$$

`unknown:  integrand is`

$$\frac{-\_th_1 - \_X + \_X\,\_th_2 + \_th_2\,\_X\,\_th_1}{\_X\,(\_th_1 + \_X)^2}$$

`int/risch/logpoly:  integrating`

$$\frac{-\_th_1 - \_X + \_X\,\_th_2 + \_th_2\,\_X\,\_th_1}{\_X\,(\_th_1 + \_X)^2}$$

`int/risch/int:  integrand is`

$$\frac{1 + \_th_1}{(\_th_1 + \_X)^2}$$

`int/risch/ratpart:  integrating`

$$\frac{1 + \_th_1}{(\_th_1 + \_X)^2}$$

`int/risch/ratpart:  Hermite reduction yields`

$$-\frac{1}{\_th_1 + \_X} + \int 0\, d\_X$$

`int/risch/int:  integrand is`

$$0$$

`int/risch/int:  integrand is`

$$0$$

`int/risch/logpoly:  result is`

$$-\frac{\_th_2}{\_th_1 + \_X}$$

`int/risch:  exit Risch integration`

$$-\frac{\ln(x)}{e^x + x}$$

What we have seen? At the first stage of the integration routine Maple tries to simplify the problem by applying the known integration rules for polynomials (rule (e) in section 2.). Next, (as any calculus student would) tries to integrate using simple table lookup processes. Then looks for other specific types of intergrands and uses appropriate methods: integration by part technique, possible substitutions (here called derivative-devides method). This heuristic method used by Maple obtains the correct answer for a suprisingly large percentage of integral problems. (And even very quickly). If the heuristic part fails then the problem is converted into the *exp-log* notation and a finite decision procedure is invoked. The result of the latter procedure will be either the integral expressed in the exp-log notation or an indication that could not find the elementary integral in which cases $\int f(x)\,dx$ returns itself. Does it mean that the elementary integral does not exist? Not by all means. In such cases some very elementary trick could help.

EXAMPLE 4.4. — Don't trust Maple's integrator too much.

```
> f:=x->log(x+sqrt(x^2+1));
```

$$f := x \to \log(x + \sqrt{x^2+1})$$

```
> Int(f(x),x);
```

$$\int \ln(x + \sqrt{x^2+1})\,dx$$

```
> value(%);
```

$$\int \ln(x + \sqrt{x^2+1})\,dx$$

Would it mean that the result is not an elementary function? Putting

$$f(x) = \ln(x + \sqrt{x^2+1}),$$

the partial integration yields

```
> Int(f(x),x) = x*f(x)-Int(x*diff(f(x),x),x);
```

$$\int \ln(x + \sqrt{x^2+1})\,dx = x\ln(x + \sqrt{x^2+1}) - \int \frac{x\,(1 + \frac{x}{\sqrt{x^2+1}})}{x + \sqrt{x^2+1}}\,dx$$

```
> value(%);
```

$$\int \ln(x + \sqrt{x^2+1})\,dx = x\ln(x + \sqrt{x^2+1}) - \sqrt{x^2+1}.$$

Exercise 4.1.
Compute $\int e^{\arcsin(x)}\,dx$.

5 Improper integrals

The notion of integral as defined above is rather restrictive, many extensions and generalizations have been given. Here two simple extensions will be presented, one to unbounded functions and another one to infinite intervals. Instead of stating general definitions and theorems we shall look at some specific examples.

Consider the function

$$f\ :\ x \to \frac{1}{\sqrt{x}}$$

on $[0,1]$. The function f is not well-defined in $x = 0$. Let us put $f(0) = 0$ (or any other real value). Now look at

```
> Int(1/sqrt(x),x=y..1) = int(1/sqrt(x),x=y..1);
```

$$\int_y^1 \frac{1}{\sqrt{x}}\,dx = 2 - 2\sqrt{y}$$

where y is a positive number less than 1. Obviously this expression has 2 as the limit when $y \to 0$. Define

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{y\to 0} \int_y^1 \frac{1}{\sqrt{x}}\,dx.$$

Then we have

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = 2.$$

The integral on the left is called an *improper integral*. Now let us see the extension to infinite intervals. Look at the following example:

```
> Int(1/x^2,x=1..y)=int(1/x^2,x=1..y);
```

$$\int_{1}^{y} \frac{1}{x^2}\,dx = -\frac{1}{y} + 1$$

The right-hand member has limit 1 for $y \to \infty$. Defining

$$\int_{1}^{\infty} \frac{1}{x^2}\,dx = \lim_{y \to \infty} \int_{1}^{y} \frac{1}{x^2}\,dx$$

we get

$$\int_{1}^{\infty} \frac{1}{x^2}\,dx = 1.$$

The left-hand side is again called an improper integral. Next we have an example of a *doubly improper* integral:

```
> Int(exp(-x)*(2*x-1)/sqrt(x),x=0..infinity);
```

$$\int_{0}^{\infty} \frac{e^{(-x)}\,(2\,x-1)}{\sqrt{x}}\,dx$$

The interval is infinite and the integrand is unbounded near $x = 0$. Now for $0 < y < z$ look at

```
> Int(exp(-x)*(2*x-1)/sqrt(x),x=y..z)=value(%);
```

$$\int_{y}^{z} \frac{e^{(-x)}\,(2\,x-1)}{\sqrt{x}}\,dx = -2\,\sqrt{z}\,e^{(-z)} + 2\,\sqrt{y}\,e^{(-y)}$$

When $y \to 0$ and $z \to \infty$ (independently), the right-hand side tends to 0. Hence

$$\int_{0}^{\infty} \frac{e^{(-x)}\,(2\,x-1)}{\sqrt{x}}\,dx = 0.$$

There are many interesting improper integrals, e.g. the Dirichlet integral

```
> Int(sin(x)/x,x=0..infinity);
```

$$\int_{0}^{\infty} \frac{\sin(x)}{x}\,dx$$

```
> value(%);
```

$$\frac{1}{2}\,\pi.$$

Maple doesn't compute this integral but *knows* it. We shall skip the proof.
Another famous one is the gamma-function, defined by

$$\Gamma(x) = \int_{0}^{\infty} t^{(x-1)}\,e^{(-t)}\,dt$$

for all $x \geq 1$.
Exercise 5.1.
Prove that

- $\Gamma(1) = 1$,
- $\Gamma(x+1) = x\Gamma(x)$,
- $\Gamma(n) = (n-1)!$ when n is a positive integer.

