13 Primitives and techniques of integration

This section is concerned with *integration* or *antidifferentiation* — the process of finding a function whose derivative is some given function.

Definition of primitive A function F is a *primitive* of a function f, or an *antiderivative* of f, if $F' = f$. The notation we use to denote this relationship is either

$$F = \int f \quad (\text{or } \textstyle\int f = F)$$

(when working with some generic function f), or

$$\int f(x) = F(x) \quad (\text{or } \textstyle\int f(x) = F(x))$$

(when working with a specific, named function, given by a certain rule).

Here is why primitives are useful:

if F is a primitive of f, on an interval that includes $[a, b]$, *and if f is integrable on* $[a, b]$[212], then by the fundamental theorem of calculus (part 2) we have

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

The expression $F(b) - F(a)$ comes up so frequently, it has a few different notations:

$$F(b) - F(a) = F|_a^b = F(x)|_{x=a}^b.$$

A number of important comments are in order about the definition of a primitive. We give a few examples, and make the comments along the way.

Example 1 $x^3 + 3x + \pi$ is a primitive of $3x^2 + 3$ (obviously!) and so

$$\int (3x^2 + 3)\ dx = x^3 + 3x + \pi.$$

But equally obviously

$$\int (3x^2 + 3)\ dx = x^3 + 3x + e.$$

The critical comment to be made here is that

[212]This is a subtle but important point. FTOC (part 2) says that if F satisfies $F' = f$ on $[a, b]$ *and if f is integrable on* $[a, b]$ then $\int_a^b f = F(b) - F(a)$. If we don't add the assumption that f is integrable, then we cannot draw this conclusion. There are examples of differentiable functions whose derivatives are not integrable. If V is such a function (I use V here because the first example of such a function was discovered by Volterra, and is called *Volterra's function*) then while it is true to say that $V = \int V'$ (V is a primitive of V'), it is *not* true to say that $\int_a^b V' = V(b) - V(a)$, since the left-hand side exists but the right-hand side doesn't. Remember that the FTOC part 2 says that if there's a function g with $g' = f$ *and* f is integrable then $\int_a^b f = g(b) - g(a)$.

the "=" in $\int f = F$ is **not** a true equality!

If it was, from $\int (3x^2+3)\,dx = x^3+3x+\pi$ and $\int (3x^2+3)\,dx = x^3+3x+e$ we would conclude the patently absurd

$$x^3+3x+\pi = x^3+3x+e.$$

It is very important to remember that "$\int f = F$" is actually shorthand for "$F' = f$", and *not* an assertion that two functions are identical. This is a clear abuse of the "=" sign, but hopefully one you can live with. We'll see some odd paradoxes that can arise when we forget this.[213]

Example 2 $-2/(1+\tan(x/2))$ is a primitive of $1/(1+\sin x)$. This is obvious, no? Probably not; but once it has been asserted, it can be easily checked, by differentiating $F(x) = -2/(1+\tan(x/2))$. After *a lot* of algebra, the derivative can be massaged into the form $1/(1+\sin x)$.

The comment to be made here is that

> unlike finding derivatives, which is a mechanical process, finding antiderivatives is often hard, always requires ingenuity and usually (see a later example) is practically impossible.

See https://xkcd.com/2117/ for a could summary of the situation!

Example 3 If F is a primitive of f, so is $F+c$ for any constant c.

The comment to be made here is that

> A function with an antiderivative, has infinitely many antiderivatives.

It is tempting to at this point try to prove a theorem, along the lines of: if F is a primitive of f, then *all* primitives of f are of the form $F+c$ for some constant c. We won't try to prove this, because it is false. (See later examples).

Another comment is in order here:

> There is no great value in writing "$\int f = F + C$".

The "$+C$" adds nothing — since "$\int f = F$" is shorthand for "$F' = f$", adding the "$+C$" (to convey "$(F+C)' = f$") is just saying "by the way, the derivative of a constant function is 0".

In the next example, we'll see that not only is there no value in writing "$+C$", it can sometimes be misleading.

[213] Spivak gets over this issue by defining $\int f$ to be the *set of all primitives* of f; so $F' = f$ translates to $F \in \int f$.

Example 4 $\int dx/x \neq \log x$. This seems strange. Of course $(\log x)' = 1/x$. The (somewhat subtle) issue here is that in an equation like "$F' = f$", asserting that two functions are identical, if there is no specific statement of domains, then our convention is to assume that both F' and f are each defined on their natural domain — the largest subset of the the reals for which the rule defining the function makes sense. In the equation "$(\log x)' = 1/x$", the domain of $\log x$ is $(0, \infty)$, and since log is differentiable at all points in its domain, the domain of $(\log x)'$ is $(0, \infty)$. On the other hand, the domain of $1/x$ is $\mathbb{R} \setminus \{0\}$. So, without qualification on the domains on which the two sides are being considered, it is incorrect to say $(\log x)' = 1/x$.

It is, on the other hand, perfectly correct to say

> on $(0, \infty)$, $(\log x)' = 1/x$, so $\int dx/x = \log x$.

What about on $(-\infty, 0)$? This is the domain of $\log(-x)$, and (by the chain rule) the derivative of $\log(-x)$ is $1/x$. So it is correct to say

> on $(-\infty, 0)$, $(\log(-x))' = 1/x$, so $\int dx/x = \log(-x)$.

This leads to some examples of primitives of $1/x$:

- the function that maps x to $\log x$ if $x > 0$ and $\log(-x)$ if $x < 0$; this can be more compactly expressed as $x \mapsto \log |x|$;
- for any real constant C, the function that maps x to $\log |x| + C$;
- the function that maps x to $3 + \log x$ if $x > 0$ and $12\pi^2 + \log(-x)$ if $x < 0$.

The comment that relates to this example is that

> if the domain of f is not an interval, then
> - one has to be careful about $\int f$, and
> - it's not true that any two antiderivatives of f differ by a constant.

We mentioned earlier that it it is not true that if F is a primitive of f, then all primitives of f are of the form $F + c$ for some constant c. In light of the current example, there is a natural modification to this statement, that is indeed true, and can easily be shown to be true:

> if f is continuous on its domain, and that domain is an interval, and if F is a primitive of f, then all primitives of f are of the form $F + c$ for some constant c. If the domain of f is a union of intervals, and if F and G are two primitives of f, then $F - G$ is constant on each of the intervals.

Example 5 It seems hard to find an antiderivative of e^{-x^2} (for more on this, see the discussion of elementary functions below). However, this function has a very simple antiderivative:

$$\int e^{-x^2}\,dx = \int_c^x e^{-t^2}\,dt$$

where c is any constant (since, by the fundamental theorem of calculus, the derivative of $\int_c^x e^{-t^2}\,dt$ with respect to x is e^{-x^2}).

The comment to be made on this example is:

every continuous function $x \mapsto f(x)$ has a primitive, namely $\int_c^x f(t)\,dt$.

Of course, this is not a particularly *useful* primitive: if we try to use it to calculate a definite integral like $\int_a^b f(x)\,dx$, we get

$$\int_a^b f(x)\,dx = \int_c^b f(x)\,dx - \int_c^a f(t)\,dt,$$

which really doesn't help.

The last example above shows that while finding primitives is easy, what we really want to know about is finding simple, compact expressions for primitives. Computer algebra systems can do this very well: for example, entering

"antiderivative of $1/(1 + \sin x)$"

into Wolfram Alpha yields the answer

$$\text{“}\frac{2\sin(x/2)}{\sin(x/2)+\cos(x/2)} + \text{constant''}.$$

(This is not quite the same as $-2/(1+\tan(x/2))$ that we mentioned earlier; but a little algebra shows that the two expressions $-2/(1+\tan(x/2))$ and $2\sin(x/2)/(\sin(x/2)+\cos(x/2))$ differ from each other by a universal constant).

Given that computers are very good at finding compact expressions for primitives, it's natural to ask why it's useful to spend time, as we will do, developing techniques to find primitives by hand. Here are three reasons why being able to **find** primitives is useful:

1. knowing something of the theory of finding compact expressions for primitives, allows one to troubleshoot when things go wrong using a computer algebra system (as it inevitably will);

2. underlying some of the techniques we describe (in particular integration by parts and integration by partial fractions) are valuable theorems, that are useful to know; and

3. questions about finding primitives they come up on exams, like the GRE.

So, out goal for a while will be to develop techniques to find compact expressions for primitives. "Compact" here means that we are looking for *elementary* functions as primitives:

- rational functions;
- exponential, log, trigonometric (and so hyperbolic) functions and their inverses;
- algebraic functions: functions g satisfying a polynomial equation with rational functions as coefficients (so, for example, functions that extract roots); and
- any function obtained from the previous functions by finitely many additions, subtractions, multiplications, divisions, and compositions.

Essentially, elementary functions are those that can be described in finite time using any combination of the functions $1, x, \sin, \cos, \tan, \arcsin, \arccos, \arctan, \exp$ and $\log$. It is a theorem (though a very hard one) that the function $x \mapsto e^{-x^2}$ does not have an elementary primitive; nor does $\sin x^2$, nor $\sqrt{1+x^3}$. In fact, "most" elementary functions do not have elementary primitives. But still, it will prove very worthwhile to think about those functions that *do* have elementary primitives; and that will be the topic of the next few sections.

13.1 Techniques of integration

There are five basic techniques of integration:

- Know lots of integrals!
- Linearity
- Integration by parts
- Integration by substitution
- Integration by partial fractions

The first two can be discussed quickly. First, know lots of integrals! Every differentiation, when turned on its head, leads to an integration formula, and the more of these you can recognize quickly, the better you will be at integration. Here are some of the integrals we have seen so far:

- $\int x^n dx = x^{n+1}/(n+1)$ for $n \in \mathbb{N}$, as long as $n \neq -1$.
- $\int dx/x = \log|x|$ (as long as $x \neq 0$).
- $\int x^a dx = x^{a+1}/(a+1)$ for real $a \neq -1$ (as long as $x \in (0, \infty)$).
- $\int \sin x dx = -\cos x$.

- $\int \cos x dx = \sin x$.
- $\int \sec^2 x dx = \tan x$.
- $\int e^x dx = e^x$.
- $\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1}(x)$ (as long as $-1 < x < 1$).
- $\int \frac{dx}{1+x^2} = \tan^{-1}(x)$.
- $\int \frac{dx}{\sqrt{x^2+1}} = \sinh^{-1}(x) = \log\left(x + \sqrt{x^2+1}\right)$.
- $\int \frac{dx}{\sqrt{x^2-1}} = \cosh^{-1}(x) = \log\left(x + \sqrt{x^2-1}\right)$ (as long as $x \geq 1$). But for this last example, it is easy to see that if $x < -1$ then, since $x + \sqrt{x^2-1} < 0$, we get $\log -\left(x + \sqrt{x^2-1}\right)$ as an antiderivative of $1/(\sqrt{x^2-1})$; so in fact

$$\int \frac{dx}{\sqrt{x^2-1}} = \log\left|\left(x + \sqrt{x^2-1}\right)\right| \quad \text{(as long as } |x| > 1\text{)}.$$

Second, linearity: if $F = \int f$ and $G = \int g$ then it is an easy check that $aF + bG = \int (af + bg)$.

