5.2 Integration by Substitution

In the preceding section, we reimagined a couple of general rules for differentiation – the constant multiple rule and the sum rule – in integral form. In this section we will develop the integral form of the chain rule, and see some of the ways this can be used to find antiderivatives.

Indefinite integrals

We begin with the following.

Example 5.2.1. Evaluate the indefinite integral

$$\int 2x\cos(x^2)\,dx. \tag{5.2.1}$$

Solution. The key fact here is that the integrand contains a quantity – namely, x^2 – whose *derivative* – namely, $2x$ – also appears as a factor in the integrand. Let's see how we can make use of this fact.

The strategy in cases like this is to give a name to the quantity whose derivative is present. We call that quantity u; that is,

$$u = x^2. \tag{5.2.2}$$

Then, of course, we have $du/dx = x^2$ or, "multiplying both sides by dx,"

$$du = 2x\,dx. \tag{5.2.3}$$

Of course we can't *really* "multiply both sides by dx," because a derivative is not actually a fraction – du and dx are not separate quantities, but rather, they are just pieces of the symbol du/dx. But let's proceed as if this strategy were justifiable – it *is* in fact justifiable, and we'll explain why at the end of this subsection.

The key thing to observe, to complete our problem, is this: if we substitute (5.2.2) and (5.2.3) into our integral (5.2.1), then that integral becomes a *simpler* one, involving the variable u (and only the variable u). And this simpler integral is straightforward to evaluate. Specifically, we have

$$\int 2x\cos(x^2)\,dx = \int \underbrace{\cos(x^2)}_{\substack{\text{this becomes}\\ \cos(u)}}\ \underbrace{2x\,dx}_{\substack{\text{this is}\\ \text{just } du}} = \int \cos(u)\,du = \sin(u) + C. \tag{5.2.4}$$

The first step here was not really necessary; we just reordered the factors in the integrand to make it more clear that, upon making the substitutions (5.2.2) and (5.2.3), this integrand (including the dx) becomes $\cos(u)\,du$.

We next observe that, in (5.2.4), the "question" was in terms of the variable x, while the answer was in terms of u. To be consistent, we should answer in terms of the question's original variable. This is easy: recall that $u = x^2$, so (5.2.4) reads

$$\int 2x\cos(x^2)\,dx = \sin(x^2) + C. \tag{5.2.5}$$

Strictly speaking, we're done, though it's a good idea to check our work, which we do as follows:

$$\frac{d}{dx}\left[\sin(x^2) + C\right] = \cos(x^2)\frac{d}{dx}\left[x^2\right] + 0 = 2x\cos(x^2), \tag{5.2.6}$$

so our solution (5.2.5) is correct. Note how the chain rule arises in our check. It's *because* of the chain rule that integration by substitution works. (See the discussion at the end of this subsection.)

The strategy that we employed above, and which works in many similar situations, is as follows:

Step 1. If possible, identify a quantity $g(x)$ in the integrand such that the derivative $g'(x)$ also appears as a factor in that integrand.

Step 2. Call the original quantity u. So $u = g(x)$, $du/dx = g'(x)$, and $du = g'(x)\,dx$.

Step 3. In your original integral, replace $g(x)$ by u and $g'(x)\,dx$ by du, to arrive, if possible, at an integral involving only the variable u.

Step 4. If possible, evaluate the integral in u.

Step 5. Into your final answer, replace u by $g(x)$, to put your solution in terms of the original variable.

Strategy for evaluating indefinite integrals by substitution

Here are some further examples, some of which entail slight variations on the above strategy.

Example 5.2.2. (a) Evaluate the following indefinite integrals.

