Integration by substitution

Let $f(y)$ be integrable on an interval J and let $F(y)$ be an antiderivative. Suppose that $\varphi(x) : I \to J$ is differentiable. Then $f(\varphi(x))\varphi'(x)$ is integrable on I and

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = F(\varphi(x)) + C.$$

A simpler way to remember this formula is by writing $y = \varphi(x)$ to get:

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int f(y)\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x = \int f(y)\,\mathrm{d}y.$$

Proof. From the chain rule (Theorem 8.3) we know that

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(F(\varphi(x))\big) = F'(\varphi(x))\varphi'(x) = f(\varphi(x))\varphi'(x).$$

By definition of the antiderivative the result follows. □

Example 10.5:

1. Determine $\int \sin(k(x - x_0))\,\mathrm{d}x$, where $0 \neq k \in \mathbb{R}$ and $x_0 \in \mathbb{R}$. We let

$$y = \varphi(x) = k(x - x_0) \qquad \text{so that} \qquad \varphi'(x) = k.$$

Then we have

$$\begin{aligned}
\int \sin(k(x - x_0))\,\mathrm{d}x = \int \sin(\varphi(x))\,\mathrm{d}x &= \int \sin(\varphi(x))\frac{k}{k}\,\mathrm{d}x \\
&= \frac{1}{k}\int \sin(\varphi(x))\varphi'(x)\,\mathrm{d}x \\
&= \frac{1}{k}\int \sin y\,\mathrm{d}y \\
&= -\frac{1}{k}\cos y + C \\
&= -\frac{1}{k}\cos(k(x - x_0)) + C.
\end{aligned}$$

2. Determine $\int xe^{x^2}\,\mathrm{d}x$. We let

$$y = \varphi(x) = x^2 \qquad \text{so that} \qquad \varphi'(x) = 2x.$$

Then we have

$$\begin{aligned}
\int xe^{x^2}\,\mathrm{d}x = \frac{1}{2}\int 2xe^{x^2}\,\mathrm{d}x &= \frac{1}{2}\int \varphi'(x)e^{\varphi(x)}\,\mathrm{d}x \\
&= \frac{1}{2}\int e^{y}\,\mathrm{d}y \\
&= \frac{1}{2}e^{y} + C \\
&= \frac{1}{2}e^{x^2} + C.
\end{aligned}$$

3. Determine $\int \tan x \, \mathrm{d}x$. We let

$$y = \varphi(x) = \cos x \qquad \text{so that} \qquad \varphi'(x) = -\sin x.$$

Then we have

$$\begin{aligned}\int \tan x \, \mathrm{d}x = \int \frac{\sin x}{\cos x} \, \mathrm{d}x = \int \frac{-\varphi'(x)}{\varphi(x)} \, \mathrm{d}x \\ &= -\int \frac{1}{y} \, \mathrm{d}y \\ &= -\ln|y| + C \\ &= -\ln|\cos x| + C.\end{aligned}$$

4. Determine $\int \frac{\varphi'(x)}{\varphi(x)} \, \mathrm{d}x$. This generalizes the previous example. Let

$$y = \varphi(x) \qquad \text{so that} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \varphi'(x).$$

Then we have

$$\int \frac{\varphi'(x)}{\varphi(x)} \, \mathrm{d}x = \int \frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x} \, \mathrm{d}x = \int \frac{1}{y} \, \mathrm{d}y = \ln|y| + C = \ln|\varphi(x)| + C.$$

5. Determine $\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x$. Here's a trick. Let

$$y = \varphi(x) = \sqrt{1+x^2} - x.$$

It follows that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \varphi'(x) = \frac{x}{\sqrt{1+x^2}} - 1 = \frac{x - \sqrt{1+x^2}}{\sqrt{1+x^2}} = \frac{-\varphi(x)}{\sqrt{1+x^2}} = \frac{-y}{\sqrt{1+x^2}}$$

which means that

$$\frac{1}{\sqrt{1+x^2}} = -\frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x}.$$

Integrating, we get

$$\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x = -\int \frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x} \, \mathrm{d}x = -\int \frac{1}{y} \, \mathrm{d}y = -\ln|y| + C = -\ln|\sqrt{1+x^2} - x| + C.$$

Since $\sqrt{1+x^2} - x > 0$ we don't need the absolute value, and we have

$$\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x = -\ln(\sqrt{1+x^2} - x) + C.$$

Remembering how to integrate by substitution

When asked to integrate $\int g(x)\,\mathrm{d}x$ there are two options:

(a) We are able identify that there exists $y = y(x)$ such that $g(x)$ has the form

$$g(x) = f(y(x))y'(x)$$

as we had done in the examples above. In this case, since $\frac{\mathrm{d}y}{\mathrm{d}x} = y'(x)$ we write $\mathrm{d}y = y'(x)\,\mathrm{d}x$ to get

$$\int g(x)\,\mathrm{d}x = \int f(y(x))y'(x)\,\mathrm{d}x = \int f(y)\,\mathrm{d}y.$$

This approach might work if g is a complicated function.