The other three techniques, integration by parts, substitution and partial fractions, require significantly more discussion.

13.2 Integration by parts

Suppose f', g' are both continuous (so all integrals below exist). We have

$$(fg)' = f'g + fg' \quad \text{or} \quad fg' = (fg)' - f'g.$$

An antiderivative of $(fg)'$ is fg. Suppose $A = \int f'g$ is an antiderivative of $f'g$. Then

$$(fg - A)' = (fg)' - f'g = fg'.$$

In other words, $fg - A$ is an antiderivative of fg'. The traditional way to write this is

$$\int fg' = fg - \int f'g$$

or

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx.$$

This identity is referred to as *integration by parts*, and allows the calculation of one integral ($\int fg'$) to be reduced to the calculation of another integral ($\int f'g$).

Integration by parts has a definite integral form: since

$$(fg)' = f'g + fg',$$

from the fundamental theorem of calculus (part 2) we get that, as long as $[a,b]$ is fully contained in the domains of both f and g,

$$\int_a^b (f'g + fg') = (fg)_a^b$$

or

$$\int_a^b f(x)g'(x)dx = f(x)g(x)|_{x=a}^b - \int_a^b f'(x)g(x)dx.$$

The key to applying integration by parts is to identify that the function to be integrated can be decomposed into the product of two functions, one of which is easy to differentiate (this will play the role of f), and the other of which has an obvious antiderivative (this will play the role of g').

As an example, consider $\int x \log x \; dx$. Here we take $f(x) = \log x$ (so $f'(x) = 1/x$) and $g'(x) = x$ (so one valid choice for g is $g(x) = x^2/2$). We have

$$\int x \log x \; dx = \frac{x^2 \log x}{2} - \int \frac{x}{2} \; dx = \frac{x^2 \log x}{2} - \frac{x^2}{4},$$

a result which can easily be checked by differentiating.

More generally, consider $\int x^a \log x \; dx$ with $a \neq -1$. Here we again take $f(x) = \log x$ (so $f'(x) = 1/x$) and $g'(x) = x^a$, so one valid choice for g is $g(x) = x^{a+1}/(a+1)$). We have

$$\int x \log x \; dx = \frac{x^{a+1} \log x}{a+1} - \int \frac{x^a}{a+1} \; dx = \frac{x^{a+1} \log x}{a+1} - \frac{x^{a+1}}{(a+1)^2}.$$

What about $a = -1$? Again taking $f(x) = \log x$, $f'(x) = 1/x$, $g'(x) = 1/x$, $g(x) = \log x$ (Note: we don't need $\log |x|$ here, since the domain of $(\log x)/x$ is $(0, \infty)$, we get

$$\int \frac{\log x}{x} \; dx = \log^2 x - \int \frac{\log x}{x} \; dx.$$

It appears that we have gone in a circle! But no: we have an (easy) equation which we can solve for $\int (\log x)/x \; dx$, that yields

$$\int \frac{\log x}{x} \; dx = \frac{\log^2 x}{2},$$

again a result which can easily be checked by differentiating.

We need to be a little careful in justifying the above, because of the previous observation that the "=" in $F = \int f$ has to be treated with care. Formally what we are doing is saying: "if A is an antiderivative of $(\log x)/x$, then from integration by parts, $\log^2 x - A$ is also an antiderivative of $\int \frac{\log x}{x} \; dx$. But now, since $(\log x)/x$ is a continuous function defined on an interval, that says that A and $\log^2 x - A$ differ by a constant, or in other words, $A = (\log^2 x)/2 + C$ for some constant C".

Other manipulations that we do with integral equalities can be just as easily be justified formally; we won't do so any more, unless there is an extra subtlety that needs to be pointed out.

Hidden inside the last example was the special case $a = 0$, where we considered $\int \log x \, dx$ — which isn't obviously the product of two functions — and applied integration by parts by "introducing" the function $g'(x) = 1$ into the picture. More generally, for any f with f' continuous, we have

$$\int f = \int f \cdot 1 = xf - \int xf'.$$

For example

$$\int \tan^{-1} x \, dx = x \tan^{-1} x - \int \frac{xdx}{1+x^2} = x \tan^{-1} x - \frac{1}{2} \log(1+x^2).$$

Integration by parts sometimes reduces a more complicated integral to a less complicated one, that still needs some non-trivial works to solve; sometimes even another iteration of integration by parts.

Example: For $n \geq 0$, $n \in \mathbb{N}$, set $I_n = \int x^n e^x \, dx$. We have $I_0 = e^x$ rather easily. For $n > 0$, we use integration by parts with $f(x) = x^n$, $g'(x) = e^x$ to get

$$I_n = x^n e^x - n \int x^{n-1} e^x \, dx = x^n e^x - n I_{n-1}.$$

This is a *reduction formula* that allows us to calculate I_n recursively:

- $I_0 = e^x$
- $I_1 = xe^x - 1 \cdot I_0 = xe^x - e^x = e^x(x-1)$
- $I_2 = x^2 e^x - 2e^x(x-1) = e^x(x^2 - 2x + 2)$
- $I_3 = x^3 e^x - 3e^x(x^2 - 2x + 2) = e^x(x^3 - 3x^2 + 6x - 6)$
- $I_4 = x^4 e^x - 4e^x(x^3 - 3x^2 + 6x - 6) = e^x(x^4 - 4x^3 + 12x^2 - 24x + 24)$

and in general (this is an easy induction) $I_n = e^x P_n(x)$ where $P_n(x)$ is a polynomial of degree n defined recursively by $P_0(x) = 1$ and $P_n = x^n - nP_{n-1}(x)$.

We'll see plenty more reduction formulae.

Notice that in this example, we had a choice: both x^n and e^x are easy both to integrate and differentiate. There's no golden rule for what to do in this case. Sometimes one choice works and the other doesn't, sometimes both do, and sometimes neither work. With lots of practice you should start to develop an intuition; but for the moment, a good rule-of-thumb is:

> if one of the functions involved is a polynomial, try to make the choice that reduces the degree of the polynomial.

This doesn't always work, but often does.

Integration by parts is sometimes symbolically written

$$\int udv = uv - \int vdu$$

Here "u" can be thought of as the part of the function that's easy to differentiate (so, f; its derivative appears as "du"), while "dv" can be thought of as the part of the function that has an easy antiderivative (so g'; its antiderivative appears as "v").

For an example in this language, consider $A_n = \int \frac{dx}{(x^2+1)^n}\, dx$, $n \geq 0$. We have $A_0 = 1$ and $A_1 = \arctan x$. For general $n \geq 2$, set $u = 1/(1+x^2)^n$, $dv = dx$, so $v = x$ and $du = -2nxdx/(1+x^2)^{n+1}$. We get

$$\begin{aligned} A_n &= \frac{x}{(1+x^2)^n} + 2n \int \frac{x^2}{(1+x^2)^{n+1}}\, dx \\ &= \frac{x}{(1+x^2)^n} + 2n \int \frac{(1+x^2)-1}{(1+x^2)^{n+1}}\, dx \\ &= \frac{x}{(1+x^2)^n} + 2nA_n - 2nA_{n+1}. \end{aligned}$$

So

$$A_{n+1} = \frac{x}{2n(1+x^2)^n} + \frac{(2n-1)}{2n} A_n$$

(valid for $n \geq 1$). For example, at $n = 1$ we get

$$\int \frac{dx}{(x^2+1)^2} = \frac{x}{2(1+x^2)} + \frac{\arctan x}{2}.$$

For the rest of this section we'll use an integration by parts reduction formula to derive Wallis' formula for π, and see the connection between Wallis' formula and the binomial coefficients.

We begin by defining, for integers $n \geq 0$, $S_n := \int_0^{\pi/2} \sin^n x dx$. We have

$$S_0 = \frac{\pi}{2}, \quad S_1 = \int_0^{\pi/2} \sin x dx = 1,$$

and for $n \geq 2$ we get from integration by parts (taking $u = \sin^{n-1} x$ and $dv = \sin x dx$, so that $du = (n-1)\sin^{n-2} x \cos x dx$ and $v = -\cos x$) that

$$\begin{aligned} S_n &= (\sin^{n-1} x)(-\cos x)|_{x=0}^{\pi/2} - \int_0^{\pi/2} -(n-1)\cos x \sin^{n-2} x \cos x dx \\ &= (n-1)\int_0^{\pi/2} \cos^2 x \sin^{n-2} x dx \\ &= (n-1)\int_0^{\pi/2} (1-\sin^2 x)\sin^{n-2} x dx \\ &= (n-1)S_{n-2} - (n-1)S_n, \end{aligned}$$

which leads to the recurrence relation

$$S_n = \frac{n-1}{n} S_{n-2} \quad \text{for } n \geq 2.$$

Iterating the recurrence relation until the initial conditions are reached, we get that

$$S_{2n} = \left(\frac{2n-1}{2n}\right)\left(\frac{2n-3}{2n-2}\right)\cdots\left(\frac{3}{4}\right)\left(\frac{1}{2}\right)\frac{\pi}{2}$$

and

$$S_{2n+1} = \left(\frac{2n}{2n+1}\right)\left(\frac{2n-2}{2n-1}\right)\cdots\left(\frac{4}{5}\right)\left(\frac{2}{3}\right)1.$$

Taking the ratio of these two identities and rearranging yields

$$\frac{\pi}{2} = \left(\frac{2}{1}\right)\left(\frac{2}{3}\right)\left(\frac{4}{3}\right)\left(\frac{4}{5}\right)\cdots\left(\frac{2n}{2n-1}\right)\left(\frac{2n}{2n+1}\right)\frac{S_{2n}}{S_{2n+1}}.$$

Now since $0 \leq \sin x \leq 1$ on $[0, \pi/2]$ we have also

$$0 \leq \sin^{2n+1} x \leq \sin^{2n} x \leq \sin^{2n-1} x,$$

and so, integrating and using the recurrence relation, we get

$$0 \leq S_{2n+1} \leq S_{2n} \leq S_{2n-1} = \frac{2n+1}{2n} S_{2n+1}$$

and so

$$1 \leq \frac{S_{2n}}{S_{2n+1}} \leq 1 + \frac{1}{2n}.$$

This says that by choosing n large enough, the ratio S_{2n}/S_{2n+1} can be made arbitrarily close to 1, and so the product

$$\left(\frac{2}{1}\right)\left(\frac{2}{3}\right)\left(\frac{4}{3}\right)\left(\frac{4}{5}\right)\cdots\left(\frac{2n}{2n-1}\right)\left(\frac{2n}{2n+1}\right)$$

can be made arbitrarily close to $\pi/2$ by choosing n large enough. This fact as usually expressed by saying that $\pi/2$ can be described by an "infinite product":

$$\frac{\pi}{2} = \left(\frac{2}{1}\right)\left(\frac{2}{3}\right)\left(\frac{4}{3}\right)\left(\frac{4}{5}\right)\left(\frac{6}{5}\right)\left(\frac{6}{7}\right)\cdots.$$

This infinite product was probably first written down by John Wallis in 1655. Wallis' other claim to fame is that he was probably the first mathematician to use the symbol "∞" for infinity.

