Chapter 5. Numerical Integration

These are just summaries of the lecture notes, and few details are included. Most of what we include here is to be found in more detail in Anton.

5.1 Remark. There are two topics with similar names:

- Reverse of differentiation

  Indefinite integral

$$\int f(x)\,dx = \text{ most general antiderivative for } f(x)$$

- Definite integral

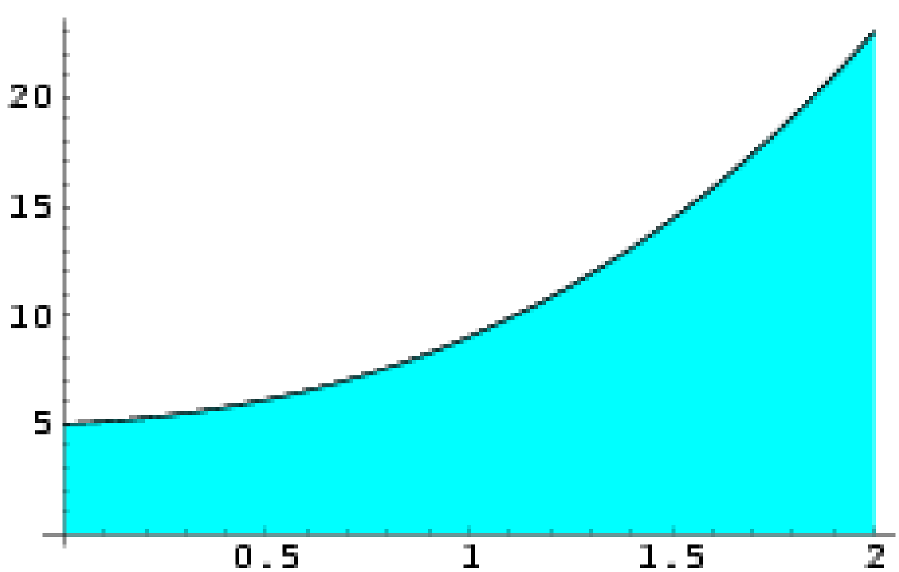
  This is related to summation (it is a limit of sums of a certain kind). The integral sign $\int$ was originally invented as a modified S (for sum).

There is no reason to expect a connection between these two different things, but there is. See course 1S1, the book by Anton or the appendix for more details.

What we need is the idea of a Riemann sum.

5.2 The Definite Integral. $\int_a^b f(x)\,dx$ (read integral from a to b of the function f). Answer is a number and does not involve x. Notation is convenient.

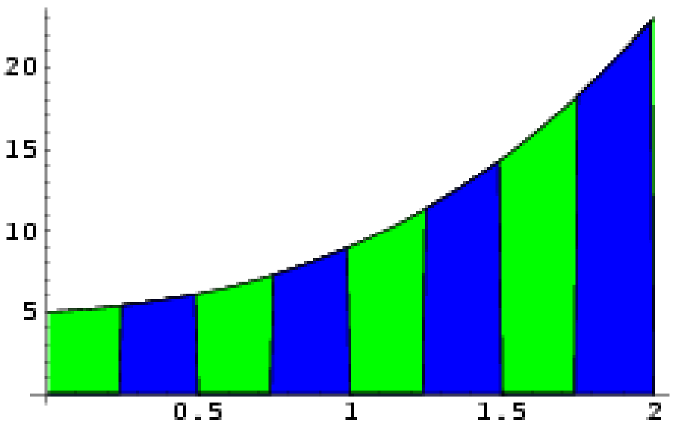
Graphical interpretation: $\int_a^b f(x)\,dx$ is the area of the region in the plane bounded by the graph $y = f(x)$, the x-axis and the vertical lines $x = a$, $x = b$.

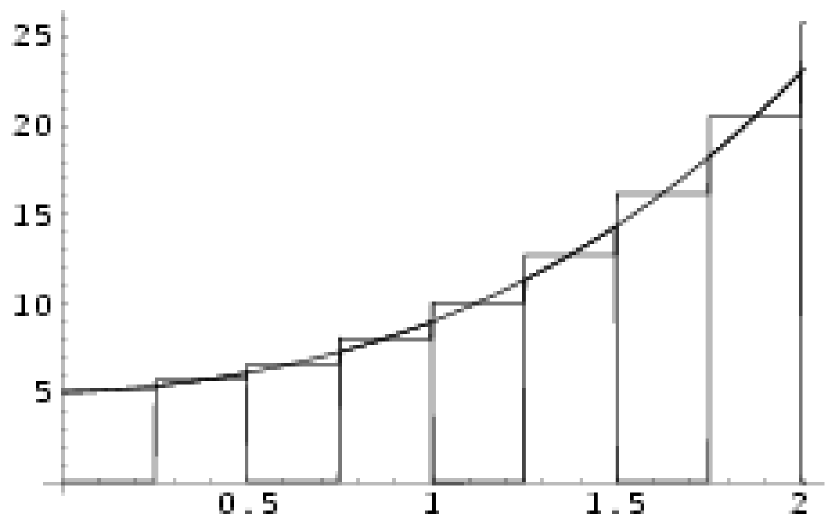


(picture good for case $f(x) \geq 0$)

Problems: need a way to compute the area of such a shape.