(i) $\displaystyle\int (2+5x^4)(2x+x^5)^{26}\,dx$ (ii) $\displaystyle\int e^{\sin(x)}\cos(x)\,dx$ (iii) $\displaystyle\int \frac{e^x}{1+(e^x+4)^2}\,dx$

(iv) $\displaystyle\int \frac{\sin(\ln(z))}{z}\,dz$ (v) $\displaystyle\int x\sin(x^2+1)\,dx$ (vi) $\displaystyle\int f(g(x))g'(x)\,dx$

(b) Solve the initial value problem

$$\frac{dy}{dt} = \frac{(2\ln(t)+3)^2}{t}, \qquad y(1) = 0.$$

Solution. (a) (i) We look for a part of the integrand whose derivative also appears as a factor there. We see that $u = 2x + x^5$ works: it gives us $du/dx = 2 + 5x^4$. This in turn gives $du = (2 + 5x^4)\,dx$, so

$$\int (2+5x^4)(2x+x^5)^{26}\,dx = \int u^{26}\,du = \frac{u^{27}}{27} + C = \frac{(2x+x^5)^{27}}{27} + C.$$

As a check on our work, we note that

$$\frac{d}{dx}\left[\frac{(2x+x^5)^{27}}{27}+C\right] = \frac{27(2x+x^5)^{26}}{27}\frac{d}{dx}[2x+x^5]+0 = (2+5x^4)(2x+x^5)^{26},$$

as required.

For subsequent parts of this example, we'll omit the "check," though it's a good idea to include it unless and until you are completely comfortable with the substitution technique.

(ii) Here we keep track of substitutions "in the margin," a process we will follow hereafter.

$$\begin{aligned} &\int e^{\sin(x)}\cos(x)\,dx = \int e^u\,du \\ &= e^u + C \\ &= e^{\sin(x)} + C. \end{aligned} \qquad \left|\; \begin{aligned} &u = \sin(x) \\ &\frac{du}{dx} = \cos(x) \\ &du = \cos(x)\,dx \end{aligned}\right.$$

(iii) As we illustrate here, one can go straight from "$u = \ldots$" to "$du = \ldots$," skipping the "$du/dx = \ldots$" step in between.

$$\begin{aligned} &\int \frac{e^x}{1+(e^x+4)^2}\,dx = \int \frac{1}{1+u^2}\,du \\ &= \arctan(u) + C \\ &= \arctan(e^x+4) + C. \end{aligned} \qquad \left|\; \begin{aligned} &u = e^x + 4 \\ &du = e^x\,dx \end{aligned}\right.$$

One other thing to note from the above example is the following. We could have put $u = e^x$ instead of $u = e^x + 4$; we would have wound up with the same du. Generally speaking, though, we want to "break off" as much as possible into the quantity we call u, to obtain a u-integral that is as simple as possible.

(iv)

$$\begin{aligned} &\int \frac{\sin(\ln(z))}{z}\,dz = \int \sin(u)\,du \\ &= -\cos(u) + C \\ &= -\cos(\ln(z)) + C. \end{aligned} \qquad \left|\; \begin{aligned} &u = \ln(z) \\ &du = \frac{1}{z}\,dz \end{aligned}\right.$$

(v) If your u-substitution gives you a du that's "off by a constant factor" from where you want it to be, you can always divide through by this factor, and proceed as usual. The following example illustrates how this works.

$$\begin{aligned} &\int x\sin(x^2+1)\,dx = \int \sin(u)\left(\frac{du}{2}\right) \\ &= \frac{1}{2}\int \sin(u)\,du = -\frac{1}{2}\cos(u) + C \\ &= -\frac{1}{2}\cos(x^2+1) + C. \end{aligned} \qquad \left|\; \begin{aligned} &u = x^2 + 1 \\ &du = 2x\,dx \\ &\frac{du}{2} = x\,dx \end{aligned}\right.$$

The point of the last line in the margin is that our substitution gives us a $2x\,dx$, while our original integral only involves an $x\,dx$. So we divide the equation $du = 2x\,dx$ by two, to get something (namely, $x\,dx$) that we can substitute directly into the integral. We then use the constant multiple rule for definite integrals to pull the factor of $1/2$ out front.

(vi)

$$\begin{aligned} &\int f(g(x))g'(x)\,dx \\ &= \int f(u)\,du. \end{aligned} \qquad \left| \begin{aligned} &u = g(x) \\ &du = g'(x)\,dx \end{aligned} \right.$$

The above example encapsulates the general idea behind integration by substitution. Of course, one must still perform the integration in u (if possible).