Note that Wallis' formula is not a particularly good way to actually estimate π; because we have $1 \leq S_{2n}/S_{2n+1} \leq 1 + 1/2n$, it turns out that to get an estimate of π correct to k decimal places, we need to take $n \approx 10^k$. This is similar to the rate of convergence of the approximation based on $\arctan 1$.

Wallis' formula can be used to estimate the binomial coefficient $\binom{2n}{n}$. Indeed,

$$\begin{aligned}
\binom{2n}{n} &= \frac{(2n)(2n-1)(2n-2)\cdots(3)(2)(1)}{(n)(n-1)\cdots(2)(1)(n)(n-1)\cdots(2)(1)} \\
&= 2^n \frac{(2n-1)(2n-3)\cdots(3)(1)}{(n)(n-1)\cdots(2)(1)} \\
&= 2^{2n} \frac{(2n-1)(2n-3)\cdots(3)(1)}{(2n)(2n-2)\cdots(4)(2)} \\
&= \frac{2^{2n}}{\sqrt{2n+1}}\sqrt{\frac{(2n+1)(2n-1)(2n-1)(2n-3)(2n-3)\cdots(3)(3)(1)}{(2n)(2n)(2n-2)(2n-2)\cdots(4)(4)(2)(2)}}
\end{aligned}$$

and so

$$\frac{\sqrt{n}\binom{2n}{n}}{2^{2n}} = \sqrt{\frac{n}{2n+1}}\sqrt{\frac{(2n+1)(2n-1)(2n-1)(2n-3)(2n-3)\cdots(3)(3)(1)}{(2n)(2n)(2n-2)(2n-2)\cdots(4)(4)(2)(2)}}$$

For large enough n, $\sqrt{n/(2n+1)}$ can be made arbitrarily close to $1/\sqrt{2}$, and the other term on the right-hand side above can (by Wallis' formula) be made arbitrarily close to $\sqrt{2/\pi}$, so the whole right-hand side can be made arbitrarily close to $\sqrt{1/\pi}$. In other words,

$$\lim_{n\to\infty} \frac{\sqrt{n}\binom{2n}{n}}{2^{2n}} = \frac{1}{\sqrt{\pi}}.$$

(Note that this is not a very helpful limit: at $n = 10,000$ the expression $\sqrt{n}\binom{2n}{n}/2^{2n}$ evaluates to around 0.564183, whereas $1/\sqrt{\pi} \approx 0.564189$).

This limit is usually written[214]

$$\frac{\binom{2n}{n}}{2^{2n}} \sim \frac{1}{\sqrt{n\pi}} \quad \text{as } n \to \infty;$$

here I am introducing the symbol "$\sim$", read as "asymptotic to", which is defined as follows:

$$f(n) \sim g(n) \quad \text{as } n \to \infty$$

if $\lim_{n\to\infty} f(n)/g(n) = 1$. The sense is that f and g grow at essentially the same rate as n grows. Note that this does *not* say that f and g get closer to one another absolutely as n grows; for example $n^2 \sim n^2 + n$ as $n \to \infty$, but the difference between the two sides goes to infinity too. It's the *relative* (or proportional) difference that gets smaller.

This estimate for $\binom{2n}{n}$ has a connection to probability. If a fair coin is tossed $2n$ times, then the probability that it comes up heads exactly k times is $\binom{2n}{k}/2^{2n}$. This quantity is at its largest when $n = k$ (some easy algebra), at which point it takes value very close to $1/\sqrt{n\pi}$ (as we have just discovered).

[214] Another way to write this is: $\pi = \lim_{n\to\infty} \frac{16^n}{n\binom{2n}{n}^2}$.

Some easy algebra also suggests that we should expect $\binom{2n}{k}/2^{2n}$ to be quite close to $\binom{2n}{n}/2^{2n}$ for k fairly close to n. If this is the case, then we might expect that the probability of getting some number of heads between $n - n_0$ and $n + n_0$ to be somewhat close to $2n_0$ times the probability of getting n heads, or somewhat close to $2n_0/\sqrt{n\pi}$. If this is true, then by the time n_0 gets up to somewhere around $\sqrt{n}$, the probability of getting some number of heads between $n - n_0$ and $n + n_0$ should be somewhat close to 1.

This intuition can be made precise, in a result called the *central limit theorem*, one of the most important results in probability. One very specific corollary of the central limit theorem is that if a coin is tossed $2n$ times, then for any constant C the probability of getting between $n - C\sqrt{n}$ and $n + C\sqrt{n}$ heads is at least $1 - e^{-C^2/3}$. For example, with $n = 1,000,000$ and $C = 5$, on tossing a coin $2,000,000$ times, the probability of getting between $995,000$ and $1,005,000$ heads is at least $1 - e^{-25/3} \approx .99976$.

13.3 Integration by substitution

Just as the product rule led to integration by parts, the chain rule also leads to an integration principle, integration by substitution.

Integration by substitution, easy case If f, g' are both continuous (so all integrals in question exist), and if F is a primitive of f, then (since $(F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x)$) we have

$$\int f(g(x))g'(x)\ dx = (F \circ g)(x) = F(g(x)).$$

The key to applying this form of integration by substitution is to recognize that the integrand $f(g(x))g'(x)$ can be written as the product of two things — one, $f(g(x))$, is a function f of some basic building block g, and the other, g', is the derivative of the basic building block. If f has a known antiderivative F, then the integral can be expressed in terms of this.

Consider, for example, $\int \tan x\ dx$. Re-expressing as

$$\int \frac{\sin x}{\cos x}\ dx = -\int \frac{-\sin x}{\cos x},$$

there are obvious choices for f ($f(x) = 1/x$, so $F(x) = \log|x|$), and $g(x) = \cos x$, with $g'(x) = -\sin x$. This leads to

$$\int \frac{\sin x}{\cos x}\ dx = -\int \frac{-\sin x}{\cos x} = -\log|\cos x| = \log|\sec x|.$$

Notice that we multiplied the integrand by -1 to "massage" it into the correct form: initially, the integrand was almost of the form $f(g(x))g'(x)$, but not quite (we will return to this issue later).

As another example, consider

$$\int \frac{xdx}{(1+x^2)^n}$$

for $n = 1, 2, \ldots$. Recognizing that the integrand is mostly a function of $1 + x^2$, and that the rest of the integral is almost the derivative of $1 + x^2$, we take

$$f(x) = \frac{1}{x^n} \quad \text{and} \quad g(x) = 1 + x^2.$$

Noting that $g'(x) = 2x$ and

$$F(x) = \int f(x)\ dx = \begin{cases} \log|x| & \text{if } n = 1 \\ \frac{-1}{(n+1)x^{n+1}} & \text{if } n \geq 2, \end{cases}$$

we get

$$\int \frac{xdx}{(1+x^2)^n} = \frac{1}{2}\int \frac{2xdx}{(1+x^2)^n} = \begin{cases} \frac{\log(1+x^2)}{2} & \text{if } n = 1 \\ \frac{-1}{(n+1)(1+x^2)^{n+1}} & \text{if } n \geq 2. \end{cases}$$

There is a shorthand for this process. If we make a *change of variables* — a *substitution* — $u = g(x)$, and (formally) write $du = g'(x)\ dx$, then, re-expressing $\int f(g(x))g'(x)\ dx$ entirely in terms of the new variable u, the integral becomes

$$\int f(u)\ du,$$

which is $F(u)$, or, going back to expressing in terms of variable x, $F(g(x))$. The message here is: if we simply make the substitution $u = g(x)$, re-express the integral in terms of u solve (as a function of u), then go back to expressing in terms of variable x, we get the correct answer — this, even though the expression "$du = g'(x)\ dx$" doesn't (yet) have any official meaning.

This shorthand allows integration by substitution to be done quite quickly, without having to explicitly identify f, g, et cetera.

As an example: via the substitution $u = \log x$ (so $du = dx/x$) we have

$$\int \frac{(\log x)^2}{x}dx = \int u^2 du = \frac{u^3}{3} = \frac{(\log x)^3}{x}.$$

In all examples so far we were extremely fortunate that either (as in the last example) one part of the integrand was easily recognizable as the derivative of another, or (as in the other examples), that became the case after a simple manipulation. We now present a more general substitution method that is far more versatile, in that in can (in principle) be used for *any* integrand. The idea is that we for (essentially) any function f that is given, and (essentially) any function g (of our choosing), we can (usually) express $f(x)$ in the form $(H \circ g)(x)g'(x)$ for a suitably constructed function H, and then use "easy" integration by substitution to complete the integration.

Indeed,

$$\begin{aligned} f(x) &= \frac{f(x)}{g'(x)}g'(x) \\ &= \frac{f(g^{-1}(g(x)))}{g'(g^{-1}(g(x)))}g'(x) \\ &= \left(\frac{f \circ g^{-1}}{g' \circ g^{-1}}\right)(g(x))g'(x), \end{aligned}$$

all steps valid as long as:

- g is differentiable (so continuous)
- g' is never 0, and
- g^{-1} exists (so, since g continuous, g must be monotone, assuming that we are working on some interval).

From this we get:

Integration by substitution, general case If f and g' are both continuous (so all integrals in question are certain to exist), and if g is invertible, with non-zero derivative, and if H is a primitive of

$$\frac{f \circ g^{-1}}{g' \circ g^{-1}},$$

then (since

$$\begin{aligned} (H \circ g)'(x) &= H'(g(x))g'(x) \\ &= \frac{f(g^{-1}(g(x)))}{g'(g^{-1}(g(x)))} g'(x) \\ &= \frac{f(x)}{g'(x)} g'(x) \\ &= f(x) \end{aligned}$$

we have

$$\int f(x)\ dx = (H \circ g)(x) = H(g(x)).$$

As a simple example, suppose that we are considering $\int x/(1+x^2)\ dx$, and we do not recognize that after a simple manipulation the integrand becomes of the form $f(g(x))g'(x)$. We have at our disposal the general method of integration by substitution, which allows us to re-express the integrand, by "substituting out" any expression that we may choose. A general rule-of-thumb to keep in mind for integration by substitution is:

> identify any awkward/prominent/annoying part of the integrand, and try to substitute that out.

Here, the awkward/prominent/annoying part of the integrand is the $1 + x^2$, so we set $g(x) = 1 + x^2$. We have to be a little careful now, since g is *not* an invertible function either on its domain, or on the domain of the integrand. It becomes invertible if we restrict it to either non-negative reals, or non-positive reals, so let us do that.