More fundamentally, need a definition of what you mean by the area.





5.3 Notation. We need notation to write down formulae for these pictures.

Dealing with $\int_a^b f(x)\,dx$

n subdivisions of interval $[a, b]$

Number division point $a = x_0 < x_1 < x_2 < \cdots < x_n = b$

Heights of rectangles $y_j = f(x_j^*)$, $x_{j-1} \leq x_j^* \leq x_j$ any point.

"Area" of jth rectangle = width × height = $(x_j - x_{j-1})y_j$

$$\begin{aligned}
\text{Total "area"} &= \text{sum of these} \\
&= y_1(x_1 - x_0) + y_2(x_2 - x_1) + \cdots + y_n(x_n - x_{n-1}) \\
&= f(x_1^*)(x_1 - x_0) + f(x_2^*)(x_2 - x_1) + \cdots + f(x_n^*)(x_n - x_{n-1})
\end{aligned}$$

(called *Riemann sums* for the integral)

5.4 Definition. $\int_a^b f(x)\,dx =$ limit of these Riemann sums as $n \to \infty$ and max width $\to 0$.

5.5 Theorem (Important Theorem)**.** *This limit makes sense if f is continuous on the finite closed interval $[a, b]$ (including end points).*

Proof of this is too hard for us.

5.6 Notation. We can use *Sigma notation* for sums to make the formulae look shorter.
For numbers $u_1, u_2, \ldots, u_n$

$$\sum_{i=1}^{n} u_i \quad \text{means} \quad u_1 + u_2 + \cdots + u_n$$

5.7 Examples.

$$\sum_{j=1}^{5} j^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 1 + 4 + 9 + 16 + 25 = 55$$

$$\sum_{k=1}^{25} 2k^2 + k = (2 + 1) + (2 \times 2^2 + 2) + \cdots + (2(25)^2 + 25)$$

5.8 Remark. Usual choice for Riemann sums: all n subintervals equally wide. Common width is then $h = \dfrac{b-a}{n}$

Then $x_j = a + jh$ for $0 \leq j \leq n$
Riemann sum becomes

$$\sum_{j=1}^{n} f(x_j^*)(x_j - x_{j-1}) = \sum_{j=1}^{n} f(x_j^*)h = h \sum_{j=1}^{n} f(x_j^*)$$

Limit of these is the integral. Limit hard to find directly as a rule, but a computer can find the sum for large n.

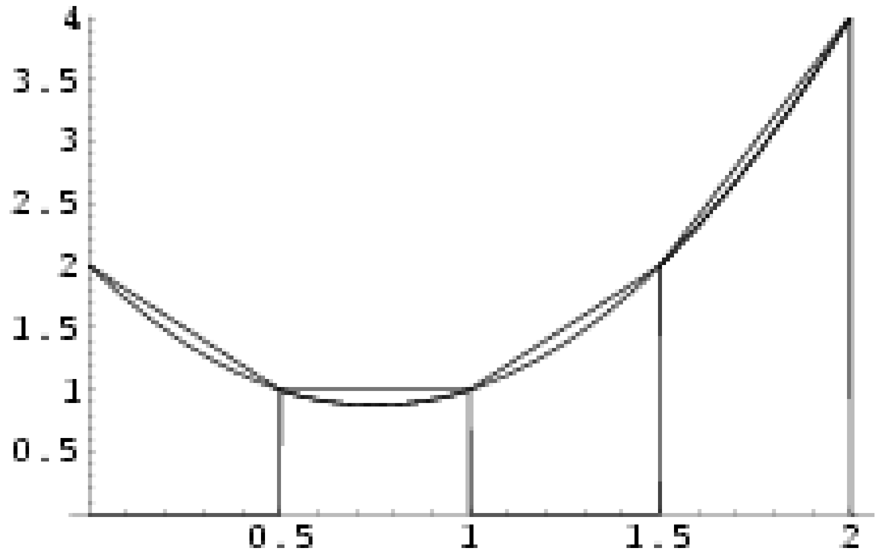
5.9 Trapezoidal Rule. The *trapezoidal rule* is a technique for finding definite integrals

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

numerically.

It is one step more clever than using Riemann sums. In Riemann sums, what we essentially do is approximate the graph $y = f(x)$ by a step graph and integrate the step graph.

In the Trapezoidal rule, we approximate $y = f(x)$ by a continuous graph made up of bits of lines.



Use n equal divisions. $h = \dfrac{b-a}{n}$. $x_i = a + ih$. Let $y_i = f(x_i)$ (for $0 \le i \le n$).