(b) If we cannot identify $y = y(x)$ as above, we try to go about it the other way around: identify x as a function of y: $x = x(y)$, compute $\frac{\mathrm{d}x}{\mathrm{d}y} = x'(y)$ and write $\mathrm{d}x = x'(y)\,\mathrm{d}y$ to get:

$$\int g(x)\,\mathrm{d}x = \int g(x(y))x'(y)\,\mathrm{d}y.$$

This approach might work if g is a simple function.

6. Determine $\int \sqrt{1-x^2}\,\mathrm{d}x$. Let's use the second approach from above. We start with
$$g(x) = \sqrt{1-x^2}.$$
Now we choose $x(y) = \sin y$ because it will simplify g, since $1 - \sin^2 y = \cos^2 y$. So we have
$$x(y) = \sin y \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}y} = \cos y \qquad \Rightarrow \qquad \mathrm{d}x = \cos y\,\mathrm{d}y$$
and we get
$$\begin{aligned} \int \sqrt{1-x^2}\,\mathrm{d}x &= \int \sqrt{1-\sin^2 y}\cos y\,\mathrm{d}y \\ &= \int \cos^2 y\,\mathrm{d}y \\ &= \frac{y}{2} + \frac{\sin(2y)}{4} + C \\ &= \frac{y}{2} + \frac{1}{2}\sin y\cos y + C \\ &= \frac{y}{2} + \frac{1}{2}\sin y\sqrt{1-\sin^2 y} + C \\ \textit{(using that } y = \arcsin x\textit{)} \qquad &= \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + C. \end{aligned}$$

7. Determine $\int \frac{1}{e^x+e^{-x}}\,\mathrm{d}x$. Using the second approach, we have

$$g(x) = \frac{1}{e^x + e^{-x}}.$$

Choose $x(y) = \ln y$ to undo the exponents. Then we have

$$x(y) = \ln y \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{y} \qquad \Rightarrow \qquad \mathrm{d}x = \frac{1}{y}\,\mathrm{d}y$$

and we get

$$\begin{aligned}
\int \frac{1}{e^x + e^{-x}}\,\mathrm{d}x &= \int \frac{1}{e^{\ln y} + e^{-\ln y}} \frac{1}{y}\,\mathrm{d}y \\
&= \int \frac{1}{y + \frac{1}{y}} \frac{1}{y}\,\mathrm{d}y \\
&= \int \frac{1}{1 + y^2}\,\mathrm{d}y \\
&= \arctan y + C = \arctan e^x + C.
\end{aligned}$$

Example 10.6 (Rational functions): Let us give some more examples of integration by substitution, this time with some rational functions.

1. The simplest case is easy:

$$\int \frac{1}{x + c}\,\mathrm{d}x = \ln|x + c| + C.$$

2. The next case follows from what we know for power functions ($r > 1$):

$$\int \frac{1}{(x + c)^r}\,\mathrm{d}x = \frac{1}{1 - r} \frac{1}{(x + c)^{r-1}} + C.$$

3. Next we consider $\int g(x)\,\mathrm{d}x$ where

$$g(x) = \frac{1}{x^2 + bx + c} \qquad \text{with discriminant } \Delta = b^2 - 4c < 0$$

so that the polynomial in the denominator is strictly positive (no real roots). We can write

$$\begin{aligned}
x^2 + bx + c &= x^2 + bx + \underbrace{c - \frac{1}{4}b^2}_{-\frac{1}{4}\Delta} + \frac{1}{4}b^2 = \underbrace{x^2 + bx + \frac{1}{4}b^2}_{(x+\frac{1}{2}b)^2} - \frac{1}{4}\Delta = (x + \frac{1}{2}b)^2 - \frac{1}{4}\Delta \\
&= -\frac{1}{4}\Delta \left[1 + \left(\frac{x + \frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}\right)^2\right]
\end{aligned}$$

Make the substitution

$$y(x) = \frac{x + \frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}$$