First, consider the problem on non-negative reals. We have $g(x) = 1 + x^2$, $g'(x) = 2x$[215] and $g^{-1}(x) = \sqrt{x-1}$[216]. So

$$\frac{f \circ g^{-1}}{g' \circ g^{-1}}(x) = \frac{\frac{\sqrt{x-1}}{x}}{2\sqrt{x-1}} = \frac{1}{2x}$$

which has primitive $(\log x)/2$ (note x non-negative here, so we don't need the absolute value). So we may take $H(x) = (\log x)/2$, and get that

$$H(g(x)) = \frac{\log(1+x^2)}{2}$$

is an antiderivative of the original function f, at least when we restrict to the domain of positive reals. A similar calculation gives that an antiderivative of f is $(1/2)\log(1+x^2)$ on negative reals (now $g^{-1} = -\sqrt{x-1}$, but the negative sign disappears in the calculation of $(f \circ g^{-1})/(g' \circ g^{-1})$, since it appears in numerator and denominator).

As with the easy substitution method, there is a shorthand way to proceed. Start with the substitution $u = g(x)$, and then re-express everything in the integrand in terms of u:

- $u = g(x)$ so $x = g^{-1}(u)$ (requiring g to be invertible)
- $du = g'(x)dx$ so $dx = du/g'(x) = du/g'(g^{-1}(x))$ (requiring g' not to ever be 0)
- $f(x) = f(g^{-1}(u))$.

The integral becomes

$$\int \frac{(f \circ g^{-1})(u)}{(g' \circ g^{-1})(u)}\, du,$$

so if H is a primitive of $(f \circ g^{-1})/(g' \circ g^{-1})$, then the integral is $H(u)$, or, in terms of x, $H(g(x))$.

As a simple example, consider $\int f(ax+b)\, dx$, where a, b are constants and where F is a known primitive of f. Via the substitution $u = ax + b$ (so $du = adx$, $dx = du/a$), we get

$$\begin{aligned} \int f(ax+b)\, dx &= \int \frac{f(u)}{a} du \\ &= \int \frac{1}{a} \int f(u)\, du \\ &= \frac{1}{a} F(u) \\ &= \frac{F(ax+b)}{a}. \end{aligned}$$

[215]There will clearly be a problem at 0, since g' is 0 there. So let's restrict the domain of g a little further, to *positive* reals.

[216]Note that restricting g to positive reals, it has range $(1, \infty)$, so that is the domain of g^{-1}, while the range of g^{-1} is positive reals. That is why we take the positive square root when computing g^{-1}.

Notice that we didn't have to think about the involved expression $(f \circ g^{-1})/(g' \circ g^{-1})$ here; simply by formally re-expressing the whole integrand in terms of the new variable u, we inevitably reach $(f \circ g^{-1})(u))/((g' \circ g^{-1})(u)\ du$.

The process of integration by substitution in the general case can be thought of as being similar to the process of integration by parts, in that it can be used to replace one integration problem with another, hopefully simpler, one. It can be done entirely mechanically. As stated earlier in more general terms, a rough guiding principle should be:

> identify a "prominent" part of the integrand, call it $g(x)$, and substitute for it by setting $u = g(x)$ and then completely re-expressing the integrand in terms of u.

This leads to a new integral that, since it makes no reference to g, is hopefully simpler to evaluate that the original. As the next few examples show, this new, simpler integral may require the application of some other integration techniques, (maybe another application of integration by substitution) to crack; or, as we will see in at least one example, the new integral may be just as hopeless as the old.

Example 1 $\int 1/(1+\sqrt{1+x})\ dx$. Here an obvious substitution is $u = \sqrt{1+x}$, which gives $du = dx/(2\sqrt{1+x}) = dx/2u$, so $dx = 2udu$. We get

$$\begin{aligned} \int \frac{dx}{1+\sqrt{1+x}} &= \int \frac{2udu}{1+u} \\ &= 2\int \frac{udu}{1+u}. \end{aligned}$$

We now do another substitution, $w = 1+u$, so $dw = du$, and $u = w-1$, leading to

$$\begin{aligned} \int \frac{udu}{1+u} &= \int \frac{(w-1)dw}{w} \\ &= \int dw - \int \frac{dw}{w} \\ &= w - \log|w|. \end{aligned}$$

Reversing the substitutions,

$$\begin{aligned} \int \frac{dx}{1+\sqrt{1+x}} &= 2\int \frac{udu}{1+u} \\ &= 2(w - \log|w|) \\ &= 2\left((1+u) - \log|1+u|\right) \\ &= 2\left(1+\sqrt{1+x} - \log\left(1+\sqrt{1+x}\right)\right) \end{aligned}$$

(with the absolute value sign removed in the last log, since $1+\sqrt{1+x} > 0$ always).[217]

[217] The substitution $u = 1+\sqrt{1+x}$ would also have worked here; as would the trick of writing $u/(1+u) = ((u+1)-1)/(u+1) = 1 - 1/(u+1)$, with obvious antiderivative $u - \log|1+u|$, instead of using a second substitution; note that this would have lead to the final answer $2\left(\sqrt{1+x} - \log\left(1+\sqrt{1+x}\right)\right)$, differing from the answer we got by a constant.

Example 2 $\int e^{\sqrt{x}}\ dx$. An obvious substitution is $u = \sqrt{x}$, with $du = dx/2\sqrt{x}$, so $dx = 2\sqrt{x}du = 2udu$, leading to

$$\begin{aligned}\int e^{\sqrt{x}}\ dx &= 2\int ue^u\ du \\ &= 2\left(ue^u - \int e^u\right) \quad \text{(integration by parts)} \\ &= 2\left(ue^u - e^u\right) \\ &= 2\left(\sqrt{x}e^{\sqrt{x}} - e^{\sqrt{x}}\right).\end{aligned}$$

We could have also tried the substitution $w = e^{\sqrt{x}}$, so $dw = ((e^{\sqrt{x}})/2\sqrt{x})dx$, or $dx = 2(\log w)/w$, which leads to

$$\int e^{\sqrt{x}}\ dx 2\int \log w\ dw,$$

and again an application of integration by parts finishes things.

Example 3 $\int e^{x^2}\ dx$. Here the obvious substitution is $u = x^2$, [218] so $du = 2xdx$, and $dx = du/(2x) = du/(2\sqrt{u})$, leading to

$$\int e^{x^2}\ dx = \frac{1}{2}\int \frac{e^u du}{\sqrt{u}}.$$

The obvious substitution here, $w = \sqrt{u}$, just returns us to $\int e^{w^2}\ dw$, and no other substitution or clever integration by parts helps matters — as mentioned earlier, e^{x^2} is a function with no elementary antiderivative.

There is a definite integral version of integration by substitution. With the notation as in the indefinite version, we have

$$\int_a^b f(x)\ dx = (H \circ g)(x)|_{x=a}^b = H(u)|_{u=g(a)}^{g(b)} = \int_{g(a)}^{g(b)} \frac{(f \circ g^{-1})(u)}{(g' \circ g^{-1})(u)}\ du.$$

So, the only change between definite and indefinite integration is that after the substitution $u = g(x)$, as well as re-expressing the integrand in terms of u, we also re-express the limits of integration in terms of u; and then there is no need to re-express things in terms of x before evaluating the integral.

We illustrate with some examples. Consider first $\int_{\pi/4}^{\pi/2} \cot x\ dx = \int_{\pi/4}^{\pi/2} \frac{\cos x\ dx}{\sin x}$. Set $u = \sin x$, so $du = \cos x dx$. At $x = \pi/4$ we have $u = \sin(\pi/4) = \sqrt{2}/2$, and at $x = \pi/2$ we have $u = 1$, so

$$\int_{\pi/4}^{\pi/2} \frac{\cos x\ dx}{\sin x} = \int_{\sqrt{2}/2}^{1} \frac{du}{u} = \log 1 - \log(\sqrt{2}/2) = (\log 2)/2.$$

[218] As with the example of $x/(1+x^2)$, we formally should split the domain of x^2 into two invertible parts to do this example correctly.

As another simple example, consider $\int_0^1 \frac{xdx}{1+x^2} = \frac{1}{2}\int_0^1 \frac{2xdx}{1+x^2}$. We set[219] $u = 1+x^2$, so $du = 2xdx$. At $x = 0$ we have $u = 1$, and at $x = 1$ we have $u = 2$. So the integral is

$$\frac{1}{2}\int_1^2 \frac{du}{u} = (\log 2)/2.$$

It's tempting here to conjecture:

> all definite integrals calculated by substitution evaluate to $(\log 2)/2$.

This is false, however.[220]

13.4 Some special (trigonometric) substitutions

To illustrate how the rough guiding principle of integration by substitution might sometimes break down, consider $\int \sqrt{1-x^2}dx$. It's tempting to try the substitution $u = x^2$, so $du = 2xdx$, $dx = du/2x = \frac{1}{2}\frac{du}{\sqrt{u}}$, making the integral

$$\frac{1}{2}\int \sqrt{\frac{1-u}{u}}du,$$

and any obvious substitution gets you right back where you started.

Alternately, one could try the completely non-obvious substitution $u = \arcsin x$ (note that the domain of the integrand is $[-1, 1]$, which is exactly the domain of arcsin), so $x = \sin u$, $1 - x^2 = \cos^2 u$, $\sqrt{1-x^2} = \cos u$ (as x ranges over $[-1, 1]$, u ranges over $[-\pi/2, \pi/2]$, where cos is positive), $dx = \cos u \; du$, and the integral becomes

$$\int \cos^2 u \; du,$$

a completely different kettle of fish, and possibly amenable to a more direct attack than $\int \sqrt{1-x^2}dx$.

There is a general principle here.

Trigonometric substitutions, 1 A function involving a square root of a quadratic expression can often be reduced to an integral involving trigonometric functions, via the following substitutions (all motivated by the identity $\sin^2 + \cos^2 = 1$ and its relatives). Note that throughout we may assume $a, b > 0$.

- If the integrand involves $\sqrt{a^2 - b^2x^2}$, try the substitution $u = \arcsin \frac{bx}{a}$, or $x = \frac{a}{b}\sin u$. With this substitution,

$$\sqrt{a^2 - b^2x^2} = a\sqrt{1 - b^2x^2/a^2} = a\sqrt{1 - \sin^2 u} = a\sqrt{\cos^2 u} = a\cos u,$$

[219]Note that $1 + x^2$ *is* invertible on the domain $[0, 1]$.

[220]This was a joke.

(so here we are motivated by $\cos^2 u = 1 - \sin^2 u$), while $dx = \frac{a}{b}\cos u \, du$. The domain of $\sqrt{a^2 - b^2x^2}$ is $[-a/b, a/b]$. For x on this domain, bx/a ranges over $[-1, 1]$ (the domain of arcsin), so the substitution makes sense. Since arcsin has range $[-\pi/2, \pi/2]$, and on this range cos is positive, we can justify the line $a\sqrt{\cos^2 u} = a\cos u$ above.

- If the integrand involves $\sqrt{a^2 + b^2x^2}$, try the substitution $u = \arctan\frac{bx}{a}$, or $x = \frac{a}{b}\tan u$. With this substitution,

$$\sqrt{a^2 + b^2x^2} = a\sqrt{1 + b^2x^2/a^2} = a\sqrt{1 + \tan^2 u} = a\sqrt{\sec^2 u} = a\sec u,$$

  (so here we are motivated by $\sec^2 u = 1 + \tan^2 u$), while $dx = \frac{a}{b}\sec^2 u \, du$. The domain of $\sqrt{a^2 + b^2x^2}$ is $\mathbb{R}$. For x on this domain, bx/a ranges over $\mathbb{R}$ (the domain of arctan), so the substitution makes sense. Since arctan has range $[-\pi/2, \pi/2]$, and on this range sec is positive, we can justify the line $a\sqrt{\sec^2 u} = a\sec u$ above.