The trapezoidal rule formula can be written

$$\int_a^b f(x)\,dx \cong h\left(\frac{1}{2}y_0 + y_1 + y_2 + \cdots y_{n-1} + \frac{1}{2}y_n\right).$$

Proof (of this formula) uses area of ith trapezoid $= \dfrac{h}{2}(y_{i-1} + y_i)$

5.10 Example. Find $\int_1^3 e^{x^2}\,dx$ approximately using the Trapezoidal rule with $n = 10$.

Solution:

i	x_i	y_i	Weight	Weight $\times y_i$
0	1.0	2.71828	1/2	
1	1.2	4.2207	1	
2	1.4	7.0993	1	
3	1.6	12.9358	1	
4	1.8	25.5337	1	
5	2.0	54.5982	1	
6	2.2	126.469	1	
7	2.4	317.348	1	
8	2.6	862.642	1	
9	2.8	2540.2	1	
10	3.0	8103.08	1/2	
			Sum	8003.95

Sum times h is 1600.79

5.11 Remark. A natural question to ask at this point is: *how accurate is the Trapezoidal rule?*

Theoretically we know that as $n \to \infty$, the trapezoidal rule approximation $\to \int_a^b f(x)\,dx$, but that does not help us to know how close we are to the limit if we use $n = 100$ or $n = 1000$. The following theorem gives a worst case scenario.

5.12 Theorem. *Let T_n denote the result of using the trapezoidal rule formula with n steps to approximate $\int_a^b f(x)\,dx$. Then*

$$\left| \int_a^b f(x)\,dx - T_n \right| \leq \frac{b-a}{12} h^2 M_2$$

where M_2 is the largest value of $|f^{(2)}(x)| = |f''(x)|$ for $a \leq x \leq b$.

Note: Can rewrite this in terms of n using $h = (b-a)/n$.

We won't prove this, or say anything about how to show it is true.

5.13 Example. (i) In $\int_1^3 e^{x^2}\,dx$ with $n = 10$ how far off can the trapezoidal rule be?

Solution: For this we need to know M_2 and that involves $f''(x)$ with $f(x) = e^{x^2}$.

$$\begin{aligned} f'(x) &= 2xe^{x^2} \\ f''(x) &= 2e^{x^2} + 2x(2x)e^{x^2} \\ &= (2+4x^2)e^{x^2} \end{aligned}$$

It is fairly clear then that (in this case) $f''(x) > 0$ always (so that $|f''(x)| = f''(x)$) and the largest value for $0 \leq x \leq 3$ is $M_2 = f''(3) = 38e^9 = 307917$.

The theorem tells us then that the approximation T_{10} differs from the actual integral by *at most* (or at worst)

$$\left| \int_1^3 e^{x^2}\,dx - T_{10} \right| \leq \frac{b-a}{12} h^2 M_2 = \frac{3-1}{12} h^2 (307917).$$

The value of h is $h = \dfrac{b-a}{n} = \dfrac{3-1}{10} = 0.2$ and so the error could be as large as $\frac{1}{6}(0.04)307917 = 2052.78$.

(ii) How large should we chose n so that the trapezoidal rule approximation T_n to the same integral is certainly within 0.5 of the right value?

Solution: It will certainly be enough to choose n so that

$$\frac{b-a}{12} h^2 M_2 = \frac{3-1}{12} \left(\frac{3-1}{n} \right)^2 M_2 < 0.5$$

Rearranging this, it says we are safe if

$$n > \sqrt{\frac{1}{6}(4)M_2/0.5} = \sqrt{\frac{4}{3}307917} = 640$$

So $n = 641$ will certainly do.

5.14 Simpsons Rule. Simpsons Rule is the next most sophisticated method after the trapezoidal rule. With Riemann sums we used approximation by step graphs (bits of constant graphs one after the other), with the trapezoidal rule we used bits of straight lines, and now we use bits of quadratic graphs $y = ax^2 + bx + c$.

The first problem is that, while 2 points determine a line, we need 3 points to pin down a quadratic graph. Then we also need a formula for the 'area under a quadratic graph' (or the integral of it) analogous to the formula $h\left(\frac{1}{2}y_0 + \frac{1}{2}y_1\right)$ we used for the area of a trapezoid.

5.15 Lemma. *If we have 3 points in the plane with different x-coordinates, then there is exactly one quadratic graph passing through them.*

Proof. (just an idea of how it works)

If the points are (x_0, y_0), (x_1, y_1) and (x_2, y_2) then one way to find the quadratic is to write it down via the *Lagrange interpolation formula*

$$q(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}y_0 + \frac{(x-x_2)(x-x_0)}{(x_1-x_2)(x_1-x_0)}y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}y_2$$

Another approach is to start with a general $q(x) = ax^2+bx+c$ and use the 3 equations $q(x_0) = y_0$, $q(x_1) = y_1$, $q(x_2) = y_2$ to find a, b, c. □

5.16 Lemma. *For 3 points equally spaced horizontally $(x_0, y_0) = (x_1 - h, y_0)$, (x_1, y_1) and $(x_2, y_2) = (x_1 + h, y_2)$ and $y = q(x)$ the quadratic graph through the 3 points*

$$\int_{x_1-h}^{x_1+h} q(x)\,dx = h\left(\frac{1}{3}y_0 + \frac{4}{3}y_1 + \frac{1}{3}y_2\right)$$

Proof. We'll skip it. You can find it in Anton. It is just a bit messy, not really complicated. It turns out to make life a lot easier if you assume $x_1 = 0$ and this you can assume by shifting the y-axis. □

5.17 Simpsons Rule. Idea: For $\int_a^b f(x)\,dx$, choose n even and preferably large. Divide the interval $[a, b]$ into n equal sections each of width $h = \frac{b-a}{n}$, division points $x_i = a + ih$ ($0 \le i \le n$), corresponding values $y_i = f(x_i)$. Pair off intervals and approximate $y = f(x)$ on each pair by a quadratic graph.

