<u>Integration in Mathematics B</u>

John A. Belward, Department of Mathematics
University of Queensland
e-mail: jab@maths.uq.edu.au, web: www.maths.uq.edu.au/~jab

This presentation concerns integration. I am not going to go through a long set of difficult manipulative examples, rather I want to raise important (as I see them) points I hope this will help people when preparing work programs and assessment instruments.

Until one does problems with applications in applied areas, the extension to topics in Mathematics C which concern integration will most likely concern freer use of the log, exponential and trigonometric functions, so a lot of what is here is relevant to Mathematics C too.

I want to make some points about

- indefinite integrals
- the definite integrals
- what is area
- the fundamental theorem of calculus
- simple substitution
- numerical integration
- manipulation packages

<u>What is integration</u>

What is it for? It enables us to navigate in space. If we know our velocity vector $\boldsymbol{v}$ at any time t then our position vector is given by $\boldsymbol{s}$ where $\frac{ds}{dt} = \boldsymbol{v}$ and if we have $\boldsymbol{s} = \boldsymbol{s}_0$ at $t = t_0$ then

$$\boldsymbol{s} - \boldsymbol{s}_0 = \int_{t_0}^{t} \boldsymbol{v}(t)dt$$

so we just have to integrate $\boldsymbol{v}$(t) ! To know $\boldsymbol{v}$(t) we integrate the acceleration vector $\boldsymbol{a}(t)$ in a similar way.

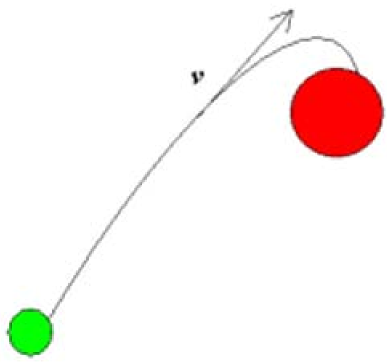


Figure 1

There are other less attention grabbing examples.

Indefinite integrals

The expression

$$\int f(x)dx$$

stands for the function which when differentiated gives $f(x)$.

It is commonly known as the indefinite integral.

It is also sometimes known as the antiderivative of $f(x)$, or even a primitive

Examples:

$$\int 2x^3 dx = \frac{x^4}{2} + c \ ; \quad \int \sin(t+\pi)dt = -\cos(t+\pi) + c \ ; \quad \int e^{u^2} 2u du = e^{u^2} + c$$

The number c can have any value, simply because the derivative of a constant is zero.

These examples are done "by inspection" which is a pompous way of saying "guess and check".

So there is a nagging question. Since I have guessed the answer someone else might make a guess too which can be verified, but is different from mine.

Example (esoteric)

$$\int 2\sec^2 x \tan x dx = \tan^2 x + c \quad \text{and} \quad \int 2\sec^2 x \tan x dx = \sec^2 x + c$$

because

$$\frac{d}{dx}\tan^2 x = 2\sec^2 x \tan x \ \text{ and } \ \frac{d}{dx}\sec^2 x = 2\sec^2 x \tan x$$

Here the theoreticians (pure mathematicians) come to our assistance proving that if two functions have the same derivative they differ from each other by a constant at the most.

Doing indefinite integration boils down to learning manipulative tricks.

Examples

$$\int 27x^2(x^3+2)^8 dx = (x^3+2)^9 + c$$

and

$$\int \cos^3\theta d\theta = \int \cos^2\theta \cos\theta d\theta = \int (1-\sin^2\theta)\cos\theta d\theta = \int \cos\theta d\theta - \int \sin^2\theta \cos\theta d\theta = \sin\theta - \frac{\sin^3\theta}{3} + c$$

I'd expect that the second of these to be set only as a challenge problem.

Definite integration

The definite integral is denoted by

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

Its geometric interpretation is the area defined by the lines $y = f(x)$, $y = 0$, $x = a$ and $x = b$, when $f(x)$ is positive from a to b.