- If the integrand involves $\sqrt{b^2x^2 - a^2}$, try the substitution[221] $u = \operatorname{arcsec}\frac{bx}{a}$, or $x = \frac{a}{b}\sec u$. With this substitution,

$$\sqrt{b^2x^2 - a^2} = a\sqrt{b^2x^2/a^2 - 1} = a\sqrt{\sec^2 u - 1} = a\sqrt{\tan^2 u} = a|\tan u|,$$

  (so here we are motivated by $\tan^2 u = \sec^2 u - 1$), while $dx = \frac{a}{b}\sec u \tan u \, du$. In the last two cases we wrote (and justified) $a\sqrt{\cos^2 u} = a\cos u$ and $a\sqrt{\sec^2 u} = a\sec u$; here we have to be a little more careful, and really need to write $a\sqrt{\tan^2 u} = a|\tan u|$. Indeed, the domain of $\sqrt{b^2x^2 - a^2}$ is $(-\infty, -a/b] \cup [a/b, \infty)$. If we are on the negative part of this domain then bx/a ranges over $(-\infty, -1]$, which is the negative part of the domain of arcsec. On this domain, arcsec ranges over the values $(\pi/2, \pi]$, and the tangent function is negative here. If we are on the positive part of the domain of $\sqrt{b^2x^2 - a^2}$ then bx/a ranges over $[1, \infty)$, which is the positive part of the domain of arcsec. On this domain, arcsec ranges over the values $[0, \pi/2)$, and the tangent function is positive here. So we get

$$\sqrt{b^2x^2 - a^2} = \begin{cases} -a\tan u & \text{if } x < -a/b, \\ a\tan u & \text{if } x > a/b. \end{cases}$$

We've already seen the example of $\int \sqrt{1 - x^2}dx$ transforming into $\int \cos^2 u \, du$ via the substitution $x = \sin u$. Here is another example, that involves the arcsec function, and so requires some care.

Example $\int \frac{\sqrt{25x^2 - 4}}{x} \, dx$. Following the discussion above, the sensible substitution is $x = (2/5)\sec u$, so $dx = (2/5)\sec u \tan u \, du$, and

$$\sqrt{25x^2 - 4} = 5\sqrt{x^2 - (2/5)^2} = 5\sqrt{(2/5)^2\sec^2 u - (2/5)^2} = 2\sqrt{\sec^2 u - 1} = 2\sqrt{\tan^2 u}.$$

Following the discussion above, we know that we have to treat separately the cases $x \geq 2/5$ and $x \leq -2/5$.

[221] The arcsec function, which somewhat weird, is discussed at the very end of Section 12.4.

- **Case of $x \geq 2/5$:** Here $2\sqrt{\tan^2 u} = 2\tan u$, so

$$\sqrt{25x^2-4} = 2\tan u,$$

and the integral becomes

$$2\int \tan^2 u du.$$

We'll discuss trigonometric integrals like this in general in a short while, but this one can be dealt with fairly easily: using $\sec^2 - 1 = \tan^2$ we get

$$2\int \tan^2 u du = 2\int (\sec^2 u - 1) du = 2\left(\tan u - u\right).$$

We would like to re-express this in terms of x, recalling $x = (2/5)\sec u$. One way is to simply write

$$2\left(\tan u - u\right) = 2\left(\tan\left(\sec^{-1}\left(\frac{5x}{2}\right)\right) - \sec^{-1}\left(\frac{5x}{2}\right)\right).$$

This can be considerable cleaned up.

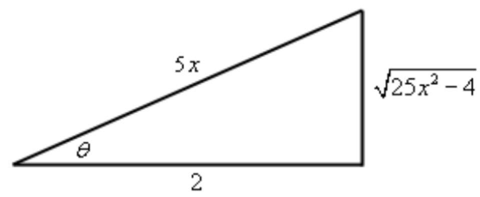
Since $x \geq 2/5$ we have $5x/2 \geq 1$, and so $\sec^{-1}(5x/2)$ is between 0 and $\pi/2$. Now we use $\sec^2 = 1 + \tan^2$ to get that $(5x/2)^2 = 1 + \tan^2(\sec^{-1}(5x/2))$, so $\tan(\sec^{-1}(5x/2)) = \pm\sqrt{(5x/2)^2 - 1} = \pm\sqrt{25x^2-4}/2$. But which is it, plus or minus? Well, since $\sec^{-1}(5x/2)$ is between 0 and $\pi/2$, and tan is positive in that domain, we must take the positive square root.

As it happens, we can also re-express $\sec^{-1}(5x/2)$ in terms of simpler (more fundamental) trigonometric functions. We have that $\sec^{-1}$ is the inverse of the composition $(r \circ \cos)$, where r is the reciprocal function $x \mapsto 1/x$. We know[222] that $(r \circ \cos)^{-1} = \cos^{-1} \circ r^{-1} = \cos^{-1} \circ r$ (because $r = r^{-1}$). So $\sec^{-1}(5x/2) = \cos^{-1}(2/5x)$ (and this calculation does not depend on $x \geq 2/5$; it works for all x in the domain). So we get, in this case,

$$\int \frac{\sqrt{25x^2-4}}{x}\,dx = \sqrt{25x^2-4} - 2\cos^{-1}\left(\frac{2}{5x}\right).$$

There is another way to do the simplifying calculation, that may be more intuitive. Recall that in the present case $\sec^{-1}(5x/2)$ is between 0 and $\pi/2$. Remembering that in a right-angled triangle with one angle θ, the secant of θ is the length of the side opposite the angle, divided by the length of the hypothenuse, we are led to the following right-angled triangle:

[222] This was a homework problem.



The angle θ is $\sec^{-1}(5x/2)$, which can clearly be expressed in terms of more fundamental inverse trigonometric functions: for example, it is $\cos^{-1}(2/5x)$; and $\tan\theta$ is $\sqrt{25x^2-4}/2$.

An issue with the "right-angled triangle" approach is that it is not so easy to implement it when the angle one is working with is negative, or greater than $\pi/2$; in this case the earlier method is probably more reliable.

- **Case of $x \leq -2/5$**: Here, following the discussion earlier, $2\sqrt{\tan^2 u} = -2\tan u$, so the integral becomes

$$\begin{aligned} -2\int \tan^2 u du &= -2\int(\sec^2 u - 1)du \\ &= -2\tan u + 2u \\ &= -2\tan\left(\sec^{-1}\left(\frac{5x}{2}\right)\right) + 2\sec^{-1}\left(\frac{5x}{2}\right) \\ &= -2\tan\left(\sec^{-1}\left(\frac{5x}{2}\right)\right) + 2\cos^{-1}\left(\frac{2}{5x}\right). \end{aligned}$$

Since $x \leq -2/5$ we have $5x/2 \leq -1$, and so $\sec^{-1}(5x/2)$ is between $\pi/2$ and π. Now we use $\sec^2 = 1 + \tan^2$ to get that $(5x/2)^2 = 1 + \tan^2(\sec^{-1}(5x/2))$, so $\tan(\sec^{-1}(5x/2)) = \pm\sqrt{(5x/2)^2-1} = \pm\sqrt{25x^2-4}/2$. Since $\sec^{-1}(5x/2)$ is between $\pi/2$ and π, and tan is *negative* in that domain, we must take the *negative* square root — $\tan(\sec^{-1}(5x/2)) = -\sqrt{25x^2-4}/2$ and

$$\int \frac{\sqrt{25x^2-4}}{x}\,dx = \sqrt{25x^2-4} + 2\cos^{-1}\left(\frac{2}{5x}\right)$$

in this case. Note the subtle difference between the two regimes: when x is positive we subtract $2\cos^{-1}(2/5x)$, while when x is negative we add that term.[223]

[223] If you ask Wolfram Alpha to evaluate the integral in this example, you get the answer

$$\int \frac{\sqrt{25x^2-4}}{x}\,dx = \sqrt{25x^2-4} + 2\tan^{-1}\left(\frac{2}{\sqrt{25x^2-4}}\right),$$

valid for all x with $|x| \geq 2/5$. Amusingly, this answer *does not* differ from our answer by a universal constant. On $[2/5,\infty)$, $\tan^{-1}(2/\sqrt{25x^2-4})$ and $-\cos^{-1}(2/5x)$ differ by a constant, while on $(-\infty, 2/5]$, $\tan^{-1}(2/\sqrt{25x^2-4})$ and $+\cos^{-1}(2/5x)$ differ by a *different* constant; see the graph below:

The above discussion covers all cases where there is a quadratic expression under a square root, *when the quadratic has no linear term.*[224] But every quadratic with a linear term can be massaged into a quadratic without a linear term, by a simple linear substitution, a process known as "completing the square". This goes as follows: first assuming that $p > 0$,

$$\begin{aligned} px^2 + qx + r &= \left(\sqrt{p}x + \frac{q}{2\sqrt{p}}\right)^2 + r - \frac{q^2}{4p} \\ &= y^2 + r - \frac{q^2}{4p}. \end{aligned}$$

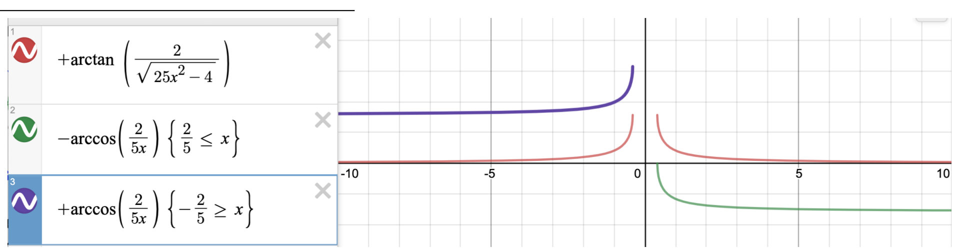
If $r > q^2/4p$ then the substitution $y = \sqrt{p}x + q/(2\sqrt{p})$ reduces $px^2 + qx + r$ to the form $y^2 + a^2$; if $r < q^2/4p$ then it reduces it to the form $y^2 - a^2$. If, on the other hand, $p < 0$, then

$$\begin{aligned} px^2 + qx + r &= -((-p)x^2 - qx) + r \\ &= -\left(\sqrt{-p}x - \frac{q}{2\sqrt{-p}}\right)^2 + r - \frac{q^2}{4p} \\ &= \left(r - \frac{q^2}{4p}\right) - y^2. \end{aligned}$$

If $r > q^2/4p$ then the substitution $y = \sqrt{-p}x - q/(2\sqrt{-p})$ reduces $px^2 + qx + r$ to the form $y^2 - a^2$. If, on the other hand, $r < q^2/4p$, then from the quadratic formula we find that the (real) range of $px^2 + qx + r$ is empty, and we are in the one case where there is no point to considering the integration.