$\int_a^b f(x)\,dx \cong$ sum of integrals of these quadratics.

The formula we get out of this is

$$\int_a^b f(x)\,dx = h\left(\frac{1}{3}y_0 + \frac{4}{3}y_1 + \frac{2}{3}y_2 + \frac{4}{3}y_3 + \cdots + \frac{2}{3}y_{n-2} + \frac{4}{3}y_{n-1} + \frac{1}{3}y_n\right)$$

5.18 Example. Find $\displaystyle\int_1^3 e^{x^2}\,dx$ approximately using Simpsons rule with $n = 10$.

Solution:

i	x_i	y_i	Weight	Weight $\times y_i$
0	1.0	2.71828	$1/3$	
1	1.2	4.2207	$4/3$	
2	1.4	7.0993	$2/3$	
3	1.6	12.9358	$4/3$	
4	1.8	25.5337	$2/3$	
5	2.0	54.5982	$4/3$	
6	2.2	126.469	$2/3$	
7	2.4	317.348	$4/3$	
8	2.6	862.642	$2/3$	
9	2.8	2540.2	$4/3$	
10	3.0	8103.08	$1/3$	
			Sum	7288.84

Sum times h is 1457.7 (and this is the approximate value for the integral given by Simpson's rule).

5.19 Remark. Now we ask about the accuracy of this method. Is it any better after the somewhat greater complication?

5.20 Theorem. *Let S_n denote the result of using Simpsons rule formula with n steps to approximate $\int_a^b f(x)\,dx$. Then*

$$\left|\int_a^b f(x)\,dx - S_n\right| \leq \frac{b-a}{180}h^4 M_4$$

where M_4 is the largest value of $|f^{(4)}(x)|$ for $a \leq x \leq b$.

Note: Can rewrite this in terms of n using $h = (b-a)/n$.

5.21 Example. (i) In $\int_1^3 e^{x^2}\,dx$ with $n = 10$ how far off can Simpsons rule be?

Solution: (details not all here) You can find that for $f(x) = e^{x^2}$, the fourth derivative is $f^{(4)}(x) = (16x^4 + 48x^2 + 12)e^{x^2}$ and $M_4 = 1740e^9 = 1.40994 \times 10^7$.

The theorem tells us then that the approximation S_{10} differs from the actual integral by *at most* (or at worst)

$$\left|\int_1^3 e^{x^2}\,dx - S_{10}\right| \leq \frac{b-a}{180}h^4 M_4 = \frac{3-1}{180}h^4(1.40994 \times 10^7)$$

The value of h is $h = \dfrac{b-a}{n} = \dfrac{3-1}{10} = 0.2$ and so the error could be as large as $\frac{1}{90}(0.2)^4(1.40994 \times 10^7) = 250.655$

(ii) How large should we chose n so that Simpsons rule approximation S_n to the above integral is certainly within 0.5 of the right value?

Solution: It will certainly be enough to choose n so that

$$\frac{b-a}{180}h^4 M_4 = \frac{3-1}{180}\left(\frac{3-1}{n}\right)^4 M_4 < 0.5$$

Rearranging this, says we are safe if

$$n > \left(\frac{1}{90}(16)M_4/0.5\right)^{1/4} = 47.3181$$

So $n = 48$ will certainly do. (Note how much smaller this is than 641 (needed for the trapezoidal rule in the same situation.)

5.22 Remark. The methods we have discussed were about finding definite integrals numerically. We will look later at methods for finding antiderivatives in a fairly systematic way. By the fundamental theorem, if we can find antiderivatives we can find definite integrals $\int_a^b f(x)\,dx$ analytically (that means we can come up with an exact formula for the value as opposed to a numerical approximate value). Our example $\int_1^3 e^{x^2}\,dx$ is one that we will not ever be able to do analytically.

5.23 Remark. In the trapezoidal rule, one rarely uses Theorem 5.12 to guarantee the desired accuracy of the estimate T_n.

A more pragmatic approach is to work with T_2, $T_{2^2} = T_4$, etc until we get to a value of T_{2^k} which has stabilised and where 2^k is reasonably big.

There is a way to avoid making the same calculations over and over again and still to follow this strategy. The point is that if we did this for the example $\int_1^3 f(x)\,dx$ (say with the $f(x) = e^{x^2}$ above), we would be calculating

$$\begin{aligned}
T_1 &= \frac{3-2}{1}\left(\frac{1}{2}f(1) + \frac{1}{2}f(3)\right)\\
T_2 &= \frac{3-2}{2}\left(\frac{1}{2}f(1) + f(2) + \frac{1}{2}f(3)\right)\\
T_4 &= \frac{3-2}{4}\left(\frac{1}{2}f(1) + f(1.5) + f(2) + f(2.5) + \frac{1}{2}f(3)\right)
\end{aligned}$$

Since we have to reuse $f(1)$ and $f(3)$ each time, we might be tempted to remember the answers. Then we need $f(2)$ every time after T_2 and so we might like to keep a record of the answer to that, and so on.