(b) We first antidifferentiate:

$$\begin{aligned} y &= \int \frac{(2\ln(t)+3)^2}{t}\,dt = \int u^2\left(\frac{du}{2}\right) \\ &= \frac{1}{2}\int u^2\,du = \frac{1}{2}\cdot\frac{1}{3}u^3 + C \\ &= \frac{1}{6}(2\ln(t)+3)^3 + C. \end{aligned} \qquad \left| \begin{aligned} &u = 2\ln(t)+3 \\ &du = \frac{2}{t}\,dt \\ &\frac{du}{2} = \frac{1}{t}\,dt \end{aligned} \right.$$

Into this solution, we now substitute $y(1) = 0$; we get

$$0 = y(1) = \frac{1}{6}(2\ln(1)+3)^3 + C = \frac{3^3}{6} + C = \frac{9}{2} + C.$$

Solving for C gives $C = -9/2$, so our solution is

$$y = \frac{1}{6}(2\ln(t)+3)^3 - \frac{9}{2}.$$

It should be noted that substitution doesn't always work. For example, consider the indefinite integral

$$\int e^{-t^2/2}\,dt.$$

We mentioned earlier (see Example 4.5.2(v)) that $f(t) = e^{-t^2/2}$ does not have a closed-form antiderivative. Not knowing this, or not believing it, we might try to integrate by substitution, putting $u = -t^2/2$, for example. The problem with this approach is that it gives us $du = -t\,dt$. And what do we do with the factor of t here? We could try dividing through by $-t$; that is, we could try writing $du/(-t) = dt$. We'd get

$$\int e^{-t^2/2}\,dt = \int e^u\left(\frac{du}{-t}\right),$$

but then we're stuck with the t in the denominator. (We can't pull the t outside of the integral; the constant multiple rule applies only to *constant* multiples, but $u = -t^2$, so t is not a constant with respect to u.)

Of course, the substitution approach *does* work in many situations – situations where some function *and* its derivative appear in the integrand. And we are still left with the question: when this approach works, *why* does it work? We've not yet answered this question because, in our arguments and computations above, we have always gone from a statement of the form $u = g(x)$ to one of the form $du = g'(x)\,dx$. And we have only justified this informally: we have argued that $u = g(x)$ implies $du/dx = g'(x)$, which in turn implies $du = g'(x)\,dx$. But the second part of this argument is not rigorous, because an integral is not a fraction – and we have not even *defined* du or $g'(x)\,dx$.

So why *does* substitution work? Philosophically, it's because a derivative is a lot *like* a fraction. Indeed, since $du/dx = \lim_{\Delta x \to 0} \Delta u/\Delta x$, a derivative is a *limit* of fractions. Because of this, it behaves much like a fraction in many ways.

Mathematically, to say that substitution "works" is to say that the "new" integral that results from a substitution really is the *same* as the original integral into which this substitution was made. In other words, we are claiming that, if $u = g(x)$, then

$$\int f(g(x))g'(x)\,dx = \int f(u)\,du$$

(cf. Example 5.2.2(vii) above). To show that this is in fact the case, let F be an antiderivative of f. Then we know that

$$\int f(u)\,du = F(u) + C = F(g(x)) + C. \tag{5.2.7}$$

If we can show that

$$\int f(g(x))g'(x)\,dx = F(g(x)) + C \tag{5.2.8}$$

as well, then the left-hand sides of (5.2.7) and (5.2.8) will be equal, and we'll be done.

To demonstrate (5.2.8), we need only show that the derivative of the right-hand side equals the integrand on the left. We do so using the chain rule:

$$\frac{d}{dx}[F(g(x)) + C] = F'(g(x))g'(x) = f(g(x))g'(x),$$

since, again, $F' = f$. This concludes our proof.

Substitution in Definite Integrals

Recall that a definite integral

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

is a signed area between the graph of $y = f(x)$ and the interval $[a,b]$ on the x-axis. It's important to note that the limits of integration a and b are x-values. So, if we make a substitution $u = \ldots$ into our integral, this substitution affects these limits as well. (Think of the substitution $u = g(x)$ as transforming the interval $[a,b]$ into the interval $[g(a),g(b)]$.) We need to account for this in our computations.