Hint: integration by parts.
Last but not least, some good advice: use your brains and think twice before computing. The integral

$$\int_{-\pi}^{\pi} \frac{\sin t}{1 + \sin^8 t} dt$$

seems to be hard to evaluate but there is also no need. You can almost immediately conclude that the integrand is odd and therefore the integral must be equal to zero.

6 Numerical integration

We have seen that some functions cannot be integrated in terms of elementary functions. This is not as bad as you may expect, because in practice it often suffices to find an approximate value of the definite integral and there exists good methods of numerical approximation. We shall now discuss the simplest and most obvious methods called *quadrature techniques*. We wish to direct special attention to the fundamental fact that the meaning of an approximate calculation is not precise unless it is supplemented by an estimate of the errors occurring.

The notations and conventions used here are:

$$I = \int_a^b f(x)\,dx,$$

where a and b are constants $(a < b)$ and $f(x)$ is given as a continuous differentiable function. Then we know that the integral exists. We have seen in Chapter 2 that we may restrict to regular subdivisions, therefore let the interval of integration be divided into n equal parts of length $h = (b-a)/n$. We denote the points of subdivisions by $x_0 = a, x_1 = a+h, \ldots, x_n = b$, and the values of the function at the points of division by $f_0, f_1, \ldots, f_n$. Similarly, the values of the function at the midpoints of the intervals by $f_{\frac{1}{2}}, f_{\frac{3}{2}}, \ldots, f_{\frac{2n-1}{2}}$. It is fairly natural to use the Riemann sum approximation to I, as we did it in section 2. The procedure **subdiv**(a,b,n) — we shall use it in the next examples — returns a regular subdivision $[a = x_0, x_1, \ldots, x_n = b]$.

EXAMPLE 6.1.

```
> f:=x->1/x: S:= subdiv(1,2,10):
> integral_plot(f,S,'riemann');
```

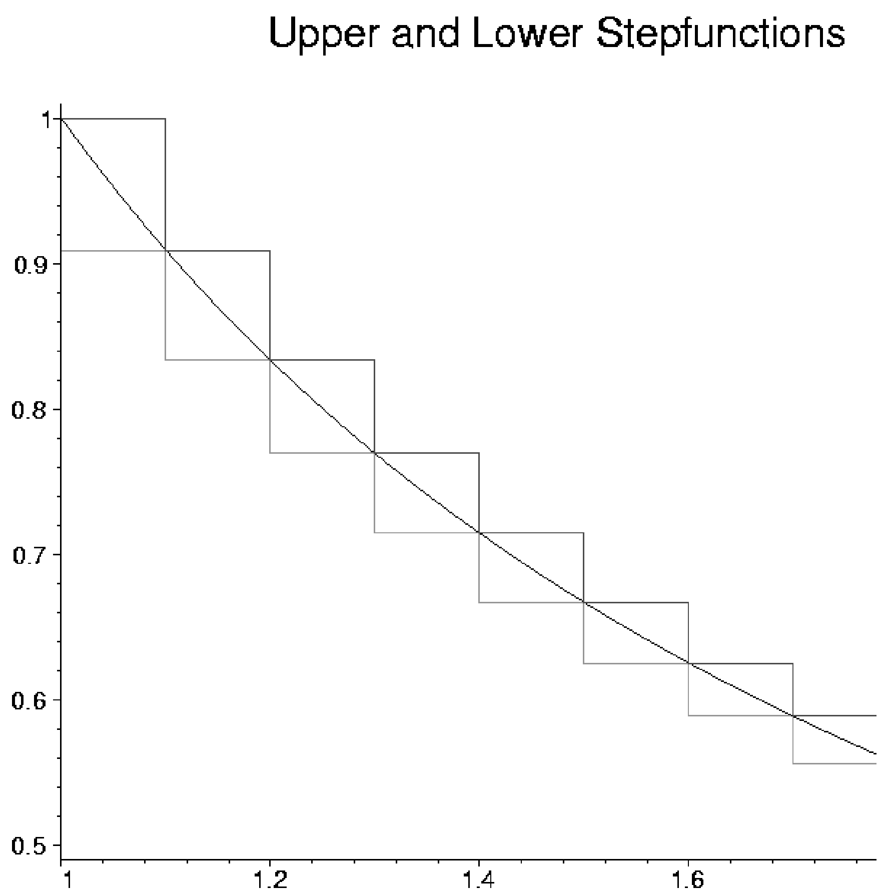


Figure 7:

```
> upper(f,S);
```

$$.7187714032$$

```
> lower(f,S);
```

$$.6687714032$$

We have a lower approximation for I as 0.66877 and an upper one, which is 0.71877. These differ by 0.05 so there may be an error greater then 7% if either of these values is taken as an estimate for the integral. "OK, let us simply choose a finer subdivision and we shall get a better approximation" would be the next idea. Observe, however, that we have to calculate the minimum (or the maximum) of the function in every subinterval $[x_i, x_{i+1}]$, which is rather costly and error-prone. This method is easy to apply if the integrand $f(x)$ is known to be monotonic. And the other cases? The technique mentioned above is not a very good one from practical point of view. We require better techniques to satisfy two important criteria:

1. The technique should give results of high accuracy.

2. The technique should require only a small number of function evaluations.

In most cases the Riemann-sum method does not satisfy the criteria 1 and 2 but it enables us to formulate other methods realizing better these desirable properties. The most obvious method for approximating to I without calculating maximum (or minimum) values on the subintervals is directly connected with the good old stepfunctions. For the value of the stepfunction on $[x_i, x_{i+1}]$ we choose f_i. Then for the integral we have got an approximate expression:

$$I \approx h(f_0 + f_1 + \ldots + f_{n-1}).$$

This method is called *rectangle rule*. (Here and hereafter the symbol $\approx$ means "is approximately equal to".)
EXAMPLE 6.2.

```
> f:=x->x^(1/2)*sin(x)*cos(x): n:=6: S:=subdiv(0,2,n):
> integral_plot(f,S,'rectangle');
> int_num(f,S,'rectangle');
```

$$.3903146893$$

We obtain a closer approximation with no greater trouble if we replace the rectangle area by a trapezoid area $h(f_i + f_{i+1})/2$. For the whole integral this gives the approximate expression:

$$I \approx h(f_1 + f_2 + \ldots + f_{n-1}) + h(f_0 + f_n)/2.$$

This is the *trapezoid formula*, since, when the areas of the trapezoids are added, each value of the function except the first and the last occurs twice.

```
> integral_plot(f,S,'trapezoid');
> int_num(f,S,'trapezoid');
```

$$.3011246599$$

The approximation becomes even better if instead of choosing the trapezoid under a chord we chose the trapezoid under the tangent to the curve at the point with the abscissa $x = x_i + \frac{h}{2}$. The area of this trapezoid is simply $hf_{i+\frac{1}{2}}$, and the approximation for the entire integral is

$$I \approx h(f_{\frac{1}{2}} + f_{\frac{3}{2}} + \ldots + f_{\frac{2\,n-1}{2}})$$

which is called the *tangent formula*. Let us see how it is work.

```
> integral_plot(f,S,'tangent');
> int_num(f,S,'tangent');
```

$$.3187318652$$

The next method is called *mid-point rule*. In this case we assume the function values between x_i and x_{i+1} are constant and have the value equal to the function value at $(x_i + x_{i+1})/2$.

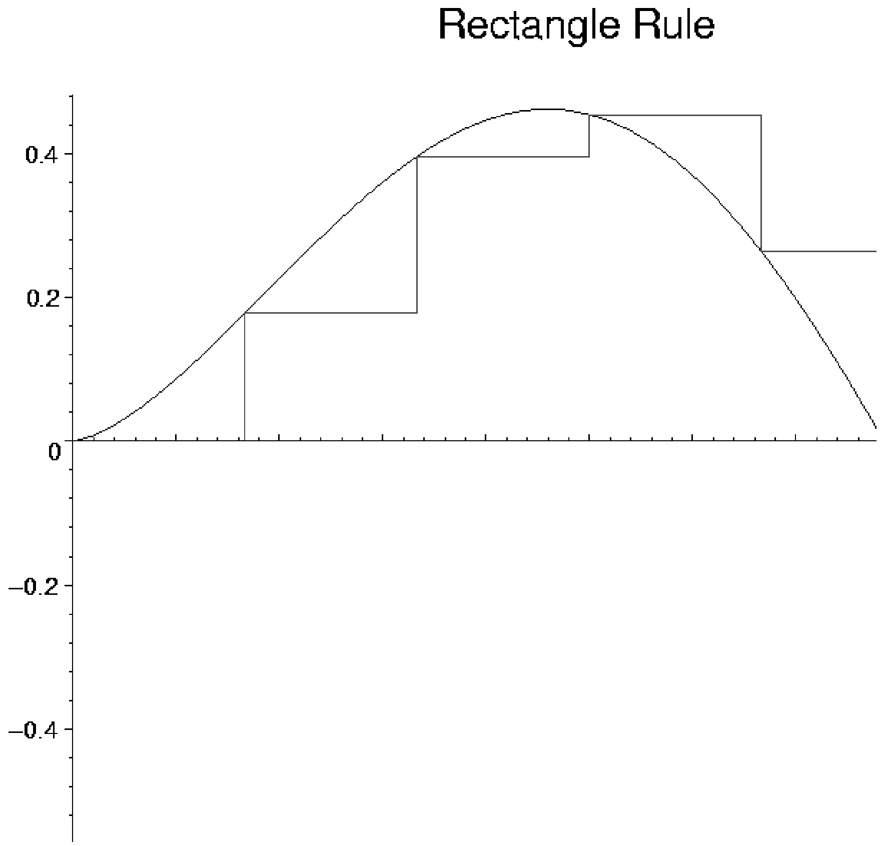


Figure 8:

```
> integral_plot(f,S,'mid-point');
> int_num(f,S,'mid-point');
```

$$.3187318652$$

Observe, that the values of the approximating integral in the case of the tangent method and the mid-point method are the same. Is it a coincidence? Explain why not.

The next method depends on estimating the subarea of the integral at two adjacent subintervals, i.e. between the abscissa x_i and $x_i + 2h = x_{i+2}$ by considering the upper boundary to be no longer a straight line but a parabola. To be more precise, the parabola which passes through the three points of the curve with abscissa x_i, x_{i+1}, x_{i+2}. The equation of the parabola is

$$y = f_i + \frac{(x - x_i)\,(f_{i+1} - f_i)}{h} + \frac{(x - x_i)\,(x - x_i - h)\,(f_{i+2} - 2\,f_{i+1} + f_i)}{2\,h^2}.$$

Exercise 6.1. Prove the latter statement.
After a brief calculation we get the area under the parabola:

$$\frac{h\,(f_i + 4\,f_{i+1} + f_{i+2})}{3}.$$

If we assume that $n = 2m$, i.e. that n is even, by the addition of the subareas we obtain the *Simpson's rule:*

$$I \approx \frac{h\,(4\,(\sum_{i=1}^{m}\ f_{2\,i-1}) + 2\,(\sum_{i=1}^{m-1}\ f_{2\,i}) + f_0 + f_{2\,m})}{3}.$$

```
> integral_plot(f,S,'simpson');
```