Here's an alternate approach to completing the square, using the quadratic formula:

$$\begin{aligned} px^2 + qx + r &= p\left(x^2 + \frac{q}{p}x + \frac{r}{p}\right) \\ &= p\left(x - \left(\frac{-\frac{q}{p} + \sqrt{\frac{q^2}{p^2} - 4\frac{r}{p}}}{2}\right)\right)\left(x - \left(\frac{-\frac{q}{p} - \sqrt{\frac{q^2}{p^2} - 4\frac{r}{p}}}{2}\right)\right) \\ &= p\left(\left(x + \frac{q}{2p}\right) - \frac{1}{2}\sqrt{\frac{q^2}{p^2} - \frac{4r}{p}}\right)\left(\left(x + \frac{q}{2p}\right) + \frac{1}{2}\sqrt{\frac{q^2}{p^2} - \frac{4r}{p}}\right) \\ &= p\left(\left(x + \frac{q}{2p}\right)^2 - \left(\frac{1}{2}\sqrt{\frac{q^2}{p^2} - \frac{4r}{p}}\right)^2\right) \end{aligned}$$



This is cautionary example that shows that you have to keep your wits about you when dealing with integrals of functions that are defined on unions of intervals.

[224] We don't consider the fourth case, $\sqrt{-b^2x^2 - a^2}$, since this has empty domain.

If $p > 0$ we have reduced to one of the the forms $a^2 + b^2x^2$ or $b^2x^2 - a^2$ (which one depending on whether $(q^2/p^2) - (4r)(p)$ is positive or negative). If $p < 0$ then $(q^2/p^2) - (4r)(p)$ must be non-negative (otherwise there are no reals in the range of $px^2 + qx + r$), and we have reduced to the form $a^2 + b^2x^2$.

Example $\int \frac{x}{\sqrt{2x^2-4x-7}}\, dx$. To make things easier, we pull out a factor of $\sqrt{2}$:

$$\int \frac{x}{\sqrt{2x^2 - 4x - 7}}\, dx = \frac{1}{\sqrt{2}} \int \frac{x}{\sqrt{x^2 - 2x - (7/2)}}\, dx.$$

We now complete the square:

$$x^2 - 2x - (7/2) = (x-1)^2 - 1 - (7/2) = (x-1)^2 - (3/\sqrt{2})^2.$$

We make the substitution $u = x - 1$, so $du = dx$, and $x = u + 1$, so the integral becomes

$$\begin{aligned} \int \frac{x}{\sqrt{x^2 - 2x - (7/2)}}\, dx &= \int \frac{u+1}{\sqrt{u^2 - (3/\sqrt{2})^2}}\, dx \\ &= \int \frac{u}{\sqrt{u^2 - (3/\sqrt{2})^2}}\, dx + \int \frac{1}{\sqrt{u^2 - (3/\sqrt{2})^2}}\, dx. \end{aligned}$$

The first of these integrals can be handled by a simple substitution $v = u^2 - (3/\sqrt{2})^2$. For the second, the earlier discussion suggests the substitution $u = 3/\sqrt{2}\sec v$. The details are left as an exercise.

The examples given above indicate that it is important to understand the integrals of functions of sin, cos, et cetera, as these pop up naturally in the study functions involving square roots of quadratics. All functions of trigonometric functions can be re-expressed purely in terms of sin and cos, so we concentrate our attention only on functions of sin and cos. In particular, we are going to focus attention on functions of the form $\cos^p \sin^q$, where p and q are integers; using linearity of the integral, virtually every function of trigonometric functions that we will need to be able to deal with, can be reduced to this form.

Our examination of this integrals will break into cases, according to the parity (oddness or evenness) of p, q.

- **Case 1**: p is odd. Here we write

$$\begin{aligned} \int \cos^p x \sin^q x\; dx &= \int \cos^{p-1} x \sin^q x\; \cos x\; dx \\ &= \int \left(\cos^2 x\right)^{\frac{p-1}{2}} \sin^q x\; \cos x\; dx \\ &= \int \left(1 - \sin^2 x\right)^{\frac{p-1}{2}} \sin^q x\; \cos x\; dx \end{aligned}$$

(i.e., we "peel off" one copy of $\cos x$, to join the dx; note that we can do this whether p is positive or negative; note also that since p is odd, $(p-1)/2$ is an integer). We now make the substitution $u = \sin x$, so $du = \cos x dx$, to get

$$\int \cos^p x \sin^q x \; dx = \int \left(1-u^2\right)^{\frac{p-1}{2}} u^q \; du.$$

If $(p-1)/2 \geq 0$ then expanding out the polynomial $\left(1-u^2\right)^{\frac{p-1}{2}}$ and multiplying through by u^q, we have reduced to integrating a linear combination of functions of the form u^k, $k \in \mathbb{Z}$ (very easy); if $(p-1)/2 < 0$, then at least we have reduced to integrating a rational function in u, a topic that we will shortly address.

Example: $\int \cos^3 x \sin^4 x \; dx$. We make the substitution $u = \sin x$, $du = \cos x dx$, so

$$\int \cos^3 x \sin^4 x \, dx = \int \cos^2 x \sin^4 \cos x \, dx = \int (1-\sin^2) x \sin^4 \cos x \, dx = \int (1-u^2) u^4 \, du.$$

- **Case 2**: q is odd. This is almost identical to Case 1. Here we "peel off" one copy of $\sin x$, to join the dx, and make the substitution $u = \cos x$, so $du = -\sin x dx$, and we reduce to a rational function in u via $(\sin^2 x)^{(p-1)/2} = (1-\cos^2 x)^{(p-1)/2} = (1-u^2)^{(p-1)/2}$.

  Example: $\int \frac{\cos^2 x}{\sin^5 x} \; dx$. After the substitution $u = \cos x$,

$$\int \frac{\cos^2 x}{\sin^5 x} \; dx = \int \frac{\cos^2 x}{\sin^6 x} \sin x \; dx = -\int \frac{u^2}{(1-u^2)^3} \; du.$$

- **Case 3**: p, q even, both non-negative, at least one positive[225]. From

$$\begin{aligned} \cos^2 x + \sin^2 x &= 1 \\ \cos^2 x - \sin^2 x &= \cos 2x, \end{aligned}$$

  we get the identities

$$\cos^2 x = \frac{1+\cos 2x}{2}, \quad \text{and} \quad \sin^2 x = \frac{1-\cos 2x}{2}$$

  which leads to

$$\cos^p x \sin^q x = \left(\frac{1+\cos 2x}{2}\right)^{\frac{p}{2}} \left(\frac{1+\cos 2x}{2}\right)^{\frac{q}{2}}.$$

  Expanding this out, and separating out the monomials in the polynomial, we get a collection of integrands of the form $\cos^{p'} 2x$ with p' non-negative, and with p' smaller than p. Any such terms with p' odd can be dealt with by an application of Case 1;

[225] Things are rather trivial if both $p, q = 0$...

any such terms with both p' even can be dealt with by *another* application of Case 3. Because the highest powers involved are strictly decreasing, this process terminates after a finite number of iterations.

Example: $\int \frac{\cos^6 x}{\sin^2 x}\, dx$. We write

$$\begin{aligned}\frac{\cos^6 x}{\sin^2 x} &= (\cos^2 x)^3(\sin^2 x) \\ &= \left(\frac{1+\cos 2x}{2}\right)^3\left(\frac{1-\cos 2x}{2}\right) \\ &= \frac{1}{16}\left(1+2\cos 2x - 2\cos^3 2x - \cos^4 2x\right).\end{aligned}$$

So (ignoring the constants) we've reduced to four integrals:

- the 1 is trivial;
- the $\cos 2x$ is easy (after the substitution $u = 2x$, which is the sort of easy substitution one should just do in one's head);
- the $\cos^3 2x$ is an instance of Case 1, and can be dealt with by the substitution $u = \sin 2x$.
- the $\cos^4 2x$ is an instance of Case 3, but with smaller powers than the original instance. We write

  $$\cos^4 2x = \left(\frac{1+\cos 4x}{2}\right)^2 = \frac{1}{4}(1+2\cos 4x + \cos^2 4x).$$

  We have three simpler integrals, the first trivial, the second an instance of Case 1, and the third and even simpler instance of Case 3 (clearly, the last one that will be encountered in this particular example).

This doesn't cover every expression of the form $\cos^p x \sin^q x$ with $p, q \in \mathbb{Z}$; for example, it omits the case where p, q are both even and non-positive, with at least one of them negative.[226] This case, and a whole many more trigonometric integrals, can be dealt with by the following "magic bullet".

The last-resort trigonometric substitution Consider the substitution $t = \tan x/2$. We have

$$dt = \frac{\sec^2 x/2}{2}dx = \frac{1+\tan^2 x/2}{2}dx = \frac{1+t^2}{2}dx,$$

so

$$dx = \frac{2dt}{1+t^2}.$$

[226] I think that this is the only non-trivial omitted case.

Also,

$$\begin{aligned}
\sin x &= 2\sin(x/2)\cos(x/2) \\
&= 2\frac{\sin(x/2)\cos^2(x/2)}{\cos(x/2)} \\
&= 2\tan(x/2)\cos^2(x/2) \\
&= \frac{2t}{\sec^2(x/2)} \\
&= \frac{2t}{1+\tan^2(x/2)} \\
&= \frac{2t}{1+t^2},
\end{aligned}$$

with everything valid exactly as long as $\tan(x/2)$ is defined. And since $\cos x = \cos^2(x/2) - \sin^2(x/2)$ and $1 = \cos^2(x/2) + \sin^2(x/2)$, we have

$$\begin{aligned}
\cos x &= 1 - 2\sin^2(x/2) \\
&= 1 - 2\frac{\sin^2(x/2)\cos^2(x/2)}{\cos^2(x/2)} \\
&= 1 - 2\tan^2(x/2)\cos^2(x/2) \\
&= 1 - 2t^2\cos^2(x/2) \\
&= 1 - 2\left(\frac{t^2}{\sec^2(x/2)}\right) \\
&= 1 - 2\left(\frac{t^2}{1+\tan^2(x/2)}\right) \\
&= 1 - 2\left(\frac{t^2}{1+t^2}\right) \\
&= \frac{1-t^2}{1+t^2},
\end{aligned}$$

again with everything valid exactly as long as $\tan(x/2)$ is defined.

The upshot of this is

> *any* integrand in the variable x that is a function of $\sin x$, $\cos x$ (and the other trigonometric functions) (not necessarily a rational function — it could involve roots, and exponentials, too) can be converted into an integrand in the variable t that does *not* mention any trigonometric functions, by the substitution $t = \tan x/2$ (though this substitution does not do away with roots or exponentials). In particular, if an integrand is a *rational* function of trigonometric functions, it can be converted into a rational function of t by this substitution.