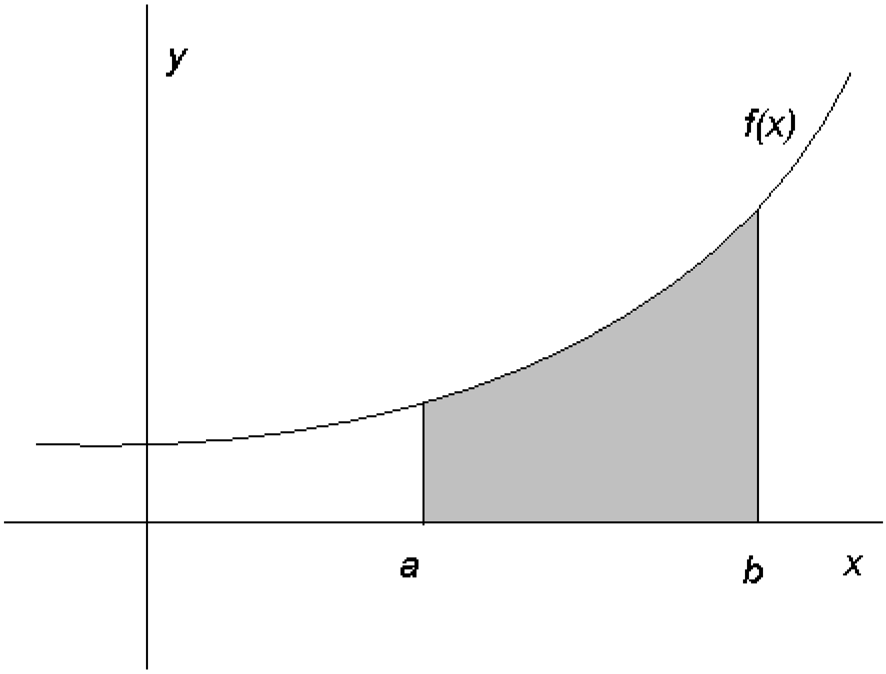


Figure 2

$$\text{Shaded Area} = \int_a^b f(x)dx$$

There is nothing to be proved in this statement if the definition of area for sets of points enclosed by curved lines has been made.

Defining area

Strict proofs probably belong to second year tertiary mathematics, but an intuitive idea can easily be grasped with the use of a couple of diagrams. One of them has an immediate practical application.

This is the tiler's problem mistakenly called the butcher's problem.

A millionaire in Brisbane has a pool included in the back yard of his architect designed home. Being architect designed, for aesthetic reasons the shape of the pool is random, as shown. The millionaire is tough and has had the man screw his price right down low. Thus the area of the pool has to be known accurately to minimise wastage. How does he find this out?
Well, he's never done any maths save for a little geometry and arithmetic. So he gets the plans of the back yard which show the position of the pool accurately and proceeds as follows.

He overlays a piece of squared paper on the plan and counts the number of squares inside the pool. Subtracting this from the rectangular area he knows that the area of tiles mustn't exceed a certain amount. And he can do another sum by counting the squares that overlap the pool with part of the square overlapping the pool. This way he finds out a total that he must exceed when placing his order.

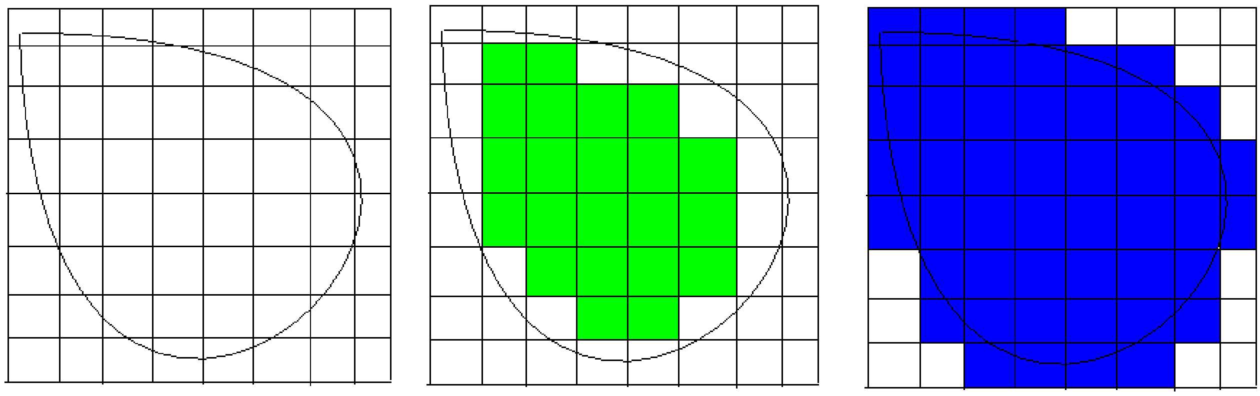


Figure 3a, b, c,

This is all very well but theses two areas differ by a lot. Can he do better? Yes if he overlays a grid of smaller squares. In this way he can reduce the grid size until the over estimate and the underestimate are near enough together.