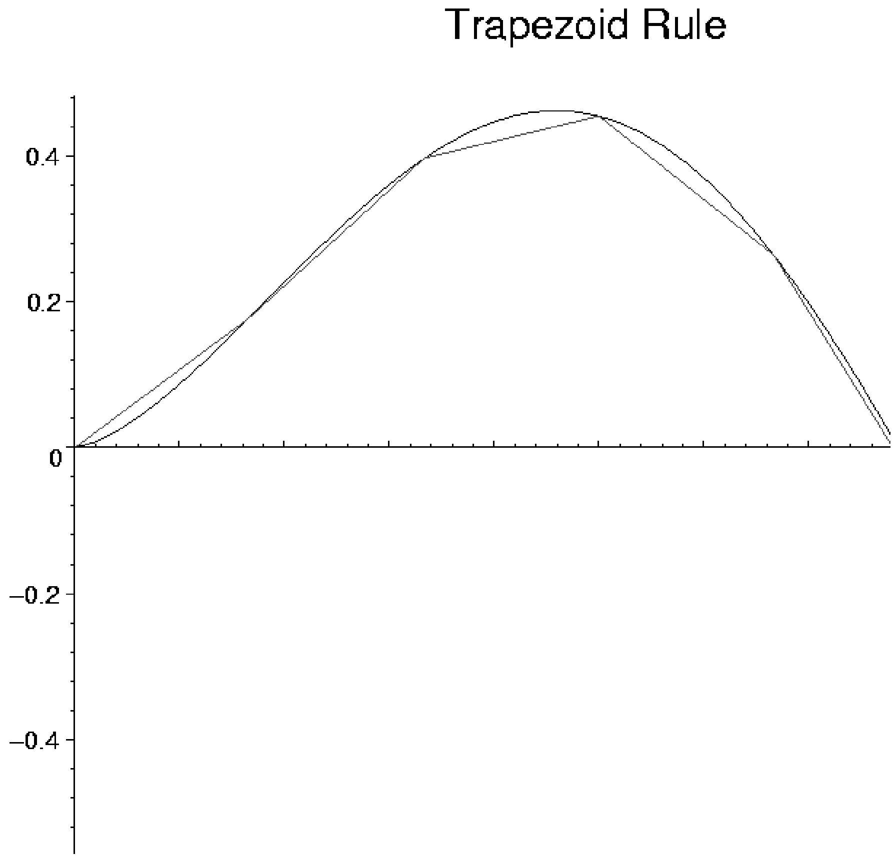


Figure 9:

```
> int_num(f,S,'simpson');
```

$$.3141154642$$

The integral of $f(x)$ can also be expressed in closed form.

```
> Int(x^(1/2)*sin(x)*cos(x),'x'=0..2)=int(f(x),'x'=0..2);
```

$$\int_0^2 \sqrt{x}\,\sin(x)\cos(x)\,dx = -\frac{1}{4}\cos(4)\sqrt{2} + \frac{1}{8}\sqrt{\pi}\,\mathrm{FresnelC}(2\,\frac{\sqrt{2}}{\sqrt{\pi}})$$

(If you don't agree, feel free to check!) The approximate value:

```
> evalf(int(x^(1/2)*sin(x)*cos(x),'x'=0..2));
```

$$.3126735508$$

It is worth to compare the results getting by the different methods:

- Rectangle rule: .3903146893
- Trapezoid rule: .3011246599
- Tangent rule: .3187318652
- Simpson's rule: .3141154642
- Maple approximation of the closed form: .3126735508

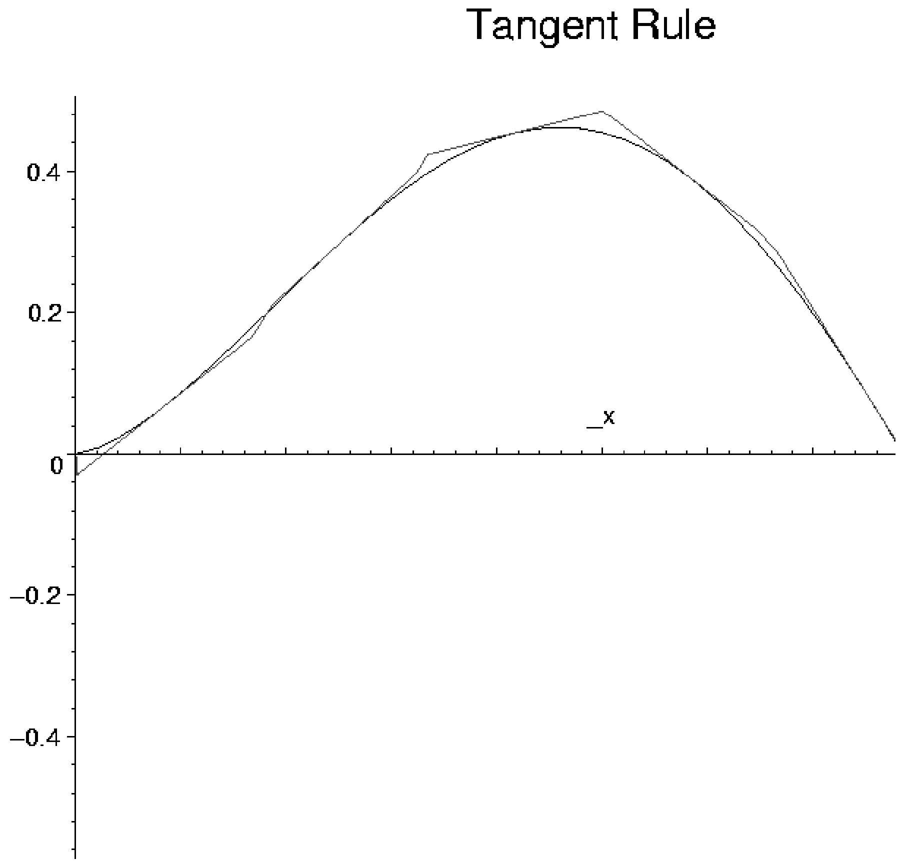


Figure 10:

In our example the Simpson's rule seems to be the best concerning the criteria 1 and 2. Is this always true or only a coincidence again? We need an estimation of the error.

Estimations can be easily given for each of our methods, if bounds for the derivatives of the function $f(x)$ are known throughout the interval of integration. Let $M_1, M_2, \ldots$ be upper bounds for the absolute values of the first, second, ... derivatives, i.e. we assume that throughout the interval $\left|f^{(i)}(x)\right| < M_i$ for all i. Then the estimation formulas are as follows:

- Rectangle rule: $\frac{M_1\,(b-a)\,h}{2}$,
- Tangent rule: $\frac{M_2\,(b-a)\,h^2}{24}$,
- Trapezoid rule: $\frac{M_2\,h^3}{12}$,
- Simpson's rule: $\frac{M_4\,h^5}{90}$.

From the last two estimates there also follow estimates for the entire integral. We see, that Simpson's rule has an error of much higher order in the small quantity h than the others, so that where M_4 is not too large, it is very advantageous for practical calculations. The proofs of the estimates are quite simple using the Taylor's theorem and can be found in every textbooks on numerical methods.