This involves a lot of recording (uses up computer memory if we do it on a computer) and there is a way to avoid so much storage. It turns out that[1]

$$\begin{aligned}
T_2 &= \frac{1}{2}T_1 + \frac{3-2}{2}f(2)\\
T_4 &= \frac{1}{2}T_2 + \frac{3-2}{4}(f(1.5) + f(2.5))
\end{aligned}$$

Thus we can calculate T_4 using only T_2 and values of $f(x)$ that were not needed for T_2.

[1]This bit is not in the book by Anton, I think.

In general,

$$T_{2n} = \frac{1}{2}T_n + \frac{b-a}{2n}\sum_{i=1}^{n} f\left(a + i\left(\frac{b-a}{2n}\right)\right)$$

and we can keep finding $T_1, T_2, T_4 = T_{2^2}, T_8 = T_{2^3}, \ldots$ without any massive storage requirement (and without recalculating any values of $f(x)$ we had to calculate previously).

This gives an *efficient pragmatic strategy*: Compute $T_1, T_2, T_4 = T_{2^2}, T_8, \ldots$ until we get to T_n with some reasonably large n and the value of T_n is roughly equal to the value of $T_{n/2}$. The chances are that both are then close to the true value of the integral $\int_a^b f(x)\,dx$.

This is an 'experimental approach' and not quite a method that is guaranteed to be 100% right.

5.24 Remark. As for the trapezoidal rule, one rarely uses Theorem 5.20 to guarantee the desired accuracy of the estimate S_n.

A more pragmatic approach is to work with S_2, $S_{2^2} = S_4$, etc until we get to a value of S_{2^k} which has stabilised and where 2^k is reasonably big.

It turns out that we can make use of the earlier efficient method of working out T_1, T_2, $T_{2^2} = T_4$, etc together with the formula[2]

$$S_{2n} = \frac{4}{3}T_{2n} - \frac{1}{3}T_n$$

This formula is 'easy' to check. For example

$$\begin{aligned}
T_1 &= \frac{b-a}{1}\left(\frac{1}{2}f(a) + \frac{1}{2}f(b)\right)\\
T_2 &= \frac{b-a}{2}\left(\frac{1}{2}f(a) + f\left(a + \frac{b-a}{2}\right) + \frac{1}{2}f(b)\right)\\
S_2 &= \frac{b-a}{2}\left(\frac{1}{3}f(a) + \frac{4}{3}f\left(a + \frac{b-a}{2}\right) + \frac{1}{3}f(b)\right)
\end{aligned}$$

$$\begin{aligned}
\frac{4}{3}T_2 - \frac{1}{3}T_1 &= \frac{b-a}{2}\left(\frac{2}{3}f(a) + \frac{4}{3}f\left(a + \frac{b-a}{2}\right) + \frac{2}{3}f(b)\right.\\
&\qquad \left. - \left(\frac{1}{3}f(a) + \frac{1}{3}f(b)\right)\right)\\
&= S_2
\end{aligned}$$

5.25 Example. Try out strategy for our example $\int_1^3 e^{x^2}\,dx$.

[2]This bit is not in the book by Anton, I think.

Appendix

5A.26 The Indefinite Integral. We will explain the idea of an indefinite integral by first giving an example. Every time you differentiate something you know the 'antiderivative' for the result. For example

$$\frac{d}{dx}(2x^4 + 2x^3 - 7x^2 + 11x - 17) = 8x^3 + 6x^2 - 14x + 11$$

and so if we happen to want to know something that has derivative $8x^3 + 6x^2 - 14x + 11$ we know that $2x^4 + 2x^3 - 7x^2 + 11x - 17$ is such a function.

We call it an an antiderivative of $8x^3+6x^2-14x+11$ — something which when differentiated gives us $8x^3 + 6x^2 - 14x + 11$.

Now what about other possible antiderivatives for the same thing? We can spot quite easily that the constant -17 is not very important. Any constant there instead of -17 would still have derivative 0 and so we realise that anything of the form $2x^4 + 2x^3 - 7x^2 + 11x + C$ is also an antiderivative for $8x^3 + 6x^2 - 14x + 11$.

In fact this is all because if two functions $f(x)$ and $g(x)$ have the same derivative, in this case if

$$f'(x) = g'(x) = 8x^3 + 6x^2 - 14x + 11$$

then

$$\frac{d}{dx}(f(x) - g(x)) = 0$$

for all x. The only functions that are defined on an interval and have derivative 0 are constant functions. So $f(x) - g(x) = C =$ a constant and so $f(x) = g(x) + C$. This means that if we find one antiderivative (in our example $2x^4 + 2x^3 - 7x^2 + 11x - 17$) for the given function (in our example $8x^3 + 6x^2 - 14x + 11$) then the most general one is of the form $2x^4 + 2x^3 - 7x^2 + 11x - 17 + C$. Since $C - 17$ is another constant we can write

$$\int 8x^3 + 6x^2 - 14x + 11\,dx = 2x^4 + 2x^3 - 7x^2 + 11x + C.$$

We can go about things a bit more systematically, by starting with the simplest rules for differentiation and turning them into rules for finding antiderivatives.