This is a "last resort" substitution, because in general if there is *any* other way to approach the integration problem, the path is almost always going to be easier that way![227]

We give some examples:

Example 1 $\int \cos^2 x \, dx$. The obvious thing to do here is to write

$$\int \cos^2 x \, dx = \int \left(\frac{1+\cos 2x}{2}\right) \, dx = \frac{x}{2} + \frac{\sin 2x}{4}.$$

Using the "last resort" substitution $t = \tan(x/2)$ we get

$$\begin{aligned} \int \cos^2 x \, dx &= \int \left(\frac{1-t^2}{1+t^2}\right)^2 \frac{2}{1+t^2} \, dt \\ &= 2\int \frac{(1-t^2)^2}{(1+t^2)^3} \, dt. \end{aligned}$$

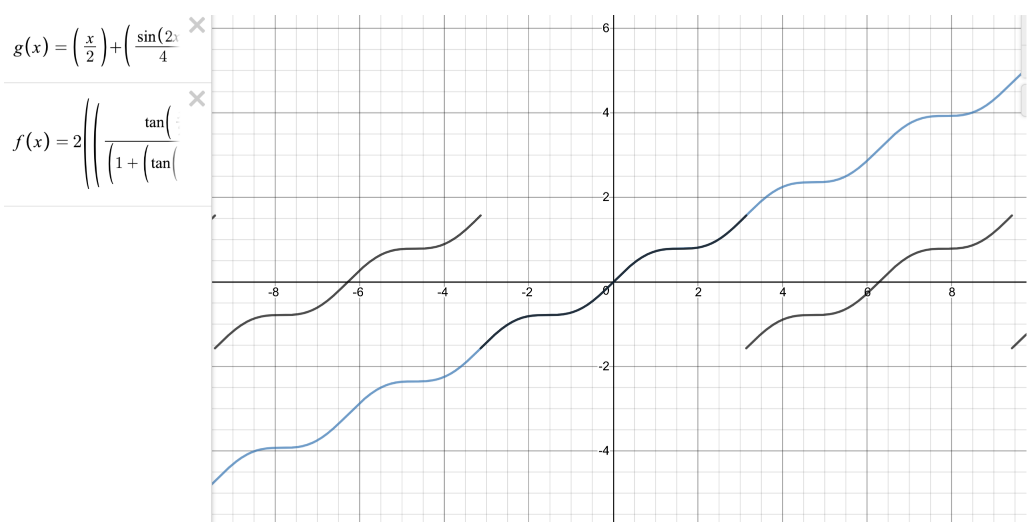
As we will shortly see, this kind of integral can be handled quite mechanically, but the result is quite hideous. `Mathemaica` gives the integral as

$$2\left(\frac{t}{(1+t^2)^2} - \frac{t}{2(1+t^2)} + \frac{\arctan(t)}{2}\right),$$

so that

$$\int \cos^2 x \, dx = 2\left(\frac{\tan(x/2)}{(1+(\tan(x/2))^2)^2} - \frac{\tan(x/2)}{2(1+(\tan(x/2))^2)} + \frac{\arctan(\tan(x/2))}{2}\right) := f(x).$$

This obviously equals $x/2 + (\sin 2x)/4 := g(x)$, right, at least up to an additive constant? Not exactly ... the Desmos screenshot below shows the two functions:



[227] Also, there is a slight issue with this method — see example 1, or (exercise) see if you can spot the issue before looking at example 1.

Notice that the domain of f is *not* all reals; it is all reals other than $\pm\pi, \pm 3\pi, \pm 5\pi, \ldots$. On the interval $(-\pi, \pi)$, f and g agree; on all other intervals they agree up to a constant. The issue here is that in making the substitution $t = \tan(x/2)$, one has to give up all values of x of the form $\pm\pi, \pm 3\pi, \pm 5\pi, \ldots$; $\tan(x/2)$ is not defined for these values. On all other values, the substitution works fine (modulo dealing with the hideous expressions that come out of it).

Example 2 $\int \frac{dx}{1+\sin x}$. Here, the only course of action seems to be to apply the "last resort" substitution, and in fact it works beautifully, leading to

$$\int \frac{dx}{1+\sin x} = 2\int \frac{1}{(1+t)^2}\,dt = \frac{-2}{1+t} = \frac{-2}{1+\tan(x/2)}.$$

13.5 Integration by partial fractions

In the last section we saw that many integrals can be reduced to integrals of rational functions, via appropriate substitutions. There is a method that, in principle at leasts, can find a primitive of any rational function. The method is, on the whole, fairly okay to understand at a theoretical level, but (unfortunately) rather difficult to implement practically except in some simple cases.

Setup

A *rational function* is a function f given by $f(x) = \frac{P(x)}{Q(x)}$ where P and Q are both polynomials, and Q is not zero.

Because our concern is with finding antiderivatives of rational functions, and this is easy when Q is a constant (in which case the rational function is just a polynomial), we will throughout assume that the degree of Q is at least 1. Recall that the *degree* $\deg(Q)$ of the polynomial $Q(x)$ is the highest power of x in the polynomial that has a non-zero coefficient.

Also, since it's easy to find an antiderivative of the zero function, we will assume that P is not zero.

By scaling Q by a constant, if necessary, we can assume that

$$Q(x) = x^n + q_1 x^{n-1} + \cdots + q_{n-1}x + q_n$$

where $n \geq 1$, and that

$$P(x) = p_0 x^m + p_1 x^{m-1} + \cdots + p_{m-1}x + p_m$$

where $m \geq 0$ and $p_0 \neq 0$.

Three key facts

To find an antiderivative of $P(x)/Q(x)$ we will need to use three facts from algebra, that we will not prove.

Fact 1 (the division algorithm for polynomials): If $\deg(P) \geq \deg(Q)$ then there are polynomials $A(x)$ (the *quotient*) and $B(x)$ (the *remainder*) with $\deg(B) < \deg(Q)$ such that $P(x) = A(x)Q(x) + B(x)$, or

$$\frac{P(x)}{Q(x)} = A(x) + \frac{B(x)}{Q(x)}.$$

A and B can be found by polynomial long division.

Running example: Consider the rational function

$$\frac{3x^6 - 7x^5 + 9x^4 - 9x^3 + 7x^2 - 4x - 1}{x^4 - 2x^3 + 2x^2 - 2x + 1}.$$

When we start long division, the first term will definitely be $3x^2$ (that's what's needed to get the leading x^4 in the denominator up to $3x^6$). Now

$$(x^4 - 2x^3 + 2x^2 - 2x + 1)(3x^2) = 3x^6 - 6x^5 + 6x^4 - 6x^3 + 3x^2,$$

and when this is subtracted from $3x^6 - 7x^5 + 9x^4 - 9x^3 + 7x^2 - 4x - 1$ we get

$$-x^5 + 3x^4 - 3x^3 + 4x^2 - 4x - 1.$$

So we continue the long division with $-x$ (that's what's needed to get the leading x^4 in the denominator up to $-x^5$). Now

$$(x^4 - 2x^3 + 2x^2 - 2x + 1)(-x) = -x^5 + 2x^4 - 2x^3 + 2x^2 - x,$$

and when this is subtracted from $-x^5 + 3x^4 - 3x^3 + 4x^2 - 4x - 1$ we get

$$x^4 - x^3 + 2x^2 - 3x - 1.$$

So the next term in the long division is 1; and when $x^4 - 2x^3 + 2x^2 - 2x + 1$ is subtracted from $x^4 - x^3 + 2x^2 - 3x - 1$ we get $x^3 - x - 2$. This has degree smaller than 4, so the remainder term has been reached, and the long division is finished:

$$\frac{3x^6 - 7x^5 + 9x^4 - 9x^3 + 7x^2 - 4x - 1}{x^4 - 2x^3 + 2x^2 - 2x + 1} = 3x^2 - x + 1 + \frac{x^3 - x - 2}{x^4 - 2x^3 + 2x^2 - 2x + 1}.$$

Examples to work out: Find A and B for the following rational functions.

1. $\frac{x^5}{x^4+x+1}$
2. $\frac{2x^7+3x^6-x^5+4x^4+5x^2-1}{x^4+x^2-x}$
3. $\frac{(1-x)^4}{(1+x)^4}$
4. $\frac{x^4+2x^3+3x^2+2x+1}{x^2+x+1}$

The point of Fact 1 is that since

$$\int \frac{P}{Q} = \int A + \int \frac{B}{Q},$$

and $\int A$ is easy to find (A being a polynomial), from here on in the method of partial fractions we need only concentrate on rational functions of the form $B(x)/Q(x)$, i.e., those where the degree of the numerator is less than the degree of the denominator.

Fact 2 (a corollary of the fundamental theorem of algebra): The polynomial deg(Q) can be factored into linear and quadratic terms:

$$Q(x) = (x - r_1)^{\alpha_1} \cdots (x - r_k)^{\alpha_k} (x^2 - 2s_1 x + t_1)^{\beta_1} \cdots (x^2 - 2s_\ell x + t_\ell)^{\beta_\ell}$$

where the r_i's, s_i's and t_i's are reals, the α_i's and β_i's are natural numbers, the r_i's are distinct from each other, the pairs (s_i, t_i) are distinct from each other (so there is no co-incidence between any pairs of factors), $s_i^2 < t_i$ for each i (so none of the quadratic terms can be factored further into linear terms), and $\deg(Q) = \sum_i \alpha_i + 2\sum_j \beta_j$.

Moreover, each quadratic term $x^2 - 2s_i x + t_i$ can be written in the form $(x - a_i)^2 + b_i^2$ with a_i and b_i real and b_i positive (this comes straight from $s_i^2 < t_i$: we have $x^2 - 2s_i x + t_i = (x - s_i)^2 + t_i - s_i^2 = (x - a_i)^2 + b_i^2$ where $a_i = s_i$ and $b_i = \sqrt{t_i - s_i^2}$).

Lurking behind Fact 2 is the *fundamental theorem of algebra*, which says that every polynomial with complex coefficients has a root in the complex numbers. Given a complex polynomial $C(z)$ with root c, using the division algorithm it is possible to write $C(z) = (z - c)\tilde{C}(z)$, where $\tilde{C}(z)$ is a complex polynomial whose degree is one less than that of C, and repeating this process we get that C factors fully into linear terms, as $C(z) = (z - c_1) \cdots (z - c_n)$ where $n = \deg(C)$. Here the c_i are complex numbers; but Q, having only real coefficients, possibly has some of these roots being real (these are the r_i above). It turns out that for a polynomial with all real coefficients, the complex roots appear in what are called *complex conjugate pairs*: pairs of the form $a + b\sqrt{-1}$ and $a - b\sqrt{-1}$. Noting that

$$(z - (a + b\sqrt{-1}))(z - (a - b\sqrt{-1})) = (z - a)^2 + b^2,$$

this "explains" the form of the quadratic factors above.

In general, it is *very difficult* to fully factor a real polynomial into linear and quadratic factors.

Running example: Consider $x^4 - 2x^3 + 2x^2 - 2x + 1$. After some trial-and error, we find that 1 must be a root, since $(1)^4 - 2(1)^3 + 2(1)^2 - 2(1) + 1 = 0$. So $x - 1$ is a factor, and long division gives

$$x^4 - 2x^3 + 2x^2 - 2x + 1 = (x - 1)(x^3 - x^2 + x - 1).$$

Again 1 is a root of $x^3 - x^2 + x - 1$, and

$$x^4 - 2x^3 + 2x^2 - 2x + 1 = (x - 1)(x - 1)(x^2 + 1).$$

The quadratic formula tells us that we cannot factor any further.