There are other kinds of very important quadrature methods based on orthogonal polynomials (Gaussian quadratures). There are known (and used) methods if the integrand $f(x)$ is not finite or one of its derivatives is infinite at some point in $[a, b]$ (singular integrals). In many cases miscellaneous techniques are also available (series expansions, Laplace or Fourier transforms).

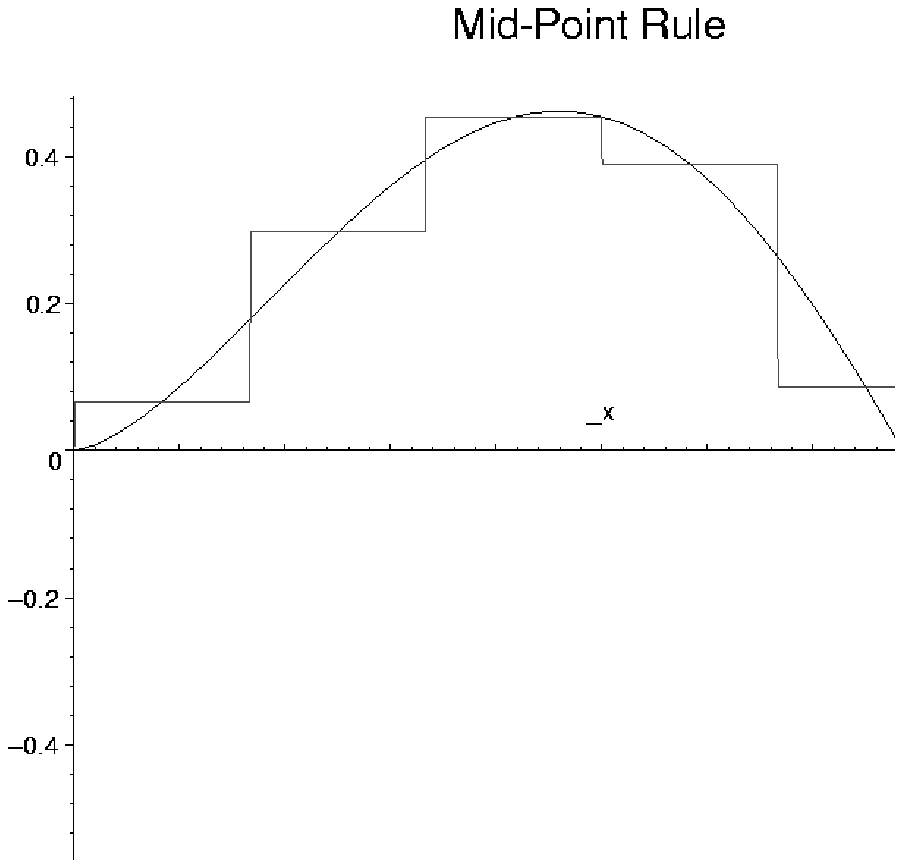


Figure 11:

Exercise 6.2.
(a) Take a larger value for n, the number of subintervals, redo the numerical integrations above and check the improvements in the approximations.
(b) The numerical computations can be improved by allow a higher value to the built-in Maple variable **Digits**. E.g.

```
> Digits:=20: int_num(f,S,'simpson'); Digits:=10:
```

$$.31411546415523086841$$

Recompute the numerical integral values of EXAMPLE 5.2 using 20 decimal "Digits".
(c) Calculate

```
> Int(exp(-x^2),'x'=0..infinity);
```

$$\int_0^{\infty} e^{(-x^2)}\, dx$$

to within 0.01.
(d) From the formula

```
> Pi/4 = Int(1/(1+x^2),'x'=0..1);
```

$$\frac{1}{4}\pi = \int_0^1 \frac{1}{x^2+1}\, dx$$

calculate π

- using the trapezoid formula with $h = 0.1$,

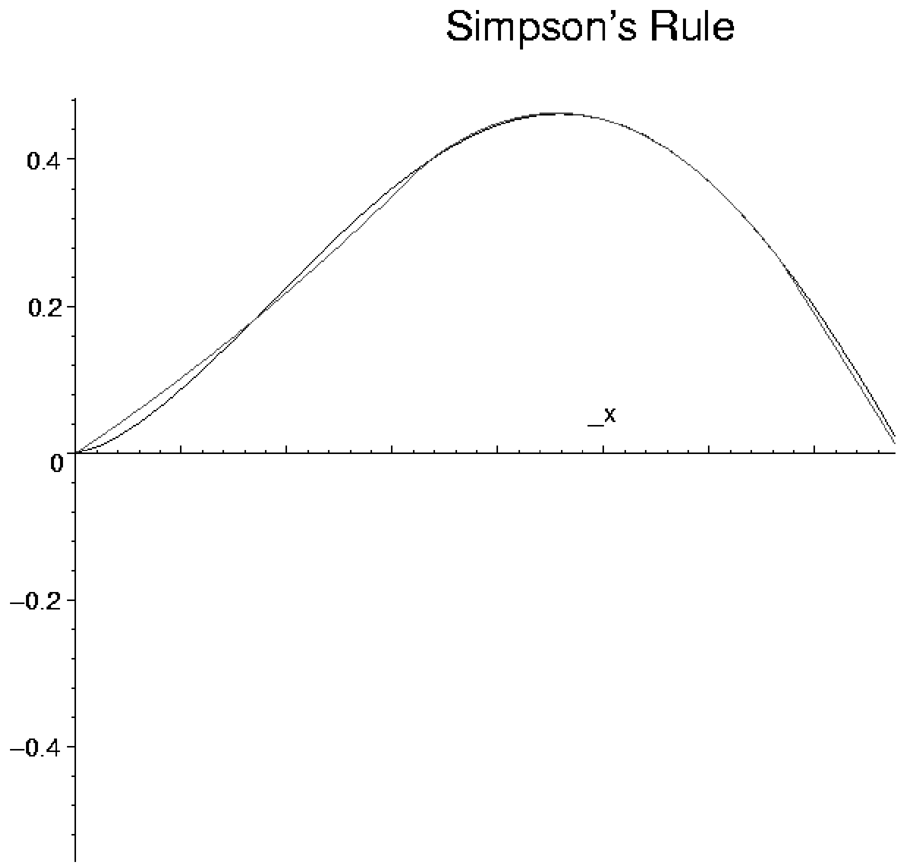


Figure 12:

- using Simpson's rule with $h = 0.1$.