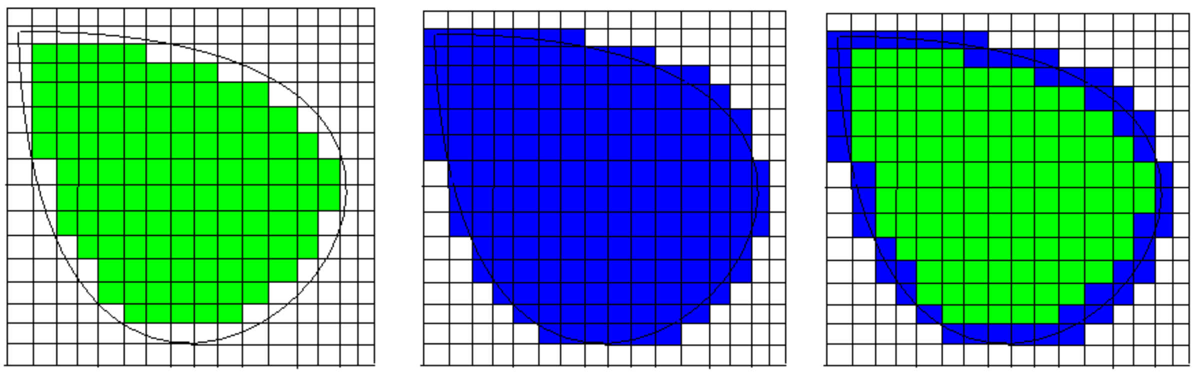


Figure 4a, b, c

In doing this he has essentially defined what we mean by area. If it is rectangular or polygonal then we find it by doing length times breadth and adding up a finite number of areas of pieces. If it is enclosed by a curve we have to overlay the grid of squares reducing the grid size to determine the limits of the inner and outer areas.

You could argue that the tiler used numerical integration for that is just what he did.

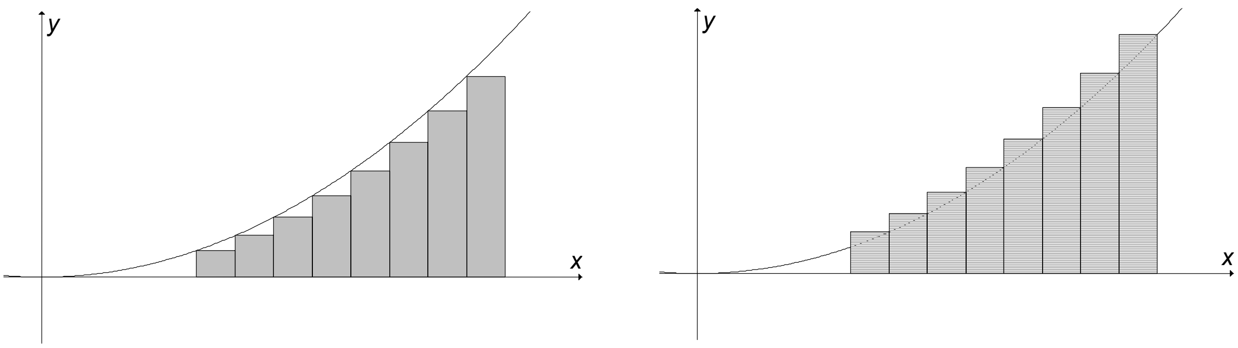


Figure 5a,b

From the diagrams above we observe that the area between the curve and the axis as shown in figure 2 is greater than the sum of the areas of the rectangles on the left and less than those on the right. By extending the horizontals and verticals it is easy to see that these areas are just the areas of a grid of underlapping squares and overlapping squares.

This is far from rigorous, but a plausible argument to show that the definition of area and the definition of the definite integral are one and the same thing.

We can write these areas as a summation using the sigma notation or write out the expression in full detail, either way is complicated but hard to avoid. For the current discussion it is enough to use the diagrams.

Nevertheless we note that

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_{i-1})\Delta x_i = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x_i = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i)\Delta x_i = \int_a^b f(x)dx$$

(t_i is a value between x_{i-1} and x_i)

This result is used endlessly in problems where the evaluation of some quantity (physical, economic, etc.) is needed.