Examples to work out: Fully factor the polynomials below into linear factors of the form $x - r$ and quadratic factors of the form $(x-a)^2 + b^2$. Start by trying a few small values of r (positive and negative) to find one with the polynomial evaluating to 0 at r; then divide by $x - r$ and repeat.

1. Factorize $x^4 - x^3 - 7x^2 + x + 6$
2. Factorize $x^4 - x^3 - 7x^2 + x + 6$
3. Factorize $x^3 - 3x^2 + 3x - 1$
4. Factorize $x^6 + 3x^4 + 3x^2 + 1$
5. Factorize $x^4 + 1$ (tricky)

The point of Fact 2 is that it feeds nicely into Fact 3.

Fact 3 (partial fractions decomposition): Let Q and B be polynomials as described above (Q has degree at least 1, and leading coefficient 1, and B has degree less than that of Q). Let Q be factored into linear and quadratic terms, exactly as outlined in Fact 2:

$$Q(x) = (x-r_1)^{\alpha_1} \cdots (x-r_k)^{\alpha_k}((x-a_1)^2 + b_1^2)^{\beta_1} \cdots ((x-a_\ell)^2 + b_\ell^2)^{\beta_\ell}.$$

Then there are real constants

$$\begin{array}{c} A_{11}, \ldots, A_{1\alpha_1}, \\ A_{21}, \ldots, A_{2\alpha_2}, \\ \ldots, \\ A_{k1}, \ldots, A_{k\alpha_k}, \\ \\ B_{11}, \ldots, B_{1\beta_1}, \\ B_{21}, \ldots, B_{2\beta_2}, \\ \ldots, \\ B_{k1}, \ldots, B_{k\beta_k}, \\ \\ C_{11}, \ldots, C_{1\beta_1}, \\ C_{21}, \ldots, C_{2\beta_2}, \\ \ldots, \\ C_{k1}, \ldots, C_{k\beta_k}, \end{array}$$

such that

$$\frac{B(x)}{Q(x)} = \begin{array}{c} \frac{A_{11}}{(x-r_1)} + \frac{A_{12}}{(x-r_1)^2} + \cdots + \frac{A_{1\alpha_1}}{(x-r_1)^{\alpha_1}} + \\ \frac{A_{21}}{(x-r_2)} + \frac{A_{22}}{(x-r_2)^2} + \cdots + \frac{A_{2\alpha_2}}{(x-r_2)^{\alpha_2}} + \\ \cdots + \\ \frac{A_{k1}}{(x-r_k)} + \frac{A_{k2}}{(x-r_k)^2} + \cdots + \frac{A_{k\alpha_k}}{(x-r_k)^{\alpha_k}} + \\ \frac{B_{11}x+C_{11}}{((x-a_1)^2+b_1^2)} + \frac{B_{12}x+C_{12}}{((x-a_1)^2+b_1^2)^2} + \cdots + \frac{B_{1\beta_1}x+C_{1\beta_1}}{((x-a_1)^2+b_1^2)^{\beta_1}} + \\ \frac{B_{21}x+C_{21}}{((x-a_2)^2+b_2^2)} + \frac{B_{22}x+C_{22}}{((x-a_2)^2+b_2^2)^2} + \cdots + \frac{B_{2\beta_2}x+C_{2\beta_2}}{((x-a_2)^2+b_2^2)^{\beta_2}} + \\ \cdots + \\ \frac{B_{\ell 1}x+C_{\ell 1}}{((x-a_\ell)^2+b_\ell^2)} + \frac{B_{\ell 2}x+C_{\ell 2}}{((x-a_\ell)^2+b_\ell^2)^2} + \cdots + \frac{B_{\ell\beta_\ell}x+C_{\ell\beta_\ell}}{((x-a_\ell)^2+b_\ell^2)^{\beta_\ell}}. \end{array}$$

The proof of Fact 3 is not very difficult, but it requires too much familiarity with linear algebra to describe here.

It is somewhat straightforward to locate the values of the constants asserted in Fact 3. Start with the equation given in Fact 3 (with all the constants unknown). Multiply both sides by $Q(x)$. The right-hand side becomes a polynomial of degree $\deg(Q) - 1$, so with $\deg(Q)$ coefficients, expressed in terms of a number of unknowns — $\deg(Q)$ unknowns, to be precise. The left-hand side becomes a polynomial with known coefficients with degree at most $\deg(Q) - 1$. Equating the constant terms on both sides, the linear terms, the quadratic terms, et cetera, one gets a collection of $\deg(Q)$ equations in $\deg(Q)$ unknowns. Using techniques from linear algebra, such a system can be solved relatively quickly to find the (unique, as it turns out) values for the constants (the A's, B's and C's).

Even without knowing linear algebra, it is fairly straightforward to perform this task, if the degrees of the polynomials involved are all reasonably small.

Running example: We seek to find the partial fractions decomposition of $\frac{x^3-x-2}{(x-1)^2(x^2+1)}$. We start with

$$\frac{x^3 - x - 2}{(x-1)^2(x^2+1)} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{Cx+D}{x^2+1}.$$

Multiplying through by $(x-1)^2(x^2+1)$ yields

$$\begin{aligned} x^3 - x - 2 &= A(x-1)(x^2+1) + B(x^2+1) + (Cx+D)(x-1)^2 \\ &= (A+C)x^3 + (-A+B-2C+D)x^2 + (A+C-2D)x + (-A+B+D). \end{aligned}$$

Equating coefficients gives

$$A + C = 1, \quad -A + B - 2C + D = 0, \quad A + C - 2D = -1, \quad -A + B + D = -2.$$

One can solve this system of four equations in four unknowns by, for example, using the first equation to write $A = C - 1$, then substituting this into the remaining three to get three equations in three unknowns, then substitute again to get two equations in two unknowns, then again to get one equation in one unknown, which is easy to solve. Plugging in that one known value, the whole system now becomes one of three equations in three unknowns; rinse and repeat. (There are systematic ways to do this process, which are very efficient, and are explored in linear algebra).

Solving this system of equations in this way gives $A = 2$, $B = -1$, $C = -1$ and $D = 1$, so that

$$\frac{x^3 - x - 2}{(x-1)^2(x^2+1)} = \frac{2}{x-1} - \frac{1}{(x-1)^2} - \frac{(x-1)}{x^2+1}.$$

Examples to work out: Find the partial fractions decompositions of the following expressions.

1. $\frac{2x^2+7x-1}{x^3+x^2-x-1}$
2. $\int \frac{2x+1}{x^3-3x^2+3x-1}$
3. $\int \frac{3x}{(x^2+x+1)^3}$
4. $\frac{1}{x^4+1}$

Finding antiderivatives of rational functions

Using the three facts above, we can reduce the task of finding an antiderivative of a rational function to that of finding antiderivatives of functions of the following types:

- polynomials — these are easy
- functions of the form $\frac{A}{(x-r)^\alpha}$ where A and r are constants, and α is a natural number. These are straightforward:

$$\int \frac{A}{(x-r)^\alpha} dx = \begin{cases} \frac{A}{(1-\alpha)(x-r)^{\alpha-1}} & \text{if } \alpha \neq 1 \\ A \log(x-r) & \text{if } \alpha = 1. \end{cases}$$

- functions of the form $\frac{Cx+D}{((x-a)^2+b^2)^\beta}$ where C, D, a and b are real constants, with b positive, and β is a natural number. To deal with these, we first write

$$\frac{Cx+D}{((x-a)^2+b^2)^\beta} = \frac{(C/2)2(x-a)}{((x-a)^2+b^2)^\beta} + \frac{Ca+D}{((x-a)^2+b^2)^\beta}.$$

Using the substitution $u = (x-a)^2 + b^2$ we get

$$\int \frac{(C/2)2(x-a)}{((x-a)^2+b^2)^\beta} dx = \frac{C}{2} \int \frac{du}{u^\beta} = \begin{cases} \frac{C/2}{(1-\beta)u^{\beta-1}} = \frac{C/2}{(1-\beta)((x-a)^2+b^2)^{\beta-1}} & \text{if } \beta \neq 1 \\ (C/2)\log u = (C/2)\log((x-a)^2+b^2) & \text{if } \beta = 1. \end{cases}$$

To deal with the $(Ca+D)/(((x-a)^2+b^2)^\beta)$ terms, we have

$$\frac{Ca+D}{((x-a)^2+b^2)^\beta} = \frac{Ca+D}{b^{2\beta}} \left(\frac{1}{\left(\frac{x-a}{b}\right)^2 + 1} \right)$$

so using the substitution $u = (x-a)/b$ we get

$$\int \frac{Ca+D}{((x-a)^2+b^2)^\beta}dx = \frac{Ca+D}{b^{2\beta-1}} \int \frac{du}{(u^2+1)^\beta}$$

In class we studied the integral $\int \frac{du}{(u^2+1)^\beta}$. If we set

$$A_\beta = \int \frac{du}{(u^2+1)^\beta}$$

then we saw that we have $A_0 = 1$, $A_1 = \arctan u = \arctan((x-a)/b)$, and for $\beta \geq 1$,

$$A_{\beta+1} = \frac{u}{2\beta(1+u^2)^\beta} + \frac{(2\beta-1)}{2\beta}A_\beta = \frac{(x-a)/b}{2\beta(1+((x-a)/b)^2)^\beta} + \frac{(2\beta-1)}{2\beta}A_\beta.$$

So we can recursively figure out an antiderivative.

Running example: We seek an antiderivative of

$$\frac{3x^6-7x^5+9x^4-9x^3+7x^2-4x-1}{x^4-2x^3+2x^2-2x+1}.$$

As we have seen, this function can be expressed as

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

Only the last of these terms requires effort. We have

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

and so the desired antiderivative of our rational function is

$$x^3 - \frac{x^2}{2} + x + 2\log(x-1) + \frac{1}{x-1} - \frac{1}{2}\log(x^2+1) + \arctan x.$$

Examples to work out

1. Find $\int \frac{2x^2+7x-1}{x^3+x^2-x-1}dx$
2. Find $\int \frac{2x+1}{x^3-3x^2+3x-1}dx$
3. Find $\int \frac{3x}{(x^2+x+1)^3}dx$
4. Find $\int \frac{dx}{x^4+1}$
5. Use the $t = \tan(x/2)$ substitution, and the method of partial fractions, to find antiderivatives for each of the following trigonometric functions:

- $\sec x$ (the answer you get is unlikely to be $\log(\sec x + \tan x)$, which is the expression that you are most likely to see if you look up a table of antiderivatives. Check that $\log(\sec x + \tan x)$ differentiates to $\sec x$, and also that the expression that you get is equal to $\log(\sec x + \tan x)$, perhaps up to an additive constant).
- $\sec^3 x$.

6. Using the "magic" substitution $t = \tan(x/2)$, and partial fractions, we see that every rational function of sin and cos has an elementary antiderivative. Show that also every rational function of e^x has an elementary antiderivative.