(e) Calculate

```
>  Int(1/(sqrt(1+x^4)),'x'=0..1);
```

$$\int_0^1 \frac{1}{\sqrt{1+x^4}}\, dx$$

numerically with an error less than 0.1.

Exercise 6.3.

Using the trapezoid rule with n subdivisions — let us denode the integral approximation by I_n — calculate the following integrals with $n = 2, 4, 8, 16, \ldots, 512$.

- $\int_0^1 e^{-x^2} dx$,
- $\int_0^1 x^{5/2} dx$,
- $\int_{-4}^4 \frac{dx}{1+x^2}$,
- $\int_0^{2\pi} \frac{dx}{2+\cos(x)}$,
- $\int_0^{\pi} e^x \cos(4x) dx$.

Exercise 6.4.

Repeat the previous exercise using Simpson's rule.

7 Applications

7.1 The area as an integral

The idea of area was our starting-point for the definition of the integral; but the connection between definite integral and area is still incomplete. The areas with which we are concerned in geometry are bounded by given closed curves. On the other hand, the area measured by the integral $I = \int_a^b f(x)\,dx$ is bounded only in part by the given curve $y = f(x)$, the rest of the boundary consisting of lines which depend on the choice of the coordinate system. If we introduce t formally as a new independent variable in the above integral writing $x = x(t)$, $y = y(t) = f(x(t))$, we have

$$\int_a^b f(x)\,dx = \int_{t_0}^{t_1} \mathrm{y}(t)\,(\frac{\partial}{\partial t}\,\mathrm{x}(t))\,dt$$

where t_0 and t_1 are the values of the parameter corresponding to the abscissa $x_0 = a$ and $x_1 = b$ respectively. Here we suppose that every point of the branch of the curve $y = f(x)$ corresponds to a single value of t in the interval $t_0 \le t \le t_1$, and conversely; furthermore $f(x)$ is everywhere positive and $\frac{\partial}{\partial t}\,\mathrm{x}(t)$ never vanishes in this interval.

We can express our formula for the area in a more elegant symmetrical form if we first transform the integral by integration by parts:

$$\int_{t0}^{t1} \mathrm{y}(t)\,(\frac{\partial}{\partial t}\,\mathrm{x}(t))\,dt = -\int_{t0}^{t1} \mathrm{x}(t)\,(\frac{\partial}{\partial t}\,\mathrm{y}(t))\,dt + x * y\,\Big|_{t_0}^{t_1}.$$

Since the curve is closed,

$$x(t_0) = x(t_1),\ y(t_0) = y(t_1),$$

and therefore

$$I = -\int_{t0}^{t1} \mathrm{y}(t)\,(\frac{\partial}{\partial t}\,\mathrm{x}(t))\,dt = \int_{t0}^{t1} \mathrm{x}(t)\,(\frac{\partial}{\partial t}\,\mathrm{y}(t))\,dt.$$

If we form the arithmetic mean of the two expressions we obtain the symmetrical form

$$I = \frac{1}{2}\int_{t0}^{t1} \mathrm{y}(t)\,(\frac{\partial}{\partial t}\,\mathrm{x}(t)) - \mathrm{x}(t)\,(\frac{\partial}{\partial t}\,\mathrm{y}(t))\,dt.$$

EXAMPLE 7.1.
As an example of the application of our formula for the area consider the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. So we can define the function

```
> f:=x-> b/a*sqrt(a^2-x^2);
```

$$f := x \to \frac{b\sqrt{a^2 - x^2}}{a}$$

In order to find its area we take the upper and lower halves of the ellipse separately and in this way we can express the area by the integral

```
> 2*Int(f(x),x=-a..a)=value(2*int(f(x),x=-a..a));
```

$$2\int_{-a}^{a} \frac{b\sqrt{a^2 - x^2}}{a}\,dx = 2\int_{-a}^{a} \frac{b\sqrt{a^2 - x^2}}{a}\,dx$$

If, however, we use the parametric representation $x = a * \cos(t),\ y = b * \sin(t)$, we find that

```
> -1/2*Int(b*sin(t)*diff(a*cos(t),t)-
> a*cos(t)*diff(b*sin(t),t),t=0..2*Pi);
```

$$-\frac{1}{2}\int_0^{2\pi} -b\sin(t)^2\,a - a\cos(t)^2\,b\,dt$$

which has the value

```
> value(%);
```

$$b\,a\,\pi$$

In this subsection we have based the definition of the area on the concept of integral and have shown that this analytical definition has a truly geometrical character, since it yields a quantity independent of the coordinate system. It is, however, easy to give a direct geometrical definition of the area bounded by a closed curve which does not intersect itself, as follows: the area is the upper bound of the areas of all polygons lying interior to the curve. The proof that the two definitions are equivalent is quite simple, but will not be given here.

7.2 Areas in polar coordinates

For many purposes it is important to be able to calculate areas using polar coordinates. Let $r = f(\theta)$ be the equation of a curve in polar coordinates. Let I be the area of the region which is bounded by the x-axis (that is, the line $\theta = 0$), the line through the origin making an angle θ with the x-axis, and the portion of the curve between these two lines. By the fundamental theorem of the integral calculus, the area of the sector between the polar angles α and β is given by the expression

$$\frac{1}{2}\int_{\alpha}^{\beta} \mathrm{f}(\theta)^2\, d\theta.$$

If $\beta > \alpha$, this expression cannot be less than zero.
EXAMPLE 7.2.
Consider the area bounded by the one loop of a lemniscate. The equation of the lemniscate is $r^2 = 2\,a^2 \cos 2\,\theta$, and we obtain one loop by letting θ vary from $-\frac{\pi}{4}$ to $+\frac{\pi}{4}$. The shape of one loop of the lemniscate with $a = 1$ is

```
> plot([2*cos(2*theta),theta,theta=-Pi/4..Pi/4],coords=polar);
```