This has nothing to do with differentiation until we learn of the fundamental theorem of Calculus.

$$\int_a^b f(x)dx = F(b) - F(a) \quad \text{where } F(x) \quad \text{is such that} \quad F'(x) = f(x)$$

There are ways of evaluating definite integrals where we do not find an antiderivative. Thus while one often learns about integration as antidifferentiation this is not the whole story by any means and it is not correct to think of integration as being the reverse of differentiation.

Demonstration of the fundamental theorem and the relation between the indefinite integral and the definite integral

If we allow either of the limits of integration in a definite integral to vary then we will have a function, so we can write

$$G(t) = \int_a^t f(x)dx$$

If the integral can be evaluated for all values of t greater than a then the function is defined. From this definition we can say

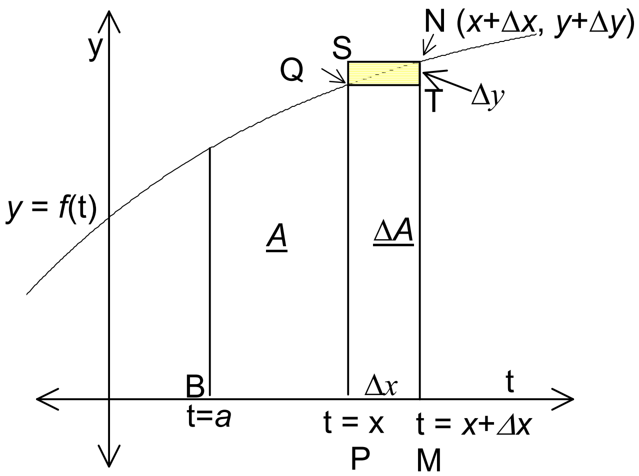
$$G(x) = \int_a^x f(t)dt$$

The roles of x and t have changed above. More on this point below.

Now we can find $G'(x)$ as follows

$$G(x+\Delta x) - G(x) = \int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt$$

and it is clear from a diagram that this is the difference of 2 areas and so itself is an area as shown below, because $G(x) = A$ and $G(x+\Delta x) = A + \Delta A$ **.**



Now $\dfrac{G(x+\Delta x) - G(x)}{\Delta x} = \dfrac{\Delta A}{\Delta x}$ and ΔA lies between (*length of PQ*)Δx and (*length of MN*)Δx.

Now lets write this down explicitly and then divide through by Δx to get successively

$$(length PQ)\,\Delta x < \Delta A < (length\, MN)\,\Delta x$$

and

$$(length PQ) < \frac{\Delta A}{\Delta x} < (length\, MN)$$

Finally, as Δx goes to zero (*lengthMN)* becomes *(lengthPQ)* which is $f(x)$ and so we've shown that

$$\text{As } \Delta x \to 0, \ \frac{\Delta A}{\Delta x} \to f(x) \ \text{ in other words } \ \frac{G(x+\Delta x)-G(x)}{\Delta x} \to f(x) \quad \text{so } \ G'(x)=f(x)$$

<u>Some notation. Which is a constant and which is a variable.</u>

Because we can add and subtract areas,

if $F(x)=\int_a^x f(t)dt,$ and $\mathrm{G}(x)=\int_b^x f(t)dt,$ then $F(x)=\int_a^b f(t)dt+\mathrm{G}(x)$ but the integral on the right hand side of this last equation is independent of x so $G'(x)=F'(x)$.

So now we can bring together the notation for definite and indefinite integrals, thus

$$\int f(x)dx=\int_a^x f(t)dt \quad \text{where } \ a \ \text{ is arbitrary}.$$

Thus it appears that a takes the part of the arbitrary constant. In fact a is <u>almost</u> arbitrary, because it must be in the range of $f(x)$.

Please note also that

$$\int f(x)dx=\int_a^x f(t)dt = \int_{t=a}^{t=x} f(t)dt = \int_a^x f(\bullet)d\bullet$$

and indeed

$$\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(cows)d(cows) = \int_a^b f(horses)d(horses) = \int_a^b f(elephants)d(elephants).$$

These statements underline the fact that the variable of integration is irrelevant as far as the value of the definite integral is concerned. The value of the integral depends on the function, f , and the interval $[a, b]$. Indeed in many advanced texts you will just see $\int_a^b f$.