We do so by noting, in our margin work, the effects of our substitution on the original limits $x = a$ and $x = b$. The following examples illustrate the main ideas.

Example 5.2.3. (i) We evaluate the definite integral

$$\int_0^{\pi/2} \frac{\cos(x)\,dx}{1+\sin(x)}$$

by making the substitution $u = 1 + \sin(x)$, into both the integrand and the limits of integration:

$$\int_0^{\pi/2} \frac{\cos(x)\,dx}{1+\sin(x)} = \int_1^2 \frac{1}{u}\,du$$
$$= \ln(u)\Big|_1^2 = \ln(2) - \ln(1)$$
$$= \ln(2).$$

$u = 1 + \sin(x)$

$du = \cos(x)\,dx$

When $x = 0$, $u = 1 + \sin(0) = 1$
When $x = \pi/2$, $u = 1 + \sin(\pi/2) = 2$

(ii) In discussing the reversal rule for definite integrals (see page 224), we mentioned that, sometimes, it's useful to be able to integrate from a larger number to a smaller one (that is, to integrate "from right to left"). The following example demonstrates how such "backwards" integrals can arise through the substitution method.

$$\int_{-2}^{1} x(5+x^2)^3\,dx = \int_9^6 u^3\left(\frac{du}{2}\right)$$
$$\frac{1}{2}\int_9^6 u^3\,du = \frac{1}{2}\cdot\frac{u^4}{4}\Big|_9^6$$
$$= \frac{6^4 - 9^4}{8} = -\frac{5265}{8}.$$

$u = 5 + x^2$

$du = 2x\,dx$

$\dfrac{du}{2} = x\,dx$
When $x = -2$, $u = 5 + (-2)^2 = 9$
When $x = 1$, $u = 5 + 1^2 = 6$

As mentioned earlier, The Fundamental Theorem of Calculus applies regardless of which limit of integration is the larger of the two. So we were able apply this theorem to the above integral from 9 to 6, without worrying especially about whether this integral denotes an area, or a signed area, or a "signed area in reverse."

Exercises

Part 1: Indefinite integrals

1. Evaluate the following indefinite integrals using substitution. Make sure to indicate your substitution clearly. (That is: what are you calling u? What is du?)

(a) $\displaystyle\int 2y(y^2+1)^{50}\,dy$

(b) $\displaystyle\int \sin(5z)\,dz$

(c) $\displaystyle\int \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx$

(d) $\displaystyle\int (5t+7)^{50}\,dt$

(e) $\displaystyle\int 3u^2\sqrt[3]{u^3+8}\,du$

(f) $\displaystyle\int v^2(3+v^3)^4\,dv$

(g) $\displaystyle\int \tan(x)\,dx$

(h) $\displaystyle\int \tan^2(x)\sec^2(x)\,dx$

(i) $\displaystyle\int \sec(x/2)\tan(x/2)\,dx$

(j) $\displaystyle\int \sin(w)\sqrt{\cos(w)}\,dw$

(k) $\displaystyle\int \frac{\sin(\sqrt{s})}{\sqrt{s}}\,ds$

(l) $\displaystyle\int \sqrt{3-x}\,dx$

(m) $\displaystyle\int \frac{dr}{r\ln r}$

(n) $\displaystyle\int e^x\sin(1+3e^x)\,dx$

(o) $\displaystyle\int \frac{y}{1+y^2}\,dy$

(p) $\displaystyle\int \frac{w}{\sqrt{1-w^2}}\,dw$

(q) $\displaystyle\int \frac{1}{1+4y^2}\,dy$

(r) $\displaystyle\int \frac{e^w}{\sqrt{1+e^w}}\,dw$

Part 2: Definite integrals

2. Use integration by substitution, together with The Fundamental Theorem of Calculus, to evaluate each of the following definite integrals. Express your answer to four decimal places. You can check your results using RIEMANN.sws if you would like. But please also show all the work required for the substitution method. (What is u? What is du? What do your limits of integration become, under your substitution? How are you using The Fundamental Theorem of Calculus?)

.