(i) We know $\frac{d}{dx}x^n = nx^{n-1}$ and so $\int nx^{n-1}\,dx = x^n + C$. For example, $\int 4x^3\,dx = x^4 + C$. We can use a little ingenuity to find that an antiderivative for x^{n-1} is $\frac{1}{n}x^n$ (as long as $n \neq 0$) and we can tidy this up to get the rule

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C \qquad (n \neq -1)$$

It is perhaps interesting to see that we can't immediately write down $\int x^{-1}\,dx = \int 1/x\,dx$. In fact the antiderivative for $1/x$ involves the natural logarithm function $\ln$ and is therefore a much more complicated thing that $1/x$. We will leave this for later.

A detail we should mention, is that when $n < 0$, x^n is not well-behaved at $x = 0$. So the domain is not an interval but two intervals $x < 0$ and $x > 0$. So, technically, the $+C$ is not adequate in these cases to describe the 'most general' antiderivative. We could switch from one value of C to another as we pass $x = 0$ and still have a valid antiderivative (when $n < 0$). However, people rarely go into this and probably you almost never should encounter this in practice. Because things blow up at $x = 0$ there should not really be a practical problem where $x > 0$ and $x < 0$ are both valid (when dealing with x^n and $n < 0$).

(ii) We know the derivative of a sum is the sum of the derivatives $\frac{d}{dx}(u + v) = \frac{du}{dx} + \frac{dv}{dx}$ and so we can see that an antiderivative of a sum is the sum of the antiderivatives.

Similarly for constant multiples.

Using these rules, we can integrate all polynomials. For example

$$\begin{aligned}\int 5x^3 - 11x^2 + 13x + 2\,dx &= 5\left(\frac{1}{4}x^4\right) - 11\left(\frac{1}{3}x^3\right) + 13\left(\frac{1}{2}x^2\right) + 2\left(\frac{1}{1}x\right) + C \\ &= \frac{5}{4}x^4 - \frac{11}{3}x^3 + \frac{13}{2}x^2 + 2x + C\end{aligned}$$

(iii) From the rules for differentiating trigonometric functions

$$\frac{d}{dx}\sin x = \cos x, \frac{d}{dx}\cos x = -\sin x, \frac{d}{dx}\tan x = \sec^2 x$$

we can write down rules for integrating some

$$\int \cos x\,dx = \sin x + C, \int \sin x\,dx = -\cos x + C, \int \sec^2 x\,dx = \tan x + C$$

With a little guesswork we can figure out some related integrals like

$$\int \cos 3x\,dx$$

We might think of $\sin 3x$ as a possible antiderivative but

$$\frac{d}{dx}\sin 3x = 3\cos 3x$$

is 3 times what we want. Since 3 is a constant, we can divide across by it and we get

$$\int \cos 3x\,dx = \frac{1}{3}\sin 3x + C$$

However, there is no easy way to do $\int \sec x\,dx$ (we will see what the answer to this is a bit later). We can write

$$\int \sec x\,dx = \int \frac{1}{\cos x}\,dx$$

and we know how to integrate $\cos x$ but that does not help. The is no good quotient rule for antiderivatives. Unlike differentiation, where we can learn a small number of rules that are enough to differentiate almost any function we can easily write down, there are easy-looking functions where antiderivatives are quite hard to find. We have seen $\int 1/x\,dx$, more recently $\int \sec x\,dx$ and the example $\int \cos(x^2)\,dx$ is one that is essentially impossible. It is not that there is no answer. There is an antiderivative but it is known that there is no way to write a finite formula for the antiderivative of $\cos(x^2)$ using the familiar functions (powers, roots, fractions, trig functions, $\ln$, e^x).

5A.27 Example. There are a very few examples which can be worked out directly from the limit of Riemann sums definition. This is one example

$$\int_2^4 x + 2\,dx$$

Solution: Take n large, n equally wide subintervals of $[a, b] = [2, 4]$ with widths $h = (b-a)/n = (4-2)/n = 2/n$. This gives subdivision points $x_i = a + ih = 2 + \frac{2i}{n}$ for $i = 0, 1, 2, \ldots, n$. To make life easy for ourselves we take $x_i^* = x_i = 2 + (2i)/n$. The Riemann sum is then

$$\begin{aligned}
\sum_{i=1}^n f(x_i^*)(x_i - x_{i-1}) &= \sum_{i=1}^n \left(\frac{2i}{n} + 2\right)\frac{2}{n} \\
&= \frac{2}{n}\left(\sum_{i=1}^n \frac{2i}{n} + \sum_{i=1}^n 2\right) \\
&= \frac{4}{n^2}\left(\sum_{i=1}^n i\right) + \frac{2}{n}\left(\sum_{i=1}^n 2\right) \\
&= \frac{4}{n^2}\left(\frac{n(n+1)}{2}\right) + \frac{2}{n}(2n) \\
&= 2\left(1 + \frac{1}{n}\right) + 4 \\
&\to 6 \text{ as } n \to \infty
\end{aligned}$$