<u>Substitution</u>

The example

$$\int e^{u^2} 2udu \ = \ e^{u^2}+c$$

is of the form

$$\int f(g(u))g'(u)du.$$

We recognize that had we differentiated $y=e^{u^2}$ we would have done so by putting $z=u^2$ then $y=e^z$ and we work out $\frac{dz}{du}$ and $\frac{dy}{dz}$ and apply $\frac{dy}{du}=\frac{dy}{dz}\frac{dz}{du}$. As said before, this is just a "guess and check" example.

If you put $z = g(u)$ and $g'(u) = \frac{dz}{du}$ then you can put

$$\int f(g(u))g'(u)du = \int f(z)\frac{dz}{du}du = \int f(z)dz\,.$$

In these manipulations we have "cancelled" du and dz ; these formal operations have to be legitimised, fortunately the pure mathematicians have done this for us.

There is a very compact way of viewing the "guess and check" structure, it is to see that we can express the integral as

$$\int e^{u^2}\,2udu = \int e^{u^2}\,d(u^2) \quad \text{and we apply} \quad \int e^{(\bullet)}d(\bullet) = e^{(\bullet)} + c \quad \text{putting } \bullet = u^2$$

Apply the same reasoning to $\int (x^2+5)^7\,xdx$, observing that a "fudge factor" of 2 (legitimate here) needs to be introduced inside and outside the integral, in the denominator and numerator respectively.

The reasons for pointing out these interpretations are two-fold. First to reduce the amount of detail, which for some will reduce the possibility of an arithmetic or algebraic blunder, and also because this way we see structure and gain some insight into what is a function. Ask the students whether they think that $\sin(x)$, $\sin(w)$ and $\sin(x^2)$ are the same functions.

My understanding of the Mathematics B syllabus is that integration by substitution is only to be done on examples which can be done by guess and check.

Certainly, applying substitution in the reverse direction where we do

$$\int f(z)dz \quad \text{with } z = g(u) \text{ so that } \int f(z)\frac{dz}{du}du = \int f(g(u))g'(u)du$$

and thus get a more tractable integral in terms of u is not in the syllabus.

<u>Numerical integration</u>

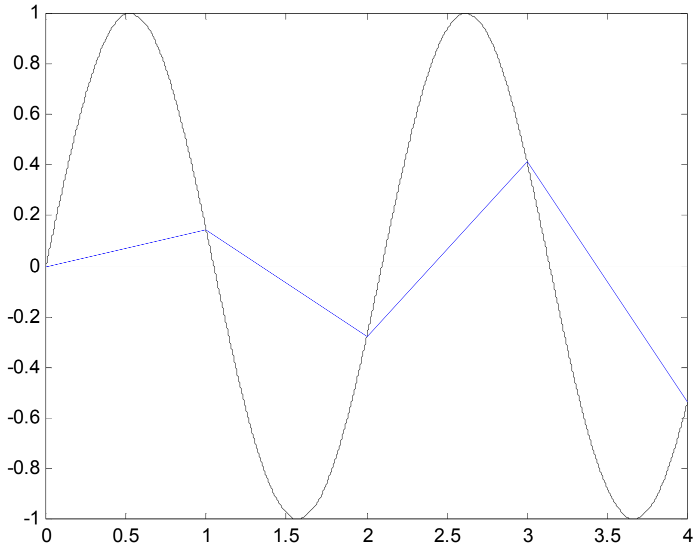
To get an idea of the accuracy of the trapezoidal rule some examples should be done in which the exact value is known. The error can be plotted and the dependence on the number of intervals investigated. Further, since the rule is obtained by integrating a piecewise linear approximation to the integrand by plotting the function and the piecewise linear approximation an understanding of the error behaviour can be obtained.

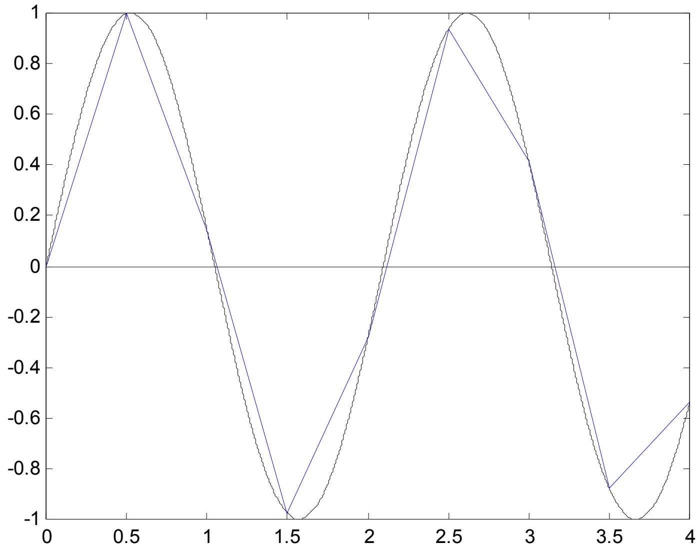
For example with $f(x) = \sin(3x)$ for the interval $[0,4]$ we get the following results

no of intervals	approximate value	exact value	absolute error	relative error
4	.00554	.052045	.04651	.894
8	.04190	.052045	.01014	.195
16	.04959	.052045	.00246	.047
32	.05144	.052045	.00061	.012