(a) $\displaystyle\int_0^1 \frac{3s^2}{s^3+1}\,ds$

(b) $\displaystyle\int_2^4 \frac{1}{x(\ln(x))^2}\,dx$

(c) $\displaystyle\int_1^3 \frac{1}{2x+1}\,dx$ (Hint: $u=2x+1$)

(d) $\displaystyle\int_0^1 \frac{\arctan(x)}{1+x^2}\,dx$

(e) $\displaystyle\int_0^1 \frac{t}{\sqrt{1+t^2}}\,dt$

(f) $\displaystyle\int_0^1 \frac{\sin(\pi\sqrt{t})}{\sqrt{t}}\,dt$

(g) $\displaystyle\int_0^2 \frac{1}{1+(x^2/4)}\,dx$

(h) $\displaystyle\int_0^{\pi/2} \sin^3(x)\cos(x)\,dx$

Part 3: Initial value problems

3. (a) Find all functions $y=F(x)$ that satisfy the differential equation

$$\frac{dy}{dx}=x^2\left(1+x^3\right)^{13}.$$

(b) From among the functions $F(x)$ you found in part (a), select the one that satisfies $F(0) = 4$.

(c) From among the functions $F(x)$ you found in part (a), select the one that satisfies $F(-1) = 4$.

4. Find a function $y = G(t)$ that solves the initial value problem

$$\frac{dy}{dt} = te^{-t^2} \qquad y(0) = 3.$$

Part 4: Miscellaneous

5. This question concerns the indefinite integral $I = \displaystyle\int \sin(x)\,\cos(x)\,dx$.

(a) Find I by using the substitution $u = \sin(x)$.

(b) Find I by using the substitution $u = \cos(x)$.

(c) Compare your answers to (a) and (b). Are they the same? If not, how do they differ? Since both answers are antiderivatives of $\sin(x)\,\cos(x)$, they should differ only by a constant. Is that true here? If so, what is the constant?

(d) Now calculate the value of the *definite* integral

$$\int_0^{\pi/2} \sin(x)\,\cos(x)\,dx$$

twice, using the two *indefinite* integrals you found in (a) and (b). Do the two values agree, or disagree? Is your result consistent with what you expect?

6. (a) What is the average value of the function $f(x) = x/\sqrt{1+x^2}$ on the interval $[0, 2]$?

(b) Show that the average value of the $f(x)$ on the interval $[-2, 2]$ is 0. Sketch a graph of $y = f(x)$ on this interval, and explain how the graph also shows that the average is 0.

7. (a) Sketch the graph of the function $y = xe^{-x^2}$ on the interval $[0, 5]$.

(b) Find the area between the graph of $y = xe^{-x^2}$ and the x-axis for $0 \le x \le 5$.

(c) Find the area between the graph of $y = xe^{-x^2}$ and the x-axis for $0 \le x \le b$. Express your answer in terms of the quantity b, and denote it $A(b)$. Is $A(5)$ the same number you found in part(b)? What are the values of $A(10)$, $A(100)$, $A(1000)$?

(d) It is possible to argue that the area between the graph of $y = xe^{-x^2}$ and the *entire* positive x-axis is $1/2$. Can you develop such an argument?

8. (a) Use a computer graphing utility to establish that

$$\sin^2(x) = \frac{1 - \cos(2x)}{2}.$$

Sketch these graphs.

(b) Find a formula for $\int \sin^2(x)\,dx$. (Suggestion: replace $\sin^2(x)$ by the expression involving $\cos(2x)$, above, and integrate by substitution.)

(c) What is the average value of $\sin^2(x)$ on the interval $[0, \pi]$? What is its average value on any interval of the form $[0, k\pi]$, where k is a whole number?

(d) Explain your results in part (c) in terms of the graph of $\sin^2(x)$ you drew in part (a).

(e) Here's a differential equations proof of the identity in part (a). Let $f(x) = \sin^2(x)$, and let $g(x) = (1 - \cos(2x))/2$. Show that both of these functions satisfy the initial value problem

$$y'' = 2 - 4y \quad \text{with} \quad y(0) = 0 \quad \text{and} \quad y'(0) = 0.$$

Hence conclude the two functions must be the same.