Here we used the fact that

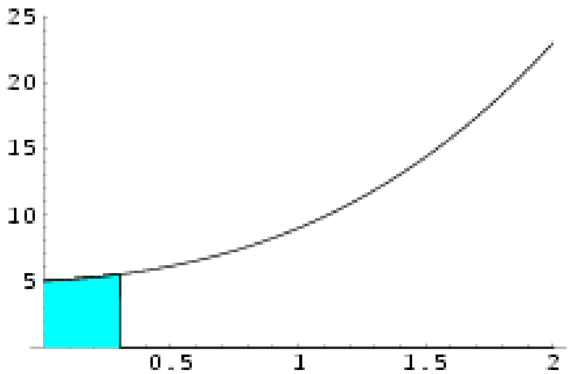
$$\sum_{i=1}^n i = 1 + 2 + 3 + \cdots + n$$

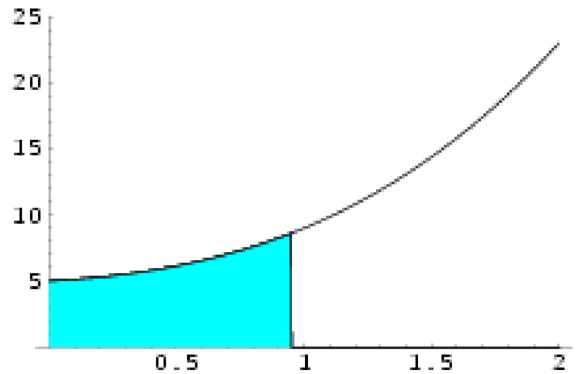
turns out to be $\frac{n(n+1)}{2}$. This can be discovered by writing the sum down backwards and adding the two formulae vertically:

$$\begin{array}{rcccccccccc}
s & = & 1 & + & 2 & + & 3 & + & \cdots & + & n \\
s & = & n & + & n-1 & + & n-2 & + & \cdots & + & 1 \\
2s & = & (n+1) & + & (n+1) & + & (n+1) & + & \cdots & + & (n+1) \\
 & = & n(n+1) & & & & & & & &
\end{array}$$

In most examples, we will not be able to write the Riemann sum as a short formula in the way we did here and so we would not find the limit.

5A.28 Remark. There is an amazing connection between definite integrals and differentiation, which we will now state. It comes about by considering not just one definite integral $\int_a^b f(x)\,dx$ but a whole infinite number of them.





Not just $\int_0^2 f(x)\,dx$ but $\int_0^x f(t)\,dt$ for $0 \le x \le 2$.

5A.29 Theorem (Fundamental Theorem of Integral Calculus)**.** *Assume that $y = f(x)$ is continuous for $a \le x \le b$. Consider $A(x) = \displaystyle\int_a^x f(t)\,dt$ for $a \le x \le b$. (A(x) is a new function, built from f and definite integration.) Then $A(x)$ is an antiderivative for f (that is $A'(x) = f(x)$ for $a \le x \le b$).*

In summary:

$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x) \qquad (a \le x \le b, \textit{ if } f \textit{ continuous})$$

This is one part of the Fundamental theorem, or one way to state it.

5A.30 Corollary. *There is an antiderivative for every continuous function f.*

5A.31 Example. Find $\displaystyle\frac{d}{dx}\left(\int_a^x \cos\left(\frac{t^3}{t^8+1}\right)\,dt\right)$

Solution: By the theorem

$$\frac{d}{dx}\left(\int_a^x \cos\left(\frac{t^3}{t^8+1}\right)\,dt\right) = \cos\frac{x^3}{x^8+1}$$

5A.32 Theorem (Other part of fundamental theorem)**.** *Assume that $y = f(x)$ is continuous for $a \le x \le b$ and suppose $g(x)$ is an antiderivative for $f(x)$ (that is $g'(x) = f(x)$ for $a \le x \le b$). Then*

$$\int_a^b f(x)\,dx = g(b) - g(a) = [g(x)]_{x=a}^b$$

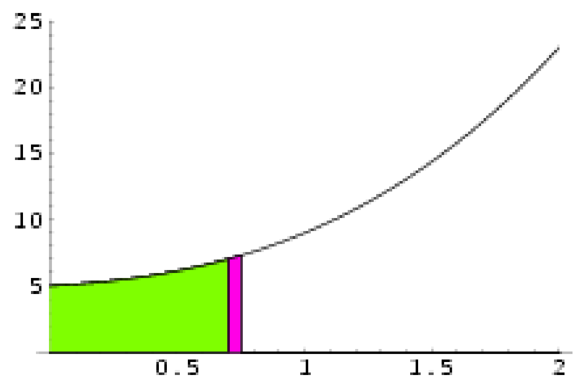
5A.33 Example. Find $\int_2^4 2x + 1\,dx$

Solution: Antiderivative $g(x) = x^2 + x$ (since $g'(x) = 2x + 1$) and so

$$\int_2^4 2x + 1\,dx = [g(x)]_{x=2}^4 = [x^2 + x]_{x=2}^4 = (4^2 + 4) - (2^2 + 2) = 14$$

Proof. (of $A'(x) = f(x)$ part of Fund. Thm.)

