Integration by Parts

The Leibniz rule for differentiation says that if \( f(x) = g(x)h(x) \), then

\[
f'(x) = g'(x)h(x) + g(x)h'(x).
\]

By the fundamental theorem of calculus

\[
\int g'(x)h(x) + g(x)h'(x) \, dx = \int f'(x) \, dx = f(x)
\]

(ignoring constants of integration). The indefinite integral (i.e., the antiderivative) of a sum of two functions is the sum of the indefinite integrals, so

\[
\int g'(x)h(x) + g(x)h'(x) \, dx = \int g'(x)h(x) \, dx + \int g(x)h'(x) \, dx.
\]

Putting it all together,

\[
\int g'(x)h(x) \, dx + \int g(x)h'(x) \, dx = f(x) = g(x)h(x).
\]

Subtracting \( \int g(x)h'(x) \, dx \) from both sides of the equation, we get the formula for integration by parts:

\[
\int g'(x)h(x) \, dx = g(x)h(x) - \int g(x)h'(x) \, dx. \tag{\star}
\]

The reason (\( \star \)) is useful is that it gives a way to integrate a product of two functions, in this case the functions \( g'(x) \) and \( h(x) \). Given a product which is to be integrated by this method, it is important to choose which factor should be called \( g'(x) \) and which one \( h(x) \). One consideration to bear in mind is that the right hand side is expressed in terms of \( g(x) \), not \( g'(x) \), which means we had better choose a \( g'(x) \) whose indefinite integral we know. Another is that it is desirable if the derivative of \( h(x) \) is in some sense a simpler function than \( h(x) \), or at least no more complicated. For instance, \( h(x) = \ln x \) makes a good choice since \( h'(x) = \frac{1}{x} \) is a more basic function than the natural log. So does \( h(x) = x \), whose derivative \( h'(x) = 1 \) is a constant (and constants are nice because we can pull them through integral signs). A function like \( h(x) = e^x \) isn’t too bad; its derivative \( e^x \) is at least no worse than the original function. The point of all this is that at the end of the day we will need to integrate \( g(x)h'(x) \), and the trick of integration by parts is only worth doing if this integral is easier to work out than the original one.

Example 1. Compute the indefinite integral \( \int e^x \, x \, dx \).

There are two factors \( e^x \) and \( x \). We can easily compute the indefinite integral of either one. Both have fairly simple derivatives, but \( x \) is especially nice because its derivative is a constant. So set \( g'(x) = e^x \) and \( h(x) = x \). Then \( h(x) = 1 \), and \( g(x) \) could be any function of the form \( e^x + C \). We might as well choose the simplest one, \( g(x) = e^x \). Now (\( \star \)) says

\[
\int e^x \, x \, dx = e^x x - \int e^x \, dx = e^x x - e^x + C.
\]
What would have happened if we had made the opposite choice? Well, \( g'(x) = x, h(x) = e^x \) implies \( h'(x) = e^x \), and \( g(x) = \frac{x^2}{2} + C \). Making the simplest choice, \( g(x) = \frac{x^2}{2} \), which means (\(*)\) says
\[
\int e^x \, dx = xe^x - \int \frac{x^2}{2} e^x \, dx.
\]
The integral on the right hand side is actually harder to do than the original integral, so we’ve lost ground by making injudicious choices.

**Example 2.** Compute the indefinite integral \( \int x^2 \sin x \, dx \).

There are several possible factorizations since \( x^2 \sin x = x \cdot x \cdot \sin x \) is a product of three terms. One reasonable choice is to take \( g'(x) = \sin x \) and \( h(x) = x^2 \). Then \( h'(x) = 2x \), and we can take \( g(x) = -\cos x \). Applying (\*)
\[
\int x^2 \sin x \, dx = -x^2 \cos x - \int (-\cos x)(2x) \, dx = -x^2 \cos x + 2 \int x \cos x \, dx.
\]
There is still an unevaluated integral, but \( x \cos x \) looks like it might be a little easier to integrate than \( x^2 \sin x \). Proceeding on that assumption, let’s try to evaluate \( \int x \cos x \, dx \) by setting \( g'(x) = \cos x \) and \( h(x) = x \). This means setting \( h'(x) = 1 \) (very promising) and \( g(x) = \sin x \). Plugging into (\*) again,
\[
\int x \cos x \, dx = x \sin x - \int \sin x \, dx = x \sin x + \cos x + C,
\]
and going back to the original problem,
\[
\int x^2 \sin x \, dx = -x^2 \cos x + 2 \int x \cos x \, dx = -x^2 \cos x + 2x \sin x + 2 \cos x + C.
\]
This is complicated enough that it’s worth taking a derivative and seeing if we get the original function. Indeed, the derivative of \(-x^2 \cos x + 2x \sin x + 2 \cos x + C\) is
\[
(-2x \cos x + x^2 \sin x) + (2 \sin x + 2x \cos x) - 2 \sin x = x^2 \sin x.
\]
Moral: sometimes applying integration by parts once isn’t enough.

**Example 3.** Compute the indefinite integral \( \int \ln x \, dx \).

This one doesn’t look like an integration by parts problem at all (at first glance). But try writing \( \ln x = g'(x)h(x) \), where \( g'(x) = 1 \) and \( h(x) = \ln x \). This means \( g(x) = x \) and \( h'(x) = 1/x \). Applying (\*),
\[
\int \ln x \, dx = x \ln x - \int \frac{1}{x} \, dx = x \ln x - \int dx = x \ln x - x + C.
\]
Problems

1. Evaluate $\int x \sin x \, dx$.
2. Evaluate $\int x \ln x \, dx$.
3. Evaluate $\int x \cos 2x \, dx$.
4. Evaluate $\int (x^2 + x + 1)e^x \, dx$.
5. Evaluate $\int xe^{-x}$.