Table 1 Results for the Trapezoidal rule with $f(x) = \sin(3x)$ for the interval $[0,4]$

Plotting the function and the approximations for 4 and 8 intervals we get the following diagrams





you can see that until we put in enough points to follow the function we will have little chance of getting an accurate answer.

<u>Using Mathematical Software</u>

I assume that everyone uses calculators to compute integrals using the trapezoidal rule. The areas that featured in earlier sections in the definition of the definite integral can be calculated easily for simple functions with very similar keystrokes or code. A spread sheet can also be used for these purposes. I use Matlab because it is accessible and I am familiar with how it works.

Using symbolic manipulation packages is not yet established as an element of first year courses at the University of Queensland. I would always use a package – demonstrate and use it in a lecture - if it is available, but I prefer to present the package as a resource which can support the course rather than be an integral part of it. The problem is the communications interface. Interpretation of the output and a realization of what to expect from the package are crucial elements in their use. These are not simple tasks to master.

I have learned from experience not to face a class without having gone through the steps ahead of time.

One day in a lecture I typed into Maple:

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

and was returned

```
-arctanh(x)
```

I had to do some talking! Of course I expected to get a logarithm as the answer.

So I then did

```
> int(1/((x-1)*(x+1)),x);
```

and got

```
1/2 ln(x - 1) - 1/2 ln(x + 1)
```

which was what I wanted. But since I wanted the students to learn that the first thing to do is to factor the expression, Maple had not helped me at all.

Here are some more examples which I found 'interesting' when doing solutions to some of the revision chapters in our book

The example was "Find the derivative of $y = \sqrt{(x+2)^3}$ "

I entered :

```
>diff(sqrt((x+2)^3),x);
```

and was returned

$$\frac{3}{2}\frac{(x+2)^2}{\sqrt{(x+2)^3}}$$

I was not impressed with that answer and had to enter

```
> diff((x+2)^(3/2),x);
```

to get

$$\frac{3}{2}\sqrt{x+2}$$

After getting $\frac{3}{2}\frac{(x+2)^2}{\sqrt{(x+2)^3}}$ I later found that I could have used

```
>simplify(");
```

Simplify does not always get you what you want, here's another sequence:

> diff((x+2)^3*((x-1)^7),x);

$$3\,(x\ +\ 2)^2(x\ -\ 1)+7(x\ +\ 2)^3(x\ -\ 1)^6$$

> simplify("); " stands for the last expression

$$10x^9 - 9x^8 - 72\,x^7 + 105x^6 + 126x^5 - 315\,x^4 + 84x^3 + 207x^2 - 180x + 44$$

> collect("",x); "" stands for the second last expression

$10x^9 - 9x^8 - 72\,x^7 + 105x^6 + 126x^5 - 315\,x^4 + 84x^3 + 207x^2 - 180x + 44$ you don't always get what you want!
- 180 x + 44

> factor("""); """stands for the third last expression !

$(10x+11)(x+2)^2(x-1)^6$ at last!

Finally, I wanted to do $\int 6x(3x^2+2)^3 dx$, expecting $\frac{1}{4}(3x^2+2)^4 + c$. Here's what I got:

>int((6*x)*(3*x*x+2)^3,x);

$\frac{81}{4}x^8 + 54x^6 + 54x^4 + 4x^2$ Well that's not what I expected, but I can factorise it into the form I want, so I did

> factor("); and got

$\frac{3}{4}\,x^{\,2}(3\,x^2+4)(9\,x^4+12\,x^2+8)$ That's not what I wanted! Then I realized the trouble was with the non-existent arbitrary constant and so I could fudge it!

> factor(""+4);

$$\frac{1}{4}(3x^2+2)^4$$

<u>Conclusion</u> There is a lot of craft needed when using manipulation packages.