Use first principles



$$A'(x) = \lim_{h\to 0} \frac{A(x+h) - A(x)}{h} = \lim_{h\to 0} \frac{\int_x^{x+h} f(t)\,dt}{h}$$

$\int_x^{x+h} f(t)\,dt \cong f(x)h$ for h small. Thus limit is $f(x)$. $\square$

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

We know $A(x) = \int_a^x f(t)\,dt$ is an antiderivative. Hence

$$\frac{d}{dx}(A(x) - g(x)) = f(x) - f(x) = 0.$$

Thus $A(x) - g(x) = c = \text{const}$.

Use $x = a$ to find constant: $A(a) - g(a) = c$. But $A(a) = \int_a^a f(t)\,dt = 0$. Thus $-g(a) = c$ and $A(x) - g(x) = c = -g(a)$ for all $x \in [a, b]$.

Use this for $x = b$ to get $A(b) - g(b) = -g(a)$, $A(b) = g(b) - g(a)$, that is

$$\int_a^b f(t)\,dt = g(b) - g(a).$$

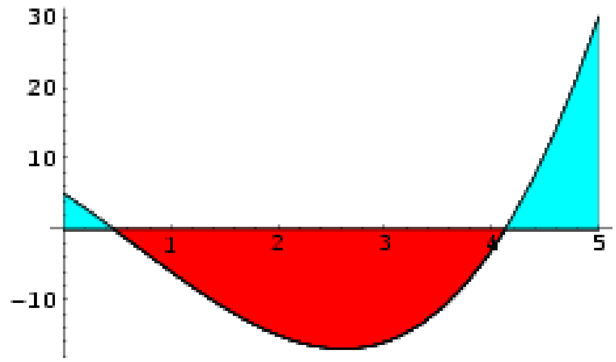
$\square$

5A.34 Correction. Correction to (or refinement of) graphical interpretation of definite integrals.

$\int_a^b f(x)\,dx =$ 'area under graph' is good when $f(x) \geq 0$ always.

When $f(x) < 0$, that part of the area is counted with a minus sign.

Consider negative terms in Riemann sum $\sum_{i=1}^n f(x_i^*)(x_i - x_{i-1})$ if $f(x_i^*) < 0$.



$$\int_a^b f(x)\,dx = (\text{sum of areas where } f(x) > 0) - (\text{sum where } f(x) < 0)$$

5A.35 Notation. There is one more thing we have skimmed over.

When $a > b$, there is a convention that allows us to write $\int_a^b f(x)\,dx$ (for example $\int_3^2 x^2 + x\,dx$) even though the limits are upside down.

By convention, the meaning for this is

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx \text{ if } a > b$$

This convention comes up in the proof of the fundamental theorem (the part where $A(x + h) - A(x) = \int_x^{x+h} f(t)\,dt$, where we have to be able to deal with $h < 0$ as well as $h > 0$).

In fact the fundamental theorem in the form

$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x)$$

is also valid for $x < a$ as long as f is continuous on an interval that includes both a and x. (The important thing is to have all points in between as well as a and x.)

The convention also fits with the other form of the fundamental theorem. If $g'(x) = f(x)$, then

$$\int_a^b f(x)\,dx = [g(x)]_{x=a}^b = g(b) - g(a)$$

remains true when $a > b$.

5A.36 Examples. (i) Find $\dfrac{d}{dx}\left(\displaystyle\int_x^0 \frac{t}{t^6+1}\,dt\right)$

Solution:

$$\frac{d}{dx}\left(\int_x^0 \frac{t}{t^6+1}\,dt\right) = \frac{d}{dx}\left(-\int_0^x \frac{t}{t^6+1}\,dt\right) = -\frac{x}{x^6+1}$$

using the above convention and the Fundamental theorem.

(ii) Find $\dfrac{d}{dx}\left(\displaystyle\int_1^{x^3+x} \cos(t^2+4)\,dt\right)$

Solution: Take $y = \displaystyle\int_1^{x^3+x} \cos(t^2+4)\,dt = \int_1^{u} \cos(t^2+4)\,dt$ where $u = x^3+x$. By the Chain rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = \frac{dy}{du}(3x^2+1)$$

and by the Fundamental theorem

$$\frac{dy}{du} = \cos(u^2+4) = \cos((x^3+x)^2+4)$$

Putting these together, we get

$$\frac{d}{dx}\left(\int_1^{x^3+x} \cos(t^2+4)\,dt\right) = (3x^2+1)\cos((x^3+x)^2+4)$$

Notice that the answer would have been the same for $\dfrac{d}{dx}\left(\displaystyle\int_{25}^{x^3+x} \cos(t^2+4)\,dt\right)$*, as the value of the lower limit of the integral does not come into the answer (as long as it is constant — and as long as the integrand is continuous wherever we go).*