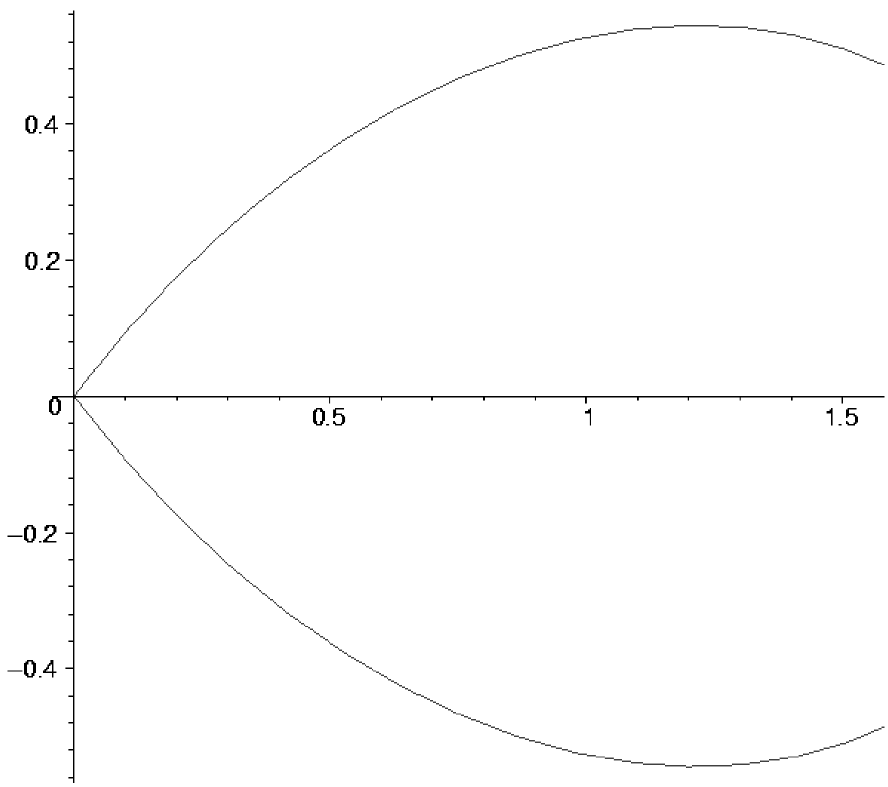


Figure 13:

The area of the loop in general is

```
> a^2*Int(cos(2*theta),theta=-Pi/4..Pi/4);
```

$$a^2 \int_{-1/4\,\pi}^{1/4\,\pi} \cos(2\,\theta)\,d\theta.$$

We find that the value of the integral is

```
> value(%);
```

$$a^2$$

7.3 Length of a curve

Another important geometrical concept associated with a curve leads to an integration. This is the length of arc. To express the length analytically by an integral, in fact, we think of the curve as represented by a function $y = f(x)$ with a continuous derivative y'. By the points $a = x_1, x_2, \ldots, x_n = b$ we divide up the interval $a \leq x \leq b$ of the x-axis, over which our curve lies, into $n-1$ subintervals of length $\Delta x_1, \ldots, \Delta x_{n-1}$. In the curve we inscribe a polygon whose vertices lie vertically above these points. The length of the curve is then defined to be the limit of the perimeters of these inscribed polygons, provided that such a limit does exist and is independent of the particular way in which the polygons are chosen. This assumption is called *rectifiability*. So the total length of the inscribed polygon is given according to Pythagoras theorem by the expression

$$\sum_{i=1}^{n-1} \sqrt{\Delta x_i^2 + \Delta y_i^2} = \sum_{i=1}^{n-1} \sqrt{1 + (\frac{\Delta y_i}{\Delta x_i})^2}\,\Delta\,x_i.$$

But by the mean value theorem of the differential calculus the difference quotient $\Delta y_i / \Delta x_i$ is equal to $f'(\xi_i)$, where ξ_i is an intermediate value in the interval Δx_i. If we now let n increase beyond all bounds and at the same time let the length of the longest subinterval Δx_i tend to zero, then by the definition of integral our expression will tend to the limit

$$\int_a^b \sqrt{1 + (\frac{\partial}{\partial x}\,\mathrm{y}(x))^2}\,dx.$$

We established the following theorem:

Theorem 3. Every curve $y = f(x)$ for which the derivative $f'(x)$ is continuous is a rectifiable curve, and its length between $x = a$ and $x = b$ ($b \geq a$) is given by the formula $\int_a^b \sqrt{1 + (\frac{\partial}{\partial x}\,\mathrm{y}(x))^2}\,dx$.

Our expression for the length of arc is still subject to the special and artifical assumption that the curve consists of one single-valued branch above the x-axis. Parametric representation frees us from this restriction. If a curve of the kind which we have been considering is given in parametric form by the equations $x = x(t),\ y = y(t)$, then by introducing the parameter t in the above expression we obtain the parametric form of the length of arc

$$\int_\alpha^\beta \sqrt{(\frac{\partial}{\partial t}\,\mathrm{x}(t))^2 + (\frac{\partial}{\partial t}\,\mathrm{y}(t))^2}\,dt,$$

where α and β are the values of t which correspond respectively to the points of the curve $x = a$ and $x = b$.
Excercise 7.1. Give the length of the arc when the curve is expressed in polar coordinates.
EXAMPLE 7.3.
Consider the parabola

```
> f:=x->1/2*x^2;
```

$$f := x \to \frac{1}{2}\,x^2$$

For its length of arc we immediately obtain the integral

```
> Int(sqrt(1+x^2),x=a..b);
```

$$\int_a^b \sqrt{1 + x^2}\,dx$$

which has the value

```
> value(%);
```

$$\frac{1}{2}\, b\, \sqrt{1+b^2} + \frac{1}{2}\ln(b+\sqrt{1+b^2}) - \frac{1}{2}\, a\, \sqrt{1+a^2} - \frac{1}{2}\ln(a+\sqrt{1+a^2})$$

EXAMPLE 7.4.
As an example for a motion along a path or trajectory consider the cycloids which arise when a circle rolls along a straight line or another circle. Here we limit ourselves to the simplest case, in which a circle of radius R rolls along the x-axis, and we consider a point on its circumference. This point then describes a cycloid. If we choose the origin of the coordinate system and the initial time in such a way that for time $t = 0$ the corresponding point of the curve coincides with the origin, we obtain the parametric representation

$$x = R(t - \sin(t)),\ y = R(1 - \cos(t))$$

for the cycloid. Here t denotes the angle through which the circle has turned from its original position. From the above equations we obtain at once that

$$\frac{\partial}{\partial t}\,\mathrm{x}(t) = R(1-\cos(t)),\ \frac{\partial}{\partial t}\,\mathrm{y}(t) = R\sin(t).$$

Hence the length of the arc is

```
> Int(sqrt(diff(x(t),t)^2+diff(y(t),t)^2),t=0..alpha)=
> Int(sqrt(2*R^2*(1-cos(t))),t=0..alpha);
```

$$\int_0^{\alpha} \sqrt{(\frac{\partial}{\partial t}\,\mathrm{x}(t))^2 + (\frac{\partial}{\partial t}\,\mathrm{y}(t))^2}\, dt = \int_0^{\alpha} \sqrt{2}\,\sqrt{R^2\,(1-\cos(t))}\, dt$$

Since $1 - \cos(t) = 2\sin(t/2)^2$ the integrand is equal to $2R\sin(t/2)$, hence for $0 \le \alpha \le 2\pi$ the equation becomes

```
> Int(2*R*sin(t/2),t=0..alpha);
```

$$\int_0^{\alpha} 2\,R\sin(\frac{1}{2}\,t)\, dt$$

The value of this integral is

```
> value(%);
```

$$-4\cos(\frac{1}{2}\,\alpha)\,R + 4\,R$$

If we consider the length of arc between two successive cusps we must put $\alpha := 2\,\pi$. Then we have

```
> eval(subs(alpha=2*Pi,%));
```

$$8\,R$$

Thus, we obtain that the lenght of arc of the cycloid between successive cusps is equal to four times the diameter of the rolling circle.

Similarly, we calculate the area bounded by one arch of the cycloid and the x-axis. If $R = 1$ then this area has the form

```
> plot([t-sin(t),1-cos(t),t=0..2*Pi]);
```

The area is

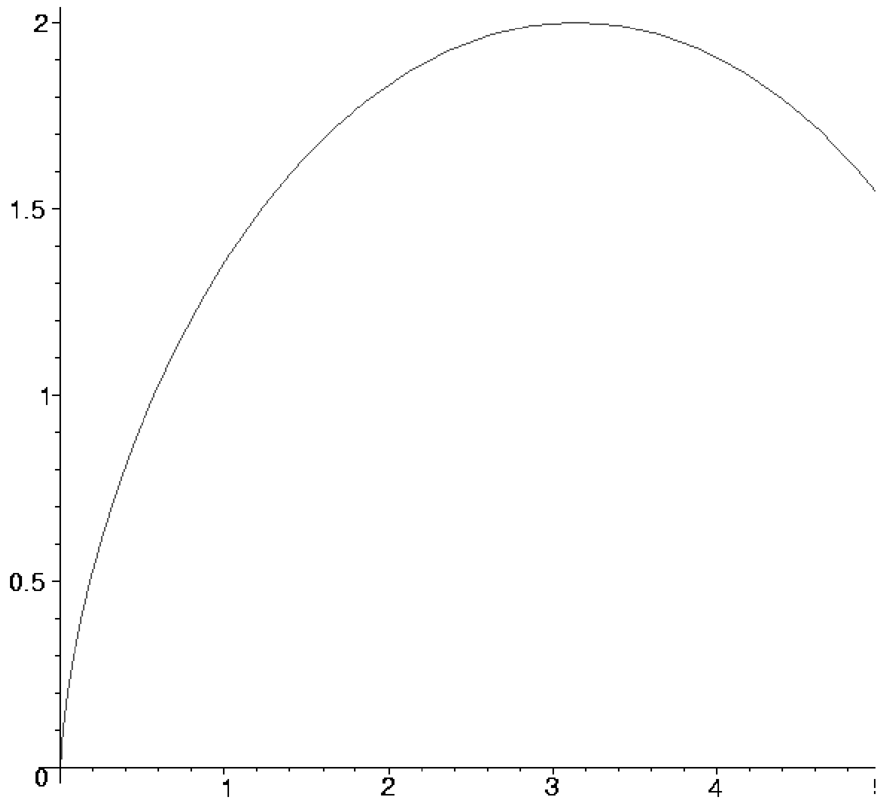
```
> Int(y(t)*diff(x(t),t),t=0..2*Pi)=
> R^2*Int((1-cos(t))^2,t=0..2*Pi);
```

$$\int_0^{2\,\pi} \mathrm{y}(t)\,(\frac{\partial}{\partial t}\,\mathrm{x}(t))\, dt = R^2 \int_0^{2\,\pi} (1-\cos(t))^2\, dt$$

```
> value(rhs(%));
```

$$3\,R^2\,\pi$$

This area is therefore three times the area of the rolling circle.



Exercise 7.2. Calculate the area bounded by the *semicubical parabola* $y = x^{\frac{3}{2}}$, the x-axis and the lines $x = a$ and $y = b$. Calculate the length of arc of it.
Exercise 7.3. Find the volume and surface area of the *torus* (or anchor ring) obtained by rotating a circle about a line which does not intersect it.
Exercise 7.4. Find the area of a *catenoid*, the surface obtained by rotating an arc of the catenary $y = \cosh(x)$ about the x-axis.

The possibilities of applications of differential and integral calculus are unbounded. In sciences and engineering mathematical models are developed to aid in the understanding of physical phenomena. These models often yield an equation that contains some derivatives of an unknown function. Such an equation is called differential equation. In order to solve these equations one requires the theory of integration. In this paper we did the first steps towards better understanding the mathematical and real world in which we live.

References

[1] Buchberger, B., *Should Students Learn Integration Rules?*, manuscript, Johannes Kepler University, Research Institute for Symbolic Computation, RISC-Linz, 1989.

[2] Courant, R., *Differential and integral calculus*, second edition, Blackie & Son Ltd., London, 1960.

[3] Kovács, A., *Komputeralgebra a tudományokban és a gyakorlatban*, Alk., Mat., Lapok, **19**, 1998, 181–202.

[4] Kovács, A., *Computer Algebra: Impact and Perspecives*, Neuw Archief voor Wiskunde, **17**/1, 1999, 29–55.

[5] Molnárka, Gy., Gergó, L., Wettl, F., Horváth, A., Kallós, G., *A MapleV és alkalmazásai*, Springer, Budapest, 1996.